Modelling Forest Growth and Yield

Applications to Mixed Tropical Forests

Jerome K. Vanclay

Professor of Tropical Forestry

Royal Veterinary and Agricultural University

Copenhagen

Currently: Professor of Sustainable Forestry,
Southern Cross University, Lismore NSW Australia

Originally published in 1994 by CAB International, Wallingford UK as
ISBN 0 85198 913 6

Note: for reasons best known to the WordPerfect people, this version has pagination similar, but not identical to the CABI edition.

Dedicated to all those working toward
the sustainable management of tropical forests,

and especially to my daughter Erica,
in the hope that she too may find the tropical forests
as interesting, diverse and extensive as I have.

Contents

Figures viii
Tables xi
Preface xiii
Foreword xv
Acknowledgements xvii

1. Introduction 1
  Role of Growth Models 2
  What is a Growth Model? 4
  Growth Modelling Approaches 5
  Components of a Model 8
  Choosing a Model 9
  Examples 10
  Exercises 13

2. Whole Stand Models 14
  Growth and Yield Tables 14
  Growth and Yield Equations 16
  More Detailed Whole Stand Approaches 22
  Synthesis 31
  Exercises 32

3. Size Class Models 34
  Stand Table Approaches 34
  Transition Matrices 43
  Cohort Models 50
  Synthesis 55
  Exercises 55

4. Single-tree and Tree List Models 57
Single-tree Spatial Models 58
Single-tree Non-spatial Models 68
Tree List Models 71
Synthesis 77
Exercises 77

5. Data Requirements 79
Differing Data Needs 80
Data for Modelling 84
Data Collection 91
Synthesis 101
Exercises 102

6. Constructing Growth Models 103
Model Design 104
Regression Techniques 112
Statistical Assumptions 119
Strategies for Grouping Data 127
Synthesis 131
Exercises 132

7. Forest Site Evaluation 134
Phytocentric Methods 136
Geocentric Methods 145
Practical Considerations 151
Synthesis 154
Exercises 155

8. Diameter Increment 156
What to Model 156
Diameter Increment Functions 162
Synthesis 170
Exercises 170

9. Mortality and Merchantability 172
Natural Mortality 173
Harvesting and other Human Factors 182
Merchantability 187
Synthesis 190
Exercises 190

10. Regeneration and Recruitment 192
Recruitment Models 193
Regeneration Models 201
Synthesis 202
Exercises 203

11. Model Evaluation and Re-calibration 204
Model Criticism 205
Benchmark Tests 211
Re-calibration 219
Synthesis 221
Exercises 221

12. Implementation and Use 223
Implementing the Model 224
Resource Data for Simulation Studies 225
Optimization Studies 229
Yield Prediction 234
Synthesis 240
Exercises 241

13. Future Directions 243
Models for all Forests 244
Different Management Needs 245
Environmental Changes 247
Integration with other Models 248
Model Construction and Use 249
Presentation of Information 249
Synthesis 250

References 252

Glossary 275
Symbols 275
Common and Botanical Species Names 277
Terminology 279

Solutions to Selected Exercises 285

Author Index 298

Subject Index 304

Figures

Fig. 1.1. The role of growth models and complementary data in providing forest management information. 2
Fig. 1.2. The role of growth models in decision making, forest management and the formulation of forest policy. 3
Fig. 1.3. Yield table, curves and equation for *Eucalyptus deglupta*. 5
Fig. 1.4. Components of forest growth and the analogous representation in a stand growth model. 9

Fig. 2.1. Yield predictions for logged dipterocarp forest. 18
Fig. 2.2. Basal area increment curves for *Callitris* stands. 20
Fig. 2.3. Possible diameter distributions generated by the Weibull p.d.f. 24
Fig. 2.4. Vector field illustrating a 2-dimensional state-space model. 28
Fig. 2.5. Markov chain representation of a model. 29
Fig. 2.6. Markov matrix and its steady state. 30

Fig. 3.1. Stand table projection with movement ratio 0.25. 36
Fig. 3.2. Smoothing an uneven-aged stand table results in smaller estimates of upgrowth for a given movement ratio. 39
Fig. 3.3. Markov chain representation of Fig. 3.1 and Equation 3.1. 44
Fig. 3.4. Single projection with a matrix model. 49
Fig. 3.5. Changes in tree rank in a pine stand. 50
Fig. 3.6. Tree record in cohort model represented by species, size and expansion factor. 52

Fig. 4.1. Competition indices include the competitive influence zone, area potentially available, size-distance, sky view and light interception approaches. 59
Fig. 4.2. Outputs from the PLATIPUS prototype included stem and branch profiles, and ring and knot patterns. 65
Fig. 4.3. Tree records representing a forest stand. 72

Fig. 4.4. Record tripling in the Prognosis model. 73
Fig. 4.5. Look-up tables can be used to find the coefficients to be used in equations to predict growth, mortality, recruitment, etc. 75

Fig. 5.1. Interpolation is safer than extrapolation. 82
Fig. 5.2. Efficient placement of ten samples. 84
Fig. 5.3. Database weaknesses revealed by comparing dynamic and static inventory data. 87
Fig. 5.4. Recommended plot layout for permanent sample plots. 91
Fig. 5.5. Example of a field form for measuring PSPs. 98

Fig. 6.1. Concepts of precision and bias. 105
Fig. 6.2. Growth and yield forms of equations are analogous. 107
Fig. 6.3. Forms of the Bertalanffy equation. 111
Fig. 6.4. Dangers of extrapolating empirical models. 113
Fig. 6.5. Graphs for screening data and models. 114
Fig. 6.6. Interpreting residual plots. 115
Fig. 6.7. Principal component and least squares fit to data. 116
Fig. 6.8. Serial correlation may be minimized by modelling growth as a function of initial size. 120
Fig. 6.9. The R^2 does not indicate how well a model fits the data. 126
Fig. 6.10. Different species groupings based on growth trends. 129

Fig. 7.1. Site index of western larch estimated from the height–diameter relationship. 139
Fig. 7.2. Height–diameter curves used to estimate site productivity in *Callitris* stands in south-eastern Queensland. 140
Fig. 7.3. Growth index equation for *Cardwellia sublimis*. 144

Fig. 8.1. Periodic increment may be a better predictor of growth at the mean diameter and competition than at initial conditions. 159
Fig. 8.2. Comparison of three "theoretical" diameter increment functions. 165
Fig. 8.3. Comparison of Bertalanffy, Vanclay and Wykoff equations for predicting diameter increment. 167
Fig. 8.4. Diameter increment probabilities in the NORM model. 169

Fig. 9.1. The concept of limiting stand density. 174
Fig. 9.2. Mortality in relation to growth rate. 177
Fig. 9.3. The logistic is better for modelling mortality than linear equations. 179
Fig. 9.4. Predicted survival patterns species in the NORM model. 180
Fig. 9.5. Probability of harvesting a *Flindersia pimenteliana* tree. 184
Fig. 9.6. Silvicultural treatment may stimulate tree growth. 185

Fig. 9.7. Probability of logging damage in Queensland rainforest. 186
Fig. 9.8. Proportion of stems remaining merchantable. 188
Fig. 9.9. Probability that a tree does not contain at least one log. . 189

Fig. 10.1. Modelling of regeneration may commence at any of several stages. 193
Fig. 10.2. Recruitment at 20 cm dbh predicted from site quality, stand composition and basal area. 195
Fig. 10.3. Recruitment estimated for different threshold diameters. 197
Fig. 10.4. Recruitment predicted from the NORM model. 199

Fig. 11.1. Matrix plot of model defined by Equations 2.13 and 2.14. 208
Fig. 11.2. Observed and predicted values plotted directly and as differences. 212
Fig. 11.3. Residuals vs initial values, length of, and age after simulation. 213
Fig. 11.4. Example of data partitioned using the duplex algorithm. 217

Fig. 12.1. Harvesting history in Queensland rainforests and forecasts based on cutting cycle analysis. 237

Tables

Table 1.1. Selected models to illustrate different scales and purposes. 7
Table 1.2. Predictions by the NORM model of future timber harvests. 12

Table 2.1. Coefficients for Weibull distributions in Fig. 2.3. 25

Table 3.1. Three approaches to Stand Table Projection. 37
Table 3.2. Characteristics of size classes in the FORMIX model. . 41
Table 3.3. Example of a tree list model showing record doubling and merging. 53

Table 4.1. Example of record tripling in the Prognosis model. . . . 73

Table 5.1. Different applications require different sampling techniques. 81

Table 7.1. Views and methods of forest site evaluation. 135
Table 7.2. Correlation coefficients between site form and subjective site quality classes in *Callitris* forest. 141

Table 12.1. Key to some alternative sampling designs showing selected selection criteria and some possible consequences. 227
Table 12.2. Diameter at which tree volume growth is maximized. 231

Table S.1. Stand table projection with different assumptions. . . . 288
Table S.2. 293

Preface

This book attempts to make growth models more accessible to foresters and others interested in mixed forests, whether planted or natural. There is an increasing interest in, and controversy surrounding the use of mixed plantations and natural forests, and rational discussion and resolution of management options require reliable growth models linked to other information systems. It is my hope that this book will help researchers to build better models, and will help users to understand how the models work and thus to appreciate their strengths and weaknesses.

During recent years, vast areas of natural forest, especially in the tropics, have been logged or converted to other uses. Well-meaning forest managers have often been over-optimistic in estimating forest growth and yields, and this has contributed to over-cutting in some forests. Growth models can provide objective forecasts, offering forest managers the information needed to maintain harvests within the sustainable capacity of the forest, and providing quantitative data for land use planners to make informed decisions on land use alternatives. In this way, I hope that this book will contribute to the conservation and sustainable management of natural forests in the tropics and elsewhere.

This is not a "How to do it" manual with step-by-step instructions to build a growth model for mixed forests. Unfortunately, modelling these forests isn't that easy. There is no single "best" way to build a model for these forests. Rather, many approaches can be used, and the best one depends on the data available, the time and expertise available to build the model, the computing resources, and the inferences that are to be drawn from the model.

So instead of writing a "cookbook" with one or two recipes, I review and illustrate some of the many approaches available, indicate the requirements of and output from each, and highlight their strengths and limitations. The book emphasizes empirical-statistical models rather than physiological-process type models, not because they are superior, but because they have proven utility and offer immediate benefits for forest management. A more comprehensive treatment of all the options is beyond the scope of this book, which is intended to serve as a ready reference manual for those building growth models for forest management. Because of my limited linguistic ability, the material covered is more-or-less restricted to English-language material. I have not attempted to review *all* the published work on growth modelling (it would be a huge task), but have tried to highlight examples that may be applicable to mixed forests in tropical areas.

I hope that the language and terminology used in this book will be accessible to all readers, especially those for whom English is a second language. The glossary may help to clarify some terms, and those that have a specific technical meaning are printed in *italics* the first time they are used. Readers should consult the glossary to clarify the meaning of these words unless they are sure of the meaning.

Exercises are given at the end of each chapter to reinforce points made in the chapter. These are simple exercises, deliberately chosen so that they can be completed quickly with pen and paper or PC and spreadsheet, but within these constraints, I have tried to keep them realistic. Some exercises (e.g. 9.1 and 10.3) require more specialized statistical analyses, but many commercial statistical packages (e.g. GLIM) are suitable. Where possible, these exercises draw on real data, but some data were simulated to create interesting exercises with few data.

Whilst my approach places more responsibility on the reader to choose and develop a suitable modelling methodology, I hope it will help readers gain a better understanding of modelling, which should in turn lead to better models and more reliable predictions. And I hope that better models will provide better information, greater understanding, and better management of mixed forests.

Jerry Vanclay
Copenhagen, 1994

Foreword

Forest managers, planners and policy makers forecast the outcomes of different types of forest use in order to make wise decisions for balanced management. Careful selection and analysis of data will always be vital to the decision-making process, but efficient and readily understandable models of growth and yield now promise to become invaluable tools. I am delighted to say this book provides the necessary access to such models. Also in a world where increasingly complex computer models are the order of the day, it is heartening to find a text which provides both a robust framework for further model development as well as a direct tool for those managers and planners with a limited modelling background. Unlike many texts that focus on esoteric aspects of modelling processes, the primary emphasis here is on robust, empirical models that can be fitted to the kinds of data that exist within many forest services.

It is widely accepted that, while growth and yield models are fundamental to management, the global community now demands more sustainable use and a more comprehensive understanding of forests and forest products. A key element of this book is the inclusion of growth and yield models for mixed-species stands. This is a notable departure from the more traditional focus on monocultural systems and opens the way to modelling the performance of species assemblages in natural or semi-natural forested lands. The demands for concepts and algorithms which can deal with complex biophysical interactions will increase with the widening perception of the forest as a resource. In each of these respects, the conceptual approaches in this text lend themselves to wider applications which may include, for example, physical environmental determination of biological diversity as well as cellulose production.

It is eminently clear that forest managers and planners will rely increasingly on access to simplified models of forest dynamics which can be used to generate options for improved resource management. This book will provide the entry point.

Forestry research has been characterized as being reductionist. In particular, forest science has been slow to respond to rapidly changing perceptions of the value of forests and to demand for changed products and markets. Scientists have acquired sophisticated knowledge of components of forest systems but have failed to integrate and synthesize information in a way that can provide adequate decision-making tools. The models outlined in this book are a potentially valuable tool to help remedy this situation.

Jeff Sayer
Director General,
Center for International Forestry Research
Bogor, Indonesia

Acknowledgements

This book was researched and written over a period of several years, and many people and several institutions have supported this work in various ways. I am grateful for the generous support of the DLH-Foundation (Denmark), the Oxford Forestry Institute and the Queensland Forest Service. I am grateful to Philip Adlard, John Horneman, Peter Kanowski (Snr), Svend Korsgaard and Dick Pegg for helping to arrange this support. Thanks also to the helpful librarians at the OFI, QFS and Royal Veterinary and Agricultural University, and to Tim Hardwick and Pippa Smart at CAB International.

Much of my knowledge of growth modelling was gained by participating in working parties and meetings of the International Union of Forest Research Organizations (IUFRO), and I thank fellow office-bearers and members for making this possible and for sharing their knowledge.

Many colleagues provided helpful comment on parts of the draft manuscript, and I thank them all for their efforts. Thomas Dean provided the data used to prepare Figure 9.1. Special thanks go to Tom Burk, Steve Candy, Oscar García, Bob Haight, David Hamilton, Jerry Leech, Brian McCormack, Ralph Meldahl, Bob Monserud, Mark Nester, Keith Rennolls, Steve Shifley, Phil West, Geoff Wood, Bill Wykoff and Boris Zeide for their helpful comments. Two anonymous referees also provided helpful comment. I owe a special debt of gratitude to my sternest critic, J.P. Skovsgaard, who read the entire manuscript and took me to task over many inaccuracies and ambiguities. I benefitted especially from our discussions on self-thinning and model evaluation. Any remaining errors and omissions are, of course, my own.

Last but not least, my sincere thanks to Carol and Erica for their patience and understanding during the difficult months of manuscript preparation.

Jerry Vanclay
Copenhagen 1994

Chapter One

Introduction

Growth models assist forest researchers and managers in many ways. Some important uses include the ability to predict future yields and to explore silvicultural options. Models provide an efficient way to prepare resource forecasts, but a more important role may be their ability to explore management options and silvicultural alternatives. For example, foresters may wish to know the long-term effect on both the forest and on future harvests, of a particular silvicultural decision, such as changing the cutting limits for harvesting. With a growth model, they can examine the likely outcomes, both with the intended and alternative cutting limits, and can make their decision objectively. The process of developing a growth model may also offer interesting new insights into stand dynamics.

There is an extensive literature on growth modelling of pure even-aged forest stands, especially of plantations. However, these ecosystems are in many respects rather simplistic, and many of these modelling approaches do not apply in forest *stands* with trees of many ages or many species. Tropical moist forests pose a special problem, as there may be hundreds of species and a great diversity of tree sizes and growth patterns. This complexity means that many techniques ideal for plantation modelling are not well suited to modelling mixed forests, especially in the tropics.

Tropical moist forests (especially rainforests) are singled out for specific attention in this book, but material is not confined to the tropics or to moist forests. The book focuses on the special problems of modelling forests comprising many species, a wide range of trees sizes, and indeterminate ages. These problems are common to most tropical forests, and to many natural and semi-natural forests elsewhere.

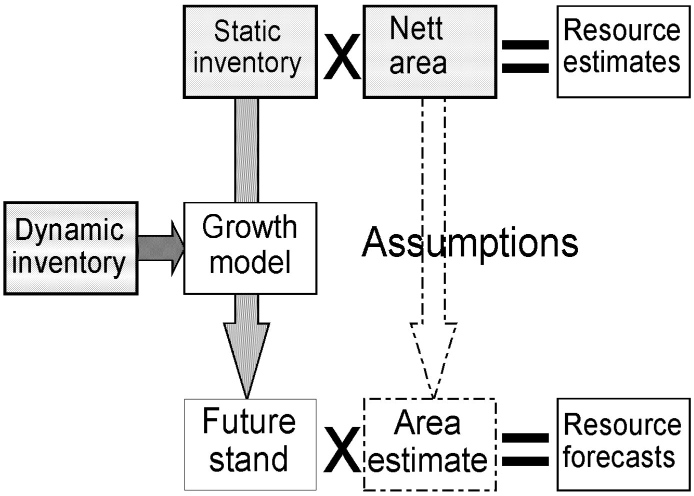


Fig. 1.1. The role of growth models and complementary data in providing forest management information.

Role of Growth Models

Growth models are of limited use on their own, and require ancillary data to provide useful information. With suitable inventory and other resource data, growth models provide a reliable way to examine silvicultural and harvesting options, to determine the sustainable timber yield, and examine the impacts of forest management and harvesting on other values of the forest.

Forest managers may require information on the present status of the resource (e.g. numbers of trees by species and sizes for selected strata), forecasts of the nature and timing of future harvests, and estimates of the maximum sustainable harvest. This information can be compiled from three sources:

1. area estimates of the existing forest,
2. stand level inventory of the present forest, and
3. growth and harvesting models based on dynamic inventory data.

The role of, and relationship between these three primary data sources is illustrated in Fig. 1.1. Static inventory ("snapshot" or point-in-time inventory using temporary plots) and area estimates will not be addressed here — readers should refer to standard inventory texts (e.g. Loetsch *et al*.

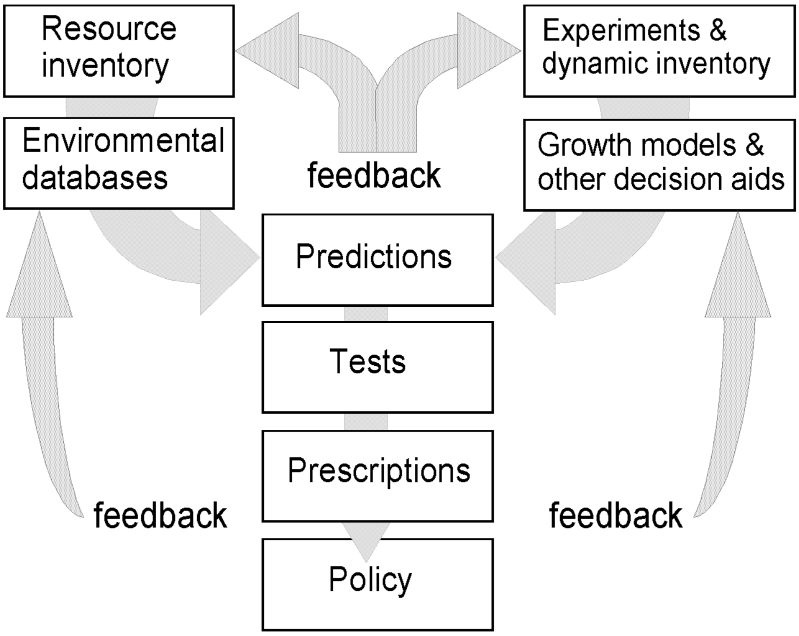


Fig. 1.2. The role of growth models in decision making, forest management and the formulation of forest policy (after Nix and Gillison 1985).

1973, de Vries 1986, Schreuder *et al*. 1993). A growth model is a synthesis of dynamic inventory data indicating growth and change in the forest. These data may be obtained from permanent plots. The construction of models from such data is the main subject of this book.

Growth models may also have a broader role in forest management and in the formulation of forest policy. Used to advantage and in conjunction with other resource and environmental data, growth models can be used to make predictions, formulate prescriptions and guide forest policy. Figure 1.2 is a simplified representation of this process. It shows that growth models are but one step in the formulation of forest policy and management prescriptions, and that supplementary data and adequate testing are also required. Some important feedbacks are indicated, but in practice, there should be many more arrows indicating interactions between the various boxes. The feedback loops are especially important. In the context of growth modelling, there should be sufficient feedback to ensure that inventory is adequate and model predictions are reliable across the range of resource conditions and management prescriptions entertained. These various roles draw on different qualities in growth models and provide a place for different kinds of models.

What is a Growth Model?

A model is an abstraction, or a simplified representation, of some aspect of reality (and should not be confused with the normative meaning of the word, something worthy of being imitated). We frequently use models unconsciously, e.g. making mental models to visualize cause-effect relationships to help explain and anticipate the behaviour of systems. Models may be stated in verbal (e.g. a description) or material forms (e.g. a scale model). A mathematical model is like a verbal model, but uses mathematical language which is more concise and less ambiguous than natural language. Computers have become indispensable as tools to assist modelling, but are not central to the process of modelling. Modelling is about making a good representation, and the computer is merely a convenient way to realize it. García (1994) likened "computer modelling" to "typewriter poetry".

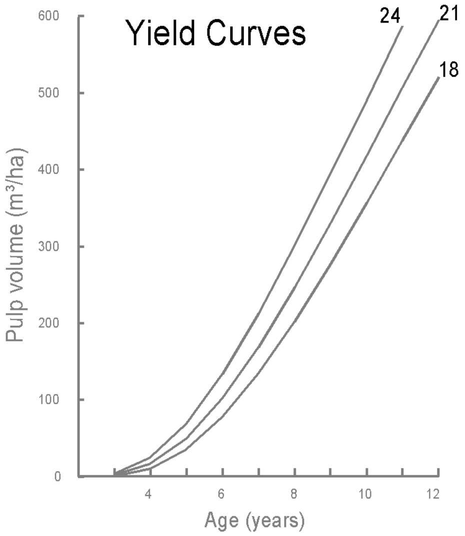
A *stand growth model* is an abstraction of the natural dynamics of a forest stand, and may encompass growth, mortality, and other changes in stand composition and structure. Common usage of the term "growth model" generally refers to a system of equations which can predict the growth and yield of a forest stand under a wide variety of conditions. Thus a growth model may comprise a series of mathematical equations, the numerical values embedded in those equations, the logic necessary to link these equations in a meaningful way, and the computer code required to implement the model on a computer. In its broadest sense, the term may also embrace yield tables and curves, which are analogous to equations, but which have been stated in a tabular or graphical form, rather than a mathematical form.

Growth refers to the increase in dimensions of one or more individuals in a forest stand over a given period of time (e.g. volume growth in m^3 $ha^{-1}y^{-1}$). *Yield* refers to their final dimensions at the end of a certain period (e.g. volume in m^3 ha^{-1}). In even-aged stands, a growth equation might predict the growth of diameter, basal area or volume in units per annum as a function of age and other stand characteristics, whereas a yield equation would predict the diameter, stand basal area or total volume production attained at a specified age. In an uneven-aged stand, yield is the total production over a given time period, while growth is the rate of production. Growth and yield are related mathematically (if yield is y, growth is the derivative dy/dt), and little distinction will be made in this book between these analogous forms (see Chapter 6).

An important but simple model in forestry is the plantation yield table, which may comprise only two columns of figures, the left column showing plantation age, and the right column indicating the expected standing volume at that age. Additional columns may indicate production for different sites (Fig. 1.3). The yield table may also be expressed graphically as a series of curves, with the horizontal axis indicating age and the vertical axis indicating volume produced. It may also be expressed more concisely as a mathematical equation.

Yield Table

Age (y)	Site Index (m) 18	21	24
3	1	2	4
4	11	17	25
5	36	50	69
6	78	102	133
7	134	169	212
8	201	246	301
9	276	330	395
10	356	418	491
11	438	507	586
12	520	594	679



Yield Equation: $\log(V+1) = 3.534 - 14.02/t + 0.2314\ S/t$

Fig. 1.3. Pulpwood yield table, curves and equation for *Eucalyptus deglupta* in the Philippines (after Tomboc 1977).

At the other extreme are complex plantation growth models which may model the spatial positions of individual trees, their diameter, height and crown size. Such models may indicate timber quality and knot size, and may be linked to conversion simulators which predict veneer and sawn timber outturn. Similarly, growth models for mixed forests may vary from the simple to the complex.

Growth Modelling Approaches

There are so many growth models in existence that it is impossible to examine the methodology used in each. Thus it is necessary to identify some commonality, and to consider just a few examples for each class of model. It is useful to classify models on the level of detail they provide. A model may be considered a whole stand model, a size class model, or a single-tree model, depending on the detail required, provided and utilized by the model.

Whole stand models are often simple and robust, but may involve complexities not possible in other approaches. Population parameters such as stocking (number of trees per unit area), stand basal area and standing volume are used to predict the growth or yield of the forest. No details of the individual trees in the stand are determined. Stem size distributions may be inferred from existing or predicted distributions.

Size class models provide some information regarding the structure of the stand. Several techniques are available to model stand structure, but one of the most widely used is the method of stand table projection which essentially produces a histogram of stem diameters. This approach is a compromise between whole stand models and single-tree models. When the class size is infinitely large and only one class exists, then the method is a whole stand approach. When the class width is infinitely small and each tree is considered a single class, then the method is the single-tree approach.

The most detailed approach is that of *single-tree models* which use the individual tree as the basic unit of modelling. The minimum input required is a list specifying the size of every tree in the stand. Some models also require the spatial position of the tree, or tree height and crown class. Single-tree models may be very complex, modelling branches and internal stem characteristics, and may be linked to harvesting and conversion simulators (e.g. Mitchell 1988, Vanclay 1988*b*).

These three classes of model will be used throughout this book, but they are only for convenience, and in practice, models form a continuum from one extreme to another and some do not fit comfortably in any of these classes.

Two other classes of models are beyond the scope of this book, but warrant a mention. *Process models* attempt to model the processes of growth, taking as input the light, temperature and soil nutrient levels, and modelling photosynthesis, respiration and the allocation of photosynthates to roots, stems and leaves (e.g. Landsberg 1986, Mäkelä 1992). These are also known as mechanistic of physiological models. These models help to provide a better understanding of growth and stand dynamics, but have not yet successfully been used for predicting timber yields for forest management. Ecosystem *succession models* (e.g. Shugart 1984, Botkin 1993) attempt to model species succession, but are generally unable to provide reliable information on timber yields. The FORCYTE model (Kimmins 1988) uses a hybrid approach and can predict forest yields undera variety of nutrient regimes. These and other approaches are contrasted in Table 1.1.

It is useful to distinguish between *models for understanding* and *models for prediction* (Bunnell 1989). Models for understanding (e.g. process models) are useful to comprehend and link previously isolated bits of knowledge and may help to identify gaps where more work is needed. The benefits come from the insights gained while developing and exploring the model, and future uses (if any), are less important. Conversely, models for prediction may sacrifice specific details of growth processes to achieve greater efficiency and accuracy in providing information for forest management. Realism is not necessarily a virtue in a model, and it may be better to abstract just those aspects that are most relevant in each instance. This book is concerned mainly with models for prediction, and their use in forest management.

Table 1.1. Selected models to illustrate different scales and purposes.

Use	Resolution	Driving variables	Example
Empirical models			
Atmospheric studies	Global primary production	Evapo-transpiration	Lieth & Box (1972)
National forest planning	Stand variables	Age, stand basal area	Clutter (1963)
Regional planning	Individual trees	Tree species & sizes	Prognosis (Stage 1973)
Silvicultural studies	Tree crowns	Tree & branch variables	TASS (Mitchell 1975)
Silvicultural & conversion studies	Wood characteristics	Branches, ring width & density	SYLVER (Mitchell 1988)
Succession & Process models			
Ecological studies	Individual trees	Tree species & sizes	JABOWA (Botkin 1993)
Nutrient cycling	Individual trees	Trees, nutrients	FORCYTE (Kimmins 1988)
Physiological studies	Mass of foliage, branches, roots	Biomass, photosynthesis, respiration	Sievänen *et al.* (1988)

Finally, irrespective of its detail, a model may be deterministic or stochastic. A *deterministic* growth model gives an estimate of the expected growth of a forest stand, in the same way that the mean indicates the expected trend for a population. Given the same initial conditions, a deterministic model will always predict the same result. However, because of natural variation in the environment, real forest stands may not grow exactly the same amount each year, but may grow more or less than the expected amount.

A *stochastic* model attempts to illustrate this natural variation by providing different predictions, each with a specific probability of occurrence. Any one of these estimates may correspond exactly to the growth under some circumstances, but may differ from the expected growth. A single estimate from a stochastic model is of little use, as a whole series of estimates is necessary to provide useful information of the variability of predictions. For instance, twenty estimates from a stochastic model will not only give a good indication of the expected growth (the mean or median), but also of its variability. The term "stochastic" is used in preference to the term "probabilistic", which is sometimes used in other contexts.

Deterministic and stochastic models serve complementary purposes. Deterministic models are effective for determining the expected yield, and may be used to indicate the optimum stand condition. Stochastic models may indicate the reliability of these predictions, and the risks associated with any particular regime. Other methods such as variance propagation techniques (e.g. Gertner 1987*a*) may provide similar information more efficiently than stochastic models. Both deterministic and stochastic predictions can be obtained from some models. Although stochastic models can provide some useful information not available from deterministic models, most of the information needs for forest planners and managers can be provided more efficiently with deterministic models.

Components of a Model

The more detailed approaches of forest stand modelling are not based on the overall growth of a forest stand, but need to discriminate several growth components in order to model these processes effectively. The nature of the components distinguished depends upon the forest type and the approach used. In mixed forests, an obvious requirement is to discriminate individual species or several species groups. In models for intensively managed plantations, mortality and recruitment may frequently be ignored. However in many natural forests, these form an important aspect of the stand dynamics, and may have considerable influence on volume yield of the stand. In addition, the components identified in whole stand models tend to differ from those of single-tree models.

In size class and single-tree models, the components usually identified are diameter (or basal area) *increment*, *mortality* and *recruitment* (Fig. 1.4). Diameter increment is a simple concept and is relatively easy to measure and to predict. Forecasts of mortality must not only estimate the number of trees, but also the species and sizes of trees dying. Another aspect of change to be modelled is the deterioration of merchantable stems, which can be modelled in the same way as mortality. Recruitment may be predicted as *ingrowth* of trees reaching breast height, but some models may simulate seedlings from germination, while other models may adopt a larger threshold size such as 10 cm diameter. Each of these components will be addressed individually in subsequent chapters.

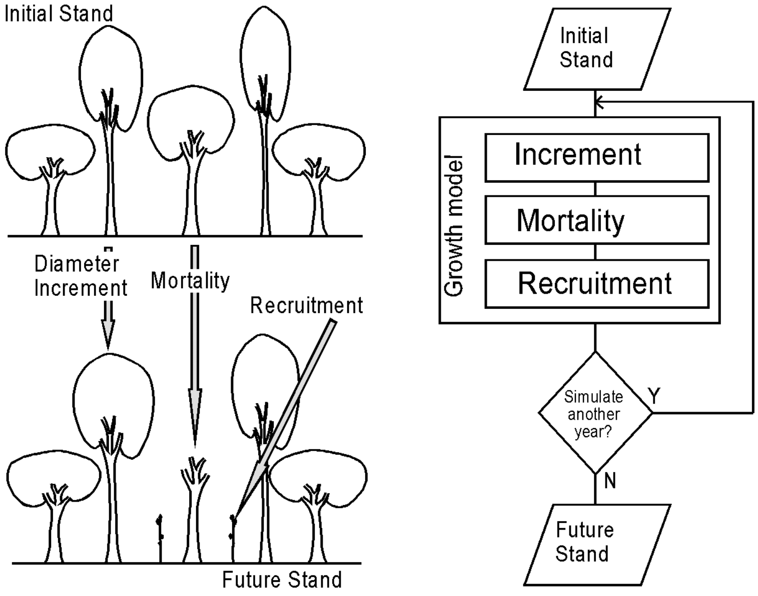


Fig. 1.4. Components of forest growth and the analogous representation in a stand growth model.

Choosing a Model

The preceding sections have hinted at several modelling approaches, and subsequent chapters will demonstrate the diversity of modelling techniques that have been used. It is appropriate at this stage to give a few guidelines for critical selection of an appropriate technique.

A map provides a convenient analogy to a growth model. To find your way with the least inconvenience, you must choose the map with the right amount of detail for your particular purpose. There is a place for national route planning maps, for topographic maps, and for urban street directories, and a map ideal for one application may prove useless in another situation. Every model is an abstraction of reality and will be wrong in some sense. Users should remember that all models may be wrong, and that some may be more useful than others.

The right model to choose is the one that is most useful for your application, and the choice should be based on your application and on your resources. It is the end use that finally determines the best approach for modelling forest stands. Some applications which may require growth models include site evaluation, testing hypotheses of growth, estimating expected yields, examining variability of yield, exploring silvicultural options, establishing optimal management regimes, investigating effects of constraints on management, and evaluating the timber quality (cf. Figs 1.1 and 1.2). These and other applications each may require a different approach to growth modelling. No single model can serve all these needs, and all demand different qualities of a model, but some general guidelines can be given.

A model should provide information that is sufficiently accurate and detailed to suit the intended purpose. Flexibility to accommodate a range of stand conditions and management options may be an important quality. It should be possible to provide estimates for a range of time horizons and to indicate species and size class combinations. Optimization studies may require that functions are continuous in the first derivative of key response variables (e.g. the diameter increment function should be smooth, not stepped) and provide reasonable extrapolations. Models which are unnecessarily complicated may incur several costs: greater computational costs, loss in precision of estimates, and difficulty in understanding and assessing the utility of the model. Intending users should ask:

1. Does the approach make sense?
2. Will the model work for my application and input data?
3. What range of data was used to develop the model?
4. Do model assumptions and inferences apply to my situation?
5. What confidence can I place in model predictions?
6. Be sceptical and demand proof!

Buchman and Shifley (1983) offered a more detailed checklist to help potential users assess the suitability of a growth model for a stated application; key items include ease of use, accuracy of predictions and biological realism.

Both simple and sophisticated models have their place. The appropriate approach depends upon the data, facilities and expertise available. What is important is what you *do* with the model. A sophisticated model is of no benefit if it is not used, while a very simple model may be of great benefit if it leads to better forest management. Thus the model should be objective, unbiased, documented and available. Users should understand how the model was derived, and should appreciate its strengths and weaknesses.

Examples

The book does not delve into specific details of models, and readers who wish to explore intimate details of models discussed here should consult the references cited. Apart from some superficial coverage of plantation models, most forest mensuration texts give few details of growth modelling, and the best source of reference may be recent conference proceedings (e.g. Ek *et al.* 1988, Pinto da Costa and Preuhsler 1994), a sign of the rapid developments occurring in this field.

Two well documented growth models that are widely used in north America are STEMS (Leary 1979, Belcher *et al.* 1982) and Prognosis (Stage 1973, Wykoff *et al.* 1982, Wykoff 1986). Good introductory and reference manuals are available freely, in English, for both of these models.

In this book, we will study the NORM (North Queensland Rainforest Management) model in some detail. This model provides a good case study because the north Queensland rainforests comprise several hundred tree species, including over one hundred of commercial importance. The original model was used to provide information for management and planning in the tropical rainforests of Queensland, Australia (e.g. Vanclay and Preston 1989), and variants now exist for all indigenous forests (managed for timber production) in Queensland. A brief overview of the use and capabilities of the NORM model is given here in the introduction, because it is important that modellers do not loose sight of their objective: for management-oriented models, this should be their use in forest management. Readers should keep this objective in mind while reviewing the possible modelling methodologies presented in the next three chapters.

The NORM model includes a growth sub-model (with equations to estimate diameter increment, mortality and recruitment) developed from permanent plot data, and a harvesting sub-model based on logging studies. These sub-models effectively summarize a mass of research and monitoring data and make it available for use by others (cf. Fig. 1.1, p. 2). The construction of these models is examined in detail in this book, but many users with only a general understanding of the models use them to examine the silvicultural alternatives. The model is generally used in conjunction with a forest resource database (comprising geographic information and static inventory data) to explore management options and examine short- and long-term simulations of future timber harvests (cf. Fig. 1.2, p. 3). Most applications rely on simulations at the management unit (or homogeneous sub-unit) level, and inventory data are required for each of these management sub-units. By repeatedly simulating the growth and harvesting of each individual plot and aggregating these results for each unit, future yields can be estimated for part of all of the forest resource.

Table 1.2 (overleaf) illustrates some of the forecasts possible from this model, including details of future harvests and an indication of the impact on the residual forest. The regional average stand basal area and merchantable volumes fluctuate and fall to a minimum (36 $m^2 ha^{-1}$ and 18 $m^3 ha^{-1}$ respectively) after about 100 years, but later stabilize at levels near the present values. This indicates that the simulated harvest does not deplete forest reserves, but may be sustainable. Several statistics concerning the simulated harvests further support the contention that the harvest is sustainable. The harvested volume declines to critical levels (13 $m^3 ha^{-1}$ is near the limit for commercial viability) after 100 years, but subsequently recovers. The average stem size of harvested trees remains relatively constant (2.7–3.0 m^3), but the proportion (by volume) contributed by trees over 100 cm diameter falls to 1%, while the bulk of the volume derives from trees 60–100 cm in diameter. The harvest continues to be

Table 1.2. Predictions by the NORM model of future timber harvests from a a Queensland rainforest, illustrating one application and some capabilities of growth models (from Vanclay and Preston 1989).

Period beginning	Regional average		Average characteristics of simulated harvest			
	Basal area m^2ha^{-1}	Merch. volume m^3ha^{-1}	Harvest volume m^3ha^{-1}	Stem size m^3	Size dist. –60–100– cm dbh	Main species† in harvest %
1990	40	24	18.9	2.9	10:75:14	Y 21, M 13
2037	38	20	18.0	2.8	10:87: 3	S 22, M 14
2077	37	25	18.8	3.0	8:90: 2	S 21, M 16
2119	36	29	17.5	3.0	8:90: 2	S 20, M 17
2166	36	23	14.7	2.8	11:86: 3	M 17, N 17
2201	37	18	13.4	2.7	13:84: 2	M 21, N 16
2231	38	19	13.3	2.7	13:85: 1	M 22, Q 21
2253	39	21	14.4	2.8	13:86: 1	Q 22, M 19
2290	40	26	17.3	2.8	14:86: 1	Q 23, M 19

† M: maple silkwood (*Flindersia pimenteliana*), N: northern silky oak (*Cardwellia sublimis*), Q: Queensland maple (*F. brayleyana*), S: silver ash (*F. bourjotiana*), Y: yellow walnut (*Beilschmiedia bancroftii*).

dominated by *Cardwellia sublimis* and *Flindersia* species, which have traditionally been an important component of the timber harvest in this area.

Table 1.2 is a condensed summary of simulation outputs, but indicates the nature and utility of information available from growth models. Such analyses require sophisticated computer systems, but useful forecasts also can be made with more simplistic approaches (e.g. Mendoza and Gumpal 1987).

Growth models offer forest managers a powerful analytical tool to investigate quickly and efficiently, the response of the forest to various management regimes. They allow foresters to determine a regime that should maximize volume or value production, or maximize the production of a particular product. It also enables them to determine the effect of a revised harvest programme to exploit a change in demand. They can investigate effects of many constraints on forest operations, and their effect on yields. But the most powerful feature is the ability of the model to assist managers to make reliable long-term forecasts, so that they can make long-term commitments to the capital intensive wood processing industry, secure in the knowledge that the forest will not be over-exploited.

In the next few chapters, we will survey the different growth modelling approaches applicable to uneven-aged mixed-species stands, and consider their merits, before we move on to examine how to construct a suitable model.

Exercises

1.1. List people or groups (at least 5) in your organization who could use a growth model. State how they might use it and how it could help them to work more effectively. What special needs would these potential users have, and how might you ask a modeller to customize a model to satisfy their needs. What impacts would this have on forest management practices (Think about improved silviculture, efficiency, cost savings)? Does your organization have a growth model; if so, does it meet these needs? Do you think these potential benefits justify the development of a growth model (or a better one if one already exists)?

1.2. Clutter (1980) suggested that many people have little basis for their choice of growth model or modelling methodology. He suggested several common fallacies that are wrongly used as selection criteria:

> 1. Among the models existing at any point in time for a particular forest type, there is one, and only one, best model and all other models are inferior to it.
> 2. Complexity is desirable for its own sake.
> 3. Model quality is directly proportional to the amount of computer time required to run the model.
> 4. Nonlinear models are inherently superior to linear models.
> 5. Models written in APL are superior to . . . models written in Fortran, which are, in turn, superior to models written in Basic.
> 6. Models which do not require a computer for their implementation are the most inferior of all and are unworthy of discussion.
> 7. Stochastic models are intellectually preferable to and aesthetically superior to deterministic models.

Discuss why these fallacies do not provide a suitable basis for model selection. Give seven criteria that do form a good basis for choosing a model.

Chapter Two

Whole Stand Models

We begin our survey of growth modelling approaches by looking at the traditional yield table, one of the oldest ways to predict forest growth and yield. We continue our survey through other whole stand approaches of increasing complexity and sophistication, and in subsequent chapters will look at size class, single-tree and other ways to model a forest stand in more detail.

Whole stand models are those growth and yield models in which the basic units of modelling are stand parameters such as basal area, stocking, stand volume and parameters characterizing the diameter distribution. They require relatively little information to simulate the growth of a stand, but consequently yield rather general information about the future stand. This chapter examines several such modelling approaches ranging from simple yield tables to sophisticated methods for predicting future diameter distributions.

Growth and Yield Tables

Yield tables are summaries of expected yields tabulated by stand age, site index, etc. (e.g. Fig. 1.3, p. 5). Growth tables are a variation more suited to uneven-aged stands, and tabulate expected growth according to various stand characteristics. Although these techniques are infrequently used in natural forests, they continue to be important in plantation management. It is appropriate to begin with a brief overview of yield tables, as they provided the foundations for forest growth modelling and demonstrate several important concepts.

Yield Tables

A yield table presents the anticipated yields from an even-aged stand at various ages, and is one of the oldest approaches to yield estimation. The concept was apparently first applied in the Chinese "Lung Ch'uan codes" some 350 years ago (Vuokila 1965). The technique as we know it today was devised in Europe in the eighteenth century. The first yield tables were published in Germany in 1787, and within a hundred years over a thousand yield tables had been published. Modern yield tables often include not only yield, but also stand height, mean diameter, number of stems, stand basal area and current and mean annual volume increments. Two classes of yield tables are distinguished, normal and variable density yield tables. In reality, these do not form discrete classes, but rather reflect a continuum from normal yield tables, through variable density yield tables and other approaches, to single-tree growth models (e.g. Leary 1991).

Normal yield tables provide estimates of expected yields tabulated by stand age and site index for ideal, fully stocked or "normal" forest stands. These were usually based on data derived from stem analyses and temporary plots, analyzed using graphical techniques. Vuokila (1965) and Spurr (1952) discussed the various approaches used in Europe and North America for the construction of yield tables. Yields were generally tabulated by age and site index, but could also be presented as alignment charts (e.g. Reineke 1927, Hamilton 1988). Normal yield tables may provide reliable estimates of potential yields for even-aged stands similar in character to those used in developing the table, but may be less satisfactory for natural stands where age may vary considerably within stands.

More sophisticated calculation and analytical techniques enabled additional variables to be included in yield calculations. Stand density was an obvious choice for a third variable as it enables data from partially stocked plots to be used and means that the yield table can be applied to any stand. Many variable density yield tables have been published, and recent examples include Edwards and Christie's (1981) tables for plantations in Britain. These tables give height, stems per hectare, mean diameter, stand basal area, mean stem volume, volume per hectare, mean annual and cumulative volume production at five year intervals for many species-site-management regime combinations.

The approach has also been applied to mixed stands, especially for selection forests in central Europe. There are several ways to build compact tables for natural forests. The basal area of the dominant species may be expressed as a percentage of total stand basal area in mixed forests (e.g. MacKinney *et al.* 1937), and a "main stand" may be identified in uneven-aged forests (e.g. Duerr and Gevorkiantz 1938).

Growth Tables and Percentages

Yield tables usually require some estimate of stand age, and thus cannot easily be applied to uneven-aged stands. Growth tables attempt to overcome this limitation by tabulating growth under various stand conditions. Variables other than age used to characterize stands for growth prediction include volume, density, height, average diameter and crown class. These variables were often used in conjunction with estimates of site productivity and time since logging.

Expected growth can be tabulated more compactly if expressed as a percentage rather than in absolute terms, and several formulae for growth percentages have been proposed. Formulae indicating percentage growth of tree diameters and stand basal areas have been used for many years to estimate growth in pure even-aged stands (e.g. Schneider 1853). Growth percentages can also be applied directly to estimates of stand volume to predict volume increment per hectare. Other formulae exist for predicting sustainable yields from normal forests.

The simplicity of these formulae is attractive, but is also a limitation. Some formulae may not account for mortality, especially if applied to individual trees. There are many alternative formulae, all with different assumptions and limitations, but the compound interest formula may be the most satisfactory for general growth predictions (e.g. Rudolf 1930). However, the method is unreliable and may be very deceptive (e.g. Wahlenberg 1941). Although some of these formulae may give reasonable estimates of short-term increment in stands for which they were derived, it is inevitable that percentages must decline in the longer term.

All these methods (including growth and yield tables) suffer inherent weaknesses including subjective bias and inability to extrapolate for long periods or to other stands. Whilst these methods may give a rough indication of growth where data and computing resources are limited, better methods are available and should be used where resources permit.

Growth and Yield Equations

Conceptually, the distinction between tables and equations is unimportant as equations can be evaluated and presented as tables. However, the converse does not hold, and equations are a concise and convenient way to express growth and yield relationships.

Growth and yield equations may be developed from data that do not identify the individual trees. Thus growth may be determined as the difference in total plot volume (or basal area) at two remeasurements, and not directly from individual tree data. Such estimates may be satisfactory where all trees are present at both measures, but where there is recruitment or mortality it may be difficult to reconcile successive remeasures. Calculations usually assume that trees do not change rank (i.e. the 2nd largest tree always remains 2nd largest, etc.). This assumption is rarely satisfactory, and it is always preferable to use data in which the individual trees are identified.

It is useful to begin with some definitions to clarify some concepts of growth (e.g. Husch *et al*. 1982). Define V_1 and V_2 as the live stand volume at measurements 1 and 2 respectively, V_m as mortality volume, V_c as harvested (cut) volume and V_i as ingrowth volume. Growth may be defined as *net growth* $\Delta V_n = V_2 - V_1 + V_c$, or as *net growth of initial volume* $\Delta V_{ni} = \Delta V_n - V_i$. If the threshold for recruitment is sufficiently small, the difference between net growth and net growth or initial volume will be negligible (i.e. $V_i \rightarrow 0$ and $\Delta V_{ni} \approx \Delta V_n$). *Gross growth* is obtained by adding mortality (V_m) to net growth (ΔV_n or ΔV_{ni}). *Survivor growth* is the growth on those trees that were alive at both measures. Gross growth may also be called *accretion*, and equals survivor growth plus the growth on trees that died between the two measures, plus recruitment. These definitions also apply to stand basal area and other stand parameters. Note that survivor growth is not the same as net growth of initial volume, because it does not deduct the volume lost through mortality.

The distinction between survivor growth and accretion is important in whole stand models because it affects the modelling of mortality. If the growth model predicts accretion then mortality must be predicted explicitly, whereas if the model predicts survivor growth then mortality has been already taken into account.

Empirical Yield Equations

MacKinney and Chaiken (1939) published one of the first yield equations fitted by linear regression. Their equation was based on the hypothesis that relative growth rate varies inversely with age ($\partial V/V \propto t^{-2}$), thus that $\ln V = \beta_0 - \beta_1 t^{-1}$. This equation contains two parameters, one (β_0) which defines the upper asymptote (i.e. volume at end of the rotation), and one (β_1) which determines the rate of growth. They assumed that these parameters were simple linear functions of site index and stand density. The equation was subsequently used in many other models. Desirable features of this model include (Clutter 1963):

> **1.** The mathematical form of the variates implies relationships which agree with biological concepts of even-aged stand development (Schumacher 1939).
> **2.** The use of ln V as the response variable rather than V will generally be more compatible with the statistical assumptions customarily made in linear regression analysis (linearity, normality, additivity and homogeneity of variance).
> **3.** The use of ln V as the response variable is a convenient way to express mathematically the interactions of the explanatory variables in their effect on V.

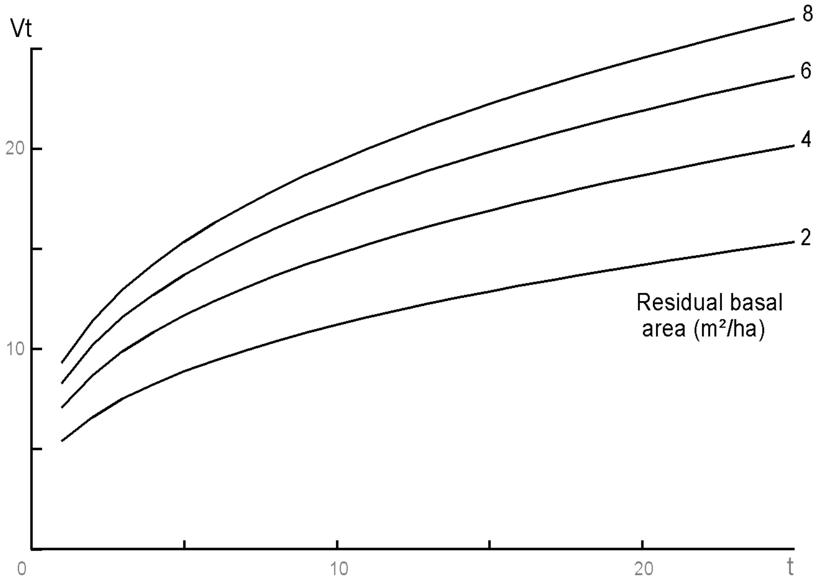


Fig. 2.1. Yield predictions for logged dipterocarp forest, with site quality 25 m (drawn from Equation 2.1).

Such equations have not been used much in mixed forests, but Mendoza and Gumpal (1987) predicted yield of dipterocarps in the Philippines with an empirical function of initial basal area, site quality and time since logging (Fig. 2.1):

$$\ln V_t = 1.34 + 0.394 \ln G_0 + 0.346 \ln t + 0.00275\, S_h\, t^{-1} \qquad (2.1)$$

where V_t is timber yield ($\mathrm{m^3\,ha^{-1}}$, 15+ cm dbh), t years after logging ($t>0$), G_0 is residual basal area ($\mathrm{m^2\ ha^{-1}}$) of dipterocarps (15+ cm dbh) after logging, and S_h is site quality (m) estimated as the average total height of residual dipterocarp trees (50–80 cm diameter). Whilst this equation cannot be extrapolated for long periods (or for $t \rightarrow 0$) or to other situations, it was useful for estimating the time to and volume yield of the next harvest, given the residual stand after the previous harvest.

Empirical Growth Equations

Yield equations have a limitation that assume a prescribed management regime throughout the period of projection. Growth equations have an advantage that harvesting or any other silvicultural treatment may be simulated at any time during the projection.

Buell (1945) estimated growth of uneven-aged mixed stands by classifying stands according to species composition and silvicultural history, and combining species into four groups according to their potential growth rate and stem form. He assumed that volume increment of a single tree could be expressed as a quadratic equation in diameter,

$$\Delta v = \beta_0 + \beta_1 d + \beta_2 d^2$$

and that the volume increment of the stand was the sum of the individual increments

$$\Delta V = \beta_0 N + \beta_1 \Sigma d + \beta_2 \Sigma d^2$$

where N is stems per hectare, and d is tree diameter.

Nelson (1963) argued that stand basal area increment (ΔG) of even-aged stands decreased asymptotically with age (t), increased with site index ($S_{h,t}$, the height at a nominated index age) and decreased as the stand basal area (G) diverged from the optimum. This led him to propose a quadratic function in basal area which also included site index and age. The statistically significant terms were:

$$\Delta G = \beta_0 + \beta_1 G\, t^{-1} + \left(\beta_2 + \beta_3 t^{-1} + \beta_4 S_{h,t}\right) G^2$$

Another equation with a similar shape was used by Vanclay (1988*a*) for uneven-aged cypress pine (*Callitris* spp.) stands in Queensland:

$$\ln \Delta G = -3.071 + 1.094 \ln G + 0.007402\, G\, S_{h,d} - 0.2258\, G \qquad (2.2)$$

where ΔG is stand basal area increment ($m^2\,ha^{-1}\,y^{-1}$), G is stand basal area ($m^2\,ha^{-1}$) and $S_{h,d}$ is site form (m), an estimate of site productivity based on the height–diameter relationship. This equation was deliberately formulated to provide sensible predictions for extreme values of stand basal area (Fig. 2.2).

Many other stand growth equations have been published, and no attempt is made to review them all. These three have been selected to illustrate some aspects involved in formulating robust equations that extrapolate safely.

Compatible Growth and Yield Equations

Some analyses fail to exploit the relationship existing between growth and yield and this may lead to instances where estimates of yield (e.g. from yield tables) differ from successive summations of growth estimates based on the same data. Buckman (1962) and Clutter (1963) are usually credited as the first to publish forest growth and yield equations explicitly formulated to give consistent estimates of growth and yield, but many researchers have been aware of this relationship for a long time (e.g. Bertalanffy 1941, 1949).

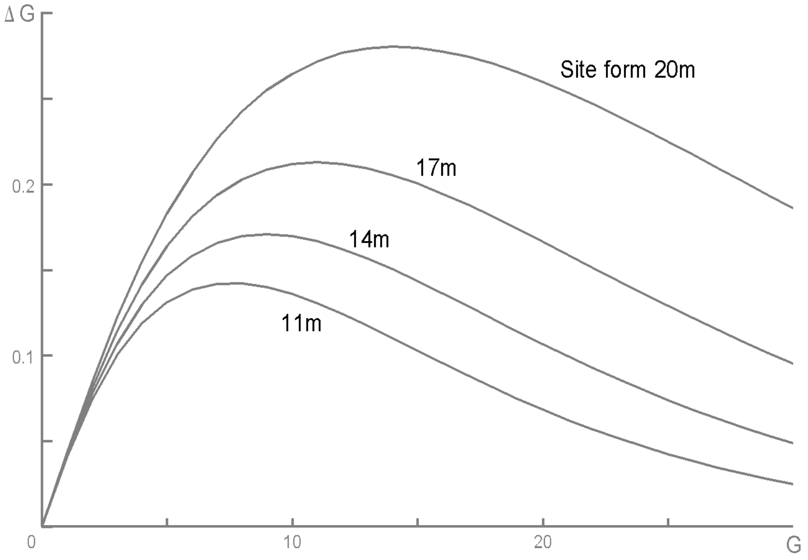


Fig. 2.2. Basal area increment curves for *Callitris* stands in south-eastern Queensland (drawn from Equation 2.2).

Clutter (1963) adapted the equation proposed by Schumacher (1939):

$$\ln V = \beta_0 + \beta_1 S_{h,t} + \beta_2 \ln G + \beta_3 t^{-1}$$

where V and G are stand volume and basal area at age t years respectively, and $S_{h,t}$ is site index. This equation can be differentiated to give the growth equation (assuming that site index is constant):

$$\frac{dV}{dt} = \beta_2 \frac{V}{G}\left(\frac{dG}{dt}\right) - \beta_3 V t^{-2}$$

Clutter expressed basal area as a differentiable yield function in age and site index, so that basal area increment (dG/dt) could be estimated from stand basal area, age and site index. This led to five compatible equations which could be used to derive estimates of present standing volume, basal area increment, volume growth, predicted future basal area and predicted volume yield. Sullivan and Clutter (1972) refined this equation to provide numerically consistent estimates of growth and yield from initial stand age, basal area, site index and future age. Burkhart and Sprinz (1984) presented a method for simultaneously estimating the parameters in the Sullivan-Clutter approach.

The Schumacher-Clutter equation has been dealt with in some detail because it is an important example illustrating some concepts of model formulation. By thinking about growth processes, formulating an hypothesis, and constructing an equation on this basis, Schumacher (1939) devised a growth and yield model that has proved to be useful, reliable and widely used for many pure even-aged stands. Fifty years of research and development resulted in better ways to estimate the parameters in the equation, but the basic equation remains unaltered, indicating the quality of the original work. This approach of "think, hypothesize and formulate accordingly" is the way that all growth models should be constructed.

Theoretical Equations

The Bertalanffy (also known as the Chapman-Richards equation, see Chapter 6) and several similar equations (see e.g. Zeide 1989, 1993) can be expressed as growth or yield equations in the same manner as the compatible growth and yield equations discussed above. However, they are sometimes thought to be superior to empirical equations as they are supposedly biologically based. The supposed biological basis is doubtful (see Chapter 6), but these equations have other merits.

Many models incorporating these equations have been published, and we examine just one study in an uneven-aged stand. Moser and Hall (1969) described stand volume as an *allometric* function of basal area ($V = \beta_0 G^{\beta_1}$), predicted basal area increment from stand basal area using the Bertalanffy equation ($\Delta G = \beta_2 G^{\beta_3} - \beta_4 G$), and thus could estimate volume increment as:

$$\Delta V = \beta_1 V (\beta_2 G^{\beta_3 - 1} - \beta_4)$$

Yield estimates may be obtained by integrating the basal area increment function, and substituting for volume. Thus the cumulative volume growth over any specified period of time can be estimated. This approach leads to compatible growth and yield equations for both basal area and volume. Although it takes no account of site, this equation has been used in many growth and yield studies in uneven-aged stands (e.g. Murphy and Farrar 1982, with *Pinus taeda–P. echinata* stands).

Systems of Equations

A better understanding of growth in uneven-aged stands may be obtained if components of growth are individually identified and expressed collectively as a system of equations to predict stand growth.

Furnival and Wilson (1971) developed a growth and yield model as a system of equations and solved all the coefficients using simultaneous estimation. They formulated their yield model for white pine stands as the following system:

$$\ln \bar{h} = b_1 + b_2 \ln t \qquad (2.3)$$

$$\ln N = b_3 + b_4 \ln d_g \qquad (2.4)$$

$$\ln G = b_5 + b_6 \ln \overline{h} \tag{2.5}$$

$$\ln f = b_7 + b_8 \ln d_g \tag{2.6}$$

$$\ln G \equiv \ln k + \ln N + 2 \ln d_g \tag{2.7}$$

$$\ln V = \ln f + \ln \overline{h} + \ln G \tag{2.8}$$

where V is volume ($\mathrm{m^3\,ha^{-1}}$), $\overline{h}$ is mean tree height (m), t is average main stand age at breast height, G is basal area ($\mathrm{m^2\,ha^{-1}}$), d_g is the diameter (cm dbh) of the tree of mean basal area, N is number of trees ($\mathrm{ha^{-1}}$), f is the form factor, and k is a constant ($\pi/40\,000$). Equation 2.4 is Reineke's (1933) stand density equation, and Equation 2.7 is the identity defining stand basal area.

Equations 2.3–2.5 were differenced to obtain the growth formulation. Of three estimation techniques considered (direct, 2-stage and 3-stage least squares), three stage least squares used in conjunction with the growth formulation gave the smallest overall standard errors of estimation. This "systems formulation" allows a variance-covariance matrix to be produced and can be used to estimate standard errors of predicted growth and yields.

This example has some limitations. Equation 2.7 is an identity, and 2.8 is effectively exact, but Equations 2.3 to 2.6 are approximations that are not entirely realistic. For example, Equation 2.4 accommodates only mortality due to limiting competition, and does not account for other sources of mortality. However, it remains one of the few examples of the simultaneous estimation of a system of equations for a whole stand model. Some other examples were discussed by Borders (1989).

More Detailed Whole Stand Approaches

Some of the models discussed above provide rather limited information about the forest stand (in some cases only stand volume) but effective management and planning also require information about sizes and species contributing to this stand volume. An early solution to this problem was to include stand or stock tables in published yield tables. Some yield tables were presented as alignment charts and included additional axes indicating average and minimum stem sizes as well as other information regarding the dimensions of the crop. A remarkable example of this is Reineke's (1927) "composite chart" which concisely expresses relationships between age, dominant height, average dbh, basal area, stocking and total volume under bark for three stand fractions (dominant trees, trees over 10 cm dbh, and trees over 17 cm dbh). The same concept may be applied in computer-based models, and here we consider three types of whole stand model which estimate the size distribution and other details of trees in the stand.

Whole-stand Distribution Models

Instead of simply predicting volume or stand basal area, a growth model may reveal some aspects of stand structure by estimating the diameter distribution. Many probability density functions (p.d.f.s) can be used to describe the stand diameter distribution; the Gram-Charlier (e.g. Meyer 1930), Beta (e.g. Prodan 1953), Weibull (e.g. Bailey and Dell 1973) and Johnson's S_B (e.g. Hafley and Schreuder 1977) distributions have all proved useful in this regard. The Weibull and S_B distributions seem to be the most appropriate for even-aged forest stands (Gadow 1984).

Bailey and Dell (1973) and Schreuder and Swank (1974) demonstrated that diameter distributions in many even-aged forest stands can be quantified by the Weibull (1951) distribution (Fig. 2.3, overleaf):

$$f(x) = \frac{\beta}{\alpha}\left[\frac{x-\gamma}{\alpha}\right]^{\beta-1} \exp\left[-\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right] \tag{2.9}$$

Since height and volume can be expressed as an allometric function of diameter, they are distributed according to the Weibull distribution whose parameters are given by a simple transformation of the parameters describing the diameter distribution (Stacey and Mihram 1965).

There are two ways that these distributions can be estimated within in a whole stand model:

1. The parameters of future size distributions (e.g. the values of α, β and γ in Equation 2.9 above) may be predicted from existing parameters and other stand variables. This is known as the *parameter prediction* approach, because the parameters of the p.d.f. are predicted directly. Future values of stand parameters such as basal areas and volumes can be estimated from the predicted p.d.f. by integration.
2. The parameters of the size distribution at any time t may be estimated in terms of stand variables (e.g. stem number, mean diameter, basal area) at time t, and forecasts can be made by estimating future values of these stand variables. This is known as the *parameter recovery* approach because the parameters of the p.d.f. are recovered by matching the moments of the p.d.f. to estimated stand level variables. The parameter recovery approach has been found to give better results than the parameter prediction method (Reynolds *et al.* 1988).

Hyink and Moser (1979) used the Weibull distribution and the parameter prediction approach to model the yield of an uneven-aged forest. They noted that given the total number of trees, a diameter distribution, and some function *g(x)* of diameter, the definite integral of the product of the function *g(x)* and the distribution could predict stand characteristics corresponding to the function *g(x)* for any component of the stand. Thus the diameter distribution could provide estimates of top height, mean height, total volume and merchantable volume removed in harvesting. They assumed that the distribution could be adequately represented by the three parameter

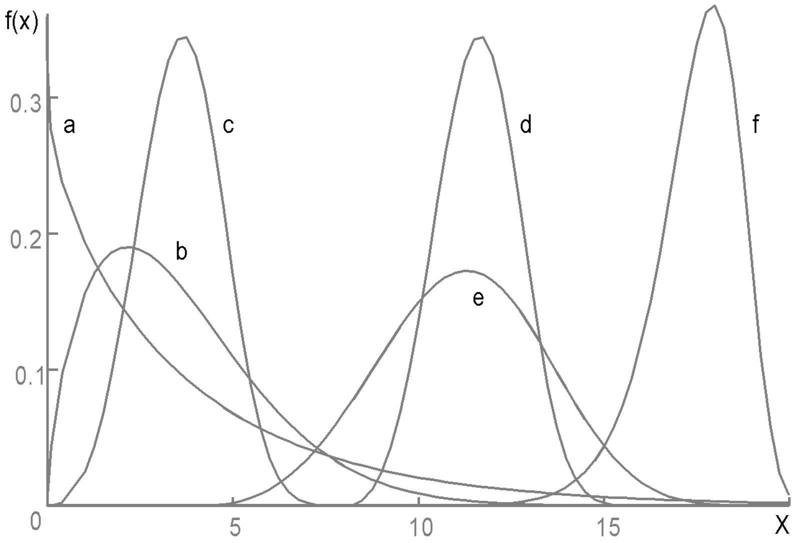


Fig. 2.3. Possible diameter distributions generated by the Weibull p.d.f. (Eqn 2.9), showing the influence of each parameter (Table 2.1) on the shape of the distribution.

Weibull distribution, and that growth of the stand could be accommodated by changing these parameters. Of these three parameters, Hyink and Moser (1979) assumed that the location parameter (γ) was a constant 17 cm dbh, the lower limit of measurement; leaving the scale (α) and shape (β) parameters, and the number of trees to be determined. They predicted these with empirical functions:

$$dN_i/dt = f_i\left(d_g, \Sigma d\right)$$

$$dN_m/dt = f_m(N)$$

$$d\alpha/dt = f_\alpha\left(N, dN_i/dt, \Sigma d, \alpha\right)$$

$$d\beta/dt = f_\beta\left(N, dN_i/dt, \Sigma d_{<21}, \alpha, \beta\right)$$

where N is stems per hectare, N_i is recruitment at 17 cm dbh, N_m is mortality, Σd is the sum of diameters, $\Sigma d_{<21}$ is the sum of diameters less than 21 cm, and α and β are the scale and shape parameters of the Weibull distribution respectively. The method may be adapted for mixed stands by using several sets of equations, one set for each species (e.g. Lynch and Moser 1986, Bowling *et al.* 1989).

Table 2.1. Coefficients for Weibull distributions in Fig. 2.3.

Example	Parameters of the Weibull function		
	α	β	γ
a	4	0.95	0
b	4	1.6	0
c	4	3.6	0
d	4	3.6	8
e	8	3.6	4
f	18	18.0	0

The approach makes more sense if the reader realizes that the parameter α is approximately equal to the stand median diameter, and β indicates the skewness of the diameter distribution. When $\beta = 3.6$, the distribution is approximately normal and symmetrical. As $\beta \rightarrow 1$, the distribution becomes skewed, with the mode (i.e. the highest point) moving left and a long tail forming on the right (Fig. 2.3). Similarly, as β increases (and $\beta > 3.6$), the distribution mode moves right and a long tail forms on the left. When β becomes very large or small, the variance of the distribution decreases.

The *parameter recovery* approach may offer a more robust alternative. The parameters of the distribution are predicted indirectly by matching the moments of the distribution to predicted stand attributes such as stand basal area and mean diameter. This approach is an efficient way to estimate the parameters of the Weibull distribution (e.g. Reynolds *et al.* 1988), but computational difficulties may be intractable with other distributions such as Johnson's S_B.

The utility of both methods depends on the ability of the Weibull distribution to characterize the diameter distribution adequately. Although the Weibull distribution can describe a great variety of shapes (e.g. Krug *et al.* 1984, Rennolls *et al.* 1985; Fig. 2.3), it is always uni-modal, and thus may not be suitable for all stands irrespective of species composition and silvicultural regime. The S_B distribution is more flexible and can represent some bi-modal distributions, but cannot readily be fitted using the parameter recovery approach. Multi-modal stands could possibly be accommodated by using a series of distribution functions, but this becomes rather complex and it may be better to adopt a size class or tree list approach for such stands.

Users should not expect a whole-stand diameter-distribution model always to give reliable information at the individual tree level. A model which gives satisfactory predictions at the stand level, may give unreliable predictions for some individual trees, especially for the largest or smallest individuals in the stand. When successive diameter distributions are estimated, they normally imply reasonable growth for most of the trees comprising the stand, but there may be some exceptions, and users should be warned that precision deteriorates if they disaggregate estimates to individual trees or size classes over short periods.

State Space Models

García (1984, 1994) used a *state space* approach to model plantations. The stand is represented by a few *state variables*, usually stand basal area, number of trees per hectare and top height. It is assumed that these state variables summarize the historical events affecting the future development of the stand, and thus that future states can be determined by the current state and future actions, and that other variables of interest, such as volume, can be derived from these state variables. This assumption is critical to this and several other modelling approaches and requires that two conditions be satisfied. The state variables must adequately describe the composition and structure of a forest stand, and should reflect all past silvicultural events, so that growth predictions do not need estimates of stand age, time since thinning, etc. These assumptions imply that growth predictions can be made simply by updating these few state variables.

These assumptions allow the system to be described by the specified state variables and a few simple *transition functions* (García 1994). Let the state at time t be specified by a list of n state variables represented as an n-dimensional state vector $\boldsymbol{x}(t)$. The inputs and outputs are also finite-dimensional vectors $\boldsymbol{u}(t)$ and $\boldsymbol{y}(t)$ respectively. Then the behaviour of the system is described by a transition function

$$\boldsymbol{x}(t) = \boldsymbol{F}\left[\boldsymbol{x}(t_0),\ \boldsymbol{U},\ t-t_0\right] \tag{2.10}$$

and an *output function*

$$\boldsymbol{y}(t) = \boldsymbol{g}[\boldsymbol{x}(t)] \tag{2.11}$$

Equation 2.10 gives the state at any time t as a function of the state at some other time t_0, of the inputs denoted by $\boldsymbol{U}$, and of the elapsed time between t_0 and t. The output function (2.11) estimates outputs as a function of the current state (e.g. it might predict volume from number of trees, stand basal area and top height).

A transition function must possess three properties:

1. There should be no change for zero elapsed time:

$\boldsymbol{F}[\boldsymbol{x}(t),\ \boldsymbol{U},\ 0] = x(t)$ for all t, $\boldsymbol{x}(t)$, $\boldsymbol{U}$.

2. The result of projecting the state first from t_0 to t_1, and then from t_1 to t_2, must be the same as that of the one-step projection from t_0 to t_2:

$$\boldsymbol{F}\left[\boldsymbol{F}\left[\boldsymbol{x}(t_0),\ \boldsymbol{U},\ t_1-t_0\right],\ \boldsymbol{U},\ t_2-t_1\right] = \boldsymbol{F}\left[\boldsymbol{x}(t_0),\ \boldsymbol{U},\ t_2-t_0\right] \text{ for any } t_0 \leq t_1 \leq t_2 .$$

3. A change of state can be caused only by inputs within the relevant time interval, and not events at other times:

$$\boldsymbol{F}\left[\boldsymbol{x}(t_0),\ \boldsymbol{U}_1,\ t_1-t_0\right] = \boldsymbol{F}\left[\boldsymbol{x}(t_0),\ \boldsymbol{U}_2,\ t_1-t_0\right] \text{ if } \boldsymbol{u}_1(t) = \boldsymbol{u}_2(t), \text{ for } t_0 \leq t \leq t_1 .$$

Transition functions generated by integration or differential equations automatically satisfy these conditions, so that the model can be stated:

$$d\boldsymbol{x}/dt = \boldsymbol{f}(\boldsymbol{x},\ \boldsymbol{u}) \tag{2.12}$$

$$\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x})$$

The state-space approach avoids the need to model directly the complex relationship over time between inputs and outputs, by describing the state of the system at a point in time and modelling the rate of change of state (Equation 2.12)

García (1984) used a multi-variate generalization of the Bertalanffy function in his state-space model for *Pinus radiata* plantations. The uni-variate Bertalanffy equation can be expressed as a linear differential equation with a power transformation of the state variable. For example, for site index curves,

$$dh_T^{\gamma}/dt = \alpha h_T^{\gamma} + \beta$$

where h_T is the state variable top height and α, β and γ are parameters to be estimated. The multi-variate generalization can be expressed:

$$d\boldsymbol{x}^{\boldsymbol{\gamma}}/dt = \boldsymbol{\alpha}\boldsymbol{x}^{\boldsymbol{\gamma}} + \boldsymbol{\beta}$$

with $\boldsymbol{x}^{\gamma}$ defined as $\boldsymbol{x}^{\gamma} = e^{\gamma \ln \boldsymbol{x}}$, where $\boldsymbol{x}$ is an n-dimensional state vector and α, γ and β are n-dimensional matrices and vectors of parameters. García (1984) found that γ was independent of site index, and that α and β changed by a constant factor with site index. Thus site index acted like a change in the time scale (i.e. growth is faster, but follows the same pattern), and could be accommodated in the model by putting $\tau = \eta t$ in place of t. However, this response with site index may not apply to all localities.

Although this approach gave good predictions and provided an effective framework for a series of plantation growth models, it offers less utility for mixed forests. Adapting the approach for mixed forests may be complex, as it is unrealistic to assume that mixed forests could be described adequately with only three state variables.

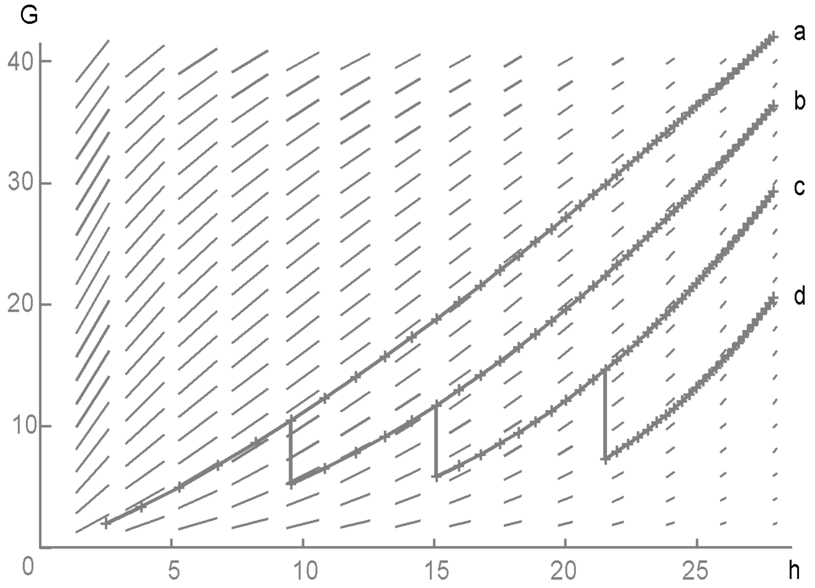


Fig. 2.4. Vector field illustrating a 2-dimensional state-space model. Line segments show annual growth for each stand condition. Trajectories trace development of 4 stands with different thinning regimes (after García 1994).

García (1994) offered a nice example of a two-dimensional state-space model, based on basal area increment and height–age equations of Clutter (1963) and Clutter and Lenhart (1968) respectively. Reformatted (and converted to metric units) as transition functions of current state, these equations become:

$$dh_T/dt = 0.0752\,h_T\left(3.59 - \ln h_T\right)^2 \tag{2.13}$$

$$dG/dt = 0.0752\,G\,\left(4.08 - \ln G\right)\left(3.59 - \ln h_T\right) \tag{2.14}$$

where h_T is top height (m) and G is stand basal area (m^2 ha^{-1}). These functions are illustrated as a vector field in Fig. 2.4. Line segments illustrate one year's growth for each stand condition (but note that those in the top left of the figure are unattainable). Without intervention, a stand will develop in the direction indicated by these line segments. The trajectories illustrate the predicted development of a stand which is 2.5 m high with 2 m^2 ha^{-1} basal area at age 5. Trajectory (a) illustrates the undisturbed development, while (b–d) are thinned to half the initial basal area after 5, 10 and 25 years respectively. The ticks (+) on the trajectories mark annual intervals.

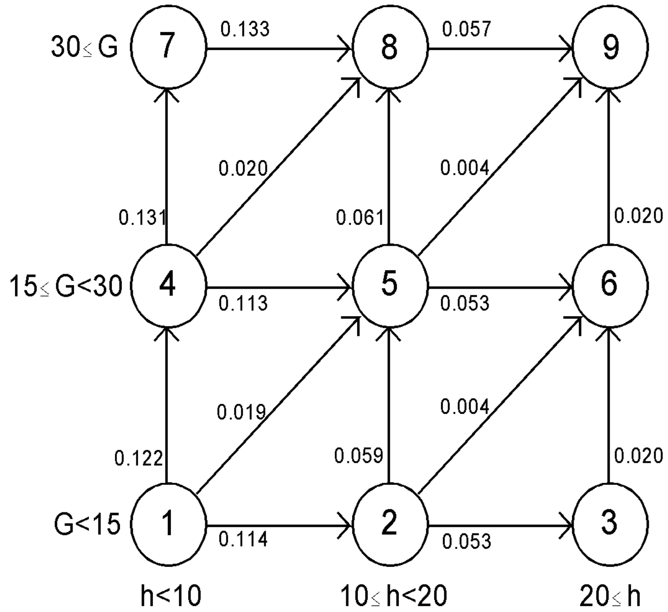


Fig. 2.5. Markov chain representation of a model based on Equations 2.13 and 2.14. Arrows indicate transition probabilities.

Whole-stand Transition Matrices

Markov chains provide a concise way to summarize the behaviour of a system, and may be used to model forest growth. We are concerned here with the whole-stand implementation, not the size class approach which is discussed in the next chapter. Here, states of the chain typically reflect stand density or species composition of the whole stand.

Consider a hypothetical system S, with n distinct states $S_1, S_2, \ldots, S_n$. If the system starts in state S_i, then in a single time interval, it has probability P_{ij} of moving to state S_j. Provided that these P_{ij} depend only on the current state S_i and not on any historic events, these probabilities can be expressed in a square matrix, termed the *transition probability matrix* or stationary Markov chain. This is a very concise way to summarize the behaviour of a system, but does little to help understand it. The basic assumptions are similar to those of the state-space approach, but the method uses a finite number of discrete classes ($S_1, S_2, \ldots, S_n$) rather than several continuous variables to indicate the state of the system. Instead of transition functions, this approach uses transition probabilities. Two nice features of these matrices are (i) that it is easy to determine the future state of the system by multiplying by the matrix (but only for multiples of the time

Matrix corresponding to Markov chain in Fig. 2.5:

$$\begin{bmatrix}
0.745 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0\dagger \\
0.114 & 0.884 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.053 & 0.980 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.122 & 0.0 & 0.0 & 0.736 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.019 & 0.059 & 0.0 & 0.113 & 0.882 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.004 & 0.020 & 0.0 & 0.053 & 0.980 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.131 & 0.0 & 0.0 & 0.867 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.020 & 0.061 & 0.0 & 0.133 & 0.943 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.004 & 0.020 & 0.0 & 0.057 & 1.0\dagger
\end{bmatrix}$$

Steady state derived after exchanging the two entries marked (†):

$$\begin{bmatrix}
0.066 & 0.066 & 0.066 & 0.066 & 0.066 & 0.066 & 0.066 & 0.066 & 0.066 \\
0.065 & 0.065 & 0.065 & 0.065 & 0.065 & 0.065 & 0.065 & 0.065 & 0.065 \\
0.176 & 0.176 & 0.176 & 0.176 & 0.176 & 0.176 & 0.176 & 0.176 & 0.176 \\
0.030 & 0.030 & 0.030 & 0.030 & 0.030 & 0.030 & 0.030 & 0.030 & 0.030 \\
0.073 & 0.073 & 0.073 & 0.073 & 0.073 & 0.073 & 0.073 & 0.073 & 0.073 \\
0.384 & 0.384 & 0.384 & 0.384 & 0.384 & 0.384 & 0.384 & 0.384 & 0.384 \\
0.030 & 0.030 & 0.030 & 0.030 & 0.030 & 0.030 & 0.030 & 0.030 & 0.030 \\
0.159 & 0.159 & 0.159 & 0.159 & 0.159 & 0.159 & 0.159 & 0.159 & 0.159 \\
0.017 & 0.017 & 0.017 & 0.017 & 0.017 & 0.017 & 0.017 & 0.017 & 0.017
\end{bmatrix}$$

Fig. 2.6. Markov matrix representation of the Markov chain in Fig. 2.5, and its steady state.

interval), and (ii) that the *steady state* (i.e. the long-term equilibrium) of the system can may be determined by repeatedly squaring the matrix.

Figure 2.5 illustrates a simple 9-state Markov chain derived from Equations 2.13 and 2.14 (cf. Fig. 2.4). The comparatively low transition probabilities on the diagonals (e.g. P_{1-5}) are due to the simplistic assumption that within any state, a stand is equally likely to have any of properties (i.e. basal area and height) defining the state. Figure 2.6 presents the same data as a Markov matrix. Note that the entries represent the probability of moving from *column* to *row*. Thus the entry 0.114 represents the probability of moving from state 1 (column 1) to state 2 (row 2; cf. Fig. 2.5). Notice that state 9 is an *absorbing* state, and that there is no movement out of this state. Since state 9 is the only absorbing state, the steady state will indicate state 9 with probability 1.0 (try it, by repeatedly squaring the matrix). Swapping the two entries labelled † in Fig. 2.6, assumes that stands are clearfelled and replanted after reaching state 9 ($h \geq 20$ m, $G \geq 30$ m^2 ha^{-1}). The steady state then indicates that 38% of the forest will be in state 6, and that 1.7% will be clearfelled and replanted each year. Notice that the present condition of the forest has no bearing on the steady state; after sufficient cycles the same steady state will be reached from any starting point.

Hool (1966) used this approach to predict the behaviour of an even-aged mixed forest under certain management regimes. He defined silvicultural activities A_k (e.g. thinning), which take the system from the state S_i to state S_j with probability $P_{ij} A_k$. With each activity induced transition there is an associated revenue $R_{ij} A_k$, and each state S_i has an associated value V_i . By valuing each state and activity, Hool (1966) could determine optimum thinning schedules. He used a two year transition interval, and identified 36 states based on silvicultural history (thinned/unthinned), standing volume (6 volume classes each 40 $m^3 ha^{-1}$) and stocking (3 classes each 170 stems ha^{-1}). He considered three activities, undisturbed growth (over the two year transition interval), thinning and harvesting. Hool did not explain the development of these matrices, but was concerned with applications of the approach, and with recommendations for management of forests in Indiana (USA). He showed that the mean transition time in any state was 2.4 to 3.8 years, depending on the state, and that if left undisturbed, the woodlands would remain in one of three states with about 200 $m^3 ha^{-1}$ in 180 stems ha^{-1} with probability 0.86.

Binkley (1980) used whole-stand transition matrices to examine succession in forest stands. The states of his matrix indicated the dominant species on each plot. His analysis suggested that assumptions inherent in the method were untenable (see next chapter), and that transition matrices were not a reliable way to predict forest stand dynamics. These assumptions and other limitations mean that Markov chains should be used only with caution, especially for examining ecological problems (e.g. Jeffers 1978).

Synthesis

Whole stand models have been very useful for modelling plantations, but they appear of limited utility for mixed forests, where the number of species and the potential for multi-modal size distributions creates difficulties in characterizing the stand with few stand-level variables.

Although this class of models does not appear particularly useful for modelling mixed forests, many of the concepts are important and applicable more widely. In particular, the transition function which forms the basis of the state-space modelling approach (p. 26), is an important concept that recurs in many alternative approaches more suited to mixed forests. Similarly, the concepts of the Markov chain (p. 29) will recur in the next chapter.

You should have noticed that the many alternative modelling approaches form a continuum rather than discrete classes, and that the classification of models under various headings is an arbitrary one for the purposes of discussion. In reality, possible modelling approaches merge seamlessly from one to another.

In the next chapter, we examine models which provide more detail by partitioning the stand into several size classes.

Exercises [1]

2.1. Which of the models, if any, reviewed in this chapter would suit the needs of your organization? List the strengths and weaknesses of the modelling approach most suited to your needs. What data and analyses would you need to build such a model? Does your organization have sufficient suitable data to construct such a model?

2.2. You are involved in the analysis of some permanent plots in which all trees were measured on two occasions. Write formulae to calculate survivor growth and accretion from these individual tree data. Assume that *s* trees are alive on both occasions, *d* trees died between remeasures, and *r* trees were recruited at the second measure. *Hint*: Check the definitions given on page 17.

2.3. Demonstrate that the vector field (e.g. Fig. 2.4, p. 28) of a state-space model is equivalent to traditional graphs of growth and yield versus age, by using Equations 2.13 and 2.14 to plot trajectories (a–d) against time.

2.4. The Markov chain/matrix approach requires discrete classes, so it better suited situations where these are found naturally. Test your understanding of the method by drawing a Markov chain for the gambling game "two-up". A person, the "spinner" tosses two coins; if they are both tails, he retires, and if they are both heads, he continues (one head and one tail is considered a "no throw", and he tosses again). Spectators wagering with the spinner collect if he throws tails; the spinner collects only if he has thrown three pairs of heads. Draw the Markov chain for the game, compile the matrix, and calculate the chance that the spinner collects his takings. Is there a forestry application where such discrete states occur naturally?

2.5. Leech (1993) suggested that past trends could be used to predict future land use changes. Use the data given below to calculate anticipated plantation areas in 10 years time. What will be the eventual area of each land use type when equilibrium is finally reached?

---

[1] Solutions to selected exercises are given on p. 285

Use 5 years ago		Current land use and area (ha)			
Land use	Area (ha)	Native forest	Plantation	Agriculture	Other uses
Native forest	14400	10900	1200	1800	500
Plantation	3350	0	3200	90	60
Agriculture	21600	0	600	20000	1000
Other uses	2400	0	0	100	2300

2.6. Successional changes observed in swamp vegetation are given below (from Jeffers 1978; the time step is 20 years). What proportions of each community would you expect at equilibrium?

Starting state	Probability of transition to end-state			
	Bog	*Calluna*	Woodland	Grazed
Bog	0.65	0.29	0.06	0.00
Calluna	0.30	0.33	0.30	0.07
Woodland	0.00	0.28	0.69	0.03
Grazed	0.00	0.40	0.20	0.40

Chapter Three

Size Class Models

Size class models employ a class of trees as the basic unit for modelling. The term "size class" is used for convenience, as size is the most common criterion for forming classes. However, other criteria may also be used to form classes (e.g. species, age, etc.).

The size class approach is a compromise between whole stand models and single-tree models. Whole stand models may be considered size class models in which a single class encompasses the whole stand. Similarly, single-tree models effectively have a single class for each individual tree. Thus the size class approach encompasses those models which divide the stand into two or more size classes, but with fewer classes than the total number of trees.

Several size class models originate from the classical method of stand table projection, which divides the forest stand into several size classes of equal size based on tree diameters. However, the stand does not need to be partitioned into standard classes: it may be formed into *cohorts* or groups of trees with similar characteristics (e.g. species and size). Some applications require equal numbers of trees in each cohort, but the method offers considerable flexibility to construct cohorts in many ways. Diameter is not the only size criterion that may be used; height classes are also commonly used.

Stand Table Approaches

A stand table is a tabular summary showing the number of trees in each of several size classes. In mixed stands, there may be rows for each species or species group. Size classes are usually diameter classes of equal width (e.g. 10–19, 20–29, . . ., 100+ cm dbh). These tables are commonly used to summarize inventory data, and provide the basis for several popular growth models for mixed forests.

Time of Passage

A simple "rule of thumb" approach for calculating yields is based on the *time of passage*, the time for a tree to grow through a diameter class. This method has been used to determine cutting cycles in natural forests, especially where minimal inventory data are available. The method can be used with data comprising only the "leading desirables" (i.e. vigorous trees subjectively assessed to form the next harvest, Dawkins 1958), which can otherwise be very difficult to analyze.

The calculations are rather simple (e.g. Osmaston 1956, Dawkins 1958, Alder 1992). Average increments are estimated for several species groups and size classes, and times of passage are estimated by dividing the class width by the average increment within each class. The length of the cutting cycle can be estimated from the time of passage and the cutting limits (Dawkins 1958).

This approach may not always give reliable estimates of yield or duration. Because of the variability and high serial correlation of diameter increments, the method may reflect the average performance of the stand, but the harvested component may perform substantially better than this (e.g. Mervart 1972). Thus time of passage calculations based on stand mean increments may underestimate yields and overstate the length of the cutting cycle. However, leading desirables comprise the faster growing component of the stand and are likely to form the bulk of the final crop. Thus time of passage calculations based on leading desirables may be more reliable, as the errors may compensate. However, selection of the leading desirables is subjective and the intensity of selection may influence results.

Stand Table Projection

Stand table projection is one of the oldest techniques used to determine the future composition of uneven-aged forests. The method predicts the future stand table from the present stand table by adjusting each entry in the table with the estimated diameter increment (and mortality). Diameter increment estimates may be obtained from several sources, ranging from guesses and simple tabular summaries to regression analyses, depending on the nature of the data available. The method dates from times when data were few and computations difficult, and several researchers offered simple formulae for estimating *upgrowth* (i.e. stem number advancing to the next class) from summarized plot data (e.g. Herrick 1938, Chapman 1942). Now that computers have eased the burden of computation, stand tables may be updated with increment equations prepared using regression analyses, but the "educated guess" remains a standby where data are deficient.

Three methods can be used to forecast the future stand. The first assumes that all trees in each diameter class are located at the class midpoint, and that all trees will grow at the same average rate, irrespective of their present size and vigour (e.g. Husch *et al.* 1982). This essentially involves projecting the class boundaries so that future classes contain the same trees (if there is no mortality), but may have different class boundaries (and widths), which may be inconvenient for some applications.

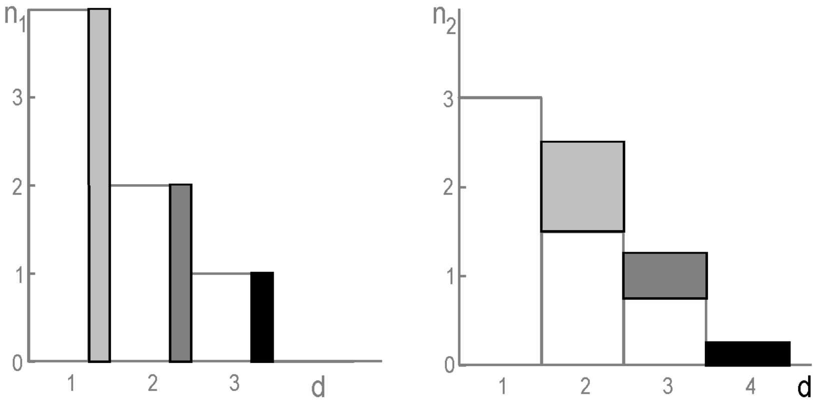


Fig. 3.1. Stand table projection with movement ratio 0.25, so that 25% of each class moves up to the next class.

It also takes no account of the variable nature of diameter increment, so that the forecasted stand will span the same number of size classes as did the initial stand.

The second alternative assumes that trees in each diameter class are uniformly distributed through the class and that each tree grows at the average rate (Fig. 3.1). For each class, a "movement ratio" is determined from the class width and average increment, and this indicates the number of trees moving to the next class (upgrowth). Projections involving stands with only a few trees require a decision regarding rounding fractions of trees (e.g. Carron 1968). In projections for large forest areas where fractions may be considered acceptable, the method may lead to bias in predicted growth of the largest trees, as the distribution of stems in the largest classes is rarely uniform. Both these methods ignore dispersion of individual increments.

The third option (Wahlenberg 1941) attempts to account for the variation in growth rate within any diameter class. This is achieved by using the actual movement of trees rather than movement ratios. An example given by Husch *et al.* (1982) illustrates one instance where the movement ratio approach predicts that 94 percent of trees will move one class and 6 percent remain, whilst Wahlenberg's method predicts that 20

Table 3.1. Three approaches to Stand Table Projection.

Projection method	Size class (cm dbh) (lower limit of class)					Mean dbh (cm)
	10	20	30	40	50	
Initial stand	40	30	20			22.8
Stand after 1 projection cycle						
Class boundary method†		40	30	20		31.8
Movement ratio (90% grow 1 class)	4	39	29	18		31.8
Wahlenberg method (20% grow 2 classes, 50% grow 1 class)	12	29	29	16	4	31.8

percent will move two classes, 50 percent will move one class, and 30 percent remains in the same class. These differences accumulate and can become substantial over several cycles. While the average growth for the whole stand may be the same in the two approaches (Table 3.1), the growth of the larger fraction will be greater with Wahlenberg's method. This may lead to considerable differences in predicted harvests, especially for selection harvesting systems where a few of the largest trees are harvested each cycle. Thus Wahlenberg's method should provide more reliable estimates of timber yields from uneven-aged stands.

One problem with the last two approaches to stand table projection is the proliferation of classes with fractional numbers of stems. This difficulty is avoided by the first approach. A related problem is the ability of some stems to move *n* classes in *n* projection periods, which may introduce bias. The first option (updating class boundaries rather than moving stems between classes) does not have these problems because it allows no variability in increments in the projected stand, but this may underestimate future yields for the harvested component of the stand. There are three ways that the proliferation of fractions can be controlled:

1. Use a longer projection interval or narrower size classes. The optimal approach is to choose the smallest class width and longest time step (consistent with user requirements) that allow upgrowth to progress only one class in any single step.

2. Allow a non-uniform distribution of stems within each class by smoothing the stand table with a single curve (e.g. Weibull function) or with a series of curves (e.g. splines).

3. Accumulate small probabilities of upgrowth until some stems can be projected, either by

(a) specifically nominating the probability or stem number required to initiate a new class, or by
(b) explicitly modelling the growth of the largest tree in the stand, and initiating a new class only when a tree attains that size.

Smoothed Stand Table Models

Campbell's (1981) model for pine plantations in Western Australia used sixteen 8 cm diameter classes. For computational efficiency, the size classes and time step were chosen so that no tree could advance more than one class in a single growth period. A series of quadratic splines were fitted to the size distribution at each time step. The splines used a quadratic equation for each diameter class, with the explicit constraints that the resulting curve must

1. be smooth and continuous,
2. encompass the appropriate number of trees in each class,
3. be positive through its range, and must
4. reach zero at the lower limit of the smallest class and the upper limit of the largest class.

The integral of these splines provides estimates of mean diameters and total basal area within each class. It also allows the computation of numbers of trees, basal area or volume between any specified diameter limits. Unlike the Weibull distributions often used for whole stand models, the spline curve approach (Smith 1979) can accommodate stands which are not uni-modal. In effect, the method is rather similar to that of the parameter recovery approach (p. 25), but is more flexible and can accommodate a wider range of size distributions.

Campbell's (1981) model predicts the increment of the mean tree in each class, and assumes a near-normal distribution of increment within each class (using a Weibull approximation to the normal distribution, so that all increments ≥ 0). Upgrowth from each class is estimated from the within-class distribution of trees and the predicted distribution of increments. Thus the heteroscedastic nature of growth is preserved in the model. Upgrowth from the largest class is restricted by specifying the minimum proportion of the total stocking required before a new class can be initiated. If the projected upgrowth is less than this critical proportion, the upgrowth remains in the existing largest class. The same critical proportion is used to absorb the smallest class into the next class, when upgrowth reduces the proportion remaining in the smallest class to a sub-critical amount. Campbell assumed that in a well managed plantation, mortality and recruitment would be negligible. The model is an integral part of the management information system for Western Australia's pine plantations (mainly *Pinus pinaster*).

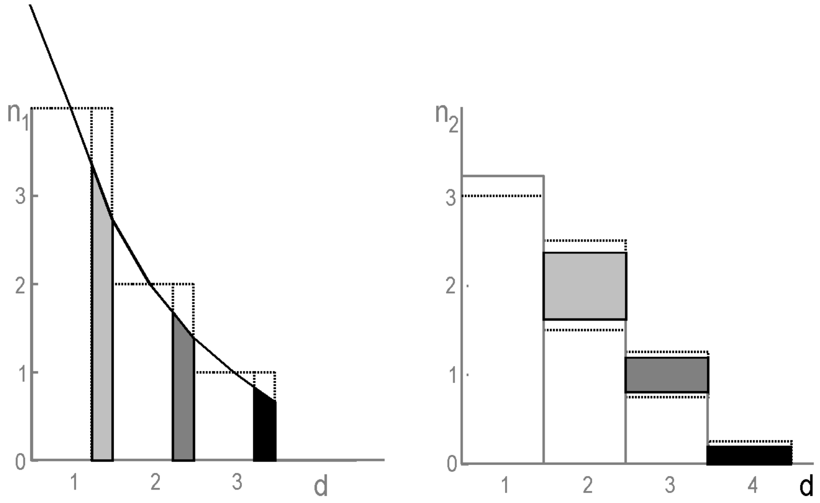


Fig. 3.2. Smoothing an uneven-aged stand table results in smaller estimates of upgrowth for a given movement ratio. Dotted lines are reproduced from Fig. 3.1 for comparison.

In 1898, the French forester de Liocourt observed that the size distribution of stems in an uneven-aged natural forest tended to form a reverse-J curve or negative exponential distribution, in which the ratio of stem numbers in adjacent size classes remains constant, provided the size classes are the same width: $n_1/n_2 = n_2/n_3 = n_3/n_4 = \ldots$ (see e.g. Philip 1994). This ratio is known as de Liocourt's quotient, q. In his stand table projection model for dipterocarp forests in Malaysia, Korsgaard (1989*a*,*b*) smoothed the stand table using such a J-curve. He found that quotients ranged from 1.3 to 1.6, and remained relatively constant over time and across stands. The J-distribution of stem sizes is more realistic than a uniform distribution within diameter classes, and when used in stand table projection, results in smaller estimates of growth (Fig. 3.2). For instance, for a given diameter increment, a uniform distribution within classes may imply that half the stems in a class would move up to the next class, whilst a J-distribution with quotient 1.5 would imply that only 0.4 of the stems move.

Smoothing the stand table affects estimates of upgrowth, and thus it is important that the method of estimating the movement ratio is compatible with smoothing, or biased estimates will be obtained. If movement ratios are determined from diameter increment estimates, then it is appropriate to smooth the stand table. However, if movement ratios are estimated from direct observations of trees advancing to the next class (i.e. Wahlenberg's method), then smoothing is unnecessary.

Korsgaard (1989*a*,*b*) also explicitly modelled the growth of the largest tree in the stand table, so that a new size class was initiated only when the largest tree attained that size. Because he was dealing with mixed tropical forests, Korsgaard modelled several stand tables, each representing different parts of stand. One implementation complied stand tables based on crown illumination classes (e.g. trees with full overhead light grow faster than those partially or fully shaded), while others use broad species groups (e.g. meranti, other dipterocarp, non-dipterocarp). Chai and Sia (1989) described the application of this model to mixed swamp forests in Sarawak.

Stand table projection was originally devised as a simple and efficient method that would allow a calculation to be completed on a sheet of paper. Computers enable several enhancements, and a typical implementation would take the stand table and the movement ratios (or a table of diameter increments) as input, and automatically update the table through the required number of steps. Thus the movement ratios normally reflect the average growth over the whole period of projection, disregarding the fact that these might decrease as stand density increases. Some implementations become quite complex, employing several tables of movement ratios for different species groups and stand densities, with additional tables for mortality and recruitment. But this sophistication seems wasted on this basically simple method, and it is appropriate to improve the methodology commensurately.

Empirical Equations

Some of the limitations of stand table projection can be overcome by using equations directly, without the intermediate step of compiling a table of movement ratios. One advantage of this is that it allows predicted increments to be adjusted for different sites, species and stand conditions. An equation also provides a more parsimonious summary of data than several tables of movement ratios.

One of the simplest applications of such equations is that of Grimes and Pegg (1979) who used only three classes (20–30, 30–40, 40+ cm dbh). They recorded basal areas and numbers in each class, and predicted the change due to growth, recruitment and mortality in these classes by simple empirical equations. Despite the crude nature of the model, it produced acceptable results in short term simulations.

Leak and Graber (1976) represented an uneven-aged beech-birch-maple stand with eleven diameter classes of different widths. They predicted diameter increment from diameter and stand basal area, and converted these to movement ratios for updating the stand table. Equations were also used to predict recruitment into the smallest class, and to predict mortality in the smaller classes. Mortality in the larger classes was assumed to be constant at about one percent. Howard and Valerio (1992) reported a similar study for mixed tropical forest in Costa Rica.

Table 3.2. Characteristics of size classes defined in the FORMIX model.

Canopy layer	Height (m)	Diameter (cm dbh)	Movement ratio
Seedlings	<1.3	<1	0.10
Saplings	1.3–15	1–10	0.08
Poles	15–25	10–25	0.05
Main canopy	25–36	25–45	0.02
Emergents	>36	>45	

Process-oriented Stand Table Models

The stand table framework can also support more detailed modelling approaches. The FORMIX model (Bossel and Krieger 1991) for Malaysian lowland dipterocarp forest uses five size classes, nominally defined as canopy layers (Table 3.2). Tree numbers and biomass in each class are modelled using process-oriented equations driven by solar radiation. Predictions are based on class means, and no species differences are accommodated. Temperature, water and nutrients are assumed to be constant and non-limiting.

Photosynthesis is modelled using a simple asymptotic saturation curve dependent on the light intercepted, adjusted for attenuation by the canopy layers above. Respiration is assumed to be a constant proportion (6%) of biomass. The model allows one of two possible mortality rates in each class, a high rate (10–50%) if the canopy layer is closed (10 000 $m^2 ha^{-1}$), and a lower rate (½–10%) otherwise. The change in biomass is computed from photosynthesis minus respiration, minus mortality (number of dead trees times mean tree biomass in the class). The mean tree diameter ($\overline{d}_i$) in each class is then computed as

$$\overline{d}_i = \left(\frac{4}{\pi f h_d} \frac{0.7\ \overline{b}_i}{0.65} \right)^{1/3}$$

where 0.7 converts tree biomass ($\overline{b}_i$) to stem biomass, and in turn, 0.65 converts this to stem volume (v). This equation is derived from the familiar relationship

$$v = fgh = \frac{\pi f d^2 h}{4} = \frac{\pi f d^2 (d h_d)}{4}$$

where f is a form factor (0.38–0.5, depending on tree biomass), g is tree basal area, h is tree height, d is tree diameter, and h_d (40–140, depending on tree biomass) converts a diameter to a height. The form factor and height–diameter conversions are read from a look-up table based on mean tree biomass. If the mean diameter within a size class exceeds the threshold diameter (Table 3.2), then the movement ratio is invoked to advance some trees to the next class. Thus the physiological basis effectively turns on or off the movement ratio in the familiar stand table approach.

The upgrowth biomass is the upgrowth number times the class mean biomass, ignoring the likelihood that the largest trees rather than mean trees advance to the next class. A look-up table was used to estimate seed production by emergent and main canopy trees, and this in turn provided estimates of recruitment into the smallest class.

Some of the parameters used in the FORMIX model have biological or physical meaning, and were obtained from published literature or estimated from empirical observations. The model requires 13 global parameters (8 of these are general parameters obtained from the literature), 4 parameters for each of the 5 canopy layers, and look-up tables for height–diameter ratios, form factors and seed production. The model seems rather insensitive to many of these parameters, and the three critical parameters are the crown diameter ratio, photosynthetic efficiency, and the biomass maintenance cost (i.e. 25, 0.5, and 0.06 y^{-1} respectively, Assmann 1961). These parameters influence the frequency of natural cycles in simulations of undisturbed development, but do not affect the general nature of the response.

FORMIX does not distinguish species, since Bossel and Krieger (1991) felt that this would add unnecessary complexity without a significant change in overall forest dynamics. However, they recognized that enhancement of the model to recognize different physiognomic groups (pioneer, shade tolerant, etc.) would be desirable. They formulated a single-gap variant of this model which modelled spatial gap dynamics. They felt that the general model gave an adequate representation of stand development during 200 years following disturbance (e.g. harvesting), but that the spatial variant was necessary when there was a significant amount of natural mortality amongst mature trees.

Simulations of undisturbed development did not gradually approach a stable stand condition, but revealed cycles with a frequency of 100 and 200 years for main canopy and emergent trees respectively. A simulation study of the Malaysian selective management system suggested that its 35 year cycle was out-of-step with these natural cycles and could not provide a sustainable harvest.

The data required to fit this model only partially correspond to conventional dynamic inventory. Efficient calibration of, and enhancements to FORMIX require data on light attenuation for various canopy structures, light response curves and tree geometry data for different physiognomic guilds, assimilate partitioning coefficients for different guilds and stages of development, and respiration rates for tree components as a function of temperature, etc. This emphasizes the interrelationship between model design and data requirements (Chapter 5).

Transition Matrices

Transition matrices are a logical and formalized extension of the method of stand table projection. They allow movement ratios and other details to be summarized in a matrix, so that growth projections can be made with a single matrix multiplication. For example, the stand table projection illustrated in Fig. 3.1 can be expressed:

$$\begin{pmatrix} 0.75 & 0 & 0 & 0 \\ 0.25 & 0.75 & 0 & 0 \\ 0 & 0.25 & 0.75 & 0 \\ 0 & 0 & 0.25 & 1.00 \end{pmatrix} \times \begin{pmatrix} 4.0 \\ 2.0 \\ 1.0 \\ 0.0 \end{pmatrix} = \begin{pmatrix} 3.00 \\ 2.50 \\ 1.25 \\ 0.25 \end{pmatrix} \tag{3.1}$$

The matrix and vector are multiplied by computing $v'_i = \Sigma_j m_{ij} v_j$ (e.g. for the second row, 0.25×4.0 + 0.75×2.0 + 0×1.0 + 0×0.0 = 2.5). Notice that the columns sum to 1.0, so that the total of the numbers in the vector remains unchanged. Mortality can be modelled by reducing the numbers in the matrix (e.g. for 1% mortality, replace 0.75 with 0.74, and 1.0 with 0.99 in the matrix of Eqn 3.1). Figure 3.2 could be summarized in the same way, but the transition probabilities would be 0.2 and 0.8 instead of 0.25 and 0.75.

Three variations of this approach exist, and differ in their underlying assumptions: Markov chains, Usher matrices and their generalizations.

Markov Chains

The construction of a Markov chain requires that

(a) at any time a system can be in any one of a finite number of states, and

(b) during the next time interval, has a known probability of moving to any other state.

The probability of movement must depend only upon the current state, and not on historic events in the system. We have already considered how a Markov chain can provide a basis for a whole stand model (p. 29). An alternative is to define the states of the system in terms of the sizes of individual trees, rather than in terms of whole stand condition.

Consider a tree in one of the diameter classes of a stand table. During the next period, it must either remain in the class, grow into another class, get harvested, or die. This allows us to draw a Markov chain for the updating of a stand table (Fig. 3.3). By definition, stand tables involve discrete states, and thus the size class implementation of the Markov chain avoids the weakness seen in the whole-stand version, where discrete states were a poor approximation to the continuum of possible stand conditions.

The stand table implementation leads to the apparent contradiction that although a Markov chain represents a stochastic process, it is generally used as a deterministic yield model. A Markov chain could be used as a stochastic model if random numbers were drawn to decide if the whole class

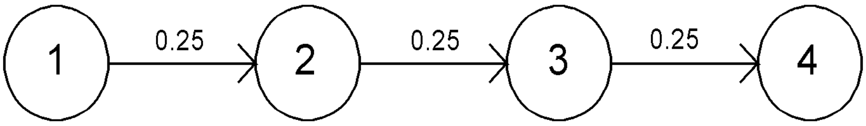


Fig. 3.3. Markov chain representation of Fig. 3.1 and Equation 3.1. Probability of remaining in a state is one minus the probability of moving. A growth model could also include states for death and harvesting.

(or members of the class) move to a new state, but the usual implementation is to move a proportion of the stems in any class to obtain a deterministic result. Some stochastic implementations of Markov chains are considered in the next chapter.

A Markov chain contains two types of states, transient and absorbing. *Transient* states are those in which the system spends a finite duration: it must eventually leave these states. In Fig. 3.3, states 1–3 are transient states, because there is an arrow with a non-zero probability of leaving these states. The size classes in a growth model are transient states, as every tree must eventually grow into the next class, get harvested, or die. *Absorbing* states are those from which the system can never leave. In our case, death and harvesting represent absorbing states.

The probabilities of movement are generally expressed as a matrix (***M***), and can be used to predict change during a single time interval:

$$\boldsymbol{V}_1 = \boldsymbol{M}\,\boldsymbol{V}_0$$

or over several time intervals:

$$\boldsymbol{V}_n = \boldsymbol{M}^n\,\boldsymbol{V}_0$$

where ***M*** represents the Markov matrix containing the probabilities of movement, and $\boldsymbol{V}_0$ and $\boldsymbol{V}_n$ are vectors describing the initial and final states respectively. In forestry, $\boldsymbol{V}_0$ is generally a list representing the initial numbers of trees in each diameter class (i.e. the stand table).

In order for these expressions to hold, two assumptions must be made. The first, called the *Markov assumption*, requires that the probability of any event depends only on the initial state, and not on any previous state (i.e. a tree's history is not required). Thus the probability that a tree grows into the next class must depend only upon the class that the tree is presently in, and not upon the characteristics of that tree (except those that define the class), upon any other tree, or upon the total number of trees in any class (or in the stand as a whole). We have already seen this assumption in conjunction with the whole-stand state-space (p. 26) and Markov-matrix models (p. 29). The second assumption, the *stationary assumption*, requires that these probabilities do not change over time. We have already made this assumption in determining the steady state of the matrix in Fig. 2.6 (p. 30).

The Markov and stationary assumptions may be quite restrictive in forestry applications. The Markov assumption makes it difficult to account for competition and suppression, since the probability of upgrowth should not depend on other trees or on the stand basal area. The stationary assumption means that the matrix parameters should not change over time, and this makes it difficult to reduce growth rates during a projection as the predicted stand basal area (or other estimate of competition) increases. Theoretical (e.g. Hulst 1979) and empirical (e.g. Binkley 1980, Roberts and Hruska 1986) studies suggest that these assumptions are untenable when modelling forest dynamics.

Despite these restrictive assumptions associated with Markov chains, many researchers have attempted to use the method for forest growth forecasting. Bruner and Moser (1973) discussed a study which used 25 states: 23 dbh classes (8–30+ inches, 20–75+ cm by 2.5 cm), and one class each for mortality and harvesting. They post-multiplied the state vector, so their matrices represent the transpose of matrices in the more common pre-multiplication approach. Their model had only one absorbing state each for harvesting and mortality, so it predicted only the numbers, but not the sizes of dead and harvested trees. The sizes of harvested trees can be recorded simply by allowing more states, one for each dead or harvest size class required (e.g. Peden *et al.* 1973, Rorres 1978). In models for mixed forests, species groups can be represented by using more states, or by compiling additional matrices and modelling each species group independently (e.g. Cassell and Moser 1974 used six matrices for different shade tolerance classes).

Predictions become unreliable if stand conditions (e.g. stand basal area) depart greatly from those conditions prevailing in the data from which the matrix was developed. This violates the stationary assumption. Provided that stand conditions remain within a narrow range similar to the development data, the stationary assumption may not be compromised, and predictions may be satisfactory. Thus the method may perform best with stands that are maintained within a narrow range of conditions (e.g. undisturbed near-climax stands, or stands that are regularly thinned to a specified basal area).

Projections can only be made in multiples of the measurement interval for the plots. However, a one-year matrix ($\boldsymbol{M_1}$) may be estimated from an n-year matrix ($\boldsymbol{M_n}$) such that $\boldsymbol{M_1}^n \approx \boldsymbol{M}_n$ (Harrison and Michie 1985). An exact solution cannot be found. The one-year matrix must have entries on at least two diagonals (i.e. bi-diagonal or Usher matrix) or no growth can occur, and this enables some stems to move n classes in n years, even though the original matrix ($\boldsymbol{M_n}$) may predict less movement.

A further disadvantage of Markov matrices is the number of parameters required and the implications this has for the likely precision of each parameter estimate. Moser's 25-state matrix involves 625 constants. Fortunately, many of these are zero, but at least 100 entries (4 each for most classes: probabilities of remaining, growing one class, death or harvest) may be non-zero and must be estimated. There are also difficulties in accommodating different site and stand conditions, and these are best overcome by customizing the matrix for each stand to be projected.

Leslie and Usher Matrices

Usher matrices attempt to reduce the large number of parameters required to form a Markov matrix. Choosing the time interval and class width so that no tree can grow more than one class during the period (for convenience, called the *Usher assumption*) allows a substantial reduction in the number of parameters to be estimated. Unlike some ways of obtaining movement ratios for stand table projection, the Usher assumption does not introduce bias because it requires that a single time-step must be sufficiently short that a tree cannot grow more than one class.

A true Markov matrix contains only probabilities, but by relaxing this restriction, recruitment can be predicted by employing non-zero values in the top line of the matrix. These values are not probabilities, but reflect *fecundity*, the number of offspring for each individual in the corresponding cell of the state vector. Fecundity values in the matrix allow the amount of recruitment appearing to vary according to the presence of trees in various classes. Thus the matrix may be reduced to four vectors: one each for growth (either a tree moves into the next class, or it does not), fecundity (recruitment), harvesting and mortality. Typical applications may contain 20–25 states and require the estimation of 80–100 parameters, more than is required in many equation approaches.

Leslie (1945, 1948) pioneered the use of these matrices for animal populations where the classes represented ages. The technique was adapted for stages of insect development by Lefkovitch (1965), and for forestry (diameter classes) by Usher (1966). A rare application of a Leslie matrix to forest stands was Bosch's (1971) study of redwoods, which used age classes and in which regeneration appeared only on the death of another tree. Most studies utilize the diameter class matrix advocated by Usher (1966), but are still commonly termed Leslie matrices. However, there is an important distinction between Leslie and Usher matrices. With a Leslie matrix of age classes, *all* surviving individuals progress ("age") to the next class each cycle. With an Usher matrix, only some of the surviving trees grow into the next class, whilst those with little or no growth remain in the same class.

Usher matrices, perhaps because of the efficiencies inherent in the Usher assumption, have been used more widely than Markov matrices. Usher (1976) used these matrices to estimate optimum yield and rotation length for *P. sylvestris* plantations in Britain. Rorres (1978) continued this analysis to prove that the optimal sustained yield harvesting regime is a cutting limit regime which removes all the stems in only one class, removes a proportion of the stems in several smaller classes, and leaves all the remaining smallest classes untouched. This is consistent with some selection harvesting guidelines but at odds with harvesting practices in many countries.

Square matrices have *eigenvalues* and *eigenvectors* which satisfy the relationship

$$\boldsymbol{M}\,\boldsymbol{V} = \lambda\,\boldsymbol{V}$$

where $\boldsymbol{M}$ is a square matrix, $\boldsymbol{V}$ is a column vector, and λ is a scalar. In general, if $\boldsymbol{M}$ is $n \times n$, there will be n eigenvalues λ, each with an associated eigenvector $\boldsymbol{V}$. The eigenvalues may reveal some interesting insights into the system governed by the matrix. For example, if $\lambda = 2$, the population $\boldsymbol{V}$ will double every time step of the model, but the proportion in each class will remain the same. Recall the whole-stand transition matrices discussed in the previous chapter. Some were Markov matrices (every column summed to 1.0) with no absorbing states, for which we determined steady states (e.g. Fig. 2.6, Exercises 2.3 and 2.4). These matrices have an eigenvalue $\lambda = 1.0$, and the corresponding eigenvector is the steady state. Because Usher matrices incorporate fecundity, the eigenvalues are more interesting. Usher (1966) argued that the largest positive eigenvalue of an Usher matrix indicates the maximum exploitation, and that its eigenvector indicates the stable stand structure.

Plant demographers also attach some importance to the eigenvalue, as it should indicate whether the population is increasing or decreasing in number. In theory, pioneer and light demanding plants in undisturbed forest should have eigenvalues less than unity, whilst shade tolerant plants should have values greater than 1.0. Values greater than one for pioneer species would indicate continuing disturbance. Hartshorn (1975) obtained a value of 1.002 for the shade tolerant climax species *Pentaclethra macroloba* in Central America. In Papua New Guinea, Enright and Ogden (1979) observed values of 1.02–1.01 for the shade tolerant *Araucaria cunninghamii*, and 0.99–1.09 for the shade intolerant mid-successional *A. huntsteinii*. Unfortunately, it is not clear whether the eigenvalues are a characteristic of the species, a characteristic of the present state of the forest, or an artifact of the method. Eigenvalues assume an exponential increase in the number of trees in each size class. Thus the eigenvector may indicate optimal stand structure, but not optimal stand density. The problem of determining the optimum level of growing stock cannot be resolved with eigenvalues.

Numerous studies have investigated the sensitivity of the eigenvalues and vectors to noise in the matrix. Usher (1966) found that small variations in the fecundity terms of the matrix had very little influence on the eigenvalues and vectors, and that estimates of growth were the most important determinants of eigenvalues and vectors.

Generalized Matrices

The Markov assumption (that the transition probability depends only on the initial state) is a severe restriction to impose. If this assumption is avoided, it is possible to construct a matrix model in which recruitment is density dependent, or which accounts for suppression of smaller trees by introducing negative values in the upper right triangle of the matrix.

Buongiorno and Michie (1980) constructed a model for a mixed hardwood forest, using the stationary and Usher assumptions, but not the Markov assumption. Their matrix was essentially a bi-diagonal matrix with probabilities of movement, plus a vector and a scalar for predicting regeneration. Their model took the form (Fig. 3.4):

$$\boldsymbol{V}_{i+1} = (\boldsymbol{M} + \boldsymbol{N})\ (\boldsymbol{V}_i - \boldsymbol{H}_i)\ +\ \boldsymbol{C} \tag{3.2}$$

where $\boldsymbol{V}_i$ is the stand table at time i, $\boldsymbol{H}_i$ is the number of trees harvested from each class at time i, $\boldsymbol{M}$ is a standard bi-diagonal Usher matrix, and $\boldsymbol{N}$ and $\boldsymbol{C}$ predict recruitment. The entries in the top row of the sparse matrix $\boldsymbol{N}$ are given by $r_i = 0.27 - 9.65 g_i$ where g_i is the basal area of a tree at the midpoint of class i (all other entries are zero). The vector $\boldsymbol{C}$ is all zero, except for the first element which contains the value 109. These parameters used to model recruitment were estimated by linear regression:

$$N_r = 109 + 0.27 N - 9.65 G$$

where N_r is the total amount of recruitment (stems $\text{ha}^{-1}\text{y}^{-1}$), N is the total stocking (stems ha^{-1}) and G is the stand basal area ($\text{m}^2\,\text{ha}^{-1}$).

Buongiorno and Michie (1980) used a separate matrix to represent the harvest, so that they could efficiently examine alternatives. They found that fixed proportion harvesting would lead to the removal of about 15 square metres of basal area every 35 years, and that simple cutting limit regimes (i.e. harvesting all trees above a given size) were economically optimal.

A similar model for Indonesian forests (Mendoza and Setyarso 1986) indicated that selection logging (i.e. harvesting a proportion of trees in each merchantable size class) would sustain higher yields than simple cutting limit based on diameter. This model also revealed that harvesting practices in Indonesia could not be sustained, as residual stockings were too low to enable the next anticipated harvest in 35 years.

One limitation of transition matrix models is the difficulty of predicting recruitment reliably. Mendoza and Setyarso (1986) assumed a constant amount of regeneration in each time step. Buongiorno and Michie (1980) predicted ingrowth from the number and basal area of trees in each size class, but negative values are possible and the stand table must be inspected at every step to intercept these. Bosch (1971) predicted regeneration only on the death of another tree. Each of these alternatives has limitations that can only be overcome by more explicit methods of modelling regeneration and recruitment (see Chapter 10).

Two major weaknesses of the matrix approaches are the stationary and Markov assumptions. These may be overcome by estimating a new matrix

$$\begin{pmatrix} 0.82 & -0.04 & -0.21 & -0.43 & -0.68 & -0.98 & -1.31 \\ 0.23 & 0.70 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.26 & 0.67 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.30 & 0.65 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.30 & 0.66 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.30 & 0.81 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.19 & 0.86 \end{pmatrix} \begin{pmatrix} 100 \\ 60 \\ 40 \\ 30 \\ 20 \\ 10 \\ 3 \end{pmatrix} + \begin{pmatrix} 109 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 130.8 \\ 65.0 \\ 42.2 \\ 31.5 \\ 22.2 \\ 14.1 \\ 6.2 \end{pmatrix}$$

$$(\boldsymbol{M} + \boldsymbol{N}) \qquad \boldsymbol{V}_0 + \boldsymbol{C} = \boldsymbol{V}_1$$

Fig. 3.4. Single projection with a matrix model (after Buongiorno and Michie 1980). The top row is fecundity, the main diagonal indicates trees remaining in a class, and the second diagonal shows upgrowth. See Equation 3.2.

for each projection interval (to account for new basal area, etc.) either directly from a subset of the database (e.g. Higgins 1977) or indirectly using equations (e.g. Solomon *et al.* 1986), but that eliminates some of the attractions of the approach.

Estimating Matrices

Michie and Buongiorno (1984) discussed four ways to compute the coefficients of a matrix. They concluded that the best method was the simple expedient of tabulating individual tree data in a matrix (i.e. size class at time t versus size class at time $t+1$) and determining upgrowth for each class directly.

This is a robust approach, but fails to utilize the fact that the growth pattern of trees is such that the movement probabilities for adjacent cells of the matrix should be similar. Thus the probability of moving from class $i-1$ to class i should be similar to that for moving from class i to class $i+1$. If not, it is likely that insufficient data were used. This fact can be exploited to estimate entries with greater precision and/or fewer data, by using logistic regression (e.g. Lowell and Mitchell 1987, Vanclay 1991*d*, see Chapter 9).

Where individual trees are not identified, several methods exist for extracting the necessary probabilities of movement (e.g. Carron 1968). However, none of these methods is entirely satisfactory as all assume that the ranking of trees does not change over time, which is an untenable assumption (e.g. Weck 1955; Fig. 3.5, overleaf). It is preferable to estimate matrices from data in which each individual tree is identified, so that the fate of each individual can be unambiguously determined.

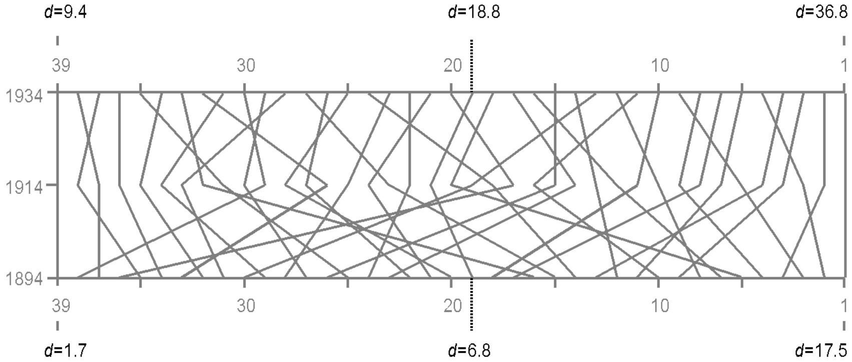


Fig. 3.5. Changes in tree rank in a pine stand aged 69 in 1934 with 820 stems ha^{-1} and mean height 15 m (redrawn from Weck 1955).

Cohort Models

The distinction made here between stand table and cohort models is based on the assignment of trees to size classes. Stand table and matrix models generally use "metric" size classes of equal and predetermined size. Cohort models allow more flexibility, so that empty classes can be eliminated, and that cohorts can be formed so that each cohort has approximately equal importance in the model. Some plantation growth models require equal stem numbers in each cohort (e.g. Alder 1979 used deciles, each with 10% of the trees by number), but most implementations do not have such constraints, and form cohorts from trees that are similar in some sense (e.g. species and size).

Various assumptions can be made about the size distribution of trees within a cohort. Some models assume that all members of a cohort are identical, some assume a uniform distribution, and fit a distribution function across several cohorts. The effect of these assumptions on predicted outcomes depends on the number of cohorts allowed by the model, and may be inconsequential if many cohorts are allowed.

Cohort models are rather logical in many respects, and eliminate many limitations inherent in alternative modelling strategies. Three main components of growth are modelled (cf. Fig. 1.4, p. 9):

1. diameter increment is modelled by incrementing the size of the representative trees;
2. mortality is simulated by reducing the *expansion factor* (the number of trees represented by each cohort); and
3. recruitment is accommodated by initiating new cohorts from time to time.

Percentile-based Models for Plantations

Clutter and Allison's (1974) model for *P. radiata* plantations in New Zealand divides the stand into 25 cohorts, each initially with an equal number of trees. The median diameter for each cohort is derived by fitting a Weibull function to the diameter distribution of the whole stand, and computing the diameters corresponding to the second, sixth, . . ., 98th percentiles. Growth estimates are based on this hypothetical median tree for each cohort. The model assumes that there is no mortality, and that trees do not change ranking, so that median trees continue to represent their cohort throughout the projection, unless thinning is simulated. The size distribution for the stand can be reconstructed at any time during simulations by fitting a Weibull distribution to the 25 hypothetical trees.

Alder's (1979) model for coniferous plantations in east Africa uses deciles, and models the development of the ten median trees corresponding to the fifth, 15th, . . ., 95th percentiles of the cumulative tree size distribution. The initial size distribution is estimated using a Weibull distribution, when the stand attains a dominant height of seven metres. The model is concerned primarily with growth after this height has been attained. Diameter increment is predicted from estimates of height increment derived from height–age curves. It is assumed that there is no mortality, so that no record of trees in each cohort is maintained: it remains constant at ten percent of the total stand stocking. All trees in each cohort are assumed to be identical to the median tree, and no attempt is made to reconstruct a diameter distribution.

Both these models have been used operationally in making predictions for commercial plantation management. These methods may be efficient for even-aged stands, but care is needed in choosing the representative trees. In uneven-aged stands, most of the trees may have small diameters, but the few large trees may be of considerable importance in terms of stand dynamics and commercial value. In such stands, it seems inefficient to form cohorts with equal numbers of trees, and alternatives should be considered.

Cohort Models for Mixed Stands

In mixed stands, the formation of cohorts may be determined by species and other attributes, instead of computational niceties such as equal numbers of trees. Many attributes may be used to form cohorts. Species, size, vigour (e.g. crown illumination) and commercial characteristics (e.g. stem straightness) are all obvious candidates for forest management models, but other factors may also be accommodated. Reed's (1980) succession model forms cohorts from trees of the same species with similar age, height, diameter, leaf biomass, etc. Trees remain in their initial cohort throughout simulation, unless damaged by browsing which initiates a new cohort for the damaged stems.

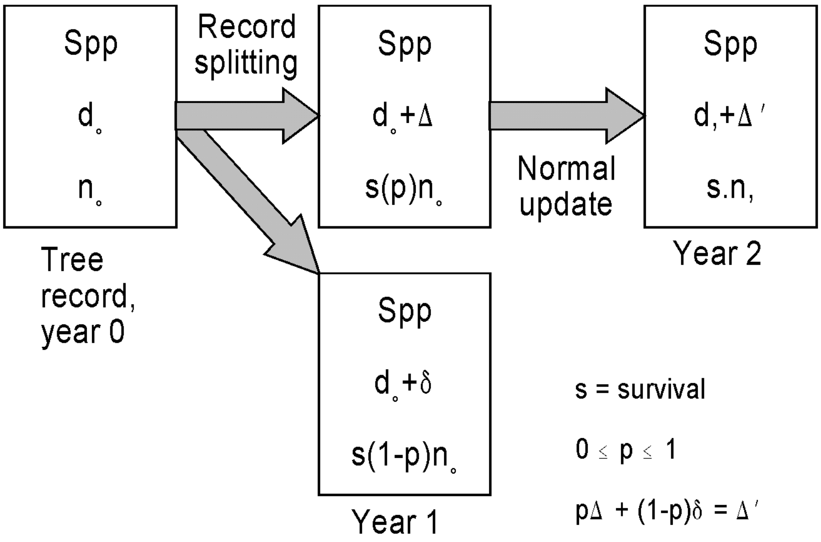


Fig. 3.6. Tree record in cohort model represented by species, size (*d*) and expansion factor (*n*). Records may be split to model variation in growth. New records should maintain mean increments and stem numbers.

Leary's (1979) model allowed varying levels of resolution. At the lowest level of resolution, it uses a single cohort for each species. At the intermediate level, it simulates three cohorts for each species. At its highest level of resolution, each cohort represents an individual tree and the model becomes a *tree list* model (see Chapter 4). Thus the user can select a level of resolution appropriate to his requirements and budget. The model predicts the sum of diameter increments for each cohort, by estimating the potential diameter increment of the mean tree in each cohort (Hahn and Leary 1979), adjusting this for stand density and competition (Leary and Holdaway 1979), and multiplying by the number of trees in the cohort. The list of individual diameters input to the model is not discarded, but is retained and at the end of the simulation each tree is updated by its share of the accumulated increment in its cohort (Leary *et al.* 1979*b*). This allows better estimates of final tree sizes.

Vanclay (1989*a*) described a cohort model which was an early prototype of the NORM model. For efficient simulation, species were grouped according to growth habit, size at maturity and harvesting guidelines (Preston and Vanclay 1988). Species groups were coded so that the appropriate growth (5 groups) and harvesting group (9 groups) could be

Table 3.3. Example of a tree list model showing record doubling and merging.

Field record†	Group identity‡	Cohort list as projected					
		Year 0		Year 1		Year 2	
		N/ha	Dbh	N/ha	Dbh	N/ha	Dbh
NSO 41	322	7.39	41.5	1.85	41.97	1.85	42.30
				5.54	41.83	1.38	42.26
						4.15	42.13
BLA 49	374	5.20	49.5	1.29	49.92	1.29	50.22
				3.88	49.79	3.86	50.09
NSO 26 defect	492	18.13	26.5	18.10	26.80	18.07	27.08
BRC 68 defect	495	2.71	68.5	2.69	68.85	2.67	69.18
MCB 42	495	7.05	42.5	6.99	42.82	6.94	43.13
MCB 36	495	9.56	36.5	9.48	36.81	9.41	37.09
MIS 24 MIS 16	495	21.21	24.5	21.04	24.76	20.87	24.99
MIS 16 MIS 16	495	140.3	16.5	139.1	16.71	138.0	16.90

† Tree species and size (cm dbh) of all trees recorded on a single point sample defined with a 10 m^{2}ha^{-1} angle gauge. Species codes are: BLA: blush alder (*Sloanea australis*); BRC: Brown cudgerie (*Canarium baileyanum*); MCB: Macintyre's boxwood (*Xanthophyllum octandrum*); MIS: various species without specific codes; NSO: northern silky oak (*Cardwellia sublimis*).

‡ The 3-digit code indicates the volume equation (4 = no commercial volume), the logging prescription (9 = not harvested), and the growth equations to be used for this species group during simulations.

identified easily. Each cohort was characterized by its species group code, diameter, and expansion factor. The model admitted a maximum of 200 cohorts for each stand simulated, and the actual number of cohorts was maintained near this limit by allowing record doubling (e.g. Fig. 3.6, and Table 3.3, first record, years 1 and 2) and merging (e.g. Table 3.3, last record, year 0). When records are doubled, one part gets a bigger-than-average increment, the other a less-than-average increment, but the proportions and increments are determined to preserve the total number of trees and the mean increment (Fig. 3.6). Similarly, when records are merged, the expansion factors are summed and the new mean tree size is a weighted average. Cohorts of small trees may contain many trees; as they attained a size of some economic importance they could split into new cohorts reflecting the actual increment distributions observed in the permanent plot data. Cohorts of overmature trees were merged as their expansion factors reduced through mortality and harvesting. Recruitment was modelled by forming additional cohorts. The model was subsequently enhanced so that each species retained its individual identity.

Process-based Cohort Models

The cohort framework has also been used for process-based models. One such model for even-aged stands of red pine in the USA, used ten cohorts and modelled tree growth in terms of diameter at one-fifth of tree height ($d_{0.2}$), tree height, and height to crown base. Sievänen and Burk (1993) examined ways to fit the model using dynamic inventory data, and how to calibrate it for various site conditions.

The model estimates photosynthesis at the tree level by assuming a linear interpolation between open growth (cf. Newnham's 1964 assumptions on p. 62) and the tree's share of closed-stand photosynthesis. Change in $d_{0.2}$ was estimated from a rather complicated equation which adjusted the balance between photosynthesis, respiration and shedding (of leaves, branches and fine roots through senescence) for the effect of tree size (e.g. bigger trees allocate more photosynthates to the non-productive stem, branches and coarse roots so that diameter increment is less for a given amount of photosynthesis).

The relative height growth within a class assumed a linear relationship with relative crown length, to account for the fact that trees with short crowns tend to have proportionally higher height growth. The crown was assumed to recede at a constant rate, determined by stand basal area at the crown base, for all trees in the stand. Mortality was predicted from the diameter increment.

The model required 15 parameters to be estimated, but these could not all be estimated empirically from the available data. This was partly due to the structure of the model (i.e. it was over-parameterized) and partly due to limitations of the data available. The most sensitive parameters were the maximum photosynthesis rate, foliage respiration, and the rate at which the crown base rises. Poor estimates for the first two of these could result in substantial bias, but suitable estimates may be obtained from other sources.

Sievänen and Burk (1993) identified several inadequate components in the model, including height growth, crown recession and tree survival. The assumption that all trees in a cohort are identical was unrealistic, especially when only ten cohorts are used. Local variations in stand density need to be accounted for if the model is to provide realistic predictions in the longer term.

The model could not be calibrated adequately from diameter measurements alone, and Sievänen and Burk (1993) observed that reliable calibration required remeasurements of numbers, diameters, heights, and crown length for individual trees on permanent plots.

Synthesis

Size class models have been used extensively to model growth and yield in many forest types ranging from pure even-aged plantations to tropical rainforests. This approach offers several advantages: it is relatively simple, computationally efficient, and provides information in sufficient detail for many forest management applications. Classical stand table projection continues to be useful where stand data are available only in summarized form and where computer resources are limited. However, the many species and wide range of stem sizes encountered in some mixed forests may require many classes, and this creates some limitations.

Matrix methods are easily implemented on computer, and produce good results where stand densities and silvicultural practices are maintained within a narrow range, but suffer from assumptions which become untenable for long projections and for wider ranges of stand conditions.

Cohort models offer great flexibility, enable projections under a wide range of conditions and provide diverse information for reporting. They offer great potential, and are likely to be used more widely in modelling mixed forests.

In the next chapter, we examine single-tree and tree list models. The latter are analogous to cohort models with a single tree in each cohort.

Exercises

3.1. The data below summarize from two measures of a permanent plot, five years apart. Using these data, predict the stand structure in 25 years time (assume that there is no mortality and no recruitment). Try at least three different methods (including stand table projection), and explain why they differ. Which is the best estimate? Why, and what are its limitations?

d_0	d_5	d_0	d_5	d_0	d_5	d_0	d_5
14.8	18.0	16.1	20.3	20.5	24.8	31.1	35.6
14.9	18.5	16.3	20.2	23.1	28.8	32.6	38.2
15.2	19.8	16.5	20.3	23.6	28.9	37.6	41.2
15.7	18.7	18.5	22.9	26.3	30.5	41.9	46.4
16.0	19.3	19.5	24.2	29.7	35.2	42.0	45.5
						47.5	50.9

3.2. Does the model in Fig. 3.4 behave in a biologically realistic way? Project the model through 100 cycles, and discuss the trends you see. Do long term predictions depend on the initial state vector? Do they converge to a stable stand structure? Summarize its strengths and weaknesses. Is it a good model? How does it compare with other approaches discussed in this chapter?

3.3. Osho (1991) used an Usher matrix to model the growth of Nigerian rainforests (below, the stand table related to a single plot 1.47 ha in area). Is it a good model? Project the stand for 400 years (the matrix represents a 18 year time step): what do you notice? Suggest a sustainable harvesting strategy for this forest.

Size class (cm dbh)	1974 Stand table	Diameter classes (cm dbh) of Usher matrix					
		4.8–9	10–19	20–29	30–39	40–49	50+
4.8–9	914	0.42	0.00	0.73	2.24	5.04	7.42
10–19	494	0.06	0.62	0.0	0.0	0.0	0.0
20–29	158	0.0	0.05	0.66	0.0	0.0	0.0
30–39	76	0.0	0.0	0.14	0.79	0.0	0.0
40–49	19	0.0	0.0	0.0	0.20	0.62	0.0
50+	26	0.0	0.0	0.0	0.0	0.21	0.89

3.4. Of the models discussed in this chapter, which do you think would be most suitable for your situation? Why? Summarize its strengths and weaknesses, and contrast it with other alternatives. What resources would you need to build (and use) this model?

3.5. Does the process-oriented basis of the FORMIX model (p. 41) provide a firmer foundation for extrapolation and inference than is provided by the other empirical models discussed in this chapter? Discuss.

3.6. Several size class models (e.g. FORMIX p. 41, Fig. 3.4 on p. 49) do not converge to a semi-stable state, but cycle (e.g. see solution 3.2). Do you think that this accurately reflects forest dynamics, or is it possible that the resolution of the models is too coarse so that natural feedback mechanisms overcorrect and cause this cycling? Discuss. How could you test this?

Chapter Four

Single-tree and Tree List Models

As the class width becomes smaller and the number of trees per cohort approaches one, the distinction between size class approaches and single-tree models becomes blurred, especially for tree list or cohort models. The distinction between *single-tree models* and *tree list models* is based on the use of an *expansion factor* which indicates the number of stems (per ha or per plot) represented by each tree record. Single-tree models are defined as those models which simulate each individual tree for a given plot (i.e. implicit expansion factor is always exactly 1.0), whereas tree list approaches may simulate any number of trees in a cohort (i.e. explicit expansion factor can be any real number >0). This distinction has significant implications for the processing of inventory data obtained from point samples (sampling with probability proportional to size which may imply fractional expansion factors), and for the modelling of mortality. Many single-tree models simulate mortality stochastically to maintain expansion factors of exactly one per plot, whereas tree list approaches may deterministically reduce expansion factors so that they represent less than one tree per plot.

This chapter will focus on three classes of model:

1. *Single-tree spatial models* which simulate individual trees or their component parts (crowns, branches, etc.) using spatial data (e.g. stem maps) indicating their position in the stand. The requirement for spatial data usually restricts these models to simulations of relatively small plots. These models are also known as distance-independent models, but the term "spatial" is preferable as they may require three-dimensional spatial data, not just the distance to neighbours.
2. *Single-tree non-spatial models* which also model individual trees, but which do not require any spatial data. These may model stand development on a per-hectare or a per-plot basis (e.g. the JABOWA model, with 10 × 10 m plots).
3. *Tree list models* which model cohorts or small "groups" of trees. A cohort may constitute any number of trees within any area (i.e. expansion

factor may be any real number), and is usually based on similarity in some sense (e.g. species, size, etc.) rather than on number of trees. The distinction between tree list models and the cohort models discussed in the Chapter 3 is one of convenience, and is based on the number of cohorts and the average number of trees in a cohort. A tree list model would normally allow more than 25 cohorts (cf. Clutter and Allison 1974, p. 51) and more than five species groups where many species are present (cf. Vanclay 1989*a*, p. 52).

Single-tree Spatial Models

Spatial models use information about the position and size of neighbouring trees to simulate the development of each individual tree in a stand (or plot). The development and practical use of these models have been hampered by a lack of suitable data, but they offer considerable potential. Development of these models has helped to formulate some important concepts of competition, and we survey some of these competition indices, before examining some examples of single-tree spatial models for pure and mixed stands.

Competition Indices

Current thinking about competition between trees in forest stands may be summarized in five axioms (Ford and Sorrensen 1992):

1. Plants modify their environment as they grow, reducing the resources available for other plants (*competition*).
2. The primary mechanism of competition is spatial interaction.
3. Plant death due to competition is a delayed reaction to the growth reduction following resource depletion.
4. Plants adjust to environmental change, responding to competition and altering the nature of the competition.
5. There are species differences in the competition process.

Many modellers have attempted to quantify these concepts concisely in an index of competition, but there is no single index that satisfactorily embodies these concepts and performs well in empirical trials. Most competition indices that have been proposed can be classified into four broad categories:

1. competitive influence zone,
2. area potentially available,
3. size–distance (including horizontal and vertical variants), and
4. sky-view and light-interception approaches.

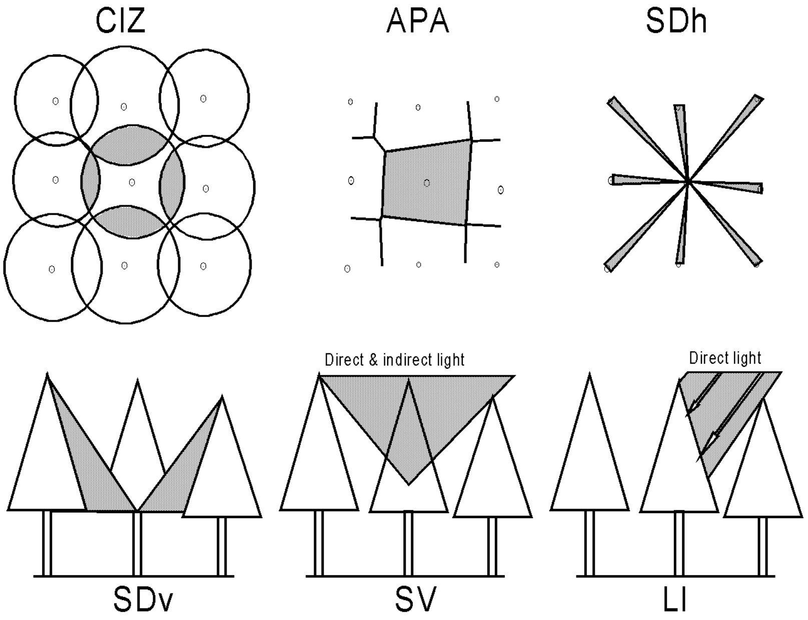


Fig. 4.1. Competition indices include the competitive influence zone (CIZ), area potentially available (APA), horizontal or vertical size–distance (SDh & SDv), sky view (SV) and light interception (LI) approaches.

The competitive influence zone (or area-overlap; Fig. 4.1, CIZ) approaches are based on the assumptions that

(a) each tree has an influence zone determined by the size of the tree, and that

(b) the competition experienced by each tree can be gauged by the potential overlap of the influence zones.

The competition index may represent the area of the potential "overlap" (e.g. m^2) or it may be scaled as a relative area (e.g. ratio between 0 and 1), and may be weighted by tree size or other factors (e.g. Ek and Monserud 1974). A non-spatial variant of this index is crown competition factor (Krajicek *et al.* 1961), which expresses the potential open-grown crown area of all trees as a percentage, so that 100% occurs about the time of crown closure and higher values indicate more competition.

The area potentially available (e.g. Moore *et al.* 1973; Fig. 4.1, APA) to each tree is usually calculated by sharing the total (simulated) plot area among the trees according to their size and position. This may be done by bisecting perpendicularly the distance between each tree and its neighbours, often using a weight according to tree size. Mixed stands may require that weights are adjusted for each species (see e.g. Tham 1989, Pretzsch 1992*a*).

In practice, it may be quite complex to ensure that the polygons do not overlap and that no unrealistic gaps remain, especially when size-weighted indices are computed for uneven-aged stands, but standard algorithms are available (e.g. Nance *et al*. 1988). Comparisons with open-grown trees are necessary to constrain the area potentially available in stands with low stocking. Three-dimensional analogues (e.g. volume potentially available to the tree crown) also exist (e.g. Pelz 1978). Empirical studies (e.g. Daniels *et al*. 1986, in *Pinus taeda* plantations) suggest that the area potentially available, weighted by tree size, is one of the better competition indices available for predicting basal area growth in plantation monocultures.

Size–distance approaches (e.g. Lemmon and Schumacher 1962, Spurr 1962*a*; Fig. 4.1, SDh and SDv) compute an index of competition based on the size of and distance to each potential competitor. The index may be based on the sum of the angles subtended by potential competitors at the subject tree (e.g. Ford and Diggle 1981), may include a threshold (i.e. competitor only if the angle is exceeded), and may be weighted for the size of competitors. Several characteristics (e.g. crown cross-sectional area, crown volume) of competitors have been incorporated in competition indices, with or without weighting for distance and relative size (e.g. Biging and Dobbertin 1992). The index may be computed in the horizontal (e.g. based on stem or crown diameters of neighbouring trees) or the vertical plane (e.g. based on heights of neighbouring trees). In either case, the reference height may be varied to obtain more meaningful estimates (e.g. breast height, crown base of subject tree, top of subject tree, etc.). To limit edge-effects and computation time, it is necessary to specify a search-radius, within which competition is appraised, and the selection of a suitable radius may be assisted by data from open-grown trees.

Ford and Diggle (1981) selected competitors as those individuals taller than a 45° angle from the top of the subject plant, and based their index on the angle subtended by competitors, but many others choose in include other characteristics in the index. For example, one index evaluated by Biging and Dobbertin (1992) defined competitors as trees within a given search radius for which the angle from the subject's crown base to the competitor's tip exceeds a specified threshold (e.g. Fig. 4.1, SDv), and based the competition index on the sum for each competitor, of crown cross-sectional areas divided by distance.

Sky-view approaches determine the proportion of the sky "seen" by each tree, sometimes weighting parts of the sky differently (e.g. more weight for the sky overhead, less for the horizons). The reference point does not need to be the top of the tree, but may be the centroid of the crown (e.g. Fig. 4.1, SV), or may involve an adjustment to account for the greater photosynthetic activity of newer foliage, so that the reference point may vary according to recent crown development of the subject tree. The most complex variant of this approach involves computing the interception of sunlight by each tree, adjusted for time of day and season of year (e.g. Fig. 4.1, LI). This may be a complex undertaking, and a simplification is to use the noon sun position, which may give a result similar to a more conventional weighted sky-view method, especially in the tropics.

Empirical studies (e.g. Opie 1968, Lorimer 1983, Martin and Ek 1984, Barclay and Layton 1990) suggest that competition indices rarely provide better estimates of increment than simple measures of stand basal area, and that the expense of determining individual tree positions in the stand is rarely warranted for the purposes of yield forecasts. This does not imply that there is no place for spatial models. On the contrary, they have offered important insights into competition and how to model it, and have been useful in researching aspects of plantation silviculture.

Some limitations are evident in most studies of competition in forest stands, and these may contribute to our present inability to define a general competition index. Some deficiencies include:

1. The performance of competition indices is intimately linked to the growth functions with which they are used, especially where it is used as a multiplier to modify a potential growth function (see Chapter 8). An inferior potential growth function may make the performance of a competition index look better if it is of the right general shape, but explains only some of the variability. Conversely, if the potential growth function is of the wrong "shape", it may understate the performance of competition indices.

2. Results may be influenced by the plot size used, and by the assumptions used in estimating the competition experienced by trees near the edge of the plot. Monserud and Ek (1974) and Martin *et al.* (1977) reviewed ways to reduce this "plot edge" bias.

3. Most studies draw on data from pure even-aged stands and do not include thinnings, and these do not provide the best test of a competition index (notable exceptions include Lorimer 1983, Tham 1989, Biging and Dobbertin 1992). Where there is no thinning, future conditions will be more-or-less like past conditions, and since present tree size (diameter, height, crown ratio, etc.) is determined by past conditions, it may be the best indicator of future growth. The only fair test of a competition index is to disrupt the correlation between tree size and growth conditions, e.g. by thinning from above, by pruning the green crown, and perhaps by removing the tops of some trees. Mixed stands may provide a good test, since a faster-growing species may become dominant and shade a formerly-dominant tree (this could be examined experimentally in pure stands by using a large umbrella to shade a dominant tree). It is unlikely that a single time interval would be sufficient to appraise a competition index properly following such experimental disturbance. Simple indices may not adequately predict recovery from competition when a competitor is removed (e.g. by thinning), and it may be necessary to include crown parameters to ensure a reasonable response.

One of only few studies in mixed stands (Biging and Dobbertin 1992) found several spatial competition indices that did perform better than stand basal area, and reported that these indices performed better for shade tolerant (e.g. white fir) than for light demanding species (e.g. ponderosa pine, in California, USA). Distance-weighted measures of crown cross-sectional area were amongst the best indices for use in predicting both height and diameter increment, especially when competitors were selected by height (i.e. a vertical size–distance approach, Fig. 4.1). Unlike many other studies, they found that increasing the search radius did not improve estimates, and that a relatively small number of competitors provided an adequate estimate. Biging and Dobbertin's (1992) analysis offered some interesting results, but some questions remain. Their data represented a single five-year growth period reconstructed using stems analysis, and did not specifically include thinning. Their potential growth function was based on site index, initial tree size and crown ratio only, and they did not compare the performance of non-spatial measures such as relative tree size, when used in conjunction with stand basal area. It is not clear if a more comprehensive basic growth function would eliminate the effect of the spatial data, or if the results would be changed by thinning.

In their present stage of development, competition indices are of limited use for management-oriented models of mixed forests. However, there remains considerable scope for further research in this area.

Models for Pure Stands

Several single-tree models have been developed for pure even-aged stands, and range in complexity from tree-level to process-based models. Here we examine just one prominent example from each class, and consider some possible implications for modelling mixed stands.

Tree-level Models

Newnham's (1964, Newnham and Smith 1964) model for even-aged stands of Douglas-fir contained three important assumptions which have provided the basis for many similar models:

1. a tree free of competition has the diameter growth rate of an open grown tree of equal diameter,
2. a tree subject to competition has its increment reduced by an amount proportional to the level of competition, and
3. mortality occurs when diameter growth falls below a threshold level.

For computational ease, Newnham assumed that competition extended only as far as eight times the initial spacing, and this led to bias in estimates of yield from closely spaced stands. Generally, Newnham's model gave reasonable results, and many other models have been based on this approach. However, Larocque and Marshall (1988) argued that whilst the overall performance of Newnham's model has been demonstrated to be satisfactory, the three basic assumptions have never been tested independently, and remain unconfirmed. Although these three assumptions are clear simple statements, they are difficult to test because competition cannot be measured directly. The third assumption regarding mortality remains controversial (see Chapter 9).

While competition indices and single-tree models allow explicit and testable hypotheses of many aspects of tree growth, they also pose several problems in accommodating unexplainable variation in growth. Ignoring this variability by using general trends may produce untenable results (e.g. a plantation at clearfall age may comprise trees all of identical size). One way to overcome these problems is to simulate some components stochastically.

Crown-level Models

Mitchell (1969) was one of the first to accommodate unexplained variation explicitly by simulating growth in a stochastic way in a single-tree spatial model. He assumed that, for even-aged stands of white spruce in Canada:

1. height of any tree can be predicted from its relative crown width (compared with open-grown trees) and the height of dominant trees,
2. longitudinal growth of the main axis of all branches is related to concurrent height growth, regardless of suppression, and thus that the radial growth of tree crowns can be predicted from height growth subject to space limitations imposed by competing trees,
3. tree diameter (dbh) and bole volume can be predicted from tree height and crown width,
4. suppression and mortality can also be predicted from relative crown width, and that
5. stochastic modelling of branch length propagates through other model components and accounts sufficiently for all natural variation in the stand.

Mitchell's model is driven by height growth. Branch length was predicted as $l_i = \beta\, h_i^{0.75}$ where l_i is the length (m) of the i^{th} branch, h_i is the distance (m) from the base of the i^{th} branch to the top of the tree, and β is normally distributed with mean 0.548 and standard deviation 0.093. This variation influences crown width, and thus propagates through subsequent height growth predictions and all other components of the model. Although this approach appears to work well for even-aged stands of white spruce (Mitchell 1969) and Douglas-fir (Mitchell 1975), it has limited utility for modelling the yield of natural forests since height growth is not a feasible driving variable in many of these forests, and stem maps are impractical in most operational inventories.

Branch-level Models

Mitchell's (1969, 1975, 1980) ideas have influenced several other models. For example, the PLATIPUS (Plantation Timber Production and Utilization System) simulation model for coniferous plantations in Queensland was designed to simulate:

1. growth of stands under a wide range of environmental and silvicultural conditions, enabling analysis of forest management options;
2. effects of most silvicultural options (site preparation, spacing fertilizing, weed control, pruning, thinning and fire) on growth and on wood characteristics; and
3. conversion of trees and logs into various products (veneer, sawnwood, chips, pulp, etc.) and determine their intrinsic value.

The PLATIPUS model simulates branch and crown dynamics in greater detail than the models of Mitchell (1969, 1975, 1980), so that knot characteristics could be inferred. Height increment was derived from standard site index (height–age) curves. Potential stem volume increment was predicted from light interception, crown area and competition (Vanclay 1988*b*):

$$\ln \Delta v_u = -6.07 + 3.38 \ln l_s + 0.382 \ln c_a + 0.879 \ln c_l - 0.0219 G \qquad (4.1)$$

where Δv_u is stem volume increment (m^3 y^{-1}, under bark), l_s is a simple light index based on the sky-view method, c_a is the horizontal projection of crown area (m^2), c_l is crown length (m) and G is stand basal area (m^2 ha^{-1}). The light index is a simple measure of the relative area of sky "seen" by the tree, determined as the mean of the cosines of the angles from the effective centre of the crown (empirically determined as tree height minus 0.24 times crown length) to the tips of the nearest neighbour in each of four cardinal directions (note that in Queensland, the sun is effectively overhead at noon during the growing season). This potential increment is modified to account for site and silviculture (e.g. fertilizing), and partitioning rules are used to determine the diameter increment at various points along the stem. Equation 4.1 is regarded as preliminary, as it was fitted with very few data. However, the data were drawn from a thinning and pruning trial which provided a reasonable range of data for each of the explanatory variables. Despite the limited database, the equation seems to give adequate predictions over a wide range of stand conditions.

Branch elongation is predicted from the stem diameter increment at the base of the branch, the number of branches in the whorl and the relative position of the whorl in the tree, but elongation ceases when a branch touches another branch. Branch diameter at base is predicted from branch

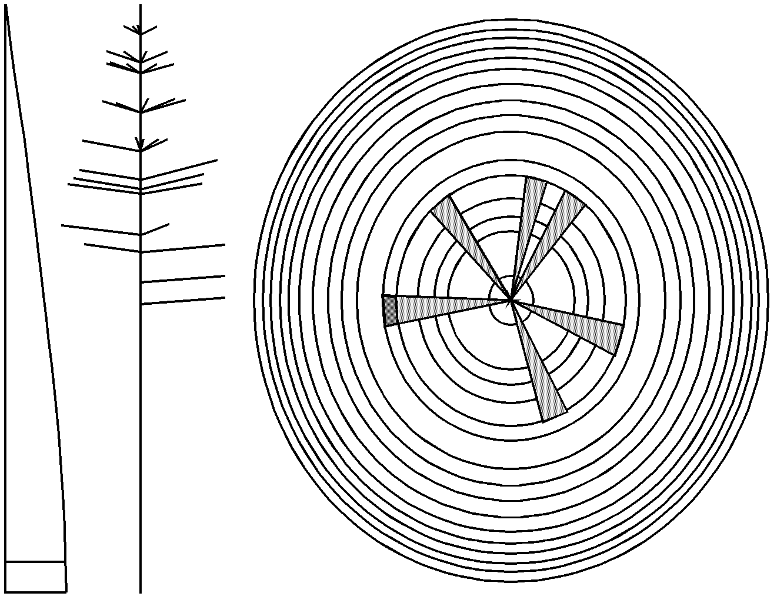


Fig. 4.2. Outputs from the PLATIPUS prototype included stem and branch profiles, and ring and knot patterns (redrawn from Bragg 1988). Notice that where pruning is not flush with the stem, dead knots may result.

length. Most functions in the model are stochastic, and the variance-covariance matrix is used to ensure appropriate correlation between stochastic components. The computational cost is minimized by modelling a three-dimensional matrix of cells, and assigning "ownership" of a cell to the tree that first occupies it (i.e. when a branch first enters a cell). Preliminary trials with a prototype version containing several subjectively-determined relationships indicated reasonable results (Nielsen 1989). The prototype (Bragg 1988) provided detailed information on each tree simulated, and could produce graphical output illustrating tree profiles and branch, knot and ring patterns for any tree in the stand (Fig. 4.2).

Mechanistic Models

West (1987) devised a framework for a spatial, process-based model for trees in pure stands. His model is an extension of the general stand-level mechanistic model of McMurtrie and Wolf (1983). West modelled stem biomass, and leaf and root biomass by annual classes, assuming that they would live for three and two years respectively. Gross photosynthesis of an individual tree is predicted from the site's potential gross photosynthetic production (per unit area), multiplied by the leaf area of the tree and an empirical modifier to account for shade within and between trees.

Respiration (a constant times leaf biomass) is subtracted from this to give net photosynthesis, available for maintenance and growth of other tree parts. The model assumes that trees die when net photosynthesis falls to zero.

The modifier is estimated from the relative leaf weight of trees within the "locality" of the subject tree, compared with leaf weights at canopy closure, and is thus analogous to the common empirical measure of crown competition factor but also takes into account crown density. The locality is defined as a circle with radius equal to the distance to the sixth nearest neighbour, consistent with the proposition that competition is restricted to a few neighbours of each tree (Cannell *et al.* 1984). Trees within this locality and with a stem biomass at least 70% of that of the subject tree are assumed to be competitors. This is consistent with the proposition that competition is one-sided (i.e. larger plants shade smaller plants, but not vice-versa), but recognizes that since tree crowns are three-dimensional, partial shading of a subject tree crown may be caused by a neighbour not quite as large as that of the subject tree. Many of the parameters required for the model were obtained from the literature, whilst others were subjectively determined. Trials with the model provided reasonable predictions during a 50-year simulation commencing with a simulated seedling stand. The model predicted realistic responses to different thinning and spacing regimes.

One of the more empirical aspects of many process-based models has been the partitioning of photosynthates between leaves, roots and stems. For example, West (1987) assumed that 20% of net photosynthates would be used for new leaves, 20% for stem and branch development, and 60% for root growth. West (1993) developed the model further to examine more realistic ways to model photosynthate partitioning in response to functional relationships between tree parts. He assumed that the general growth strategy of trees is to maximize leaf production subject to a few constraints. The constraints he examined included:

1. Stem diameter and height are related by the need for the stem to maintain structural stability and to support the crown. West assumed an allometric relationship:

$$d = \alpha \left(\frac{b_c}{b_s}\right)^{\beta} b_t^{\gamma}$$

where d is stem diameter, and b_c, b_s and b_t are the biomass of crown (leaves plus branches), stem and in total (above ground). He also assumed a similar relationship between height and biomass components.

2. Sapwood area must maintain water supply to the leaves in accordance with the "pipe-model" theory, so that

$$b_l = \beta g_s c_s$$

where b_l is leaf biomass, g_s is the sapwood area, and c_s is sapwood conductivity.

3. Leaf area within a tree cannot exceed a given maximum density (i.e. the size of the crown dictates the maximum biomass of leaves).

4. Branches must support the weight of leaves they bear, so that

$$b_b = \alpha b_l^{\beta}$$

where b_b and b_l are branch and leaf biomass respectively.

5. Height growth cannot exceed that defined by site index curves. Simulations without this constraint would allow excessive height growth in some trees that would become emergents (this strategy is exploited by some species in mixed stands, e.g. *Araucaria*, but is unrealistic in a pure stand).

West (1993) used linear programming to maximize leaf production subject to these constraints. Partitioning coefficients were about 0.8, 0.1 and 0.05 for stems, leaves and branches, but varied with tree status. These coefficients are not directly comparable to those used in other studies, because the model did not include roots. The model required several other assumptions, including

(a) a maximum crown radius,
(b) a maximum attainable increment in crown volume,
(c) a light threshold below which leaves could not survive, and
(d) that the horizontal crown radius could not exceed the vertical radius.

The model provided reasonable predictions for an even-aged *Eucalyptus regnans* plantation in Victoria (Australia). The general principles adopted in the model appear well founded and warrant further development.

Models for Mixed Stands

Fewer single-tree spatial models have been proposed for mixed stands. Ek and Monserud's (1974) model was one of the first spatial models for mixed forests. Like Mitchell's model, their model used height rather than diameter, as the key variable. Potential height increment of any tree was assumed equal to that of a dominant tree, and potential diameter increment was the corresponding diameter increment of an open grown tree of the same height. These potential increments were reduced for individual trees according to their crown ratio and competition experienced. Stochastic variation introduced into both functions was assumed to be normally distributed, without heteroscedasticity or serial correlation. Mortality was modelled using a threshold increment dependent on tree size. Regeneration was modelled from seed in a sub-model and recruited to the main model when it reached breast height.

Pretzsch (1992*a*) used the cellular matrix approach (p. 65) to model mixed beech-spruce stands in Germany. He modelled height growth, crown development (width and length), diameter increment and survival of all trees on a 5×30 m plot. Estimates are based on potential height growth (i.e.

site index curves) and the corresponding open-grown crown width, adjusted for competition and tree status. Crowns are assumed to take a standard conical (spruce) or dome shape (beech). The principal variables in predicting the modifiers are the lateral crown restriction (a competitive influence zone index based on crown width), and shading by each of the two species in the model. Crown surface area and volume are also used in calculating diameter increment and tree survival. The height increment function for spruce (unthinned) is (Pretzsch 1992*a*):

$$\Delta h = \Delta h_{\max} \times c_r^{0.1} \times 0.84e^{0.378c_{iz} - 0.208c_{iz}^2} \times e^{-0.629l_s - 0.465l_b}$$

where Δh is height increment, Δh_{max} is the potential height growth, c_r is crown ratio, c_{iz} is an influence-zone competition index based on crown width, and l_s and l_b are shading due to spruce and beech respectively. The shading is calculated using a sky-view (or cone-of-light) approach. An additional modifier is included following thinning to account for changes in competition and shading.

At their present stage of development, single-tree spatial models may not be well suited to modelling natural forests, but they are interesting in that they represent the state-of-the-art of growth model development, at least for empirical models, and so may indicate future directions if certain limitations (e.g. availability of spatial tree data) can be overcome. In planted stands, it is often possible to overcome the absence of spatial data by using the nominal initial spacing or stochastically generating tree co-ordinates, but this option is not viable in mixed forests. Unfortunately, the requirement for spatial data will not be easily overcome, and this may be why many physiological modellers have chosen size class models, rather than single-tree models, as a point of departure from empirical modelling. Nonetheless, this approach may be poised for further development, especially as computing and remote sensing developments overcome the limitations of this strategy.

Single-tree Non-spatial Models

Spatial growth models offer potential for detailed investigations of silvicultural alternatives in intensively managed plantations, including aspects not possible in other modelling approaches. However, there are some limitations restricting the utility of the approach for mixed forests:

1. permanent plot data rarely contain the detailed measurements necessary for formulating such models;
2. the cost of obtaining such detailed data restricts the application of such models to research applications rather than yield predictions and other practical applications; and

3. competition indices which provide the basis for most of these models rarely perform better than stand level measures of competition such as stand basal area.

In short, the expense of such a detailed methodology may not be warranted, and non-spatial methods offer a feasible alternative for the provision of growth forecasts and other information for forest management.

Deterministic Models

Opie's (1972) model for even-aged *Eucalyptus regnans* is an example of a deterministic single-tree model, which comprised two parts. The first fifteen years were modelled using a whole stand approach, after which individual tree diameters were estimated and subsequently modelled using a single-tree approach. The annual cycle of diameter increment (allowing for heteroscedasticity and serial correlation), tree death, and optional thinning was implemented through seven key functions. These included a tree height–age function, a basal area increment function, an increment allocation rule, a height–diameter function, and a stocking guide (Reineke's line). The model was subsequently enhanced (Campbell *et al.* 1979), and continues to form the basis of forest management models in Victoria and elsewhere in Australia (Rayner and Turner 1990). Other applications of this approach have been confined largely to even-aged stands.

Stochastic Models

The concept of the mean and its variance are well established in experiment methodology, and few researchers would formulate policy recommendations based solely on the mean experimental result. Most decisions are based not only on the mean but also on the variability of estimates. Decisions may also be based on appraisals of risk and of economic criteria. However, in growth modelling for timber yield prediction, this appraisal of variability and risk is generally lacking. Most modelling methodologies developed for timber yield prediction are deterministic. The few models that are stochastic seem to have been developed because the modeller was unable to resolve some function satisfactorily, and not primarily to evaluate variation in estimates of timber yield. This seems naive on the part of modellers and resource managers, particularly when it is not uncommon for forest owners to make long term commitments for timber supply, and where a failure to meet this commitment may be embarrassing to both parties.

Many models use a single stochastic function in one component of the model, most commonly the mortality function (e.g. Moser 1972, Shifley *et al.* 1982). Others use a stochastic diameter increment function. Stage (1973) assumed that the only significant stochastic effect was on the diameter increment; he felt that this would have sufficient repercussions on all other functions in the model. The one stochastic component in the JABOWA model (Botkin *et al.* 1972, Botkin 1993) was the number and species of trees recruited each year. Alder *et al.* (1977) and Vanclay (1991*d*) suggested models in which all relationships derived from regression analyses contain a stochastic component. The difficulty with these approaches is that care must be taken to preserve the appropriate correlation between stochastic elements. For example, it is likely that the random variation in diameter increment is correlated in time, and with other components that may be modelled (e.g. height increment). However, Meldahl (1979) found that varying the correlation between predicted height and diameter increment errors had little effect on simulated stand means and distributions.

The JABOWA model (Botkin *et al.* 1972, Botkin 1993) was one of the first stochastic non-spatial single-tree models. It is not concerned with accurate estimates of timber yield, but with evaluating the modelling concept and providing reasonable predictions of species succession. JABOWA models forest stands on 10×10 metre plots, so that all trees on the plot are competitors, avoiding the need to use tree co-ordinates. Simulating a plot size approximately the same dimensions as the crown of a mature tree, allows the model to reproduce the *gap dynamics* that occur after a large tree dies, especially the establishment of seedlings in the "gap" created where the tree was. The growth function is entirely deterministic (and rather subjective), but mortality and regeneration are stochastic. Growth estimates are based on a "fundamental growth equation" for the change in volume of an individual tree:

$$\delta(d^2h) \propto a_{leaf}\left(1 - \frac{d}{d_{max}}\frac{h}{h_{max}}\right)$$

where d and h are tree diameter and height, d_{max} and h_{max} are the maximum diameter and height attained by that species, and a_{leaf} is the leaf area of the tree. Two assumptions, that a simple quadratic relationship exists between tree height and diameter, and that a tree attains two-thirds of its maximum size at half its maximum age, allow this equation to be expressed:

$$\Delta d = \frac{\beta_0\, d\left(1 - \frac{d}{d_{max}}\frac{h}{h_{max}}\right)}{\beta_1 + \beta_2 d - \beta_3 d^2} \qquad (4.2)$$

where the βs are $\beta_0 = 5h_{max}\Delta d_{max}/d_{max}$, $\beta_1 = 2h_b$, $\beta_2 = -6(h_{max} - h_b)/d^2_{max}$, and $\beta_3 = 4(h_{max} - h_b)/d^2_{max}$. Note that Δd_{max} is the maximum diameter increment attained by the species, and h_b is breast height. This means that the fundamental growth equation can be formed from estimates of maximum diameter increment, diameter and height attained by each species, and that no other data are required. This fundamental equation was further modified by a specific light response, a soil fertility index, and other site factors.

Mortality predictions were based on two functions, one for suppression and one for random deaths. Both functions predicted the probability of an individual tree's death, and a pseudo-random number generator determined whether the tree actually died in that year. Random numbers were also used to decide both the number and species of trees recruited each year. The expected outcome of the model was based on the mean of 100 simulation experiments.

The JABOWA model was originally devised for mixed hardwoods in the New England region of the USA, but has been adapted for many other forest ecosystems (Botkin 1993). Two applications are relevant to tropical moist forests, the Kiambram model (Shugart *et al.* 1980) for sub-tropical rainforest in Australia, and Doyle's (1981) model for central America.

Alder *et al.* (1977) proposed a stochastic non-spatial model (GROPE) which could be used to model any forest for which plot measurements on at least two occasions exist. The approach assumed that all functions used in the model could be transformed and expressed as simple linear equations with normal error distributions. It was hoped that the cycle of model fitting, testing and application could be automated, so that the naive user could obtain forecasts based on any collection of permanent plot data. It is not surprising that this model has not yet come to fruition, because several aspects of model fitting remain as much an art as a science (see Chapter 6), and may be difficult to automate reliably.

Deterministic models will not be replaced by stochastic models; the efficiency and usefulness of deterministic models in providing information for forest management have been demonstrated and cannot be matched by stochastic models. Deterministic models are more efficient at predicting the mean response, and can be used to determine the optimum management strategies for forest stands in a way not possible with stochastic models. Deterministic and stochastic models are complementary, and used in concert, may both prove useful in forest management. Alternatively, variance approximation may be used to estimate the variance of predictions in a deterministic way (e.g. Mowrer and Frayer 1986, Gertner 1987*a*).

The use of single-tree models for forest management applications other than in even-aged forests has been limited by the difficulty of modelling mortality efficiently. Tree list models, which are in some senses a hybrid between single-tree and size class approaches, offer a sensible compromise, so it is important to reconsider their capabilities when implemented near the resolution of single trees.

Tree List Models

Although tree list approaches (e.g. Stage 1972, Leary 1979) may be considered size class approaches, in many respects the tree list approach when employed with suitable resolution, may be considered an enhancement of the non-spatial single-tree approach. In effect, the single-tree model maintains a list of attributes (species, dbh, etc.) for each individual tree. The tree list approach does all this, but also simulates the number of trees per hectare represented by each tree record. This simplifies the deterministic prediction of mortality, as fractions (i.e. expansion factors < 1.0/plot) can be accommodated (Fig. 4.3).

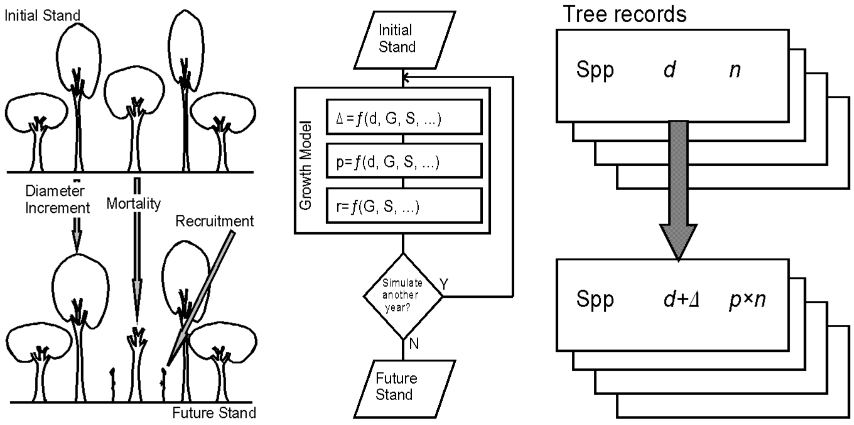


Fig. 4.3. Tree records representing a forest stand. Growth is modelled by incrementing the diameters in each record ($d+\Delta$) and mortality is accommodated by reducing expansion factors ($p\times n$).

Brand's (1981) implementation of Leary's (1979) tree list model allowed the user to specify if the mortality should be implemented in a deterministic or stochastic way. When the number of trees in a cohort falls to one (per plot) or less, mortality is always stochastic. Deterministic mortality prediction reduces the expansion factor by the predicted probability of mortality. With stochastic simulation of mortality, a random number is drawn, and if it is less than the computed probability of mortality, all the trees in the cohort "die" (i.e. the expansion factor is set to zero, and the tree record is removed from the tree list).

Stage (1973) assumed that the stochastic diameter increment function in the Prognosis model would account for sufficient variation in the stand so that the remaining functions in the Prognosis model could be deterministic. Furthermore, he assumed that provided there were sufficient trees in the stand, the effect of the stochastic single-tree function would not influence the stand total, and that predictions could be assumed to be deterministic. Where there were fewer trees, a "record tripling" procedure was used to simulate heteroscedastic growth. This "swindle" (Simon 1976) enables the model to make a deterministic prediction which emulates the average of many stochastic replications, without actually making the replications. In tripling (Table 4.1) each tree record becomes three records with 15, 60 and 25 percent of the original expansion factor and representing an increment of $\mu - 1.549\sigma$, $\mu - 0.1423\sigma$ and $\mu + 1.271\sigma$ respectively. This apportioning of increment derives from the normal $N(\mu, \sigma^2)$ distribution (Fig. 4.4), but the proportions 15:60:25 were subjective and other modellers have used different ratios (see e.g. Vanclay 1991*d*).

Table 4.1. Example of record tripling in the Prognosis model, assuming $\mu=0.1d$ and $\sigma=0.125\mu$.

Tree list in year *t*		Tree list in year *t+1*	
Stem size (cm dbh)	Expansion factor (stems ha^{-1})	Stem size (cm dbh)	Expansion factor (stems ha^{-1})
10	100	10.3	15
		10.9	60
		12.6	25
20	30	20.8	3
		21.8	12
		24.2	5

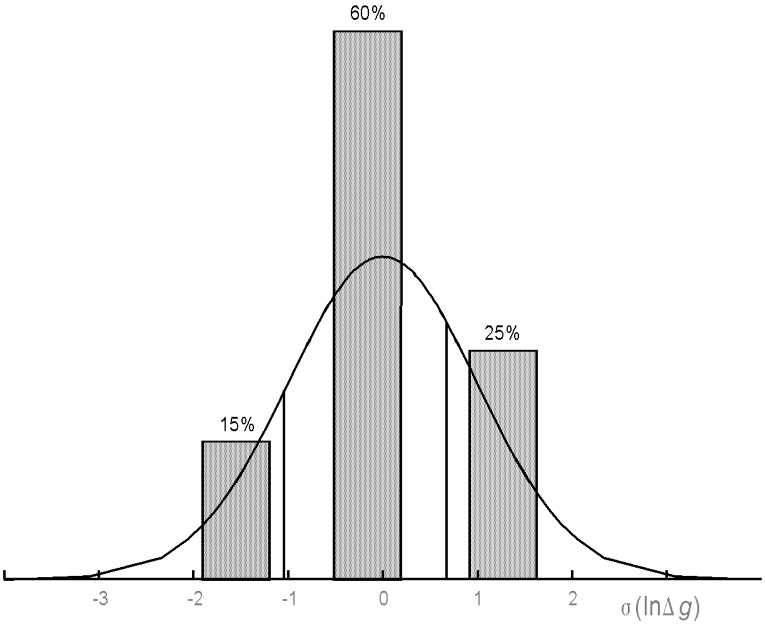


Fig. 4.4. Record tripling in the Prognosis model based on the normal distribution. Increments are predicted as ln(Δg), and show unequal variance when converted to diameter increment (Table 4.1).

Stage and Wykoff (1993) discussed how to estimate and incorporate serial correlation and stochastic effects in the model. When the number of tree records becomes very large, the most similar records are merged to maintain computational efficiency (Stage *et al.* 1993). The Prognosis model has been progressively refined (e.g. Wykoff *et al.* 1982, Wykoff 1986) and is used extensively to provide information for the management of natural forests in the Rocky Mountains region of the USA.

Natural forests, especially those in the tropics, may have many species. One way to construct a parsimonious model is to amalgamate these species into a few species groups and to simulate just these groups, but this excludes the possibility of reporting species composition. Another alternative is to retain the identity of individual species groups (where known) throughout the simulation, while using prediction equations based on species groups. This can easily be done using look-up tables to determine the equations and coefficients to be used for each species (Fig. 4.5). One advantage of this approach is that species may be grouped differently for the prediction of growth, mortality and recruitment. The same form of equation may be used for all species, but the coefficients employed in these equations may be specific to each species group.

Multi-resolution Models

Ideally, users should be able to choose the resolution level used in a model. They may require the efficiency of whole stand models for long-term planning, the precision of tree list approaches for short-term planning, and the additional information from stochastic models for risk assessment. The information provided at all these levels should be compatible. This multi-functionality is feasible, and has been implemented in several models (e.g. Leary 1979, Daniels and Burkhart 1988).

The cohort or tree list approach enables growth models to be formulated so that they can operate at any of several levels of resolution, to provide whole stand, size class or single-tree predictions according to the user's requirements. The STEMS model (Leary 1979) could be used at various levels of resolution. In the simplest case, one equation could be used for whole stand-growth estimates for stands of single species composition. Mixed stands were modelled using two or more equations of the same form. The system was designed so that it could also function as a tree list model or as a single-tree model.

Daniels and Burkhart (1988) described a model framework for pure stands which should enable compatible estimates at four levels of resolution: spatial and non-spatial single-tree levels, the size class and whole stand levels. Careful choice of variables enabled equations to be collapsed as resolution decreased. For example, area potentially available (Moore *et al.* 1973) was used at the spatial level whilst the inverse of stocking was used at the non-spatial levels. At the size class level, class

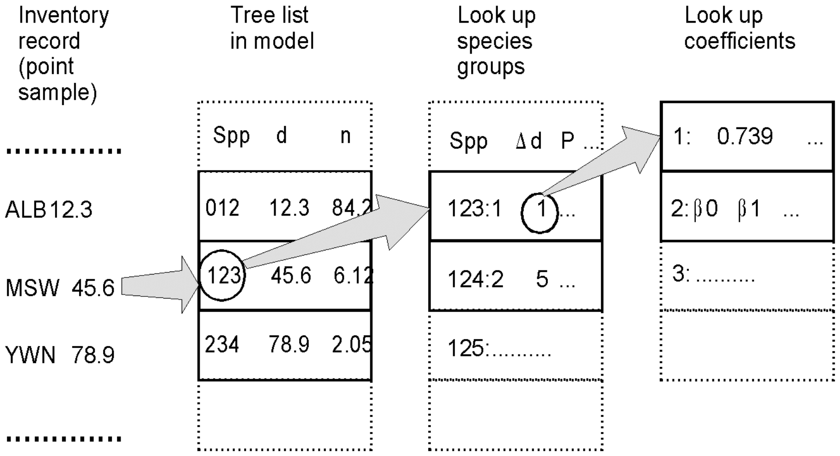


Fig. 4.5. Look-up tables can be used to find the coefficients to be used for a given species in equations to predict growth, mortality, recruitment, etc.

mean values of tree size are used in place of the individual tree values. Stand mean values could also be used, in which case size distributions could be estimated from the normal probability density function. The system relied on a common mathematical structure for models at different levels of resolution, which should ensure compatible predictions at all levels. Unfortunately, the framework as outlined by Daniels and Burkhart (1988) was specific to pure even-aged plantations, and the approach is more difficult to implement in mixed forests.

Compatible stochastic and deterministic forecasts can be obtained from tree list models by using probabilistic functions to predict diameter increment, mortality and recruitment (e.g. Vanclay 1991*d*). In stochastic mode, the predicted probabilities are compared with one or more random numbers (depending on the expansion factor of the cohort), and the fate of the entire cohort is determined accordingly. In deterministic mode, the predicted probability determines the proportion of the cohort which is incremented by one centimetre, unless the expansion factor is small, in which case probabilities are accumulated and the whole cohort is incremented when the accumulated probability reaches or exceeds unity.

Empirical Comparisons

Few empirical comparisons between various types of single-tree models, or with other types of models have been published. Clutter *et al.* (1983) drew the distinction between *explicit* and *implicit* prediction of yields. Explicit prediction systems are those which include equations to predict volume per unit area directly (i.e. some whole stand models), whilst implicit systems predict basic information on stand structure and stand volume is obtained indirectly (e.g. from tree or class mean diameters in single-tree and size class models respectively). Lenhart (1988) compared explicit and implicit (Weibull distribution) approaches for modelling coniferous plantations, and concluded that the explicit formulation provided the more accurate estimates of total timber yields. However, implicit approaches may offer more detailed information about stand structure.

Daniels *et al.* (1979) compared the predictive ability of two whole stand models and a single-tree model. The most accurate yield estimates (in terms of minimum mean square error) yield estimates were provided by the whole-stand distribution model. However, all three models provided estimates of sufficient accuracy for most plantation management uses. The relative costs of the predictions were 1:25:1400 for the whole-stand yield model, the whole-stand distribution model and the single-tree model respectively. Although the single-tree approach appears sub-optimal in both respects (accuracy and cost), it still may have a place in providing more detailed information than is available from the alternatives. The suitability of a model depends on user requirements, and whether a model can satisfy those requirements (see Chapter 11).

Ek and Monserud (1979) compared a deterministic size class model (with 5 cm dbh classes) and a stochastic single-tree spatial model (taking the average of 4 predictions), using independent data from a range of stand densities and management histories. Both models showed close agreement with reality for short term predictions (5–26 years). The single-tree model appeared to be slightly but consistently more reliable. Tests suggested that the single-tree model was not significantly different from reality, and that the size class model differed significantly from both reality and the single-tree model. Both models gave comparable predictions for long term (120 year) predictions, but no data were available to confirm the reliability of these predictions.

Mowrer (1989) demonstrated that computational efficiency is but one cost of complex models, and that complex models may propagate greater variances than more simple whole stand models. This means that any error in the inventory of initial stand condition may be magnified by methods such as single-tree models, whereas they may remain comparatively unaltered by less complex models such as whole stand models. The implication is that models should not be unnecessarily complex, but should be designed to provide specific information needs.

Synthesis

Compared to the large variety of modelling methodologies following whole stand and size class approaches, the single-tree and tree list modelling approaches are characterized by few variations in the general approaches (but many variations in the details e.g. competition indices). Seemingly, as the approach becomes more mechanistic, the logical way to implement the model becomes more apparent, and fewer alternatives are pursued. Thus the emphasis has been on adapting existing models to more forests rather than making more models. For instance, the Prognosis model (p. 72) is now in its fifth version (Wykoff 1986) and regional variants have been implemented for about two-thirds of the USA and for parts of Canada (Stage and Wykoff 1993). Several variants of the STEMS model (p. 74) also exist (e.g. Belcher *et al.* 1982, Goodwin 1988, Hilt and Teck 1988, Miner *et al.* 1988, Swain and Turner 1988).

The tree list approach has been demonstrated and widely used for uneven-aged mixed forests in temperate and tropical climates, and appears to offer the greatest versatility of the alternatives reviewed. It is likely to form the basis of many mixed-forest models, and this general model formulation will be assumed in our examination of model components and their estimation (Chapters 6–10).

This completes our survey of modelling techniques, and we now turn our attention to the collection and analysis of data for modelling. You should now have a good idea of the type of model most suited to your particular needs, and some idea of the data required to construct such a model. In the next chapter, we will examine ways to collect data suitable for model construction.

Exercises

4.1. Discuss the strengths and weaknesses of four categories of spatial competition indices (p. 58). Suggest non-spatial analogues in each category and contrast the utility of the spatial and non-spatial forms. These have mainly been used in pure stands. What enhancements may be required for use in mixed forests? Design an experiment to test these competition indices.

4.2. Discuss the strengths and weaknesses of Equation 4.1 (p. 64). What happens with predictions from this equation when the stand is thinned or trees are pruned? Is there any evidence of thinning or pruning in Fig. 4.2 (p. 65)? *Hint*: Remember that thinning affects the sky-view as well as basal area, and that pruning affects the crown centre (for computing l_s) and crown area, as well as crown length.

4.3. Criticize Equation 4.2 (p. 70), stating its strengths and weaknesses. Can reasonable estimates of Δd_{max}, d_{max} and h_{max} be obtained for all forest tree species in your locality? Estimate Equation 4.2 for these species and contrast the resulting relationships with other increment functions available for these species.

4.4. Write an outline for a tree list model for mixed forest. Write computer code or pseudo-code if you can, otherwise write clear explicit instructions in a form suitable for your computer programmer. Give special attention to the way you might deal with the many species that you expect to find. *Hint*: Study Figs 4.3 and 4.5 again.

4.5. Re-consider your answer to Exercise 1.1 (p. 13). Have you thought of any additional needs that potential model users might have? What type of model would be best suited to meet these needs? State any special adaptations to the basic approach that might be required. What data are needed to calibrate the model?

Chapter Five

Data Requirements

Models and data are like chickens and eggs; it is not obvious which comes first. Ideally, modelling and the definition and collection of data should form an iterative process, commencing with the model formulation. The model design should dictate the data requirements, so that field work can provide necessary and sufficient data efficiently. However, it may take several years to obtain the necessary data from permanent plots, and few of us can wait that long. Most modelling efforts commence with any data available, and the modelling approach often may be dictated by limitations of the data. Many models owe much of their success to the foresight and dedication of our forebears who established permanent plots and maintained both plots and measurement records carefully.

The life cycle of a datum spans its definition, collection, validation, storage, analysis and synthesis. All stages are equally important, and an efficient data management system requires a healthy balance between them. The first step is to define information needs and devise data collection procedures to satisfy those needs.

Fortunately, the data requirements of many modelling approaches are similar and allow a set of minimum data requirements to be defined and standard procedures to be established. The procedures discussed here relate to the requirements for developing growth and yield models for forest management. Additional details may be necessary if plots also are to serve ecological studies and other uses. The following is a discussion of principles, and readers seeking a manual for permanent plots in mixed forests should refer to Alder and Synnott (1992). Hutchinson's (1982) manual may be useful for researchers dealing with tropical moist forest, as it gives detailed guidelines for measuring stems which are broken, fallen, coppice, parasitic, etc.

Stem analyses do not provide reliable growth data for many tree species in tropical moist forests, so data must be obtained from remeasurements on permanent sample plots. Many anomalies may be found in the growth rings of tropical tree species (e.g. Mariaux 1981). Some evergreen trees (e.g.

Swietenia spp.) may form rings while deciduous trees (e.g. some *Ficus* spp.) may not. Some species (e.g. *Hevea braziliensis*) may form several growth rings each year, while other species (e.g. *Shorea robusta*) may form only one ring but not necessarily in the same month each year.

Permanent plots can never be completely replaced by temporary plots even for species amenable to stem analysis, because only permanent plots (i) allow satisfactory statistical comparisons within and between plots to check the adequacy of models, and (ii) provide reliable and consistent data on mortality, crown dynamics and stand level variables.

Differing Data Needs

Inventory serves many purposes, but different procedures are required to satisfy various needs of different data users in an efficient way. Some typical information requirements and corresponding sample plot procedures are summarized in Table 5.1, and include:

Resource Inventory ("What is the present nature and extent of the resource?"): Typically many plots (or point samples) will be required to achieve the desired precision. Precision can be gained by orienting plots across environmental gradients to maximize within plot variation and thus reduce between plot variance. Cost considerations usually dictate that temporary inventory plots (or point samples) are most efficient for resource inventory. Specialized techniques for timber cruising offer great efficiencies (see e.g. Schreuder *et al*. 1993), but may not provide data suitable for input to yield forecasting systems.

Continuous Forest Inventory for yield control: Some systems of yield regulation monitor the forest growth and harvesting by remeasuring a series of permanent plots, established systematically so that each plot represents an equal area of forest (see e.g. Leuschner 1990). One such system is known as continuous forest inventory (CFI). It is important that CFI plots are representative and established in various forest types and stand conditions in proportion to their area. As with resource inventory, precision is gained by minimizing between plot variance. Plots should be marked so that they can be relocated for remeasurement, but should remain inconspicuous so that they receive unbiased management.

Growth Modelling: The need to provide reliable data for growth modelling demands three qualities of permanent plots that are not necessary in CFI and some other permanent plot systems:

Table 5.1. Different applications require different sampling techniques.

Plot characteristics	Principal objective of permanent plot system			
	Resource inventory	Continuous forest inventory	Growth modelling	Site monitoring
Permanence	Temporary	Permanent	Permanent	Permanent
Area	Variable, ∝ tree size	Fixed	Fixed	Fixed
Within-plot variance	Heterogeneous	Homogeneous	Homogeneous	Homogeneous
Placement	Stratified random	Systematic	Stratified random	Purposive or systematic
Sample unit	Plot	Plot	Tree	Plant parts

(a) individual trees must be unambiguously identified,
(b) the plots should be homogeneous, and
(c) the plots should sample extremes of site and stand condition.

Trees should be identified using permanent marks and unique numbers. This not only offers more options for modelling, but also is the only sure way of detecting measurement errors. Growth modelling requires homogeneous plots, and this means minimizing within plot variance: the ability of the permanent plots to quantify the present resource is irrelevant. Thus the same plot series cannot be efficiently used for both resource inventory and growth model development. If the growth model is to be used to investigate silvicultural and management alternatives, the database must include experimental data with paired treatment and control plots, both with adequate isolation. In contrast to continuous forest inventory plots, it is not necessary for the permanent plots to be representative or numerically proportional to forest type areas, but it is essential that they sample the full range of stand conditions.

Long Term Monitoring of Environmental Change: Several researchers (e.g. Adlard 1990, Dawkins and Field 1978, Watson and Nimmo 1992) have described permanent plot systems designed to monitor subtle long term changes in a forest. Whilst such studies are desirable, few organizations have the resources or need to establish such detailed plots on the scale required to provide suitable data for growth models for forest management. Such detailed plots should be reserved for special studies. For growth modelling, it is better to sample the full range with conventional permanent

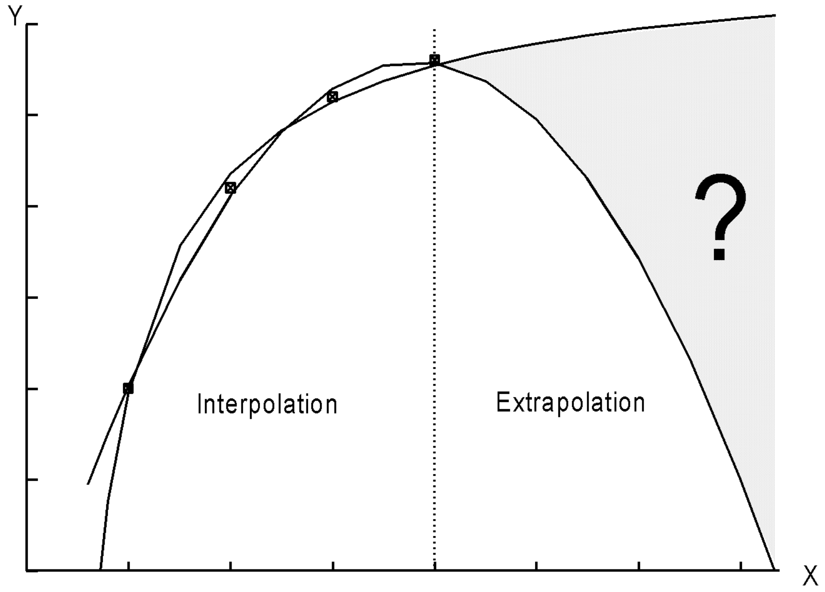


Fig. 5.1. Interpolation is safer than extrapolation. Both these lines have an R^2 better than 0.996, but provide no basis for making a prediction outside the range of the data.

plots than to have a few very detailed plots. However, quantity is no substitute for quality.

Permanent plots established to provide data for growth modelling should be designed to satisfy this primary need, and should not be compromised in order to satisfy secondary needs. They do not need to provide resource inventory data efficiently, as alternative sampling procedures can fulfil that need. This chapter focuses on permanent plots intended primarily to provide data for developing and evaluating growth models.

Development, Evaluation and Use of Growth Models

Growth modellers need data to develop models, to test models, and to use models, and each of these three activities may require data of a different nature. The initial and most obvious requirement for data is during model development when data are used to fit the basic functions comprising the model, but it is equally important to set aside some data to allow through testing of the model (see Chapter 11).

Model Development: Interpolations are always safer than extrapolations (Fig. 5.1), so permanent plot systems should be designed to sample the widest possible range of site and stand conditions. Limited but reliable data at each extreme and at the mean are more useful than copious data clustered about the mean. Remeasurements are necessary to detect growth and change, and must span a sufficient time period to incorporate climatic variation. The measurement interval should be long enough to ensure that growth patterns are not obscured by measurement error. Some statistical difficulties may arise if there are many remeasures of each plot, unless the number of plots should be large relative to the number of plot remeasurements (see Chapter 6).

Model Evaluation: Testing of a model (Chapter 11) is an important part of model construction, and should not be neglected. To provide a convincing demonstration of the adequacy of the model, the data used to test a model should not have been used to develop it, and may be drawn from a different population. Independent data are not always available, and it is common to partition the data into two subsets, one for development and the other for testing. It is important that the subset used for model evaluation should contain at least some data collected over very long periods to allow detection of possible subtle but cumulative errors in the model. In this way, plots with very many remeasures may be more useful for benchmarking models, than for initial development.

Monitoring may be viewed as on-going evaluation of a model by checking operational predictions. It involves comparing projected and realized yields to identify any discrepancies. Such discrepancies may be due to changes in management regime (especially harvesting practices), changes in site productivity, inaccurate resource data, or corruption of the parameters defining the model. Unlike the data required to develop and test growth models, data for monitoring may be collected at any time, and often may be drawn from operational resource inventories conducted as part of routine forest management activities.

Applications: A growth model is of relatively little use on its own, and is normally used in conjunction with other resource data obtained during operational resource inventory (Fig. 1.1, p. 2). Such data should provide area estimates for each forest unit, and details of its species composition, stand condition, and site productivity. Resource data have been integrated in this way since inventories began, but recently attempts have been made to define principles and procedures for efficient integration (e.g. Lund 1986). Integration simply implies combining data obtained in different places, for different reasons, by different agencies or at different times. This may involve combining regional inventories to provide national or global statistics, finding correlations between timber inventory and soil or fauna survey data, or using growth models to extrapolate static inventory data and estimate sustainable yields. Integration does not necessarily mean that you have to measure everything in every inventory. On the contrary, it is better to do a few things well than to do a lot inadequately. However, in designing

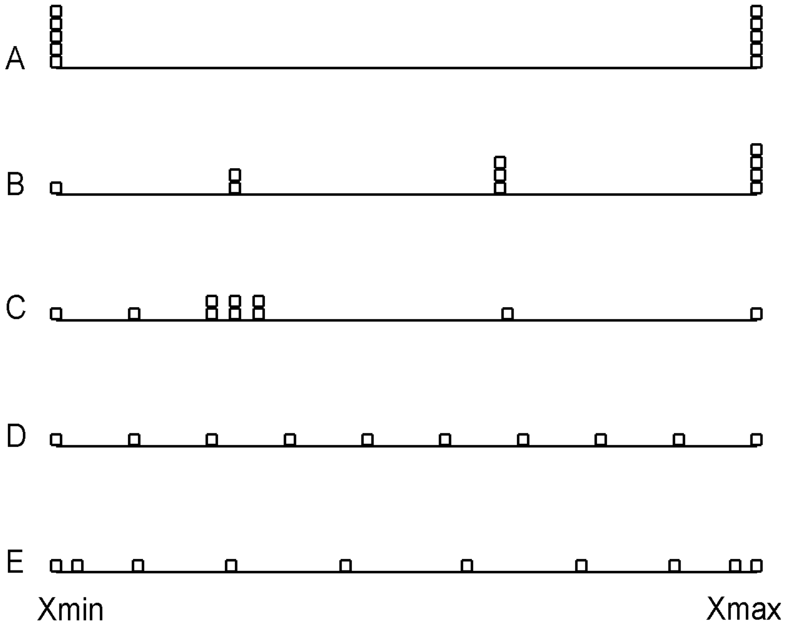


Fig. 5.2. Efficient placement of ten samples to (A) estimate slope of a straight line, (B) detect curvilinearity where variance ∝ X, (C) calibrate an optimum, (D) detect a threshold, and (E) fit a curved relationship.

a permanent plot system, it is necessary to be aware of the information requirements of other researchers and other disciplines, and to consider how these requirements can be efficiently accommodated in the design. It is not necessary that all these requirements are satisfied immediately, but rather that the design accommodates these needs so that they can be phased in as required and when feasible.

Data for Modelling

One of the main principles in collecting data for growth modelling is to sample the full range of site and stand conditions, so that model predictions may be interpolations rather than extrapolations (at least in the sense of stand conditions; all forecasts are an extrapolation in time). Further optimization of sampling design depends on what you are sampling for, but some guidelines can be explained for the simple case with a single explanatory variable (Fig. 5.2). If you know that a simple straight-line relationship exists (e.g. a volume tariff line), the optimal strategy is to gather half the data from each extreme (A); this gives the best estimate of slope. At the other extreme, if you are looking for a threshold (e.g. the onset of density-dependent mortality), the best strategy is to space samples equally across the region of interest (D). In practice, it is rarely so simple. Often, we may need to test for a curved relationship (e.g. quadratic or asymptotic), and may have a situation where the variance increases with the explanatory variable (e.g. stem volume). In such cases, it is helpful to sample at several points through the range, with sampling intensity increasing with the variance (B). Some optimum may be suspected (e.g. growth curve), and increased sampling in the vicinity of the optimal may improve the estimate (C). Studies with spline curves suggest that a useful strategy in many situations is to sample the full range with intensity increasing towards the extremes (case E with samples at $cos(\pi i/n), i=0, 2, . . ., n$).

These principles of sampling apply to the four factors that should be considered in setting up and maintaining a permanent plot system for growth modelling:

1. *Temporal distribution*: Growth varies from year to year, fluctuations can be extreme, and mortality tends to be clustered in both time and space. Short time periods may give rise to biased growth estimates, and longer periods of observation offer a better basis for modelling. This requires a long-term commitment of resources.

2. *Spatial distribution*: The uncertainty associated with extrapolations (Fig. 5.1) applies equally to geography as to other aspects of the database. Permanent plots should sample an adequate geographical range, including latitude, longitude, elevation and other topographical features such as ridge and valley locations. Although a systematic grid sample will provide a good range of latitude and longitude, some form of stratified sampling may offer a better way to incorporate other topographic aspects.

3. *Site factors*: Many factors that influence growth cannot easily be manipulated experimentally (e.g. soil type and depth), and the sampling system should ensure that the full range of these factors is included in the permanent plot system.

4. *Stand conditions*: Growth is also influenced by stand structure and composition, and these can, and should be manipulated experimentally to provide the best database for modelling. This means that in some locations, clusters of plots should be established, with some plots left undisturbed, some managed routinely (i.e. usual forest management operations), and some subject to a range of experimental treatments.

Placement of Plots

Ideally, permanent plots should sample the geographic range over which the growth model will be used, and encompass the full range of forest types, site productivity and topography. A broad range of stand basal area and tree sizes should be sampled for each tree species. This should be achieved by some form of stratified random sampling, as subjective location of plots may introduce bias. Strata may be based on standing volume, species composition, soil type or any other objective means. Care with the placement of the plot boundary is necessary to avoid bias when establishing plots at the forest edge (see e.g. Fowler and Arvanitis 1979), but such locations must also be sampled. The need for random location based on a thoughtful stratification cannot be over-emphasized.

There is some evidence that gains in the precision of growth predictions can be achieved by sampling more large trees (e.g. Gertner 1987*b*), and it may be desirable to establish some plots around subjectively selected large trees. Such subjective selection of plots may introduce bias, but this may be an acceptable trade-off to reduce the variance associated with growth predictions from large trees. To minimize bias, these plots should constitute a small proportion of the total, and should be selected within strata based on site productivity and stand density (e.g. stand basal area).

Data for growth modelling should span a range of site and stand conditions at the stand level, and a range of tree size and competition at the tree level. One way to appraise a growth modelling database is to compare it with other resource data indicating conditions found in the population of interest (the whole forest estate). Beetson *et al.* (1992) compared scatterplots of site quality versus stand basal area, and of tree size versus competition, arguing that these were the most significant variables for growth modelling and prediction. The graph in Fig. 5.3 contrasts the range of site quality and stand basal area in the permanent plot database (solid lines) with those recorded during static inventory (dotted lines) within the same climate-soil stratum. Ideally, the envelope (i.e. convex hull) enclosing the permanent plot data should also enclose all the other resource data, so that no extrapolation is necessary to make inferences about these data. Beetson *et al.* (1992) found that their database fell short of this ideal, but found that only four additional plots (chosen from temporary plots, marked × in Fig. 5.3) would stretch the envelope to satisfy this need. The existing permanent plot database included 25 plots within this stratum, but many of these plots sampled similar stand conditions. Beetson *et al.* (1992) suggested that if five new plots were established (×), then all but four of the existing plots (+) could be abandoned, and resources could be redirected to other less well sampled strata. No existing data would be discarded, and measurement records (up to 50 years) from these discontinued plots could still be used in growth modelling. The nine plots chosen to represent this stratum were chosen not only on the basis of Fig. 5.3, but also on an analogous scatterplot of tree diameter and competition (the basal area in larger trees, $G_{>d}$, one of the best non-spatial predictors of tree growth; see

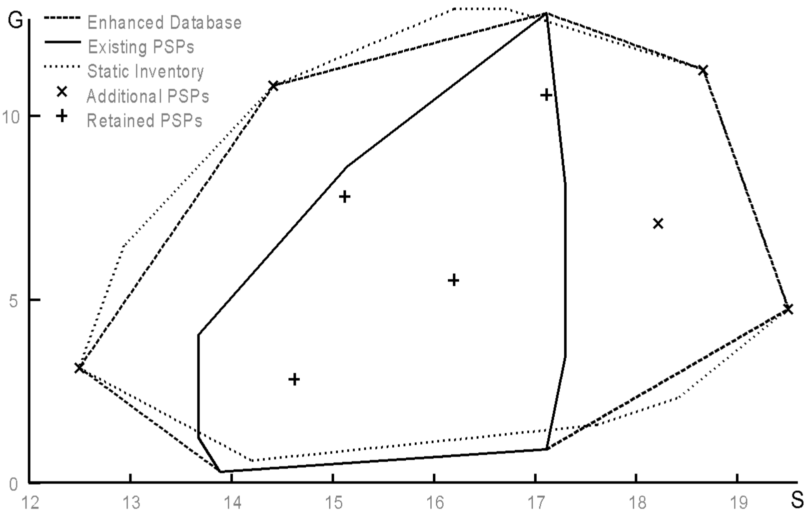


Fig. 5.3. Database weaknesses revealed by comparing dynamic and static inventory data. Five new PSPs (×) would improve the database for modelling (*Callitris* forest in Queensland, redrawn from Beetson *et al.* 1992).

Chapter 8). The addition of these new plots would mean that only negligible extrapolation (not only in terms of stand basal area and site form as seen in Fig. 5.3, but also in terms of tree diameter and competition) would be required to make forecasts for all the *Callitris* forest in the region. Computer programs also are available to aid the optimal design of sampling schemes (e.g. Kennard and Stone 1969).

New permanent plot installations should be established only as part of a carefully planned series designed to give reasonable coverage of some defined range of site, geography, stand condition and treatment. The primary objective should be to provide data for defining response surfaces, thus studies should involve many locations with minimal replication at each location. Satisfactory growth models are dependent upon the availability of high-quality data from a wide range of stand conditions and treatments. Some plots should be monitored without disturbance (i.e. no harvesting) for long periods to enable the most exacting evaluation.

Experiments

Both *passive monitoring* data (i.e. survey data from forest areas under routine management) and *treatment response* data (i.e. from paired treatment and control plots) from designed experiments are necessary. In a designed experiment, all the variables are supposedly held constant except those varied according to the design; hence all the variables are accounted for. Data collected without the aid of an experimental design may be suspect for a variety of reasons. The explanatory variables are often highly correlated and frequently the region of interest is inadequately sampled. Important variables may be omitted because their relevance was not realized. These factors may limit the ability of the model to make reliable predictions (see Chapter 6).

Although harvesting and other silvicultural treatments in passive monitoring plots will influence stand density, other unknown factors may also determine stand density and composition. Thus, there is a very real danger that attempts to describe the behaviour of the stand as a function of stand density, for instance, will be confounded by the effects of site, pest and disease occurrence, and past history. To find out what happens to a system when you interfere with it, you have to interfere with it, not just passively observe it (Box 1966).

Snedecor and Cochran (1980) reported a survey which revealed the unexpected result that the application of farmyard manure reduced the yield of potatoes by half a tonne per hectare. In contrast, in controlled, randomized experiments, manure increased the yield by three to six tonnes per hectare. The discrepancy may arise because those who had manure were livestock farmers with little interest in growing potatoes, and those who were most skilful at growing potatoes had no manure. Can we be sure that a similar problem in our data is not hampering our attempts to develop growth models (e.g. stand density and site productivity interaction)?

Passive monitoring data may indicate greatest growth on the best sites with high stand basal areas, and little growth on poorer sites with low basal areas. A growth model constructed from such data could suggest that greater increments accrue in stands with greater competition, since the effects of site quality and stand density could be confounded. Thus a model constructed from such passive monitoring data would predict a reduction in diameter increments following thinning, whilst a model from experimental data (e.g. thinning studies) would show an increase in diameter increment after thinning.

Controlled experiments (i.e. with paired treatment and control plots) thus provide a useful way to improve a database and to sample extremes of stand condition. Consideration should be given to establishing a series of plots in homogeneous tracts of each forest type. Some should be left at maximum stocking to allow expression of density-dependent mortality and natural basal area, some should be harvested and treated as a managed stand, and others should be heavily thinned to allow expression of open-grown development and regeneration. It does not matter that extreme treatments may never be applied in practice; they remain essential to properly define the response surface for growth models.

Plots which are intended to remain untreated, for example to allow expression of density-dependent mortality and natural basal area, should be clearly marked and excluded from any harvesting operations. Such plots, like all experimental plots, should have adequate buffers to eliminate edge effects. The appropriate size of the buffer depends on potential tree size, but generally should be wider than the mature tree height. Alder and Synnott (1992) recommended a 100 m buffer for all logging and thinning experiments. Other plots intended to receive routine management should be marked in such a way as to be invisible to forest workers so as to ensure representative treatment. A monitoring plot should be difficult to recognize for those who do not know where it is, and easy to recognize for those who do and are looking for it. Plots must have unambiguous addresses and this requires precise grid references and detailed instructions for relocation of each plot. In some areas, plots may suffer excessive trampling through high visitation, and to overcome this problem, Dawkins and Field (1978) marked their plot locations clearly, but 50 m away from the plot, and used buried steel markers at all four plot corners.

Number of Plots

The number of plots may be dictated by the resources available. There is little point establishing more plots than can be maintained. It is better to have few plots providing reliable data, than many plots with inadequate management. The number of plots will also be determined by the variability of the forest estate, and the need to sample the full range of forest conditions. Alder and Synnott (1992) suggested one permanent plot per 1000 ha of forest, subject to a minimum of 50 and a maximum of 1000 plots. This may serve as a guide, but the recommended number depends on local conditions and resources. The quality of the plot is paramount, and if resources are limiting, it is better to reduce the number of plots and maintain standards, than to compromise data by attempting to do too much with insufficient resources.

A database comprising a few plots each with many remeasurements violates statistical assumptions of independence, and may require special analyses (West *et al.* 1984, 1986, West 1994). This violation becomes significant when the number of remeasures is large relative to the number of plots. An alternative is to use partial replacement, abandoning plots after several remeasures and establishing new ones (e.g. Tennent 1988). However, some plots must be retained for long periods with many remeasures to allow convincing tests of model performance.

Size and Shape of Plots

A general guide to the choice of plot shape is to minimize the plot edge to area ratio, and reduce the number of corners. This leads to the choice of point samples, or circular, triangular, or four-sided plots according to the emphasis attached to corners and edges. Triangular plots are rarely used (outside Denmark), perhaps because of the high edge to area ratio, and four-sided plots are generally rectangular (or square) to help relocate corners and boundaries.

Point samples have an advantage in being defined by a single point and an angle (often expressed as a basal area factor), but they are inconvenient when dealing with recruitment, and create difficulties for some single-tree models. Circular plots are also defined by a single point and a radius, but the plot boundary becomes more difficult to define as the plot becomes large, as unlike polygonal plots, sight lines cannot be established along boundaries. As these plots are defined by a single marker (the centre), they may be more difficult to relocate if the marker is damaged or removed. Because of these and other limitations, circular plots and point samples should not form the basis of a dynamic inventory system for growth modelling in mixed forests.

Rectangular plots are more versatile. Plots marked by four corner pegs may be less likely to be lost than circular plots marked by only one peg. However, a more important reason for the choice of rectangular plots is their straight edges, few corners, and convenience. Square plots have a smaller edge to area ratio than other quadrilaterals, and are recommended for dynamic inventories.

Ideally, the plot size should be sufficiently small that the plot is homogeneous, at least with respect to forest type and site productivity, and sufficiently large to provide a representative sample of the forest stand. If a spatial model is contemplated, the plot should be large enough to allow estimates of competition to be determined for several trees on the plot. Larger plots offer greater flexibility, and plots of one hectare are recommended for growth studies in mixed forests (e.g. Lanly 1981, Campbell 1989, Alder and Synnott 1992). Some ecological studies may warrant even larger plots. One way to appraise the adequacy of a sample is to see if it approaches an asymptote on the species–area curve (McGuiness 1984, Campbell 1989). In Amazonian *terra firme* forest, the species–area curve (for trees over 10 cm dbh) may approach an asymptote between two and three hectares (Campbell *et al*. 1986).

Unless circular plots or point samples are adopted, the orientation of the plots needs to be considered. This may be inconsequential for square plots, but may be significant with elongated rectangular plots. Three possibilities exist. The plots may be randomly oriented, may be oriented according to the cardinal direction (e.g. long axis north–south), or may be oriented according to topography or other environmental gradients. In view of the need for plots which are homogeneous with regard to site productivity, the last of these is likely to be preferable. If plots are elongated, they should, for

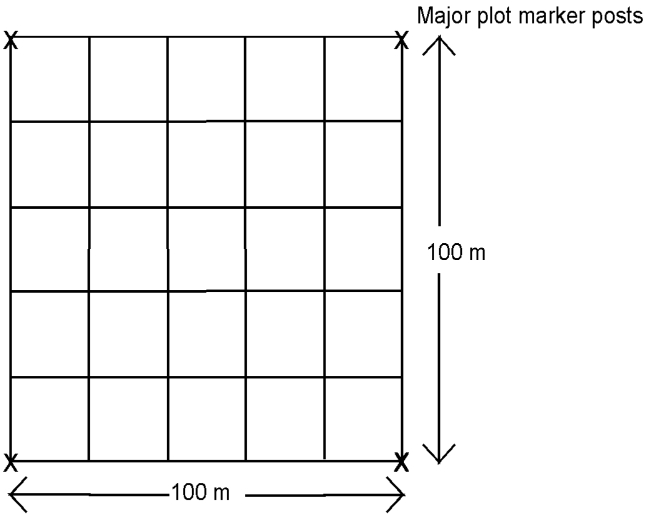


Fig. 5.4. Recommended plot layout for permanent sample plots.

growth modelling purposes, be oriented with their long axis perpendicular to the slope (i.e. parallel to contours), or any other perceived gradient of site productivity to minimize within plot variation. In contrast, static inventory with temporary inventory plots may attempt to maximize within plot variation (to reduce the between plot variation and thus the sampling error) by orienting elongated plots along environmental gradients. Dynamic inventory for growth model development has a different goal, and thus within plot variation should be minimized.

Data Collection

The following is a very brief review of data collection procedures. Readers seeking more detailed information should refer to one of the standard manuals (e.g. Hutchinson 1982, Curtis 1983, Alder and Synnott 1992).

Measurement Procedures

Data used for growth research must be of a higher quality than that generally obtained in static inventory. For example, a diameter measurement of 50 ± 0.5 cm may seem precise enough for many purposes, but if a remeasure indicates 51 ± 0.5 cm, the growth estimate will be 1 ± 0.7

cm, which is not sufficiently precise to distinguish growth from measurement error. Diameters should be measured and recorded to the millimetre, and conventions regarding rounding to the nearest or lower millimetre should be maintained from one measure to the next. The same principles apply to heights and all other parameters.

Providing that plots continue to provide useful information, existing standards and procedures should be maintained to ensure uniformity. The continuity of standards is critical, and any changes to standards should be considered carefully to ensure compatibility. If compatibility cannot be ensured, dual procedures should be continued for at least one remeasurement cycle.

When new plots are established or procedures for existing plots are revised, the following requirements should be accommodated (Whitmore 1989):

1. Divide each plot into subplots of maximum size 20 × 20 m. Mark corners permanently. Brightly coloured plastic pipe is durable, visible and easy to transport, but may be stolen or destroyed by elephants or primates. Locally made wooden posts are often used, but must be made from durable wood of suitable dimensions to resist decay and simplify plot relocation. A good precaution is to dig trenches in case posts are lost, but to minimize disturbance, these should be dug only outside the plot at the four external corners (× in Fig. 5.4). Since logging can completely destroy all plot marks, buried steel markers may be used in conjunction with these.

2. Trees must be numbered and permanently marked so that no confusion regarding identity can occur. Never use the same number twice. Do not re-use the number of a tree which dies but give ingrowth trees new numbers. Paint numbers on trees, or use an embossed aluminium tag fastened with an aluminium nail. Numbers may also be scribed (cut) onto trees, but care must be taken to avoid callus growth. Nails should be long enough to be firmly secured in the wood of the tree, while leaving sufficient room for growth. Plastic tags should not be used as they may deteriorate rapidly under tropical conditions. The nails may need renewing about once every five years. Nails should be below the merchantable section, and should always be on the same side of the tree (e.g. northern or uphill) for ease of relocation. If theft of nails or tags occurs, use paint or nail the tag near ground level and cover it with litter.

3. Make a map to show the position of every tree to the nearest 1 m or better. Work one subplot at a time. A maximum subplot size of 20 × 20 m makes mapping easy. Without a map, confusions always occur because of death, ingrowth or lost number tags. Place measuring tapes along two adjacent sides of the subplot, and estimate the coordinates of each tree. Other alternatives include pentaprisms (e.g. Reed *et al.* 1989) and laser rangefinders. It is important to make a sketch map as well as a record of coordinates, so that any errors may be detected and rectified while the measure crew is still in the forest.

4. Specify the minimum girth for the smallest tree to be included. This becomes important later when ingrowth occurs. For the sake of clarity, note when surveys originally using imperial measure have been converted to metric measure. For example, 12 inches girth = 9.7 cm diameter, but the metric minimum is commonly 10 cm.

5. Since girth can be measured sufficiently accurately only if always done at the same place, it is best to mark (paint) the point of measure on each tree. Paint marks need renewing every 3–4 years. In wet weather emulsion (acrylic) paint is easier to apply than oil-based paint. Some paints (especially oil-based paints) may cause abnormal bark shed and callus growth (e.g. on *Eucalyptus maculata* and *Flindersia pimenteliana*), so paints should be tested before general use. A less accurate alternative for species that shed their bark, is to measure girth a fixed distance above the number-tag nail; at least 20 cm away because some species develop swollen callus tissue around the nail. Make sure the girth measurement point is well above all buttresses, for these may grow upwards. Take care that blazes made to identify the tree are not too close to the point of measurement.

6. The measuring tape must be of metal or fibreglass because cloth tapes stretch when wet. All measuring tapes should be checked periodically against a reliable standard.

7. Since errors occur in recording girths, measure every tree twice independently. If the second reading differs, make a third one. Loose bark (except on species with corky or flaky bark), epiphytes and climbers must be removed from the line of measurement. The booker should always repeat the measurement for the measurer to confirm. Make sure that your helpers can read and can use all instruments (e.g. tapes, callipers, hypsometers, etc.) correctly.

8. If, despite precautions at initial survey, a buttress grows up into the point of measurement, the point will need to be moved further up the bole. This alters the measurement base for that tree, so measure at both the old and new heights, and record that a change in measurement height has occurred.

9. Trees with more than one trunk at the height of measurement should be given separate tree number and girth records. Hutchinson (1982) gave detailed guidelines for the measurement of broken, fallen and coppice stems, and these may be useful to researchers working in the tropics.

What to Measure

Plot location should be described and its coordinates (e.g. grid coordinates of the SW corner) and orientation (e.g. direction of long axis) should be recorded. Topographic features (altitude, slope, aspect, distance to ridge), climate (rainfall amount and distribution), indicator plants, and soil physical characteristics (depth, texture) should be documented. Uniformity of the site should be assessed. These variables need only be recorded at plot establishment, and at occasional remeasures.

Factors which may cause fluctuations in the observed growth include drought, heavy seed crops, pest populations, disease outbreaks and fire damage. Such information may guide the interpretation of outliers in regression analysis. Thus the occurrence of such events should be recorded on the plot measure record.

It is impossible to measure all individuals of all sizes and species, and is irrelevant to measure all seedlings and saplings since most will die. All trees (of all species including useless and dead stems) exceeding 10 cm dbh should be measured (e.g. Prance 1984, Campbell 1989, Alder and Synnott 1992). This is necessary to enable estimates of stand basal area, an important variable in predicting tree growth and stand dynamics. Limits as small as 3 cm have been adopted in natural coniferous forests (e.g. Arney 1985, Vanclay 1988*a*), but may be impractical elsewhere. Reliable ingrowth data require that the measurement limit is less than the desired ingrowth size (e.g. for recruitment at 10 cm dbh, measure all trees exceeding 8 cm). Subplots within the main plot may be used to record data on stems smaller than this measurement limit.

Voucher specimens should be collected for each tree, even if sterile, and entered into a permanent herbarium collection (Campbell 1989), unless it is very common in the locality and you are certain of its identity. Voucher specimens should, if possible, include several leaves joined to a branchlet, and any fertile material (buds, flowers and fruits), and any other distinctive parts (e.g. bark). These should be pressed, dried and mounted on paper, or should be preserved in alcohol. All parts should be clearly labelled with tree and plot numbers, location, date and collector.

Tree diameter and status (alive/dead/felled and erect/leaning/fallen with cause if evident) should be recorded at every measure. Diameters should be measured at 1.3 metres (from the ground on the uphill side of the tree) or above any buttress, and should be measured perpendicular to the axis of the tree (e.g. Avery and Burkhart 1994, Philip 1994). Every tree present at the previous measure should be accounted for. Natural mortality should be discriminated clearly from harvesting, treatment and other removals. Trees which appear lost should be recorded as such, and should not be attributed to death unless there is evidence to support this assumption.

Height, crown parameters and estimated defect should be recorded at establishment, and periodically at remeasurements. The vantage point for height measurement should be carefully chosen to allow good visibility and a sighting angle of around 45 degrees (Romesburg and Mohai 1990). The Dawkins (1958) system of classification (emergent, full overhead light, some overhead light, some side light, no direct light) is a simple, proven system which may be a good predictor of increment (e.g. Wyatt-Smith and Vincent 1962, Alder and Synnott 1992, Silva *et al.* 1994), and should be recorded for all trees on permanent plots.

In summary, the following variables should be measured:

1. At the initial enumeration (and occasionally remeasured as new technology improves the precision that can be attained):
  (a) plot location, dimensions, orientation and area,
  (b) species and coordinates of all trees on the plot,
  (c) topographic details, including altitude, aspect, slope, position on slope,
  (d) forest type and floristic attributes,
  (e) physical soil characteristics (depth, texture, colour, parent material), and
  (f) uniformity of the site;
2. At the first measure, immediately after any harvest, and periodically (e.g. every second or third measure):
  (a) sufficient tree heights for the determination of site productivity (or data necessary for alternative estimates of site productivity),
  (b) merchantable heights and defect assessments of all stems (including non-commercial species, as utilization standards may change with time), and
  (c) crown characteristics (position, length, width, form, etc.);
3. At every measure, assess all stems (including non-commercial; every stem from the previous measure must be reconciled) for:
  (a) diameter (over bark, breast high or above buttress), height to measure point, and validity (to indicate defects at measure point and anomalous but correct increments),
  (b) status (alive, dead, harvested, treated) and stance (erect, leaning, fallen, broken), and
  (c) tree coordinates (recruits only);
4. As necessary, record the occurrence of:
  (a) logging, treatment and other activities, and the prescription used,
  (b) scars and other damage which may affect measurements or growth,
  (c) meteorological phenomena (drought, flood, etc.),
  (d) mast years (heavy seed crops),
  (e) pests, diseases, fire, or any other aspect which may affect growth.

The measure crew must check doubtful items and make sure that the current measurements are correct. They should record that such checks have been made. Decrements and other anomalies in the data should not be altered once the measure crew has departed the plot. Although these data may at times look unrealistic, editing the database to alter these data may cause significant loss of information. At best, this practice may result in unrealistically low estimates of standard error associated with any functions developed. At worst, it may exclude the opportunity to investigate some originally unsuspected event (e.g. weather patterns and climate change) or unforeseen topic. There may be good reasons to edit or omit data from some specific analysis, but the main database should never be altered.

The quality and cost of data available for analysis may be improved substantially through the use of electronic data recorders (e.g. Fins and Rust 1987, Wood 1990). Electronic hand-held devices enable basic checks of the input data to eliminate simple errors (e.g. transposition) at their source, comparisons with previous measures, and speedy transfer of data to a central database. They can ensure that the measurer does not progress to the next tree or plot until all necessary variables have been recorded.

When to Remeasure

Theoretically, the frequency with which plots should be remeasured is influenced by two factors: ease of re-locating and identifying trees, and the rate of growth or change relative to the measurement error. The increment should be substantially greater than the error associated with measurement if the remeasure is to be useful; shorter intervals lead to excessive variance in regression functions. If the precision of measurement is low relative to the standard deviation of the fitted model (i.e. $\delta/\sigma>0.1$) (Aitkin *et al.* 1989) inflated estimates of variance may be obtained, unless this is allowed for explicitly in the model. To provide sufficiently precise diameter increment measurements, trees on permanent plots should be measured to the nearest millimetre, and the remeasurement intervals should be sufficiently long to allow at least one centimetre of increment to accrue. One theoretical study in Nordic forests indicated that longer intervals gave more reliable estimates of volume increment, and concluded that it was reasonable to measure once every five years (Ranneby and Rovainen 1994).

Remeasurements must be sufficiently frequent to ensure that the location of the plot and/or identities of stems are not lost; in some forest types this may be as frequent as every two years. Conversely, cost efficiencies demand that remeasurements should not be unnecessarily frequent. In the tropics, an interval of two to five years may be appropriate for the plot remeasurements.

As annual increments are generally required, measurements should be taken on the anniversary of the previous measure whenever possible, especially for annual or biennial measurements. Always try to measure during the same season, as trees in the seasonal tropics may exhibit marked seasonal fluctuations in girth due to changes in xylem water tension (e.g. Leigh *et al.* 1982, Lieberman 1982). Remeasurements should avoid periods of rapid change (e.g. bark shed, rapid growth), should aim to measure during dormancy where it occurs, and should try to replicate environmental conditions at the previous measure (e.g. avoid remeasuring immediately after rain if the previous measure was after a drought).

Remeasurements should also be taken at the time of (preferably both before and after) harvesting (or silvicultural treatment), and as soon as possible after wildfire, cyclone or other disturbance. Knowledge of impending harvesting or treatment is required, particularly if plot boundaries are concealed, so that the necessary measures can be arranged. This requires good cooperation with land owners and forest managers. It may be desirable to mark plots with buried steel pegs prior to harvesting to insure against the loss of plot markers.

Administration

Administrative and office procedures associated with maintaining measurement records are often neglected and can be a major cause of loss of information. Aspects to consider include the design of field forms, copying and storing the completed forms, and transferring the data to computer.

Forms should be designed specifically for plot measurements, and every column should be clearly marked with the data to be recorded (Fig. 5.5). The form should have no redundant fields, and staff should become accustomed to completing every field. Every form should indicate plot identity, date, page number (e.g. page 1 of 3) and the name of the assessing officer, and should be completed before departing the plot. If you create a new form, make a mock-up and test it thoroughly in the field before you commence normal plot measurements. Forms should be completed clearly and legibly (use a sharp dark pencil), and no alterations should be made after departing the plot. Forms should not be transcribed, as this invites transposition and other errors. Any duplicate copies required should be photocopied (where photocopiers are not available, transcriptions should be clearly marked as such, and should be carefully checked by a third person). Forms should be filed securely and unambiguously, preferably with one plot per folder, with the forms arranged in chronological order. It is a wise precaution to have a copy of the data stored in a remote location (e.g. district and head office).

Details from the previous measure should be available during plot remeasures. New remeasure forms can be printed by computer and may include printed details of previous measurements. Alternatively, parts of the previous measure record can be photocopied to form the new measure sheet, or the previous data can be downloaded into an electronic data recorder.

As the process of data entry may detect many illegible characters, errors and omissions, data should be entered onto computer as soon as practicable after collection, while the measure crew still recall some details of the plot. Data should be verified (i.e. re-entered independently and compared) by a different operator to detect any errors in data entry. Electronic data recorders offer several advantages, including partial validation at the point of data collection where checks can be made. When data entry is completed, further validation should screen the data for errors and omissions, and summary reports should be produced for the information of assessing officers and forest managers. Copies of the data should be made and stored in secure remote locations.

Obvious errors and omissions in the computer data file should be amended, if possible, but the temptation to manipulate the data so that it all looks consistent must be avoided. The data on the computer must accurately

Plot No ..... Subplot ..... Page ... of ...
Subplot Dimensions ...... x ...... Date ../../....
Orientation ..... Coordinates . . . .
Location . . . .
. . . .
Assessing Officer . . . .

Tree number Coordinates					
Family Genus Species Common name					
DBH Point of measure Valid/approx					
Alive/dead/cut/missing Erect/leaning/fallen Broken/injury					
Tree height Bole height Crown position Crown form Crown diameter					
Merchantable length Stem straightness Stem defects					
Notes: Flower/fruiting Pests/disease					

Fig. 5.5. Example of a field form for measuring PSPs.

reflect the field measurements. An anomalous measurement may or may not be due to measurement error in the field, and the database administrator's adjustment remains a guess rather than a fact. Any alteration inserted by the database administrator should be clearly indicated as such (in the *validity* field), and these alterations should be kept to a minimum. It is much safer to let users edit their own copy of the data as necessary for their own analyses, than to alter the master copy.

Data can be stored on computer in many ways ranging from simple text files to proprietary database management systems. No system is superior in all situations, and the appropriate system should be flexible (can you still access your data if the software fails?) and understood by the database administrator and by database users. Byrne and Sweet (1992) contrasted the database systems of 36 USA-based organizations and found that none satisfied all qualities that they deemed essential to a good system for a permanent plot database.

An effective system requires a considerable commitment in staff and resources to initiate and maintain the permanent plots, and this commitment must be on-going. The quality of the data is critical. Competent, well motivated and supervised staff are needed for measurement and for data management. Sufficient resources and trained staff are essential, or the quality and utility of data will deteriorate.

Reappraisal

Periodic reappraisal of data collection policy and practice is necessary to ensure that the data being collected are fulfilling current and perceived future needs. Reappraisals should address specifically two questions concerning data quality: "Is the specified quality adequate for current and perceived future needs?" and "Is the specified quality being attained?".

The need to sample extremes of forest condition has already been discussed. However, the concept of what is extreme changes over time. Thus it is necessary to consider if the extremes being sampled are sufficient, and if not, new plots should be established. The cost of data collection and handling is high, so plots should be abandoned when no longer useful. The decision to terminate a plot with a long measurement history should not be taken lightly, as these plots may be the most valuable for benchmarking model performance (see Chapter 11). However, it is inevitable that natural (and human) perturbations (e.g. lightning strike, landslip, insect or fungal attack, unauthorized felling) will extensively modify some plots. Such plots may no longer provide useful tree growth data, but may provide useful regeneration and other ecological data.

Perception of future requirements will change over time, so the data collection policy should be periodically amended to conform with these perceived needs. These amendments may require the termination of some plots, establishment of others, addition of new variables to be measured, or the deletion of others. However, changes in measurement procedures (especially deletion of variables) should not be undertaken lightly; stable, consistent measurement procedures are essential for growth research. There should be substantial and continued resistance to changing the plot measurement system.

Dynamic inventories should satisfy the data requirements for growth models for decades ahead. In order to provide for this next generation of growth models, it is appropriate to critically appraise the utility of the present dynamic inventory, and to establish new plots specifically directed at collecting data for such future growth models. Such a series of elite plots should sample the range of forest conditions (and include thinning studies), but should be established in limited numbers so that appropriate care and attention can be given to detail and accuracy. The emphasis with these plots must be quality, not quantity.

Existing Data

The greatest problem facing many agencies is that the data necessary for growth model development are not available. Plots may not have been established, may have been neglected or abandoned, and measure records may have been lost. As there is little that can be done to salvage such lost data, it is imperative that care and attention are devoted to existing plots and their measure records. Other problems which severely restrict the utility of data include unreliable measurements, changes to procedures, and mistaken or undetermined species identities.

Data for the development of growth models may derive from plots which were established for purposes other than growth modelling. Such plots may sample a restricted range of stand conditions, omitting very poor and exceptionally productive sites, and avoiding extremes of stocking. Thus these data may not provide an efficient means to estimate response surfaces by regression equations to predict the behaviour of the forest under various conditions. Records concerning the establishment of many plots are sketchy or unavailable, and the reasons for the placement of these plots are frequently not clear. Some plots may have been randomly or haphazardly located in defined strata, but others may have been subjectively located. Any departure from a stratified random approach in establishing these plots requires some soul-searching on the part of the modeller, in considering the possible effects of personal bias in choosing plot locations, particularly where site quality cannot be reliably quantified.

Permanent plots should receive representative management (harvesting, silvicultural treatment, etc.), except for experiments which sample extreme stand conditions. This may be assured where plots are marked with subterranean or other invisible markers, but intentional or unintentional bias in logging, treatment and other management may become significant when the plot is visible. Such management bias may not be a problem where it is reflected in the stand structure (e.g. removal of trees), but the effects of differences in logging damage and climber cutting may be more insidious. Differential management should be reflected in stand structure, but tests of some permanent plots established for 50 years in Queensland failed to detect differences between permanent and temporary plots established adjacent to them.

Growth in the tropics is often highly variable, and this variation may be attributed, at least in part, to factors such as weather, seed years, pest populations, disease outbreaks, fire damage, etc. Such information may be useful for interpreting apparently anomalous data detected during analyses, but is infrequently recorded and rarely transferred to the database. Evaluation of site productivity is a major obstacle in predicting yields in mixed forests, and development of a method for reliable site evaluation, and acquisition of the necessary data should be a priority.

How serious are these deficiencies so often present in data available for growth modelling? It is impossible to predict what difficulties these and other deficiencies may introduce, until the data are actually used in earnest. No data set can be perfect, but many will be found to contain deficiencies that will frustrate future analyses. Although plot remeasurement may appear to the measure crew to be unrewarding, collection and management of dynamic inventory data is vital to the informed forest management.

Synthesis

Permanent plots provide the basis for growth modelling, yield prediction and sustained yield management, and the reliability of these data is crucial to these and many other aspects of forest management. It is appropriate that this chapter appears near the middle of this book; data collection is central to model construction. To obtain reliable data, it is necessary to:

1. ensure consistent standards,
2. sample a wide range of stand and site conditions,
3. provide both passive monitoring and experimental plots,
4. number, mark and map all trees on all plots,
5. remeasure frequently enough to enable relocation of plots, but allow enough time for growth to exceed measurement errors, and
6. check that measurement records are unambiguous and secure.

In the next chapter, we assume that suitable data are available, and move on to consider how a growth model should be constructed.

Exercises

5.1. Design a form to record field measurements during the initial enumeration of a permanent plot in a forest near you. Take it to the field and try it ! Enter the data from the form into a text or spreadsheet file on a computer. What problems did you detect in the field and during data entry, and how would you improve your form next time? Would you use the same form when the plot was re-measured; if not, what changes would you make? How could you include some of the data from the initial measure on the remeasure form, so that field crews could cross-check these details?

5.2. Document ways that the data collected as part of Exercise 5.1 could be validated on the computer. What additional checks could you make when remeasured data become available? What errors might remain undetected by these procedures? Could these procedures be implemented on an electronic data recorder so that these checks could be made automatically in the field during plot remeasurement?

5.3. How do your field form and validation procedures compare to those in use by your organization (or by your local forest service or research institute)?

5.4. You are a research officer at a field station and have funds to investigate growth response of *Callitris* to different stand basal areas. Your data will contribute to a larger growth modelling project. You have a large area of homogeneous *Callitris* forest with 17 m^2 ha^{-1} stand basal area and 17 m site form (see Fig. 2.2, p. 20), and funds for five plots (one control and four treatments). How would you manage your project? Are five plots enough, or would you prefer to have twice as many plots of half the area? What area and dimensions should the plots be? How wide should the buffer be? What complications would be introduced if your forest contained more than one species, and how would you overcome them? Discuss. *Hint:* check Fig. 5.2 (p. 84) again.

5.5. You are involved in a project to adapt the FORMIX model (p. 41) to your region. Could you use the data collected in Exercise 5.1? What additional data would you need, and how would you go about collecting it?

Chapter Six

Constructing Growth Models

Constructing a growth model is not easy, even if suitable data are available. Some guidelines can be given, but a universal step-by-step procedure is not feasible because too much depends upon specific characteristics of each case. There are many issues involved, and this chapter aims to provide a general understanding rather than an intimate knowledge of how to deal with the problems. But after reading this chapter you should realize where problems are likely to occur, and be able to seek more detailed advice from one of the standard texts indicated.

A good model does not simply happen; it is planned that way. You cannot simply combine several haphazardly formulated relationships and expect to get reliable predictions. Instead, you should give careful thought to the design of the model at the outset of model construction. You should consider

1. what the model will be used for,
2. what inputs will be provided,
3. what outputs are required,
4. the data available for fitting the model, and
5. the resources available to construct, test and use the model.

These important factors may vary from case to case, so few universal guidelines can be given. However, two general principles apply:

1. Parsimony or *Occam's razor*: Entities should not be multiplied beyond necessity. In other words, do not include unnecessary variables and parameters in your model.

2. Keep it simple. Unnecessary complexity does not improve a model, and may create many problems. Simplicity may be defined in many ways, but a useful definition is *ease of disproving* (Oderwald 1991).

Model Design

At this point, it is necessary to know the capabilities of the different modelling approaches, and the requirements of the particular situation for which the model is being constructed (see Chapters 2–4). The first step in model construction is to prepare an outline of the model (cf. Exercise 4.1, p. 77), formulate the functional relationships required, and fit the functions to data. Here we consider some aspects of model formulation, including what variables to include, what equation to use, and how to fit the equation to the data.

Choice of Variables

Variables used in growth models should not be an arbitrary collection of those correlated with growth or yield in a forest stand, but should be carefully chosen to ensure biologically realistic predictions across the whole range of possible conditions. To understand how to achieve in a model, we need to begin by defining some types of variables that may be used in modelling.

The *response* variable (Y variate) is the thing that we are attempting to predict or explain, and is sometimes called the dependent variable. The *explanatory* variables (X variates) are those used as the basis for the prediction. These are sometimes called independent variables, but I avoid this term because they are rarely independent in the statistical sense. *Binary* variables take the value zero or one, and may be useful as a "switch" to alter the response for particular situation (e.g. habitat or soil type). They can be included in regression analyses just like any other explanatory variable. The multi-dimensional analogue of the binary variable is the *qualitative* variable, which may take a given range of integer values (i.e. 0, 1, . . ., *n*), and is equivalent to a set of *n* binary variables. If a regression package does not allow the use of qualitative variables, the same result can be obtained using *n* binary variables (Z) with $Z_i = 1$ when the qualitative factor (e.g. soil type) is i, and zero otherwise.

There are several factors which may guide the selection of possible explanatory variables. Most obvious is the need for reliable data. Obviously, if a certain variable is not present in the data available for model development, then that variable cannot be included in regression analyses leading to the development of a growth model. Less obvious is the availability of information at the time of application of a growth model. Research plots may record numerous measurements concerning stem form, crown size and shape, etc. However, if the principal use of a growth model is to project operational inventory data in order to schedule timber harvests, then only the variables which are (or can be) routinely measured in inventory assessment should be included in the model as explanatory variables.

As in any application, the results are only as reliable as the inputs. Thus a growth model which employs a variable that cannot be accurately

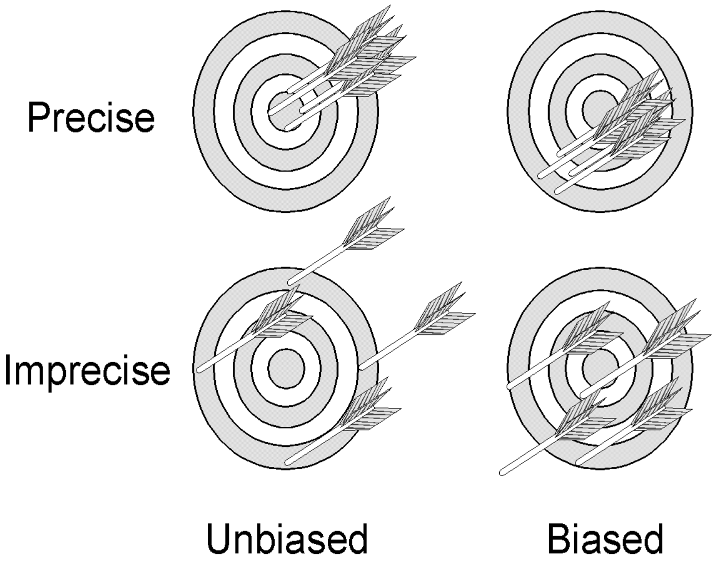


Fig. 6.1. Concepts of precision and bias. The target is a useful illustration, but the true value is usually not known when sampling. Thus accuracy is judged from precision, but care is needed to avoid bias.

determined (i.e. precisely and without bias, Fig. 6.1) is unlikely to provide good estimates of growth and yield. This applies not only to development data, but also to data used in applications. For example, internal stem defect may be measured with painstaking accuracy in research data, but since it cannot yet be determined with reasonable accuracy in the field, it should not be used as an explanatory variable in growth models.

An understanding of the biological processes affecting growth and change in trees and forests can help modellers in two ways. It should help to formulate hypotheses and select potentially useful explanatory variables and relationships for the development of models. It is also necessary for proper evaluation and interpretation of alternative model forms. Any relationship that violates accepted biological principles should be rejected, even if it results in efficient predictions for a particular data set. For example, a negative coefficient for site index may suggest that there is less growth on better sites, and this is not tenable with the concept of site index. Such conflicts are often indicative of an error (e.g. some data incorrectly recorded or entered on computer) or other anomaly in the database (e.g. excessive correlation between explanatory variables). Special techniques such as ridge regression may alleviate some of these problems, but anomalies in model coefficients should cause careful re-examination of the principles, the model and the data.

Passive monitoring plots (i.e. without experimental treatments) may reveal that the highest growth rates can be found on the best sites with high stand basal areas, and that little growth occurs on poorer sites with low basal areas (see p. 88). A growth model constructed from such data might suggest that greater increments accrue in stands with more competition, as the effects of site productivity and stand density would be confounded. Thus a haphazardly constructed model derived from such passive monitoring data would predict a reduction in diameter increments following thinning, whilst a model from experimental data (e.g. thinning studies) would show an increase in diameter increment. Even if no experimental data are available, the modeller can improve on the haphazard model by deliberate inclusion or exclusion of terms in the model (e.g. by excluding the contradictory site productivity term, even if significant, to ensure sensible extrapolation).

Growth and Yield Equations

It is important to understand the relationship between growth and yield. *Growth* refers to the increase in size of a population or an individual over a given period of time (e.g. growth in volume of a stand, in $m^3\ ha^{-1}\ y^{-1}$). *Yield* refers to the final size of a population or individual at the end of a certain period (e.g. total volume produced by a stand, in $m^3\ ha^{-1}$), and usually includes any removals (e.g. thinnings). A growth equation for even-aged stands predicts the growth of diameter, basal area or volume in units per annum, whilst a yield equation predicts the diameter, stand basal area or total volume production attained at a specified time (i.e. age). Thus a growth function may indicate that at age t a stand is growing at $\mathrm{d}y/\mathrm{d}t$ units per annum, whereas an equivalent yield function indicates that a stand at age t has produced y units. The notation $\mathrm{d}y/\mathrm{d}t$ simply means the change in y observed during a very short period of t, so that if $t_2 - t_1$ spans a very short time, Δt becomes very small, and

$$\frac{dy}{dt} = \frac{\Delta y}{\Delta t} = \frac{y_2 - y_1}{t_2 - t_1}$$

This is known as differentiating, and the converse, integration, is the analytical equivalent of summing the increments observed over many short periods, or alternatively, calculating the area under the growth curve. You need to understand these concepts for efficient modelling, and should refer to an introductory calculus text for clarification and further details.

Figure 6.2 illustrates growth and yield curves derived from a single equation. The symbols on both curves represent equal 5-year intervals, despite their unequal spacing on the growth curve. The growth function in Fig. 6.2 is expressed as a function of stand status y, but could equally well be given as a function of time t. The two equations are intimately related;

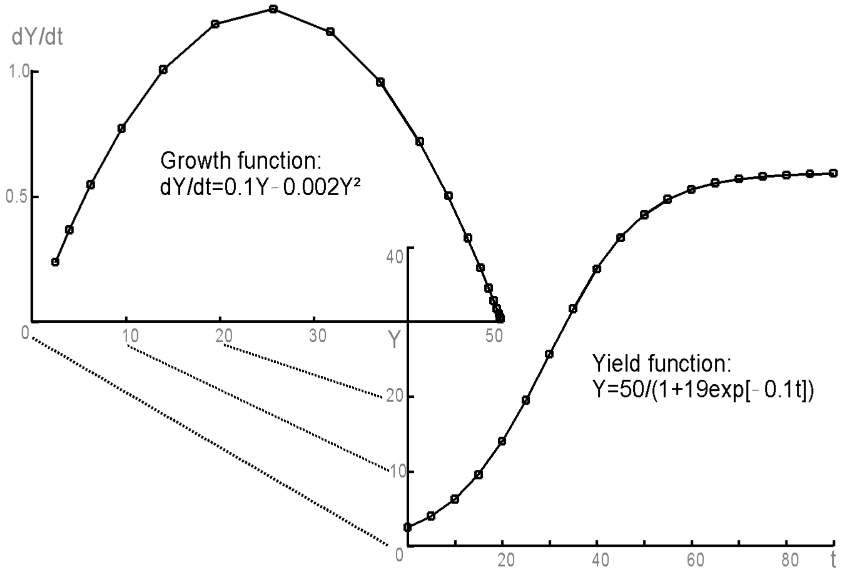


Fig. 6.2. Growth and yield forms of equations and curves are analogous. This is the logistic equation, but the principle applies to all growth and yield equations. Points on both curves represent 5-year time intervals.

either one can be derived from the other (if Y_0, the value of Y when $t = 0$, is given).

A yield equation may be produced by integrating the growth equation, and conversely, a growth equation may be obtained by differentiating the yield equation. Yield may also be determined by projecting the growth model, one step at a time, through the required number of years and summing the increments. The latter is the more flexible approach, as yield models generally do not allow for changes in management regime. In uneven-aged forests, models which express the status of the forest at some future time are termed yield models, and those which express increment are growth models.

Empirical Equations

Empirical equations are expressions which describe the behaviour of the response variable without attempting to identify the causes or to explain the phenomenon. This does not mean that empirical functions do not provide biologically realistic predictions, nor does it mean that they are inferior to supposedly biologically-based equations. They can and should be formulated to behave in a biologically realistic way across a wide range of possible conditions. The most widely used form of empirical equations, and the only form considered here, is the linear equation.

Linear equations are not necessarily straight lines, but are equations in which the explanatory variables enter in a linear (i.e. additive) fashion:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + e \tag{6.1}$$

where Y is the response variable (e.g. diameter increment), the Xs are explanatory variables (e.g. tree diameter, site productivity, stand basal area), the βs are parameters to be estimated, and e is the error term. For convenience, the error term is often omitted from equations, but is assumed to be present in its additive form unless explicitly shown otherwise, or the left-hand side of the equation refers to the expected response $\hat{Y}$ (note that $Y = \hat{Y} + e$). An explanatory variable X_i may be a transformation of X_j (e.g. X_j^2 or $\ln X_j$) or an interaction term (e.g. $X_j X_k$). The response variable Y may also be transformed to alter the shape of the relationship or to emulate a multiplicative relationship between the explanatory variables (e.g. $\ln Y$).

Linear equations are widely used in growth and yield studies, and offer several advantages. Most computers and many pocket calculators have reliable packages to fit such equations to data. The solution to the equations is unique, easily obtained, and rather robust, even when assumptions implicit in the method are violated. However, there are also disadvantages with these and other (e.g. non-linear) empirical equations. They often do little to further understanding of the processes involved in the system being modelled. While they may describe the data satisfactorily, they may give anomalous estimates for values lying outside the range of data on which they are based. Careful choice of the variables used and transformations applied will help to overcome these problems.

Theoretical Equations

In contrast to empirical equations, theoretical equations have an underlying hypothesis associated with the cause or function of the phenomenon described by the response variable. There are few theoretical equations formulated specifically for forestry applications. Most theoretical equations have been borrowed from other disciplines, and as a result may be rather empirical for some forestry applications. However, some general principles govern the behaviour of many systems, and provide the basis for these theoretical equations.

Bertalanffy (e.g. 1941, 1942, 1949, 1968) hypothesized that the growth of an organism could be represented as the difference between the synthesis and degradation of its building materials. He assumed that the processes of anabolism (synthesis) and catabolism (degradation; these assumptions were attributed to Pütter 1920) could be expressed as allometric functions of mass (Y), and thus growth ($\mathrm{d}Y/\mathrm{d}t$) would approximate

$$dY/dt = \beta_1 Y^{\alpha} - \beta_2 Y^{\gamma}$$

where α and γ are the constants of anabolism and catabolism respectively. He argued that catabolism was generally directly proportional to weight in animals, and that in any case, small deviations in γ from unity would have little effect on the form of the relationship. The parameters α and γ are highly correlated, causing problems when fitting the model, and the simplified form with $\gamma = 1$ is more convenient and fits most data almost as well as the more general form of the equation.

Bertalanffy (1949, 1968) argued that α could vary between b and 1, but would generally be b for animals. He did not propose any specific value of α for plant individuals or populations. The generalized form of the equation (i.e. value of α not constrained) is sometimes called the Chapman-Richards equation because of studies by Richards (1959) on plant growth, and by Chapman (1961) with fish populations, but the equation had been reported by Mitscherlich (1919) and was used by German foresters during the 1950s (e.g. Weck 1951).

The Bertalanffy equation may be used at the stand level or the tree level. Here we consider the stand level implementation (tree level aspects will be considered in Chapter 8). At the stand level, the derivation of the equation is rather different, because catabolism may not remain proportional to stand volume or basal area (heartwood does not respire), and anabolism is asymptotic to an upper limit (reached at canopy closure). Consider the following argument. Assume that living tissue may be proportional to girth times height $\Sigma dh \propto G^{\alpha}$ (where G is stand basal area, and $\alpha \approx 0.75$). Expressing stand volume increment as the difference between gross stand anabolism (A) and stand catabolism ($\beta_0 G^{\alpha}$), put $dV/dt = A - \beta_0 G^{\alpha}$. Assume that an allometric relationship exists between stand volume and basal area ($V = \beta_1 G^{\gamma}$, $\gamma \approx 1.0$), so that

$$\frac{dG}{dt} = \frac{A}{\beta_1 \gamma} G^{1-\gamma} - \frac{\beta_0}{\beta_1 \gamma} G^{\alpha+1-\gamma}$$

Anabolism is related to the amount of sunlight intercepted by the photosynthetic surfaces of the trees in the stand, and may exhibit an allometric relationship for low basal areas, but should be asymptotic as basal area increases, and a constant may be sufficient for well stocked stands. For an allometric relationship $A = \beta_2 G^{\eta}$ with η small ($\eta \rightarrow 0$),

$$\frac{dG}{dt} = \frac{\beta_2}{\beta_1 \gamma} G^{1-\gamma+\eta} - \frac{\beta_0}{\beta_1 \gamma} G^{\alpha+1-\gamma} \approx \beta_3 G^{\eta} - \beta_4 G$$

The value of $\alpha + 1 - \gamma$ should be close enough to 1.0 for practical purposes, but several empirical studies using the Bertalanffy equation at the stand level have reported coefficients with signs other than anticipated (i.e. $dY/dt = -\beta^{3} Y^{\eta} + \beta_4 Y$, e.g. Moser and Hall 1969, Murphy and Farrar 1982).

García (1974) examined an alternative theoretical equation for stand volume growth in plantations. He assumed that

1. gross increment in stem volume is proportional to the increase in dry weight $\Delta V \propto \Delta B$,
2. stand density remains sufficiently high to ensure full occupancy of the site and that assimilation per hectare is constant,
3. respiration is proportional to the quantity of live tissue which was in turn proportional to the surface area of the stems and thus the product of mean diameter ($\overline{d}$) times the height ($\overline{h}$) and the number of stems (n).

Thus stand volume increment (ΔV) could be predicted as $\Delta V = \beta_0 - \beta_1 n \overline{h}\,\overline{d}$. Assuming a constant form factor (f) so that $V = f G \overline{h}$ with stand basal area $G = \pi n \overline{d}^{\,2}/4$,

$$\overline{d} = \frac{2}{\sqrt{\pi f}} \sqrt{\frac{V}{n\overline{h}}} \quad \text{and thus} \quad \Delta V = \beta_0 - \beta_2 \sqrt{Vn\overline{h}}$$

Despite this nice theoretical development, the model did not fit the data well. Both these examples serve to illustrate some limitations in our present ability to quantify biological principles and reconcile them with empirical data. The "theoretical" equations presently available remain rather empirical in application.

Empirical Analogues of Theoretical Equations

The Bertalanffy equation is a member of a family of asymptotic, nonlinear growth equations (Fig. 6.3) including the:

1. Monomolecular growth equation when $\alpha = 0$:

$$Y = Y_{max}\left(1 - e^{-\gamma t}\right) \qquad dY/dt = \gamma(Y_{max} - Y)$$

2. Gompertz growth equation when $\alpha = 1$:

$$Y = Y_{max}\, e^{-\beta e^{-\gamma t}} \qquad dY/dt = \gamma Y \ln\left(Y_{max}\, Y^{-1}\right)$$

3. Autocatalytic growth (or logistic) equation when $\alpha = 2$:

$$Y = Y_{max}\left(1 + \beta e^{-\gamma t}\right)^{-1} \qquad dY/dt = \gamma Y (Y_{max} - Y) A^{-1} = \gamma Y - \eta Y^2$$

4. Bertalanffy equation (b $\leq \alpha \leq 1$):

$$Y = Y_{max}\left(1 - \beta e^{-\gamma t}\right)^{\eta} \qquad dY/dt = \beta_1 Y^{\alpha} - \beta_2 Y = \gamma Y\left[\left(Y_{max} Y^{-1}\right)^{1-\alpha} - 1\right]$$

$$= \beta_1 Y^{\alpha} e^{-\gamma t}$$

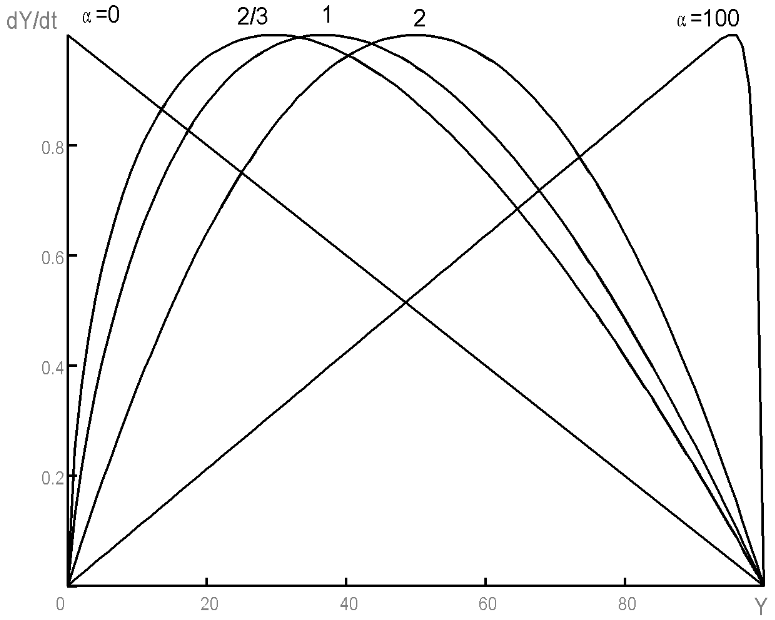


Fig. 6.3. Forms of the Bertalanffy equation $dY/dt = \beta_1 Y^{\alpha} - \beta_2 Y$.

where Y is the size of the organism, Y_{max} is the asymptotic maximum size, t is time, and α, β and γ are constants. There are some difficulties in fitting the generalized three-parameter Bertalanffy equation to data with non-linear regression (Ratkowsky 1983, 1990), and these can be avoided by using one of the two-parameter relatives in which α is fixed. However, Zeide (1993) showed that the Bertalanffy equation can be converted to a linear equation:

$$\ln(dY/dt) = \beta_0 + \alpha \ln Y + \beta_2 t$$

where t is age, $\beta_0 = \ln(\beta_1)$ and $\beta_2 = -\gamma$. In a comprehensive empirical study, Zeide (1993) examined many alternatives to the Bertalanffy equation, and concluded that it, and two variants,

$$\ln(dY/dt) = \beta_0 + \beta_1 \ln Y + \beta_2 \ln t$$

$$\ln(dY/dt) = \beta_0 + \beta_1 \ln Y + \beta_2 Y \tag{6.2}$$

were the best descriptors of height, diameter and volume growth. Equation 6.2 was slightly better than the two alternatives for predicting height and diameter growth of individual trees.

It is not surprising that these equations are equally efficient at predicting increment of tree diameter, height and other parameters. We can express many equations in a form which indicates relative growth, e.g. the Bertalanffy equation for diameter increment is $\partial d/d = \beta_1 d^{\alpha-1} - \beta_2$. Then for any parameters for which there exists an allometric relationship $v = \alpha d^{\gamma}$ (e.g. for basal area, $\gamma = 2$ and $\alpha = \pi/40\,000$), we have $\partial v/v = \partial(\alpha d^{\gamma})/\alpha d^{\gamma} = \gamma\, \partial d/d$, so that

$$\frac{\partial d}{d} = \frac{\partial g}{2g} = \frac{\partial v}{\gamma v} = \ldots = \beta_1 d^{\alpha-1} - \beta_2$$

Thus the Bertalanffy and similar equations should apply equally well to all tree and stand parameters which exhibit an allometric relationship.

Bailey (1980) suggested another equation which has no biological basis, but which is an extremely flexible function, describing a family of asymptotic curves that encompass both the Bertalanffy (when $\gamma = 1$) and the Weibull equations (when $\eta = 1$):

$$Y = Y_{max}\left(1 - \exp\left[-\beta t^{\gamma}\right]\right)^{\eta}$$

However, it cannot be converted to a linear form, and must be fitted using non-linear methods.

An empirical study by Martin and Ek (1984) found that carefully formulated empirical equations could be more accurate than theoretical equations for a wide range of data. However, theoretically based equations may be more reliable for predictions which involve extrapolations beyond the range of the data. Figure 6.4 illustrates how the Bertalanffy equation may provide safer extrapolations than an empirical polynomial equation. However, polynomial equations are notorious for their poor extrapolation, and other linear equations would perform better in this instance (e.g. an asymptotic relationship such as $Y = \beta_0 + \beta_1 X^{-1}$).

Equations can be empirical or otherwise, but should be carefully formulated to provide sensible predictions across the full range possible for all explanatory variables. Explanatory variables should not be chosen haphazardly, but should be selected deliberately to provide realism and robustness in the model. Mere goodness of fit is no justification for adopting a given model since several functions may fit the data equally well.

Regression Techniques

There are many techniques available for fitting equations to data, and the appropriate one to use depends on the relationships chosen to represent the system, the nature of the data, and on the resources available to fit the model. There may be only one guideline that holds for all approaches: plot the data, the fitted model and the residuals to check and compare them.

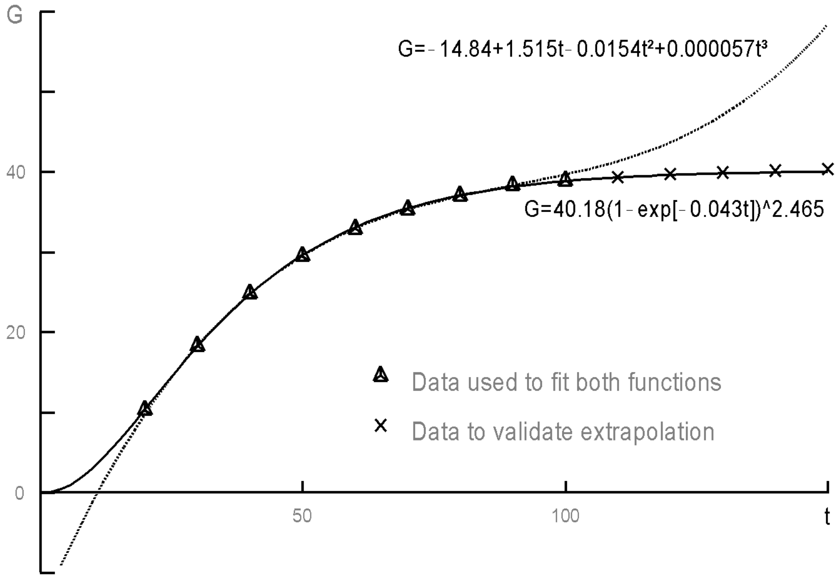


Fig. 6.4. Dangers of extrapolating empirical models (after Payandeh 1983).

Examining data and models

Careful visual examination is one of the best ways to begin to understand a set of data, to choose a model to represent the data, and to check for errors in the model and in the data. Four series of graphs should be prepared and examined:

1. scatterplots of the raw data showing the response variable plotted against possible explanatory variables;
2. graphs showing both the raw data and the fitted model plotted against the chosen explanatory variables;
3. plots of the residuals versus the fitted values, versus alternative explanatory variables, and of standardized residuals versus normal order statistics;
4. graphs showing the fitted model evaluated for a wide range of explanatory variables.

Figure 6.5 illustrates some of these graphs. Scatterplots A–C show the raw data, with the response variable (diameter increment Δd) plotted against three potential explanatory variables, tree diameter (d), basal area in larger trees ($G_{>d}$), and stand basal area (G). These allow the modeller to check that

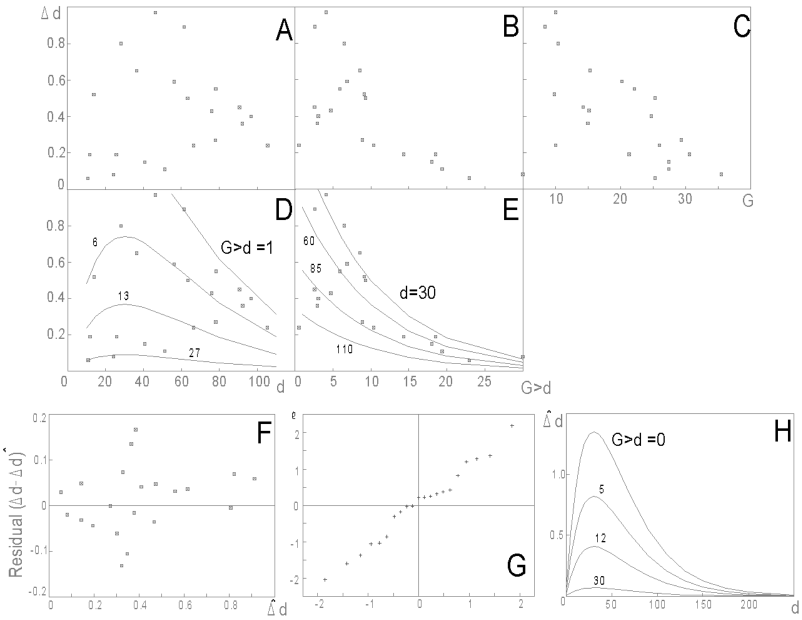


Fig. 6.5. Graphs for screening data and models. Scatterplots A–C show raw data, D-E include the fitted model, F plots residuals vs fitted values, G shows that residuals are normally distributed, and H evaluates the model at extremes.

there are no serious data errors, and give some ideas on potential relationships. Graphs D and E are the same as scatterplots A and B, but include the fitted model evaluated for the range of d and $G_{>d}$ in the data, and confirm that the fitted model is appropriate. The plot (F) of residuals $(\Delta d - \Delta\hat{d})$ versus fitted values $(\Delta\hat{d})$ may reveal trends not explained by the model (see below, and Fig. 6.6), and thus may indicate transformations that should be considered for inclusion in the model. In Fig. 6.5, the residuals are reasonable, except that the variance is rather large for increments of 0.3–0.4 cm y^{-1}. The analogous plot with studentized residuals indicates the quality of the fit in terms of standard errors rather than in the units of measurement (not shown, but left to the reader in Exercise 6.4). Residuals may also be plotted against other variables not in the model (e.g. Fig. 6.6, D). The graph of studentized residuals versus normal order statistics (Fig. 6.5, G) shows that the residuals have a near-to-normal distribution. Finally, graph H shows the fitted model evaluated for extremes of the two predictor variables in the model to demonstrate that the model gives sensible predictions under all circumstances. Readers are encouraged to fit a model to these data, given in Exercise 6.4.

Figure 6.5 illustrates just some possibilities, but there are many alternatives that can offer further insights. Residual plots may reveal non-

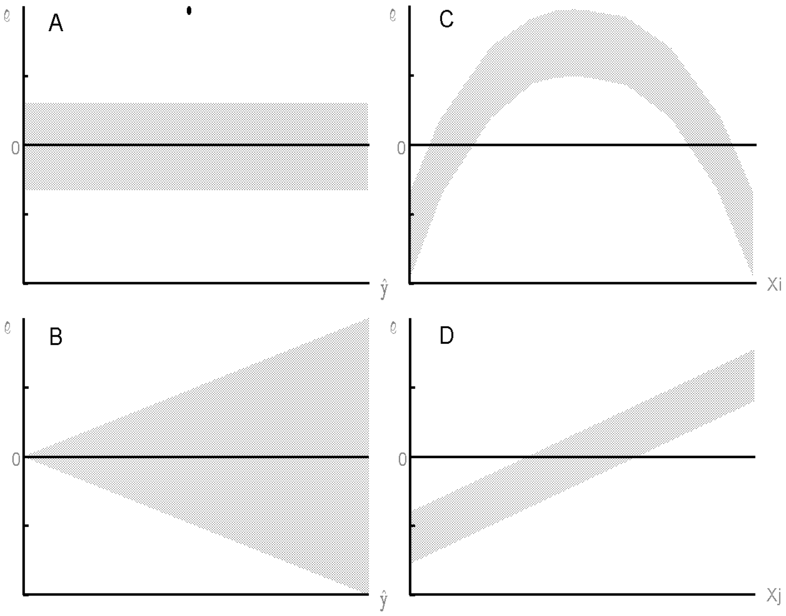


Fig. 6.6. Interpreting residual plots: (A) an outlier, (B) non-constant variance, (C) transformation required, and (D) variable X_j should be included in the model.

constant variance (Fig. 6.6, B) and outliers which warrant further investigation (A). It is useful to plot residuals against explanatory variables to look for transformations that may be required (C), and to check for additional variables that should be included (D). Residuals may be standardized and plotted against standard normal deviates to check the validity of model assumptions. Other plots (e.g. of cross-validation residuals) and criteria may offer further insights and readers should consult standard statistics texts (e.g. Weisberg 1985) for details of how to compute and interpret these. If possible, the fitted model should also be compared with data not used in fitting the model. Graphs like these, especially when coupled with comparisons using additional data, remain one of the best ways to detect errors in and inconsistencies between the raw data and the fitted model.

Hand-drawn Curves

Most models are fitted to data using regression techniques, but hand-drawn curves continue to be useful, especially where few data are available. However, when drawing curves, care should be taken to represent the data adequately. There is a tendency for hand-drawn lines to follow the first

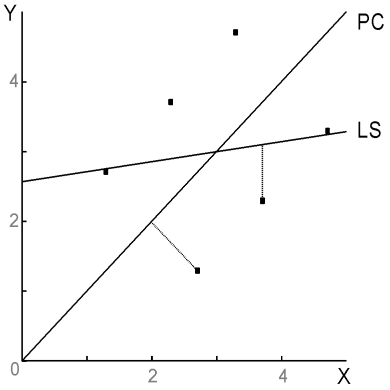


Fig. 6.7. Principal component (PC) and least squares (LS) fit to data. The naive hand-drawn fit may follow the principal component rather than the least squares fit.

principal component rather than the least squares fit (i.e. minimize the perpendicular distance from the data points to the fitted line, rather than the vertical distance as in least squares, Mosteller *et al.* 1981). This may overestimate the gradient, thus underestimating the response for small values of the predictor variables and overestimating for large values (Fig. 6.7). However, the effect may be small if the modeller is aware of regression techniques and conscious of the potential for bias.

If regression analysis is subsequently used to estimate parameters of an equation describing the hand-drawn line, remember that the goodness-of-fit statistics reflect the points chosen from the hand-drawn curve, not the goodness-of-fit to the original data. This will provide an over-optimistic estimate of the quality of the fit.

Linear Regression

Linear regression implies that explanatory variables enter the objective function in a linear or additive way (i.e. $\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots$; see Eqn 6.1 on p. 108, 111). It in no way implies that the resulting relationships are straight lines. This form of regression is widely used for fitting equations to data, and linear regression packages are available on most computers and many calculators. Three alternatives exist.

In the first, the user specifies the response and all the explanatory variables, and the computer merely determines the values of the required coefficients. This encourages a logical approach to model construction, as the user must decide which explanatory variables will best describe the response variable. Plotting residuals from preliminary models against other potential explanatory variables may assist in the formulation of the final model.

Stepwise regression methods automate the procedure. The user identifies the response variable and provides a list of possible explanatory variables. The stepwise algorithm chooses the explanatory variable most highly correlated with the response variable, then iteratively includes additional (or deletes previously included) explanatory variables until no further significant correlations can be found. There are two minor variants, forward selection (inclusion of variables only) and backward elimination (starts with all variables in the model and iteratively deletes non-significant ones). These alternatives do not necessarily select the same, or the best subset of possible explanatory variables. Problems may arise if there is much correlation between the explanatory variables in the data. Regression by leaps and bounds (Furnival and Wilson 1974) overcomes some of these problems, but is not often available in statistical analysis packages.

Combinatorial screening (Grosenbaugh 1967) fits every possible combination of the explanatory variables nominated. It will find the best combination of variables specified, but the possibility always exists that a better model could be constructed by using different transformations of, or interactions between, the explanatory variables. These automated methods focus on goodness-of-fit and overlook the importance of the deliberate selection of variables and transformations in ensuring sensible predictions.

Non-linear Regression

Non-linear regression enables the fitting of models in which the estimated parameters do not enter in a linear additive manner. Many theoretical and asymptotic models are of this form. For example, in the asymptotic equation $Y = \beta_0 + \beta_1 (\beta_2 + X)^{-1}$ the coefficient β_2 enters the equation in a non-linear way (compare the derivative $dY/d\beta_2$ for this equation with those in Eqn 6.1 on p. 108, 111). Whilst non-linear regression allows great flexibility in formulating models to ensure sensible extrapolation, it does have some limitations. One problem is that, unlike linear regression, non-linear regression does not necessarily provide a unique best unbiased solution for a given set of variables. Non-linear solutions are determined iteratively, and may be influenced by the estimation method and the starting conditions specified by the user.

Ratkowsky (1983, 1990) discussed why problems arise in iterative estimation of the parameters of non-linear models. These can usually be overcome by "reparameterizing" the model so that the model behaves in a "near to linear" fashion. The simplest method of iterative estimation, the Gauss-Newton method, can then be employed and the resulting parameter estimates are unbiased, normally distributed, minimum variance estimators. If the model does not behave in a near to linear fashion, the parameter estimates will not have these desirable properties and more complex estimation techniques may be necessary. In such cases, the use of analytical derivatives rather than computational approximations usually result in more efficient and more precise parameter estimation, but some sort of sensitivity analysis remains necessary to ensure that the global minimum has been reached, as some situations may lead to a local minimum rather than to the "best" fit.

Reparameterizing involves expressing the parameters of one model as a function only of the parameters of another model. For example, the models

$$Y = \frac{\alpha_1 X}{X+\alpha_2} \quad \text{and} \quad Y = \frac{X}{\beta_1 X+\beta_2}$$

where α_1, α_2, β_1 and β_2 are parameters, are reparameterizations of each other since it can be shown easily that $\alpha_1 = 1/\beta_1$ and $\alpha_2 = \beta_2/\beta_1$. The model is not changed in any way by this reparameterization and its fit to the data will be exactly the same in either form. However, the statistical properties of the parameter estimates may be very different depending on the form that is used.

Robust Regression Techniques

Robust regression refers to some techniques that attempt to overcome the problems of a non-normal error distribution and correlated explanatory variables. Unusual data points (e.g. Fig. 6.6, A) often occur in growth data, and although they may have an excessive effect on the least squares estimates of coefficients, they cannot validly be omitted from the analysis. Robust estimators offer one way to minimize the impact of such outliers in the database.

Schreuder *et al.* (1979) discussed three classes of robust estimators, and observed that of these, maximum likelihood estimation was most amenable to regression analysis. They demonstrated the effect of using three approaches of maximum likelihood estimation, with both relatively clean and noisy data. All three techniques gave estimates similar to ordinary least squares with the clean data, but the estimated parameters varied considerably when the techniques were applied to the noisy data.

Robust estimators are a compromise between including all the noisy data, and using an edited subset of the data from which the outliers have been removed. One advantage of editing data is that it forces the researcher to evaluate the data critically and decide if they come from the population of interest. Choosing the parameters which determine the degree of robustness in maximum likelihood techniques does not force the same decisions upon the researcher. Since the outcome from robust techniques depends largely upon the subjective determination of scale parameters, these techniques remain an investigative tool for the data analyst and are not a cure for bad data.

For most growth modelling applications a robust and reliable result can be obtained by fitting a carefully formulated linear equation using ordinary least squares linear regression. While this approach makes some assumptions unsatisfactory to theoreticians, it is relatively robust, repeatable and readily available.

Statistical Assumptions

Linear regression and other least squares methods are normally robust enough to give good estimates under adverse conditions, but it is important to understand the assumptions implicit in the method and to realize the possible implications of violating these assumptions.

Linear regression assumes that the response variable Y can be predicted from the explanatory variables $\boldsymbol{X}$:

$$\boldsymbol{Y} = \boldsymbol{X}\beta + e$$

where the *errors* (e) reflect natural variation and factors not included in the model. This is Equation 6.1 (p. 108, 111) written in matrix notation. The least squares method minimizes the sum of the square of the errors ($\sum e^2$), so the nature of these errors has a big influence on the qualities of the fitted model. The least squares method makes many assumptions made about these errors and the variables in the model, so the following discussion is necessarily brief and intended as an introduction to the selected references. In many cases, there are two ways to deal with departures from these assumptions:

(a) to employ a more sophisticated (and sometimes complex) analysis, or

(b) to improve the database and the model.

Sometimes the former option is the only one available, and thus references to analytical alternatives are given. However, the second option is usually preferable, so discussion here concentrates on ways to improve the database and the model formulation to overcome these deficiencies.

Independent Observations

The basic assumption of least squares analysis is that the errors are independently and identically normally distributed with zero mean and constant variance. The independence of observations may be compromised when several measurements are taken on the sample unit (e.g. remeasures of a permanent plot, many trees on a plot, or several measurements on a single tree). These repeated measurements on the same trees tend to be correlated, and are thus not statistically independent.

When ordinary least squares techniques are used with such data, the parameter estimates are unbiased, but the estimators of the covariance matrix of the parameter estimates and the residual variance of the regression equation may be underestimated. This means that it is not possible to carry out properly the normal statistical hypothesis tests that are usually necessary to apply the results of regression analysis.

West *et al.* (1984) reviewed how and when this problem arises and discussed some solutions that have been suggested. None of these alternatives is entirely satisfactory and only experience can suggest the most appropriate solution for a particular data set. In some cases such as when the number of multiple measurements is small in comparison with the number

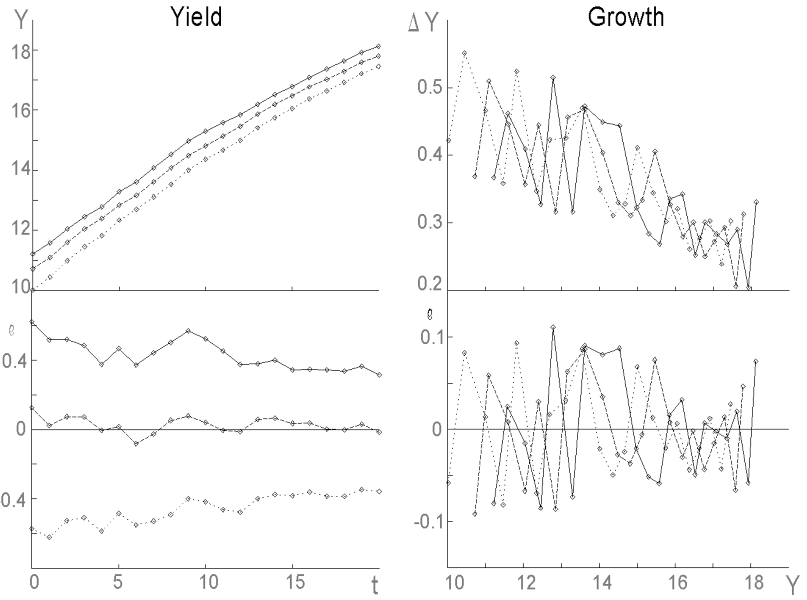


Fig. 6.8. Serial correlation may be minimized by modelling growth as a function of initial size (r=−0.30), rather than yield as a function of time (r=0.92) (simulated data with random perturbation for trees, time and measurements).

of sampling units, the problem may not be serious. The simplest solution may be to use only one observation from each sampling unit. Ordinary least squares regression is then appropriate and the problem is avoided altogether, but information is lost by using only part of the full data set. An alternative may be to use only one observation per sampling unit to formulate the model and test for the significance of terms, and having resolved the model, to fit it again using the full data set (West 1994).

In an empirical study of yield models, Borders *et al.* (1988) found no serial correlation when data were derived from non-overlapping growth intervals, and that problems associated with serial correlation may be model dependent. The specific problem of serially correlated errors appears worst when yield (i.e. total standing volume) is predicted from age. When growth is predicted from initial size (cf. predicting yield from age), these problems may be reduced and the conventional estimation methods may be acceptable (Fig. 6.8). If ordinary least squares are used despite serially correlated errors, unbiased parameter estimates will be obtained, but the variance may be underestimated. The implication is that parameter estimates may appear to be significant, when in reality they are not, possibly leading to the acceptance of a more complex model than is warranted.

Homogeneous Variance

In least squares methods, the errors e are assumed to have an identical distribution, i.e. to be *homogeneous*. In many biological applications, the variance is not constant, but may increase as the expected value of the response variable increases. This characteristic is known as *heteroscedastic*. The extreme residual values in heteroscedastic data have a big influence on the least squares estimates, and may disproportionately influence the form of the equation. Heterogeneous variance may be detected in residual plots (e.g. Fig. 6.6, C, p. 115), and tests for homogeneity of variance are discussed in standard texts (e.g. Weisberg 1985).

Two solutions to the problem exist. A variance stabilizing transformation can be used, or a variance estimating function can be used to weight the regression (e.g. Parresol 1993). Weighted regression assumes that the errors are distributed normally with variance ($v\sigma^2$), and that a function to predict v can be determined. The weights used in the regression are the inverse of this function ($w = v^{-1}$).

Suitable transformations can be determined graphically or analytically (e.g. Aitkin *et al.* 1989). If the variance is plotted against the mean, a horizontal line ($\sigma^2 = c$) indicates that no transformation is required; a straight line increasing with the mean ($\sigma^2 = \mu$) suggests a square root transformation and a line curving upwards ($\sigma^2 = \mu^2$) indicates a logarithmic transform. Alternatively, the slope of the line fitted to $\ln(\sigma^2)$ on $\ln(\mu)$ also indicates the appropriate transformation. If the line passes through the origin, a slope of 1 indicates that a square root, slope 2 a logarithmic, and slope 4 that a reciprocal transformation should be used.

If such a transformation has been used, predictions will contain transformation bias, the magnitude of which depends upon the variability of the data. Often this bias may be small enough to be ignored. However, where a poor fit is obtained, an adjustment should be made for this transformation bias when performing the back transformation (see any standard statistics text, e.g. Weisberg 1985). Weighted regression avoids the need for these transformations and corrections.

Transformations not only affect the distribution of errors, but also influence the explanatory variables. For example, a linear model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + e$$

assumes an additive interaction between the explanatory variables (X_i), whereas the transformed model

$$\ln Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + e$$

implies a multiplicative relationship between the explanatory variables and the errors

$$Y = \alpha X_1^{\beta_1} X_2^{\beta_2} \, . \, . \, . \, e$$

which may or may not be appropriate.

Normally Distributed Errors

The least squares method assumes that the errors e are normally distributed with constant (but unknown) variance $e = N(0, \sigma^2)$. Many standard statistical tests are sensitive to long-tailed distributions, and asymmetrical distributions may lead to biased estimates. Non-normal data may lead to inappropriate confidence regions about the regression surface if the usual formulae are used. Standard tests (see e.g. Weisberg 1985) may be used to check for departures from normality, and a useful visual indicator may be obtained by plotting studentized residuals versus normal order statistics, which are available in many packages.

Transformations may be used to correct departures from normality in data (e.g. Weisberg 1985). Count data (e.g. number of recruits) may be distributed in a Poisson fashion, which can be stabilized by a square root transformation. Data concerning proportions are often binomial, and can be stabilized by an angular (arc sine square root) transformation. The truncated distribution of bounded count data (e.g. mortality data, where k out of n trees die) may be improved by the logit transformation:

$$Y' = \ln \left(\frac{Y}{Y_{max} - Y} \right)$$

where Y lies between zero and Y_{max}, the upper bound.

Outliers

Outliers are observations that deviate greatly from the general trend (Fig. 6.6, A, p. 115). They may have a great influence on least squares estimates and, if due to errors, may lead to biased predictions. Since not all outliers are caused by error, they should not be deleted from the data set without careful investigation of their source and validity.

Increment data need special care when examining outliers, as negative values caused by measurement error may be associated with an excessively large value in the preceding or subsequent period. The negative outliers should not be removed from the analysis without removing the corresponding large positive value, or bias will result. Not all decrements are caused by measurement error. Some may be caused by natural variation in stem size or by bark shedding, and should not be removed from the analysis. These factors re-emphasize the importance of using a remeasurement interval sufficiently long that increments are large relative to measurement errors and any fluctuations in stem size.

Not all erroneous data are revealed as large outliers. Extreme values may exert great influence over regression results without producing large residuals (e.g. Fig. 6.9). Influence values and leverage plots (e.g. Weisberg 1985, Aitkin *et al.* 1989) enable such values to be detected. Observations with large residuals or influence values should be carefully examined and checked against the original field sheets. They should not be automatically rejected, but should always be carefully evaluated.

Independent Explanatory Variables

The explanatory variables $\boldsymbol{X}$ in the model are assumed to be independent, but are often correlated. If two (or more) variables are correlated so that $\beta_1\boldsymbol{X}_i + \beta_2\boldsymbol{X}_j = \alpha$, a singular matrix will result and no parameters can be estimated. This situation may occur in lists of potential explanatory variables for stepwise regression. For example, $\ln(d)$ is perfectly correlated with $\ln(g)$, since

$$\ln(g) = \ln(kd^2) = \ln(k) + 2\ln(d)$$

Stand basal area and crown competition factor may also be highly correlated (e.g. $G = \Sigma d^2 \pi/40000$ and $C_{cf} = 100\Sigma(\beta_0+\beta_1 d^{\alpha})^2 \pi/40000$, so $G \propto C_{cf} \propto \Sigma d^2$ if $\beta_0 \approx 0$ and $\alpha \approx 1$).

Collinearity may also exist between explanatory variables gathered in an opportunistic way without a deliberate sampling strategy. For example, data gathered only from undisturbed natural forest would not provide a good basis to fit the equation in Fig. 6.5, because tree size (d) would be highly correlated with basal area in larger trees (i.e. small trees always have large $G_{>d}$, and vice-versa). The nature of the problem can be gauged by imagining a child's play-house made by stretching a sheet between some stakes. If the stakes are arranged around the perimeter of the sheet, we will have a good model. The model is improved by some additional stakes in the middle, to give the "roof" the right shape. If the stakes form a single line (i.e. high correlation between X_1 and X_2) we will have problems draping the sheet. Even if we use rigid material (e.g. plywood; analogous to making some assumptions about the shape of the response surface), we will still have problems balancing it. The amount of "wobble" in the plywood is analogous to the variance of parameter estimates and model predictions. There is less wobble close to the stakes (i.e. within the sampled region) than there is away from them (i.e. extrapolations).

Such correlation leads to numerical problems in parameter estimation. The absolute value of the estimated parameters may be too large, the sign may be wrong, and parameter estimates may change substantially after the addition or deletion of a single data point. This situation exists where the computations are exact, and the result of even small rounding errors in the computer may make parameter estimates even more unreliable. Correlated explanatory variables may lead to substantial overestimates of the F-ratio and of t-tests, and thus to the acceptance of models containing more parameters than warranted. Variance inflation factors (Marquardt 1970) may indicate the severity of correlations between explanatory variables, and large values indicate variables that may contribute to poor extrapolations.

Fortunately, multiple collinearity does not inhibit our ability to obtain a good fit, nor does it affect inferences about mean responses or predictions from new observations, provided that these inferences are made within the region of observations (consider the play-house analogy and think about how the plywood wobbles at different distances from the stakes). However, it does mean that when formulating models, modellers should be careful to avoid the unnecessary inclusion of correlated terms in the model. The problem can also be minimized by deliberately sampling to include the full range of all explanatory variables (i.e. a good range of X_1 for every X_2, and vice-versa, e.g. see Fig. 5.3 on p. 87, and recall the dangers of passive monitoring on pp. 88 and 106).

Accurate Variables

Standard least squares regression assumes that the explanatory variables are accurate and have no error and no bias (in practice, bias may not matter, provided that it is the same in both development and applications data). However, in forestry applications, the explanatory variables may contain error, and may be rounded. The effect of this is to increase the variance of the estimates beyond the values predicted by the usual formulae. Thus the calculated confidence region will be conservative. Fortunately, the implications are minor provided that the range of each explanatory variable is large relative to its error, and the usual least squares analysis may be performed (Draper and Smith 1981). But caution is required when interpreting regression results that incorporate categorical data (e.g. crown classes coded as 1, 2 and 3), especially if they are included in the model as quantitative rather than qualitative variables.

It is also assumed that the response variable is continuous, when in practice it may be measured with finite precision and may take only discrete values. If the precision of measurement is low relative to the standard deviation of the fitted model (i.e. $\delta/\sigma > 0.1$) (Aitkin *et al.* 1989), this discreteness should be allowed for explicitly in the model. Maximum likelihood methods can be used to fit such models, but are considerably more complicated than the usual procedures. Failure to account for the discrete nature of the response variable will lead to inflated estimates of variance and the adoption of simpler models than may otherwise be indicated. The best solution is to maintain the highest possible precision in the response variable, by ensuring that all plot and tree parameters on permanent plots are measured in such a way to ensure that average growth (or other changes) is substantially greater than the likely measurement error (e.g. measuring tree diameters to the millimetre, at intervals long enough to allow 1 cm dbh increment on average).

Comparing Equations

It is not always easy to make meaningful comparisons between alternative functions. Simple F-tests provide a good basis for comparison where models have been developed from the same database, but it is more difficult to compare models fitted with different data sets. The most widely-used (and abused) criterion is probably the coefficient of determination (R^2) and variants such as the correlation coefficient (R) and the ratio of residual mean squares (R'^2). The coefficient of determination is given by $R^2 = 1 - RSS_m / RSS_\mu$ where RSS_m is the residual sum squares of the model under consideration, and RSS_μ is the residual sum squares about the mean. This expression illustrates two advantages of R^2: it is a ratio, and thus independent of units, and it is an indicator of how the model compares with a simple average ($R^2 = 0$) and with the perfect fit ($R^2 = 1$). However, the R^2 does not discriminate between "pure error" or natural variation, and lack of fit by the model. An R^2 close to one does not mean that it is the best possible model, or that it will provide good predictions.

Figure 6.9 (overleaf) illustrates a simple linear regression $\hat{Y} = \beta_0 + \beta_1 X$ fitted to four data sets formulated so that estimates of β_0, β_1, σ^2 and R^2 are identical (i.e. $\beta_0 = 3.0$, $\beta_1 = 0.5$, $\sigma^2 = 13.75$ and $R^2 = 0.667$, Anscombe 1973). Despite the similar R^2s, the graphs in Fig. 6.9 illustrate that the quality of the fit varies greatly between the four data sets. They reveal "pure error", an outlier, use of the wrong model (a quadratic term may be needed), and a case where the estimate of β_1 relies entirely on a single point with high leverage. Without further information, it is impossible to judge which of these models is suitable, but all except the first case warrant further investigation. A fuller discussion of this illustration was given by Anscombe (1973) and Weisberg (1985), but the importance of plotting both the data and the model remains obvious.

There are other disadvantages of the R^2 coefficient. It takes no account of the number of terms in the model, and more terms, even if non-significant, result in a higher R^2. Mallows' (1973) C_p is a measure of the expected variance associated with predictions from the fitted model and may be a better basis for comparison. Because Mallows' C_p accounts for the error in estimating each parameter, it leads to the acceptance of simpler models than many other decision criteria.

The R^2 is also influenced by transformations, and yield equations (e.g. $d + n.\partial d = f_1(d)$) and basal area increment equations ($\partial(kd^2) = 2kd.\partial d = f_2(d)$), both of which include initial size (d) in the response variable, nearly always indicate a higher R^2 than the corresponding diameter increment model ($\partial d = f_3(d)$), even though the functional relationship and the data may

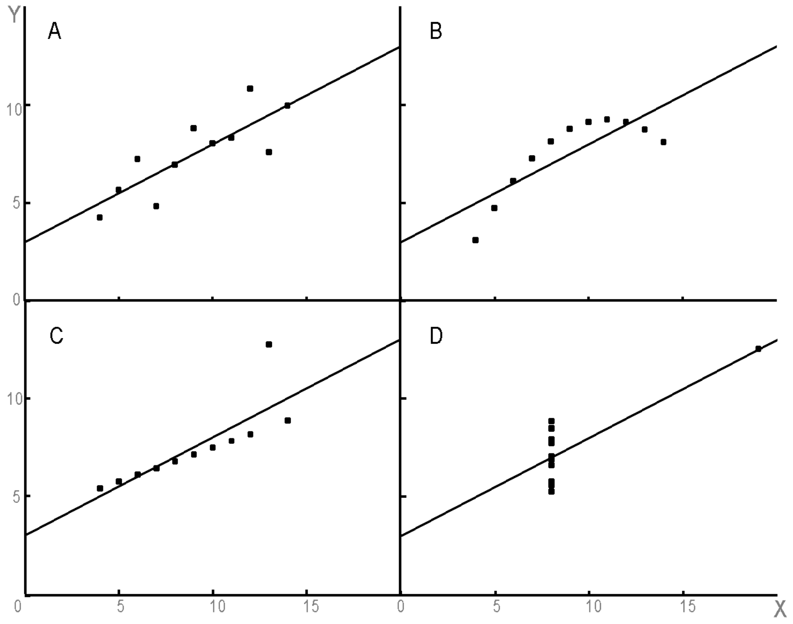


Fig. 6.9. The R² does not indicate how well a model fits the data. Here the plots reveal (A) pure error, (B) the wrong model, (C) an outlier, and (D) a point with high leverage, but all cases have R²=0.667 (redrawn from Anscombe 1973).

be the same. The Furnival (1961) index overcomes this problem by expressing the average standard error in the original untransformed units, and is derived from the maximum likelihood estimate. It can be computed as $I = s\,[f'(Y)]^{-1}$ where $f'(Y)$ is the derivative of the transformed response variable with respect to the untransformed variable, and s is the usual estimate of standard error. Furnival (1961) gave an example where the combined variable volume equation $v = \beta_0+\beta_1 d^2 h$ yields an R² of 0.96 and a Furnival index of 19.2, and an alternative formulation $v/(d^2 h) = \beta_1 + \beta_0 /(d^2 h)$ yields R² = 0.72 and Furnival index 9.4. The alternative formulation is the preferred model, as indicated by the distribution of residuals and by the Furnival index (smaller is better, since the index is based on the back transformation of the standard error), but this is not suggested by the R².

The R² also gives an over-optimistic indication of the model's predictive ability. The Prediction Sum of Squares (PRESS, see e.g. Weisberg 1985, Aitkin *et al.* 1989) is available in many statistical packages, and indicates the predictive ability of the equation by cross-validation. This entails omitting each observation in turn from the data, fitting the model to the remaining observations, predicting the response for the omitted observation, and comparing the prediction with the observed value. Since PRESS is analogous to RSS, small values are desirable, and the cross-validation R^2_{CV} is analogous to the usual R^2 coefficient.

When comparing alternatives, consider the Furnival index and PRESS as well as R^2. Do not be pre-occupied with these indices, but plot the data and fitted model to appraise the fit, consider if predictions will be biologically reasonable over a wide range of possible values for the explanatory variables, and see if the coefficients are reasonable estimates of the effects of the individual terms.

Strategies for Grouping Data

Data available for developing models for mixed forests, frequently do not contain sufficient observations to allow reliable relationships to be established for all species. In this situation, two approaches may be adopted. One can guess at the appropriate relationship, and employ this subjective relationship in the model, or data can be aggregated so that a relationship can be developed for a group of species. Subjective relationships are often used in succession models where it is important to maintain individual species identities, but unnecessary to predict yields accurately. Subjective relationships pose problems of bias, and it is preferable to use an objective method for yield prediction models. Aggregation also enables the number of equations required for modelling to be reduced to a manageable number (Occam's razor). Such an approach should aggregate species that are in some sense similar so as to minimize within-group variability and maximize the variation between groups.

Subjective Classifications

Many simple models use two or three groups based only on economic criteria. Although this may be useful in assessing the present value of a forest stand, it is a potentially dangerous approach in modelling as it may bear little relevance to the dynamics of the species, and it is preferable to group species on ecological criteria.

In the tropics, there are often a few species that are difficult to discriminate and which may be confused by inventory personnel. Combining these similar species may reduce overall errors by eliminating across-group species confusions. For example, in Queensland the common name blush silky oak (coded BSO) was used in resource inventory (but not in permanent plot work) for two similar species *Bleasdalea bleasdalei* and *Opisthiolepis heterophylla*, so these species were also assigned to the same group for growth modelling. Not only does this aid the rapid collection of data, but it also eliminates a serious source of error in the data. If the species comprising any group are in fact, frequently confused, then the variation in the grouped data will not be appreciably greater than in the individual species data, even if the growth patterns vary considerably.

It may be possible to classify trees according to their regeneration strategy and growth rate. The light-demanding pioneer species require large gaps in the canopy to become established, and tend to have rapid growth rates and small long-lived seeds. The shade tolerant species require only small gaps in the canopy to regenerate, and are often slow-growing with large short-lived seeds. Swaine and Whitmore (1988) proposed grouping into pioneer and climax guilds based on seed germination and seedling establishment strategies. They recommended that further classification should be based on height at maturity (pygmy, small, medium and large). However, these characteristics may not be known for all species, and may not provide a good indication of growth responses critical in modelling (e.g. diameter increment).

Shifley (1987) aggregated species subjectively on taxonomy and silvics. Where little is known about the characteristics of a species, the botanical affinity of a species may give a useful indication of the possible behaviour of a species, and the logical grouping to adopt. However, taxonomy may not provide a reliable basis for grouping; for example the genus *Eucalyptus* includes both the world's tallest hardwood trees (*E. regnans*) and shrubs (e.g. mallees such as *E. vernicosa*) which may barely attain 2 metres height at maturity. Many other examples exist.

The many species occurring in the tropics pose some problems for subjective grouping: determining how many groups are necessary, and how little-known species should be assigned. Meldahl *et al.* (1985), Leech *et al.* (1991) and Vanclay (1991*b*) examined objective procedures to resolve these questions.

Cluster Analysis

Meldahl *et al.* (1985) argued that the grouping should reflect the dynamics of growth, and that this could be best expressed through the coefficients of a regression equation on diameter increment. They attempted cluster analysis on these coefficients, but found that reasonable results could only be obtained when the regression analysis was constrained to a single explanatory variable. They investigated several possible explanatory variables, and obtained best results when diameter increment of each individual tree (Δd) was predicted from the basal area in larger trees ($G_{>d}$):

$$\Delta d = \beta_0 + \beta_1 G_{>d}$$

Cluster analysis, weighted by the inverse of the significance level of β_1 (e.g. Fig. 6.10, B), formed twenty clusters from 110 species-type equations. The number of data assigned to each cluster varied greatly, and the outcome was subjectively adjusted to provide the final grouping. The adequacy of final groups was tested by fitting a multiparameter linear function and examining the total (across clusters) residual sum of squares, on the assumption that a

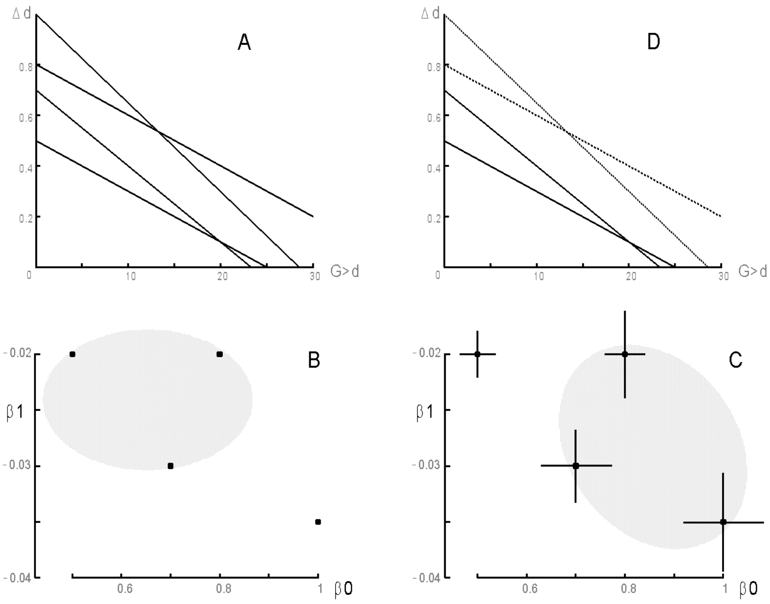


Fig. 6.10. Different species groupings based on growth trends (A), using unweighted cluster analysis (B), T^2 distance (C; lines are s.e. of βs), and pairwise comparisons (D). The "best" is context-dependent.

better grouping would result in a better fit. Whilst the method provided a satisfactory grouping of similar elements, it did not provide a unique solution.

Principal Co-ordinate Analysis

Leech *et al.* (1991) studied ways to group 27 species for the construction of volume equations. They used the so-called Behrens-Fisher analogue of Hotelling's T^2 as a measure of difference between species. This effectively measures the distance between parameter estimates in terms of their standard errors (e.g. Fig. 6.10, C). They used a polynomial equation to predict tree volume (v) from tree diameter (d) for tree i:

$$v_i = \beta_{0i} + \beta_{1i} d + \beta_{2i} d^2 + \ldots + \beta_{ni} d^n$$

Then, representing the vector of coefficients as

$$u_i^{/} = \left[\beta_{0i}, \beta_{1i}, \beta_{2i}, \ldots \beta_{ni} \right]$$

Hotelling's T^2 between two species *i* and *j* can be defined as

$$d_{ij}^2 = (u_i - u_j)' \; S^{-1} \; (u_i - u_j)$$

where S^{-1} is the combined covariance matrix of regression coefficients for species *i* and *j*. By calculating all possible combinations a symmetric matrix with zero diagonal elements can be formed. Principal coordinate analysis was used to group species on the basis of this matrix. Useful results were obtained only when the order of the polynomial, and the sign of the highest term, were the same for each of the individual species equations being compared.

Pairwise Comparisons

Vanclay (1991*b*) used pairwise comparisons between species specific diameter increment regressions. His method involved the following steps:

1. Ranking species in order of increasing number of observations.
2. Assigning the species of highest rank the founding species of group 1.
3. For each species in decreasing order of rank, conducting pairwise F-tests with the founding species of higher rank. If the incoming species was significantly different ($P<0.01$) from all existing founding species, it became the founding species of a new group. Species not significantly different from founding species remained ungrouped.
4. After identifying all founding species, those species remaining ungrouped were compared, in order of rank, with all existing groups, and grouped with the most similar group. Similarity was determined as the grouping that lead to the smallest increase in residual sum squares when the incoming species was amalgamated with the group. These comparisons were made with the whole group, not just the founding species.

This approach overcomes some of the difficulties associated with the alternatives above. Instead of comparing all possible pairs, initial comparisons are made between species with many data, reliable parameter estimates and homogeneous variance. Species with few data are only later compared with one of these major groups. It also avoids the need to arbitrarily select a subset of the more numerous species to define the groups. This selection is by no means intuitive as in Vanclay's (1991*b*) study, the species ranked 186 with only 13 observations initiated a new group. There is, unfortunately, no guarantee that the outcome is optimal, and the grouping is specific to the particular data set and increment function used. Despite these weaknesses, it provided a useful classification of 237 species into 41 groups for the development of diameter increment functions for the NORM model.

The appropriate approach to use depends on the quantity and quality of data. The methods reviewed here all have weaknesses which have yet to be overcome. The three numerical alternatives may provide different groupings (e.g. Fig. 6.10), and there may be no decisive way to determine the "best" grouping. Where few data are available, a subjective classification based on growth habit may be best. If sufficient data are available, one (or more) of the numerical alternatives may be attempted.

An aggregation based on diameter increment may not be suited to modelling mortality, and while it may be necessary to aggregate species to formulate equations, it may be best to retain individual species identities during growth and yield simulations (see Fig. 4.5, p. 75; Vanclay 1991*c*).

Synthesis

The design, construction and implementation of growth models can be complex and involves many skills. Few of us can expect to become experts in all these areas. However, a few guidelines may be adequate to ensure that reasonable models are produced:

1. Think about the design of the model first; try to keep it simple, and ensure that it meets your needs.
2. Understand the limitations of your data, and any implications they have for your model and analyses.
3. Always plot the data and the fitted model to visually examine the quality of the fit.
4. Know your own limitations, and if you don't understand something, seek help.

Success in developing models depends on carefully identifying the needs, selecting the important variables, formulating a suitable model, collecting good data (both quantity and quality), using reliable coefficient estimation procedures, and carefully evaluating the model (Chapter 11). Good modellers should rely as much on their knowledge of silviculture and on biological principles of growth, as they do on statistical tests when selecting models and developing algorithms.

Ultimately, the method of obtaining a model is irrelevant. The important thing is whether or not the model will provide useful predictions, assessed by an appropriate suite of diagnostic tests. Prominent among these criteria is the requirement that the model provides biologically reasonable predictions for the whole range of possible conditions.

In the next few chapters, we examine some specific functional relationships required to predict tree growth, mortality and recruitment in size class and single-tree models. We begin by examining ways to appraise site productivity, as the first model required in a growth modelling study may be one to quantify site characteristics.

Exercises

6.1. Demonstrate the equivalence of the growth and yield equations in Fig. 6.2. Calculate the equation which gives growth (current annual increment) as a function of time. Also calculate the equation for mean annual increment, and determine where they intersect. Can you give an equation for periodic annual increment (e.g. for a 10-year period)?

6.2. Create a data set with 10 pairs of points from the relationship $Y = X^2$ for $X = 1, 2, \ldots, 10$. Fit the simple linear model $Y = \beta_0 + \beta_1 X$, and calculate the R^2. Discuss the size of the R^2 and the quality of the fit.

6.3. Synthesize some data using the relationship $Y = X_1X_3 + (1-X_3)X_2$ with $0 < X_3 < 1$, and then fit the linear model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$. What value of R^2 do you think you will get, when the linear model is obviously irrelevant and misleading? Generate 10 or more random numbers for each of X_1, X_2 and X_3. Repeat the exercise several times to see that the result is no co-incidence. This exercise is taken from Warren (1981).

6.4. The data used to draw Fig. 6.5 are given below. These data were synthesised for this exercise, but represent tree growth in a tropical rainforest. Fit a model to these data to predict diameter increment (Δd) from tree diameter (d), basal area in larger trees ($G_{>d}$) and stand basal area (G). You may need to use transformations of some of these variables. Are all these explanatory variables necessary for a good model? Defend your model with a discussion of its strengths and weaknesses.

d	$G_{>d}$	G	Δd
13.9	9.11	9.82	0.52
28.2	6.51	10.37	0.80
46.3	4.14	10.04	0.97
61.5	2.66	8.39	0.89
75.9	4.71	15.13	0.43
90.6	2.61	14.23	0.45
105.4	0.54	10.04	0.24
10.8	22.94	25.26	0.06
25.8	14.31	21.29	0.19
36.4	8.53	15.30	0.65
56.2	6.87	20.19	0.59
66.4	10.34	25.93	0.24
78.1	5.91	22.14	0.55
92.2	2.95	14.94	0.36
11.8	18.54	30.53	0.19
24.2	29.98	35.48	0.08
40.7	18.00	27.35	0.15
51.2	19.37	27.37	0.11
63.3	9.28	25.23	0.50
78.0	8.90	29.34	0.27
96.8	3.09	24.70	0.40

6.5. Calculate the correlation coefficients (r) for each of the potential explanatory variables listed above. Do they give a good indication of the utility of the variable for modelling? Why does tree size (d) have such a low correlation? Calculate the correlation between the explanatory variables (d, $G_{>d}$, G). How would you manage your field plots to reduce this correlation?

6.6. Expand on your answer to Exercise 4.1 (p. 77), giving more details about the explanatory variables and equations that you might use in your model. What data would you need to fit these equations, and how would you collect these data?

Chapter Seven

Forest Site Evaluation

One of the first tasks facing the growth modeller may be to find a way to quantify site differences. Some sites support luxuriant forest whilst others are capable only of supporting "poor" forests. These differences may be due to soil (fertility, drainage, etc.), climate (temperature and rainfall patterns), topography (elevation, aspect, etc.) and other factors, and may be reflected in the species composition and growth patterns. Meaningful growth and yield forecasts require some evaluation of these site differences. Estimates of *site productivity* must be accurate as any bias may propagate through growth, mortality and recruitment functions to affect all modelling results.

Any single estimate of site productivity must be approximate, because it summarizes several multi-dimensional factors of the environment as a single index. The challenge is to find the most useful index. No single index based directly on environmental parameters has shown sufficient precision to be useful for forest management and modelling. The vegetation itself reflects most of the important site factors, and the growth of pure even-aged stands may provide a good measure of site productivity. Wood production may be the best indicator of site productivity for forest management purposes. However, volume production is difficult to measure, and it is convenient to use an alternative which is easier to measure. In even-aged stands of a single species, the most common alternative is *site index*, the expected height at a nominated index age (Baur 1877). Site index is often used as a proxy for volume production: height is measured at a known age, and is converted to a yield class (a measure of wood volume production) using an equation or look-up table, in the same way that girth measurements may be used to estimate tree diameter.

There are many ways to estimate the productivity of a site, and a comparison of their merits requires some definitions and a classification of the major alternatives. It is useful to classify site assessment procedures by methodology and viewpoint (Table 7.1). A phytocentric view assumes that total stand volume production or phytomass production is the ultimate

Table 7.1. Views and methods of forest site evaluation (after Leary 1985).

View	Method	
	Direct	Indirect
Phytocentric	Wood volume	Tree height
Geocentric	Soil moisture & nutrient status Photosynthetically active radiation	Climate Land form Physiography Plant indicators

measure of a site's productivity, while the geocentric view asserts the dependence of site productivity upon the soil and climatic factors. Although direct methods are preferable, they are more difficult to quantify, and this has led to the proliferation of indirect methods (Leary 1985).

> The status of indirect phytocentric methods is so inflated that some speak of direct and indirect methods, not of site productivity estimation, but of site index estimation. This appears to be an unhealthy situation; what began as an interim solution (site index) to a difficult problem (geocentric approach) should not now be called the solution to the original problem.

It is also useful to classify methodology as predictive or descriptive, and qualitative or quantitative. However, these distinctions are not clear-cut, and methodologies range from one extreme to the other. *Descriptive* systems assess site productivity on past performance by drawing on historic data, usually two or more measurements several years apart. These are easy to estimate, but their data requirement limits their utility. *Predictive* systems assess site productivity at a single point in time, using measurements of some site or crop characters made on a single occasion. These are more difficult to establish, but are more practical for most applications. *Quantitative* systems use a continuous variable, frequently height, as a measure of site, which may be expressed as a real number, or grouped into classes.

Qualitative systems usually have only a few classes, which may be labelled numerically (Classes I, II, etc.) or otherwise (e.g. poor, good). Classes need not have cardinal numbers (i.e. yield class II may not be twice as productive as yield class I). Qualitative classes require that border-line cases be resolved. Expansion of the system to recognize more classes (either to include new extremes or to encompass more classes) is difficult. Thus careful consideration must be given to the number of classes to be identified. Provided the difference in growth rate is significant, there is no advantage in having fewer classes than can be reliably recognized. Quantitative systems are generally flexible and infinitely expandable, eliminate the need to resolve border-line cases, but may give an inflated impression of precision.

Predictive approaches require measurement of some character of the site or stand and modifying it to derive an estimate of site productivity. Commonly such methods involve determining height at a known age, and transforming it to estimate site index or expected volume production. Site characteristics may also be used to estimate site productivity, and geographic regions, geology, vegetation types and crop appearance have also been used. For forest management purposes, predictive systems which enable site productivity to be estimated after a single visit, are preferable to descriptive systems that require the maintenance and remeasurement of permanent plots. However, in mixed forests, descriptive systems may form a prerequisite necessary to enable the development and evaluation of efficient predictive systems.

Much research and many publications have been devoted to site productivity assessment in plantations (even-aged, single-species stands), and no attempt is made to review these here. This chapter gives an overview of the more complex issue of site assessment in mixed forests. There has been comparatively little work done in this area, but some emerging techniques are reviewed below. As we review these alternatives, remember that we are looking for a quantitative measure of site that is:

1. reproducible and consistent over long periods of time;
2. indicative of the site, and not unduly influenced by stand condition or management history;
3. correlated with the site's productive potential; and
4. at least as good as any other productivity measures available.

Phytocentric Methods

Phytocentric measures such as site index are widely used as measures of plantation site productivity because they are easy to measure and are good predictors of the utilizable production from the site. Site index normally requires an even-aged stand of uniform development, but several authors (e.g. Duerr and Gevorkiantz 1938) have tried to adapt it for mixed forests by identifying a main even-aged stand in the forest. Stage (1963) proposed a method based on height, age and the rate of early diameter growth to compensate for early suppression. Careful selection of subject trees may overcome some problems, but many difficulties remain, and the techniques are of little use where age cannot be determined.

Visual Assessment

The appearance of the stand may give an indication of site productivity. Lewis *et al.* (1976) reported that *P. radiata* plantations develop recognizable stand differences in general vigour and form, in crown density, in needle length and colour, in tightness and colour of bark, in green level and in degree of canopy formation at age of assessment, particularly prior to thinning. The South Australian *site quality* classes (7 classes) can be recognized by experienced assessors from these qualitative characteristics.

Vanclay (1989*a*) reported the use of visual assessment to classify Queensland rainforests into two site quality classes (good, poor). Subjective assessments were generally reliable, and could be confirmed using a scoring procedure based on soil, species present, bole height and standing volume. Regression analyses of stand basal area increment on stand basal area suggested that additional classes offered no advantage. This does not reflect on the range of site productivity in these forests, but rather on the ability of assessors to classify sites reliably.

Natural Basal Area

Pienaar and Turnbull (1973) observed that even-aged stands with initial stocking above a certain lower limit, converge towards an identical stand basal area, determined by the capacity of the site. If the premise that undisturbed sites tend toward equilibrium is accepted, then the equilibrium or *natural basal area* may be assumed to be an expression of the site's productivity. This assumption is implicit in many growth models (e.g. Botkin *et al.* 1972), and the concept has been used as an indicator of site productivity for natural forests in Western Australia (Havel 1980*b*). However, as stand basal areas may fluctuate over time (especially on small plots) even when undisturbed, the approach may be liable to error. In logged stands, remeasurements over long periods are needed to estimate the equilibrium basal area.

The concept is not entirely consistent with other theories of limiting density (Chapter 9). For example, Reineke's (1933) stand density index assumes that

$$\ln N + 1.6 \ln d_g \Rightarrow constant$$

whereas the concept of natural basal assumes that stand basal area (kNd_g^2) tends towards a constant, and thus that

$$\ln N + 2 \ln d_g \Rightarrow constant$$

Since these equations conflict, these concepts cannot co-exist.

In mixed stands, natural basal area may depend upon species composition and stand structure. Sterba and Monserud (1993) found that maximum basal area may be higher in pure even-aged stands than in uneven-aged mixed stands of the same dominant height on comparable sites. They found that the difference depended on the skewness of the $d^{1.5}$ distribution, and diminished as the dominant height approached the maximum for the site. The natural basal area of any given site may be lower for light demanding and crown-shy species than for shade tolerant species. Thus natural basal area may depend on the successional status of the stand, and should be used with caution.

Stand Height

The height attained by trees at the cessation of height growth (assuming that such an asymptote exists; see e.g. Robichaud and Methven 1993) is, in theory, a good indicator of site productivity (Kramer 1967, attributed this observation to Öttelt in 1765). Stand height may be used as an estimator of site productivity if there are trees present in the stand that are sufficiently large to reflect the maximum potential height that the nominated species is likely to attain on that site. The concept is analogous to a site index with a very large index age.

The average total height of dominant and co-dominant trees remaining after logging has been used as an indicator of site productivity of dipterocarp forests in the Philippines (Canonizado 1978, Mendoza and Gumpal 1987).

One difficulty of using stand height or total tree height is that the tree tops may be difficult to see. In such cases, useful results may be obtained using height to crown break or merchantable height. Other problems include the presence of emergent trees (e.g. *Araucaria*), harvesting of the larger trees, and wind damage to tree tops.

Where suitably large trees are not available, height–diameter curves can be used to estimate the asymptotic height. This can be done by fitting an equation such as

$$h^{-1} = \beta_0 + \beta_1 d^{-1}$$

to several pairs of height (h) and diameter (d) measurements from individual trees, and estimating the maximum stand height as $h_{max} = \beta_0^{-1}$. However, extrapolation like this can be misleading, and care needs to be taken in interpreting the results.

Some non-linear equations are constrained so that they extrapolate more safely than simple asymptotic equations (e.g. Meyer 1940):

$$h = h_b + h_{\max}\left(1 - e^{-\beta d}\right)$$

where h_b is breast height. Stout and Shumway (1982) argued that the parameter β was constant for any given species in even-aged hardwood stands in the eastern USA, and suggested that this equation could be used to estimate maximum stand height from a few pairs of height and diameter measurements on dominant or codominant trees, using the relationship

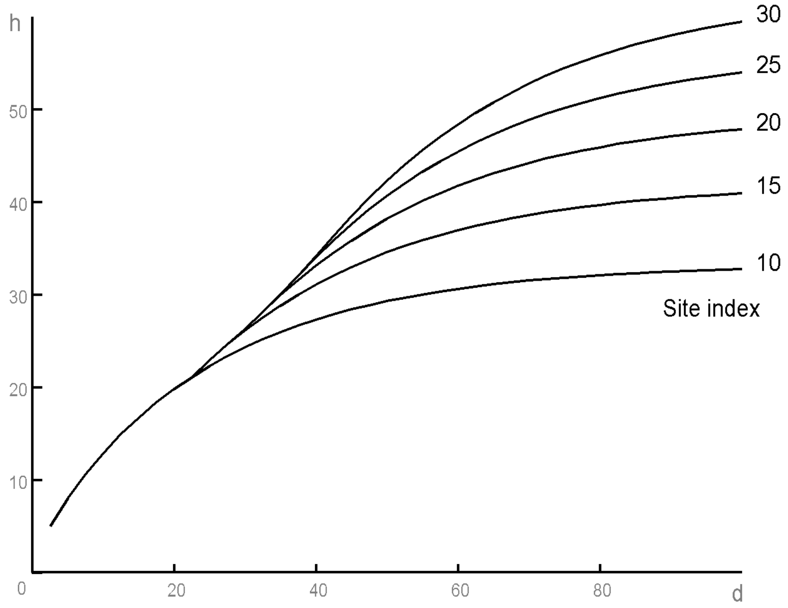


Fig. 7.1. Site index of western larch estimated from the height–diameter relationship (drawn from Equation 7.1).

$$h_{max} = \frac{h - h_b}{1 - e^{-\beta d}}$$

They also used the height–diameter relationship to estimate site indices compatible with published height–age equations, but their approach may be relevant only in even-aged stands.

Height–Diameter Relationship

To avoid the need to extrapolate the height–diameter relationship, the height at a nominated index diameter can be used as a measure of site productivity; it has been suggested that this measure could be called site form (Vanclay and Henry 1988) to avoid confusion with site index derived from the height–age relationship. The height–diameter relationship allows not only efficient evaluation of site in the field, but also the assessment of site from stereo aerial photographs by estimation from crown widths and tree heights measured on the photographs.

Reinhardt (1982) investigated the height–diameter–site relationship in western larch in the USA, and found a polymorphic trend (Fig. 7.1):

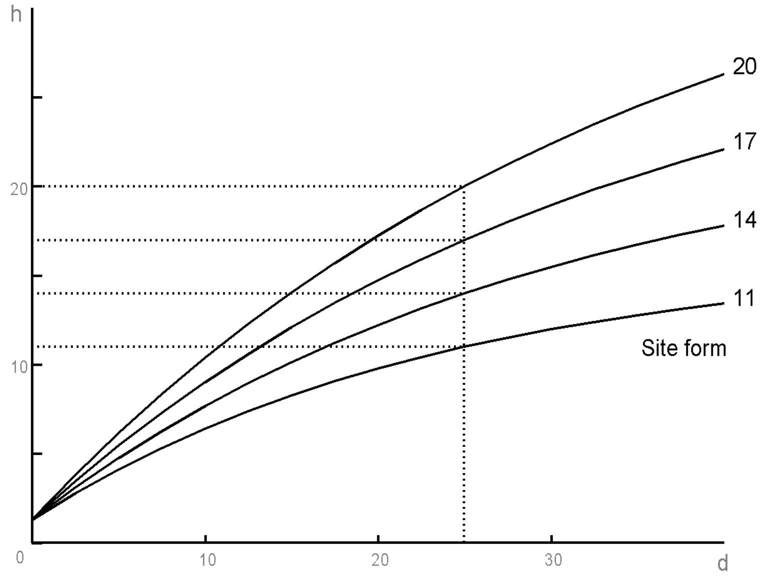


Fig. 7.2. Height–diameter curves used to estimate site productivity (site form) in *Callitris* stands in south-eastern Queensland (drawn from Equation 7.2).

$$h = 1.3 + 8.23\, S_{h,t}^{0.59} \left(1 - e^{-0.04d}\right)^{0.092\, S_{h,t}} \tag{7.1}$$

where h is tree height (m), d is diameter (cm) and $S_{h,t}$ is site index (m) at 50 years. Reinhardt (1982, 1983) worked with data from pure and mixed stands of western larch, and used the height–diameter curve to predict site index compatible with published height–age equations. The curves are not well differentiated for trees less than 40 cm dbh, and data from trees exceeding this diameter are necessary to establish a reliable relationship (Fig. 7.1).

Vanclay and Henry (1988) used the height–diameter relationship to characterize site productivity in uneven-aged coniferous (*Callitris*) forests in Queensland, using the monomolecular or Mitscherlich equation (Fig. 7.2):

$$h = h_{max} - \left(h_{max} - 1.3\right) \left(\frac{h_{max} - S_{h,d}}{h_{max} - 1.3}\right)^{0.04d} \tag{7.2}$$

Table 7.2. Correlation coefficients between site form and subjective site quality classes in *Callitris* forest in south-eastern Queensland.

Indicators of Site Productivity	Site Form (m)	Site Quality (4 classes)
Subjective site quality estimate (4 classes)	0.29	1.00
Maximum stand height (m)	0.70	0.40
Natural basal area ($m^2 ha^{-1}$)	0.46	0.21
Periodic annual volume increment ($m^3 ha^{-1} y^{-1}$)	0.40	0.34

where h is tree height (m), h_{max} is the asymptotic stand height (m), $h_{max} = -10.87 + 2.46\ S_{h,d}$, and $S_{h,d}$ is site form (m), the expected height of a 25 cm dbh tree. An adequate estimate of site form could also be obtained from a simple linear regression of height on diameter ($h = \beta_0 + \beta_1 d$) for trees 20 to 30 cm dbh, and Equation 7.2 was required only when trees of this size did not occur in the stand. Site form estimates for *Callitris* forests were relatively insensitive to harvesting, and remained constant over long periods of time. Harvests may cause a perturbation in the estimate of site form for a few years, but the estimate stabilizes in a few years when the stand remains undisturbed. Site form is positively correlated with stand basal area increment and diameter increments of individual trees (Vanclay 1988*a*), and with several other indicators of site productivity (Table 7.2). Routine field application of this method suggests that best results are obtained in well-stocked monospecific stands.

The height–diameter relationship also showed promise as a measure of site productivity in mixed eucalypt forest and exhibited a strong correlation with volume production (e.g. 44 plots had a correlation of 0.77 between site form and periodic annual volume increment). The use of several *Eucalyptus* and other Myrtaceous species to estimate site form on any plot did not appear to influence estimates.

Choice of index diameter may influence the precision of site form estimates (e.g. Goelz and Burk 1992). The height–diameter relationship can be determined with least error if the index diameter is within the range of diameters normally observed in the stand and if diameters are sufficiently large to allow height differences to be manifested (cf. Fig. 7.1). Equations may be required to predict heights in a stand of known site form, or to estimate the site form of a stand in which heights and diameters are known, and the correct form of equation should be used in regression analyses (e.g.

Curtis *et al.* 1974). The response variable about which the errors are minimized should be the variable of interest, in this case, height at the reference diameter. Some equations can be inverted and enable appropriate prediction functions for site form and for height to be formed. Parameters estimated for these alternative formulations usually differ (e.g. Monserud 1984). Omule and Macdonald (1991) offered an efficient methodology for fitting a series of compatible height–diameter curves, but remember that the shape of these curves may vary with site (e.g. Beck and Trousdell 1973).

Volume Production

Since volume production is usually the growth parameter of greatest interest to the forest manager, an evaluation of site productivity in terms of volume is desirable, but the method of measuring volume must be standardized. Utilizable volume is inadequate because utilization standards vary in time and place. Estimates of sawn volume are even more unreliable, because they depend on assumptions regarding conversion efficiency. Assmann (1961) recommended the use of solid wood (*derbholz*) volume defined as the volume under bark of all stem and branch material not less than 7 cm diameter under bark. This is convenient for conifers as it reflects utilization standards in common usage in many places. However, in trees with a deliquescent habit, this may entail the measurement of branch volume, which may be difficult.

As early as 1888, the Association of German Forest Experiment Stations adopted total volume production at 100 years of age as a standard measure of site productivity. Others (e.g. Bradley *et al.* 1966) advocate that mean annual volume increment at culmination provides a better basis for comparison, but these measures apply only to pure even-aged stands. Difficulties arise in mixed forests where there may be many different species with great variations in growth habit, growth rate and wood density. Biomass production (t $ha^{-1}y^{-1}$ dry weight) may provide a suitable basis for comparison, but should it refer to just one species, or to a specified mixture of species? In mixed forests, an index of the site is probably more useful than an index of a species on that site. Notwithstanding these limitations, the historic volume production of a well stocked and well managed forest may provide a good measure of site productivity, and may serve as a benchmark with which to test other more practical measures of site.

Direct measurement of volume production requires measurement of permanent plots over many years. Schmoldt *et al.* (1985) attempted an alternative approach by fitting yield equations to permanent plot data, and examining both the maximum growth rate and the asymptotic volume. They found that aspect and soil nutrients were significantly correlated with asymptotic basal area, maximum basal area increment and asymptotic volume, but not with maximum volume growth rate or site index. They suggested that site index is unreliable in mixed hardwood forests in North America, and suggested that coefficients from yield equations fitted to permanent plot data may provide practical alternative measures of site productivity.

Growth Index

If a simple growth model is compared with remeasures of several permanent plots, the residuals will indicate the relative site productivity of the plots. Large positive residuals indicate a better-than-average site, small residuals indicate an average site, and negative residuals indicate a poorer-than-average site. The simplest method is simply to compare mean basal area increments, but better estimates may be obtained if basal area increment is adjusted for stand basal area. More sophisticated analyses may use individual tree increments, and may take stand composition into account (see Exercise 7.2). This requires measurements over a period of several years before site productivity can be estimated for the site, but if a correlation can be discovered between the residuals and some easily measured site or crop parameter, this may provide the first step toward a predictive system.

Vanclay (1989*c*) developed an index of site productivity based on individual tree growth adjusted for stand density. The index was initially estimated from the historic measurement record of permanent plots. Later, biotic and abiotic variables correlated with the index were used to predict the growth index for other sites.

The index was based on the diameter increment of individual trees of 18 reference species. These reference species were subjectively selected because their widespread distribution and common occurrence, to ensure that all plots contained at least one tree of half these species. The 18 species were *Acronychia acidula, Alphitonia whitei, Argyrodendron trifoliolatum, Cardwellia sublimis, Castanospora alphandii, Cryptocarya angulata, C. mackinnoniana, Darlingia darlingiana, Elaeocarpus largiflorens, Endiandra* sp. aff. *E. hypotephra, Flindersia bourjotiana, F. brayleyana, F. pimenteliana, Litsea leefeana, Sterculia laurifolia, Syzygium kuranda, Toechima erythrocarpum* and *Xanthophyllum octandrum*.

An increment function (Fig. 7.3) was fitted simultaneously for all 18 reference species, to all non-overlapping remeasures on 80 permanent plots (for some plots measured more frequently, selected remeasurements with approx. 5 year intervals were used). The plot identifier was included as a qualitative variable (see p. 104):

$$\begin{aligned}\ln(\Delta d_{ijk}+\alpha) = {} & \sum_{i=1}^{18}\beta_{0i}Z_i + \sum_{i=1}^{18}\beta_{1i}Z_i d_{ijk} + \sum_{i=1}^{18}\beta_{2i}Z_i \ln d_{ijk} \\ & + \sum_{i=1}^{18}\beta_{3i}Z_i \ln G_k + \sum_{i=1}^{18}\beta_{4i}Z_i G_{>dijk} + \sum_{k=1}^{80}\gamma_k Z_k \ln d_{ijk}\end{aligned} \tag{7.3}$$

where Δd_{ijk} , d_{ijk} and $G_{>dijk}$ are the diameter increment (cm y^{-1}), initial diameter (cm dbh) and basal area in larger trees (m^2 ha^{-1}) respectively for

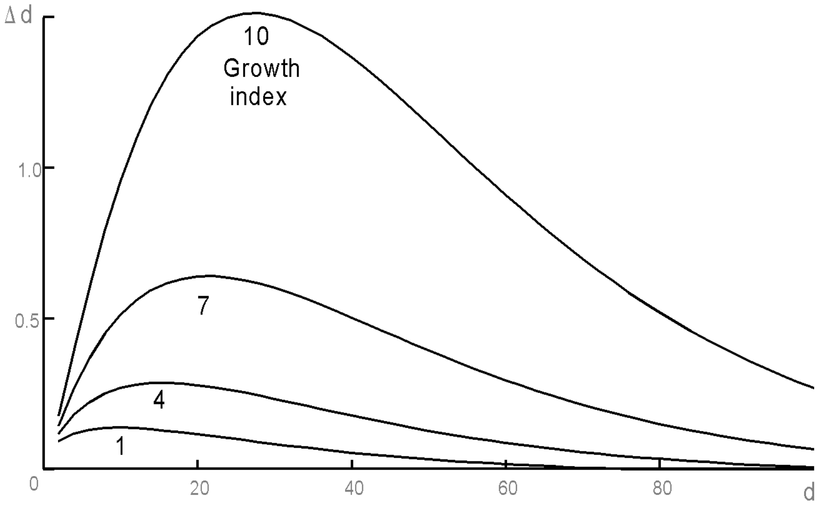


Fig. 7.3. Growth index equation for *Cardwellia sublimis* trees in stands with 30 $m^2 ha^{-1}$ basal area with 10 $m^2 ha^{-1}$ basal area in larger trees.

tree *j* of species *i* on plot *k*, G_k is the stand basal area ($m^2 ha^{-1}$) on plot *k*, Z_i is a binary variable (see p. 104) which takes the value 1 for trees of species i and zero otherwise, Z_k is a binary variable which takes the value 1 for trees on plot *k* and zero otherwise, and α, β_i and γ_k are parameters to be estimated. The parameter α was assigned the value 0.02 after inspection of residuals and examining the residual mean squares from a range of values.

This equation (7.3) can be re-arranged to estimate the growth index for a given permanent plot:

$$S_{\Delta d} = \frac{11.35}{\sum_{ij} \ln d_{ij}} \times \left\{ \sum_{ij} \ln(\Delta d_{ij} + 0.02) - \sum_{ij} \left(\hat{\beta}_{0i} + \hat{\beta}_{1i} d_{ij} + \hat{\beta}_{2i} \ln d_{ij} + \hat{\beta}_{3i} \ln G + \hat{\beta}_{4i} G_{>dij} \right) \right\} \qquad (7.4)$$

where $S_{\Delta d}$ is the growth index of the plot, d_{ij} is the diameter (cm) of tree *j* of species *i*, Δd_{ij} is its diameter increment (cm y^{-1}), $G_{>dij}$ is the basal area of trees within the plot that are bigger than tree *ij* ($m^2 ha^{-1}$), *G* is the plot basal area ($m^2 ha^{-1}$), and the $\hat{\beta}$s are estimates of the βs in Equation 7.3. The value 11.35 (in Equation 7.4) was subjectively determined to scale the growth indices into the range 0–10. This equation provides an estimate of site productivity based on the diameter increment adjusted for tree size and competition, of all trees of 18 reference species using all available remeasures for the plot. This may look very complicated, but is analogous to taking the average of growth indices read from Fig. 7.3 (or similar curves for the other species) for each remeasure of each tree of any of the 18 reference species that occur on a plot.

This procedure yielded a descriptive measure of site, in effect an objective ranking of permanent plots, and does not provide a predictive measure that can be applied to temporary plots. However, the growth index can be predicted from other measurable or observable biotic and abiotic factors.

Geocentric Methods

Phytocentric methods of site assessment are based on forest measurements, and thus cannot be used where suitable forest stands are not present. Alternative methods are important for evaluating plantation viability, and the prediction site index from site attributes has been researched thoroughly. Natural forest management may not require site assessment of bare land, but geocentric methods of site evaluation may be useful even where forest is present. Descriptive measures such as growth index require remeasured plots, and it is desirable to have a predictive method that provides an estimate of site productivity after a single visit. Many site characteristics can be recorded during a single visit, and may provide an efficient way to estimate site productivity if a relationship can be established with growth index or other proven measures of site.

One problem in appraising geocentric methods of site evaluation is that it is difficult to test them against the true site productivity, and that they are usually tested against some phytocentric estimate of site productivity (e.g. site index), which may or may not be reliable. This problem is evident in Grey's (1979) attempt to relate geocentric site factors to site index: he found that three measures of site productivity (two from alternative height–age curves and an estimate of the mean annual volume increment), each were predicted by a different subset of the site factors considered, and that no one factor was common to all three. This suggests that these three measures of site productivity were not sufficiently reliable to enable a consistent relationship to be established. Schmoldt *et al.* (1985) also found that site index was not well correlated with site factors, and suggested that alternative phytocentric estimates should also be considered in appraisals.

Climate

The best known climatic index of forest growth is Paterson's CVP index which was designed to predict the maximum growth potential in terms of volume production over large areas (Johnston *et al.* 1967). It is based on evapotranspiration, annual temperature range, mean annual precipitation, length of growing season and mean monthly temperature of the warmest month. Although it has been adopted on a national scale by several countries, it is probably only useful for economic geography and general forest statistics where estimates of potential production are required for large inaccessible and non-inventoried areas. Similar indices of net primary production for global atmospheric studies have been based on evapotranspiration (e.g. Lieth and Box 1972) and on temperature and precipitation (e.g. Esser 1984).

Czarnowski (1964) developed an equation with three climate parameters, three soil parameters and four species properties to predict the productive capacity of a species independently of age, anywhere on earth. Trials with his equation indicate that predictions were close to observed values for three species on four continents. The model was subsequently refined to predict site index (at age 20) of *P. radiata* from three climatic and nine soil characteristics, with a mean error of about 10 percent (Czarnowski *et al.* 1976). However, the equation lacks general utility and involves considerable effort in determining soil nutrient status.

Degree-days (number of days during which the mean temperature exceeds a specified temperature, usually 5 or 10 degrees C) and precipitation during growing season have been used to estimate site productivity in temperate forests (e.g. Farr and Harris 1979). The distribution of the rainfall is probably more important than the actual amount (e.g. Jackson *et al.* 1975). In temperate regions, rainfall during the growing season is important, whilst in the seasonal tropics the dry season rainfall may be critical. These variables may be used directly, or predicted as a function of latitude and elevation.

Topography

Climatic variables can only give a general indication of site productivity because they do not account for any local variations in site. Thus estimates may be improved by incorporating topographic details for the each site. An advantage of using only climatic and topographic information is that these details can easily be obtained from topographic maps (or air photos) and climatic records.

In areas of marked relief, topographic effects may be the dominant force controlling site productivity. Evans (1974) found that height at age twelve in *P. patula* plantations in Swaziland was highly correlated with elevation. Site index of oak in Ohio (USA) can be predicted from aspect, slope shape and position on slope (Carmean 1967). Stage (1976) demonstrated the interacting effects of slope and aspect on the site index of white pine, and showed that the favoured aspect can be detected by including both the trigonometrical functions *sin*(aspect) and *cos*(aspect) in equations.

Most investigations into the relationship between topography and site productivity have used simple variables such as elevation, aspect and slope.

Whilst these have produced some promising results in temperate forests, their predictive ability has not been demonstrated in the tropics. In an analysis of vegetation types, Moore *et al.* (1991) found that "steepness" (the diversity of elevation within a 200 m neighbourhood, indicative of erosional/depositional environments) and "exposure" (the average angle to the northern horizon, indicating shading by the topography; their study was in the southern hemisphere) provided better predictions than slope and aspect. Vertical height above the nearest stream may also provide good discriminations, often better than the conventional position on slope. This suggests that there remains scope for further research in this area.

Soils

Geocentric estimates of site productivity may be further refined by including details about the soil. Soil depth, colour and texture are easily determined, but soil moisture and nutrient status are more difficult to quantify and may vary in time and space. Despite these difficulties, numerous studies involving soil analysis have been made in temperate regions (see e.g. Carmean 1973). Some studies report high correlations between site productivity and soil nutrient status (e.g. Mader 1976, with white pine) while others found no useful correlation (e.g. Fralish and Loucks 1975, with aspen). Grey (1979) found little correlation between site index of *Pinus patula* in South Africa and soil structure and chemistry, but found a significant correlation with slope, distance from ridge, etc. These findings may be species-specific, and effective rooting depth may be the most influential predictor of site productivity (e.g. Schönau and Aldworth 1991, with *Acacia* plantations).

Wright and van Dyne (1971) studied 50 equations predicting site index from site factors for several species in the USA, and found that on poorly drained sites, texture and depth to impermeable layers were crucial, while on other sites, topography, available water and soil depth were important. Carmean (1979) suggested that important soil features include surface soil depth, depth to mottling, depth to impermeable layer, effective soil depth, structure, drainage and subsoil colour. This is good news, because these parameters can be determined in the field, and do not need laboratory analyses.

Webb and Tracey (1967) found that site index of *Araucaria cunninghamii* in Queensland could be predicted from surface geology and land form, within broad climatic zones. Surface geology reflects soil mineral status, as acid rock produces soils of low fertility while more basic parent material yields soils of high nutrient status. Slope, soil depth and drainage were also important factors.

Turner *et al.* (1990) formulated a soil classification for intensive *P. radiata* plantations, designed to reveal potential nutritional deficiencies and other management limitations from physical soil parameters that did not require laboratory analysis. The system was based on characteristics that indicated innate nutrient supply, the development of roots, moisture supply, and other growth and management factors, and included parent rock, texture profile, depth to and nature of impeding layer, texture and condition of the uppermost 10 cm of soil, character of horizons and the condition and colour of the subsoil. Parent rock sets upper limits to the total amounts of clay, primary quartz and many plant nutrients that can be released through mineral weathering (Brewer 1954). Turvey *et al.* (1990) found that parent rock, depth to, and nature of impeding layer explained most of the variation in wood volume production. Parameter estimates from their equation could be summed to provide estimates of wood volume production at age 11 years. The Turner *et al.* (1990) soil classification system provided better estimates of productivity than did other more widely used classifications.

Biotic Factors

Many site properties such as available water and nutrient concentrations are not easily measured, so an alternative is to measure indicative variables such as composition of ground vegetation. No causal relationship is implied, but it is assumed that both ground vegetation and wood production are influenced by the same properties (e.g. Cajander 1909, 1949).

Daubenmire (1976) drew on six basic principles to argue that vegetation is the best method for assessing site productivity:

1. Vegetation reflects the sum of all the elements of the environment which are important to plants.
2. The species with the highest competitive powers are the best indicators.
3. Forests consist of superimposed groups ("unions") which occur in different combinations over the landscape.
4. Each union is sensitive to certain special aspects of environment.
5. Many characters of vegetation have potential significance as ecologic indicators.
6. Types of environment ("habitat types") are the most basic ecologic units of landscapes.

Two main approaches can be distinguished: the classification approach (also known as subdivision, European or Braun-Blanquet) which uses the (potential) climax vegetation, and the ordination approach (also known as the Anglo-American or Clements) using indicator plants.

Classification

There are several variations on classification, but all use the potential climax vegetation to classify areas into habitat or site types, which are considered to be effectively uniform in many respects (Havel 1980*a*). The classic example of floristic classification for site evaluation is Cajander's (1909, 1949) use of various associations of ground vegetation to predict site productivity in Finland. This approach has been used with only minor modifications in Europe and north America.

Ure (1950) used this system in New Zealand to estimate *P. radiata* site class from habitat types. Reliable estimates require that the vegetation must not have been burnt, ploughed or otherwise interfered with, and that road edges are avoided. Daubenmire (1961) also found that floristic classification was a useful way to predict height growth and disease susceptibility of *P. ponderosa* in the USA.

Webb *et al.* (1970) classified Queensland rainforests on physiognomic and structural characters, and found this an effective way to indicate environmental conditions. Structural features offer greater ease and speed of data collection than floristic classifications. The approach also appears to have greater applicability over diverse geographic regions. One shortcoming is that structural features are generally defined in an informal manner, and delineation of habitat types is not necessarily unique especially when prepared by workers not familiar with the approach. This work has focused on general environmental conditions rather than site productivity, and the value of this classification for growth prediction has not been demonstrated.

One danger in shifting emphasis from assessing forest site productivity to a system which fulfils broader multidisciplinary objectives is that in doing so the value of the classification for site assessment may be reduced. However, classification into habitat types may be useful, and Monserud (1984) found that each type exhibited site index curves of different shapes. Classification may provide accurate estimates of site productivity when developed specifically for the purpose, but more general approaches that can be used for a variety of purposes may not predict the site productivity so well.

Unfortunately, floristic classification remains a rather imprecise science. Non-mathematical classifications are unavoidably subjective, whilst the outcome of mathematical approaches depends heavily on the algorithm chosen. Of the many algorithms available, only single linkage cluster analysis emphasizes the separation of clusters; other algorithms may maintain clusters without regard to the possibility that two similar units may be assigned to different major clusters (e.g. Gower 1967, Jardine and Sibson 1971).

Ordination

Two approaches to ordination exist. The first and most widely used is to use the presence (and occasionally abundance) of certain plants as an indication of site productivity; the other uses physiognomic characters such as size and shape of leaves, and the height of the indicator plants. These methods are not mutually exclusive, and may be used in conjunction.

The presence and abundance of each plant indicator express a set of environmental conditions favourable to that species. A community of such plants may reveal many biologically relevant factors and interactions of a site. Thus the use of indicator plants may reflect the integration of factors more flexibly than the climax vegetation approach (e.g. Webb *et al.* 1967). The relative abundance of plant species may be used an indicator (e.g. Corns and Pluth 1984), but ordination based on presence-absence seems to be influenced less by disturbance.

Carleton *et al.* (1985) examined the influence of temporal factors such as stand density and succession on understorey vegetation in northern Canada, and found that these have minimal influence on vegetation. They found that the understorey vegetation was most influenced by soil, and concluded that the understorey vegetation should provide a reliable indicator of site productivity.

Webb and Tracey (1967) gave a list of pioneer species indicative of good and poor sites for *Araucaria cunninghamii* plantations in Queensland; these pioneers should be present following disturbance. However, indicator species need to be chosen carefully, as even apparently stable rainforest may have a relatively high species turnover rate. Swaine *et al.* (1987) reported a species turnover of around one percent per year in an undisturbed tropical forest in Ghana. Schönau (1987) argued that plant indicators are more useful in temperate regions where there are fewer species, and suggested that vegetation on its own generally does not provide satisfactory site productivity estimates.

In Queensland rainforests, the growth index may be estimated from the presence of several tree species (Vanclay 1989*a*). If the correct taxonomy of indicator trees was known, geology contributed no further improvement. However, identification of rainforest trees is often difficult, and a single common name may refer to more than one species. Good estimates of growth index can be obtained if geology is used in conjunction with common names:

$$S_{\Delta d} = \begin{pmatrix} 4.528\, S_{al} \\ 5.934\, S_{bv} \\ 5.164\, S_{av} \\ 6.174\, S_{cg} \\ 3.837\, S_{fg} \\ 4.980\, S_{sm} \end{pmatrix} \begin{array}{l} +1.144\, Z_{BLO} + 1.286\, Z_{SBN} - 1.020\, Z_{VTX} - 0.673\, Z_{RAP} \\ +1.027\, Z_{BUA} + 1.008\, Z_{RBN} - 1.223\, Z_{CLL} + 1.516\, Z_{BGR} \end{array}$$

where all variables are binary variables which take the value one if the geology (S) or species (Z) is present on the plot, and zero otherwise, and *BLO* is blush silky oak (*Bleasdalea bleasdalei* and *Opisthiolepis heterophylla*), *SBN* is salmon bean (*Archidendron vaillantii*), *VTX* is vitex (*Vitex acuminata*), *RAP* is rapanea (*Rapanea achradifolia*), *BUA* is buff alder (*Apodytes brachystylis*), *RBN* is rose butternut (*Blepharocarya involucrigera*), *CLL* is cinnamon laurel (*Cryptocarya cinnamomifolia* and some affiliated species), and *BGR* is brown gardenia (*Randia fitzalanii*), and where the geology *al* is alluvial, *bv* is basic volcanic, *av* is acid volcanic, *cg* is coarse granite, *sm* is sedimentary-metamorphic, and *fg* is fine-grained granite. Note that only one geological type (S) may occur on a site, but that any number of species may be present and used to evaluate the growth index. None of these species are short-lived pioneer species, and the presence or absence of these species should be relatively independent of successional status and disturbance.

Keenan and Candy (1983) examined the utility of floristic data for predicting height growth of *Eucalyptus delegatensis* in Tasmania. Ten species or species groups were formed by amalgamating some less frequent taxa, and omitting species which occurred on fewer than three of the 52 plots. A binary matrix indicating relative abundance was compiled by assuming a taxon was abundant (1) if occurred over at least 30% of the plot area (0.01 ha) covered by non-eucalypt vegetation, and non-abundant (0) otherwise. Matrices comprising presence/absence (any occurrence of 28 species) and percentage cover data provided similar but inferior results. They found that the first principal component of this matrix explained 29% of the total variation in growth, more than the non-floristic site factors (slope, elevation, exposure, soil parent material, soil pH, soil drainage) which, although significant ($P<0.01$) collectively explained less than 9% of the total variation. Their analysis suggested that plant species were better predictors of height growth than human appraisals of non-floristic factors.

Principal components offer some advantages for investigations such as this. They provide more than one linear combination of the (0,1) data and can thus reflect more than one environmental gradient. Secondly, component correlations of the principal components are unchanged by adding or subtracting other explanatory variables to the model. However, one disadvantage is that principal components are specific to each set of data: additional data may give rise to different principal components, and thus subsequent studies and predictions must use the original component correlations.

Practical Considerations

The accuracy of any attempt to model a forest system depends largely upon the precision with which the site can be classified into homogeneous units (e.g. see simulation studies by Gertner and Dzialowy 1984, Smith and Burkhart 1984). Thus the ability to stratify a forest resource with respect to site productivity may make a significant contribution to the accuracy of yield predictions. Remote sensing and geographic information systems may be helpful in mapping site productivity and stratifying areas of forest.

Mapping Site Productivity using Remote Sensing

Forest types have been mapped using remotely sensed data for many years, and broad site productivity classes can often be defined. Aerial photographs may allow objective measurement of some measures of site productivity. Bonnor and Morrier (1981) used aerial photography to classify temperate mixed forest in Canada into several 5-metre site index classes with 76% success; the remaining instances underestimated by one class. Goodwin (1988) reported the use of aerial photography to determine mature stand height in temperate *Eucalyptus* forest in Tasmania. However, the ground must be visible through the canopy for these methods to work, and this may not be possible in dense forests.

Digital remote sensing from satellite (e.g. Landsat) offers potential for objective algorithm-based site productivity mapping. Fox *et al.* (1985) reported that broad site productivity classes could be estimated from aspect (computed from digital topographic data) and vegetation classes obtained through supervised classification of Landsat data. Vanclay and Preston (1990) reported that growth index in Queensland rainforests could be estimated directly from Landsat thematic mapper (TM) data and geology, although prediction equations may need to be re-calibrated for each Landsat scene. Vanclay and Preston (1990) found that the ratio of band 4 (near-infrared) and band 5 (mid-infrared) provided reasonable estimates of growth index, especially when used in conjunction with geological data:

$$S_{\Delta d} = \begin{pmatrix} 15.75\, S_{av} \\ 18.13\, S_{bv} \\ 18.42\, S_{cg} \\ 17.59\, S_{sm} \end{pmatrix} + 3.017 \frac{R_4}{R_5} - 0.284\, R_1$$

where *av, bv, cg* and *sm* are binary variables representing acid volcanic, basic volcanic, coarse granite and sedimentary-metamorphic geology respectively, and where R_1, R_4 and R_5 are reflectances observed in the blue, near-infrared, and mid-infrared Landsat TM bands respectively. However, this equation should be re-calibrated for each image, and could not be extrapolated to other passes (i.e. different dates). The ability to predict site productivity remotely may depend on timing, especially in the seasonal tropics.

Mapping Site Productivity using Geographic Information Systems

Geographic information systems have become an important and useful tool in forest management, and offer some potential for mapping site productivity. Several studies (e.g. Moore *et al.* 1991) have demonstrated the utility of these systems for mapping vegetation types, and these studies suggest potential for mapping site productivity.

Turvey *et al.* (1990) found that soil parent material and soil depth were the primary determinants of *Pinus radiata* site productivity. Suitable soil parent material classes can be derived from published geological maps, whilst soil depth may be inferred from a digital elevation model. Moore *et al.* (1991) found that steepness (elevation diversity) was easily calculated with a digital elevation model and indicated erosional and depositional areas which may serve as a suitable proxy for soil depth.

Multiple Estimates

Different methods of site productivity assessment may give rise to differing estimates, and the forest manager may have no basis for resolving these differences. Choosing the most popular or well established technique is one alternative, others may include choosing the method that intuitively seems right, or taking the mean or median of all available estimates. If permanent plot data are available, the alternative techniques can be tested using standard procedures (e.g. Freese 1960, Gregoire and Reynolds 1988, see Chapter 11), but where no such data are available, the "true" value cannot be determined and alternative selection procedures are required. Reed and Jones (1989) suggested an objective approach based on psychometrics (e.g. Campbell and Fiske 1959) to help reconcile different estimates of site productivity. The method involves conceptualizing the relationships among the estimates, standardizing the estimates, calculating correlations and identifying correspondences between the approaches. The most consistent approach is assumed to be the most reliable.

Changing Species of Estimation

Where a crop parameter is used as a measure of site production, it may refer to a specific species in the crop, or to a stand of specified composition. It may be possible to gauge the potential of a site for another species or composition by a transformation of the measured parameter. Another application of this technique is that it enables extensive areas of mixed forest to be evaluated in terms of one standard species, even if that species is not present over the whole area.

Foster (1959) found a curvilinear relationship between the site indices of eastern white pine and red maple. Red maple has rapid early height growth and may exceed the height of white pine during the first 46 years, after which the trend reverses. Red maple is more sensitive to site than white pine, being taller on good sites and shorter on poor sites.

Shoulders and Tiarks (1980) examined the influence of rainfall, slope and available soil moisture on the height at age 20 of four species of pines on the USA Gulf Coast. Relative heights of these species were affected by all three factors. *P. echinata* is the tallest where annual rainfall is less than 1300 mm. Elsewhere, *P. elliottii* or *P. taeda* may be taller, depending on soil, slope and rainfall distribution.

Equations comparing site indices may be useful for assessing the potential performance of species on sites where they are not present. However, care should be taken with interpretation, as most comparisons published to date compare only the height at index age. Such comparisons should be made only after comparing the methods of determining site index for both species, and comparing the shape of the height growth curves. Differing growth habits of various species mean that height or site index of different species may not provide a reliable basis for comparing site productivity. Periodic annual volume increment realized under a specified management regime may provide a better basis for comparison.

Uniformity of Site

A further complexity in assessing site productivity arises where the site is not uniform, but is perforated by physical obstructions such as rocky outcrops. An insidious feature of this phenomenon is that conventional measures of site such as site index may indicate the potential of the better pockets, and fail to indicate the true average productivity of the site. However, top or predominant height is often determined as the mean of the tallest tree on each of several adjacent non-overlapping plots, and this should reduce bias from such phenomena.

MacLean and Bolsinger (1973) proposed the use of a stand density index (Curtis 1970) predicted from indicator plants to adjust yield estimates. An alternative approach is to examine the physical attributes of the site and determine an arbitrary reduction, but this approach is not without difficulty.

Synthesis

This chapter has neatly compartmentalized the many options available for site evaluation into discrete categories. This is convenient for the present study, but irrelevant in application. It is likely that the best approaches to site evaluation may employ a combination of several of these options.

The development and evolution of an efficient method of site evaluation for mixed forests will rely on comparisons of alternatives with long term growth recorded on permanent plots. This will require considerable amounts of quality data. Rayner (1992) found that several classifications based on climatic, edaphic, landform and vegetative characteristics did not help to explain forest stand dynamics, and attributed this to inadequate selection of site attributes, measurement techniques and sampling intensity. Others report more promising results, but it appears to be more difficult to demonstrate an environment–productivity relationship, and there remains ample scope for further research in this area.

Indices such as growth index show promise, but in practice would normally be estimated from stand and environmental variables, including indicator species. Measures of stand height such as maximum stand height, canopy height and the height–diameter relationship may also prove useful where visibility allows unimpeded estimation of tree heights. Any measure of site productivity should be tested to ensure that it is reproducible and consistent over long periods of time, and not unduly influenced by stand condition or management history.

Exercises

7.1. Suggest a measure of site productivity that might prove suitable for your forests. How would you test this measure to ensure that it was a good indicator of site productivity? Could you do this with existing data, or would you need to collect new data? Design an experiment to calibrate and test this measure of site productivity.

7.2. Stand basal areas ($m^2\,ha^{-1}$) recorded during successive remeasures at 10-year intervals on six permanent plots are given below. Derive an estimate of site productivity for each of these six plots. Define three productivity classes and choose two plots for each class. *Hint*: read the section on Growth Index (p. 143) again.

Plot	$G_0, G_1, G_2, \ldots$	Species present						
		A	B	C	D	E	F	G
1	11.2, 13.3, 15.5, 17.6	1	0	0	1	0	0	1
2	16.3, 17.6, 18.8	0	1	0	0	1	1	1
3	9.7, 12.2	0	0	1	1	0	1	1
4	19.5, 21.2, 22.8	0	0	0	0	1	1	1
5	12.4, 15.0, 17.6, 20.0	0	0	1	1	0	1	1
6	9.3, 11.0, 12.7	0	1	0	0	1	0	0

7.3. The data above also indicate some of the species that were found on the six plots of Exercise 7.2. Are some of these species useful as potential indicator species to predict the site productivity? Which ones?

Chapter Eight

Diameter Increment

Most models suitable for mixed forests predict individual tree growth explicitly, often with equations to estimate diameter increment from tree diameter and other variables. These equations are usually fitted with data from remeasured permanent plots on which all trees have been individually identified. Some alternatives exist where such data are not available, but these approaches are generally less accurate. This chapter focuses on single-tree diameter increment equations that can be used in tree list and size class models, and assumes that suitable permanent plot data are available for developing these equations (see Chapters 3–5).

It is often convenient to model diameter increment, but this is not the only alternative, and the principles discussed in this chapter apply equally to other measures of stem dimensions, including cross-sectional area, height and volume (e.g. see p. 112). These alternatives should be considered, even if the model utilizes diameter increments, because it may be more efficient to estimate one of the alternatives, and convert it to a diameter increment. For instance, the FORMIX model (Bossel and Krieger 1991) for Malaysian dipterocarp forest estimates volume increments from nett photosynthesis and converts these to diameter increments (see p. 41).

What to Model

The increase in stem size of individual trees can be modelled as

(a) diameter increment,
(b) basal area increment,
(c) future diameter, or
(d) future basal area.

Although arguments have been advanced in favour of all these options, all four are mathematically related and there should be little difference between the alternatives, provided that the assumptions concerning the error term (p. 119) are satisfied.

Some models (e.g. Mitchell 1969, Ek and Monserud 1974, Alder 1979) simulate height development, and derive diameter increment from the predicted height growth. This method has been useful for modelling regeneration and in modelling pure even-aged coniferous forests, but the difficulty of measuring heights of mature trees is a major limitation in many forests, and the approach will not be considered further in this chapter.

We begin by examining the options for modelling increment, and address the alternative approach of estimating future diameter or basal area later.

Diameter or Basal Area Increment

Individual tree growth may be predicted as basal area increment or as diameter increment. Some argue that modelling basal area increment is preferable, since basal area increment supposedly resembles more closely the volume growth achieved by the tree than does diameter increment. Others justify the use of basal area increment as the response variable because it usually has a higher R^2, but this claim is invalid since R^2 does not provide a valid comparison between these alternatives (see Chapter 6, p. 125). Both these arguments are unfounded.

Tree diameter increment and basal area increment are related mathematically (i.e. $\partial g = \partial(kd^2) = 2kd\,\partial d$), and any apparent differences in the goodness-of-fit may be due to differences in the error structure and implied functional relationship, rather than the superiority of one model over the other. Many researchers (e.g. Bella 1971, Johnson 1973, West 1980) have observed that the use of basal area increment as the response variable tends to give higher values of R^2. This has more to do with the limitations of the R^2 statistic (e.g. p. 125) than the suitability of the response variable, and is unlikely to be reflected in other more suitable statistics (e.g. Furnival index, p. 126). Empirical studies (e.g. West 1980, Shifley 1987) offer no evidence of any difference in the precision of estimates of future diameter from diameter and basal area increment equations.

This means that the decision to model diameter increment or basal area increment may be based on convenience. One of these alternatives may be desirable because it offers a more satisfactory error distribution, but the same effect may be obtained through weighted regression and other techniques (p. 119). This chapter focuses on diameter increment, but the prediction of basal area increment is equivalent in most respects as the two are intimately related.

Future Diameter or Diameter Increment

The increase in stem diameter can be expressed as a yield function that estimates future diameter ($d_n = F(d_0, t) + e_1$), or as a growth function which estimates the increment over a specified period ($\Delta d = f(d) + e_2$, where e_1 and e_2 are the errors associated with the estimates). Yield equations can be differentiated to form growth equations and thus to provide compatible growth and yield estimates. Where there is a single measurement interval of the same duration for all observations, both formulations should, in theory, yield the same result. However, where measurement intervals vary and where there are repeated measurements from the same plots, these alternatives may give rise to different results, because the error structure will be different (i.e. errors e_1 in the yield formulation have linear units whereas the errors e_2 in the growth form are per unit time, e.g. mm and mm y^{-1} respectively).

The growth equation may pose some problems not encountered with the yield equation, especially if the measurement interval is long or varies greatly. If remeasures are fairly frequent, initial values of various tree and stand variables (d, G, etc.) can be used in regression analyses. However stand conditions may change significantly during longer intervals, so in such cases, some account should be taken of the change in the initial values. The simplest alternative is to use the value corresponding to the middle of the interval, estimated as mean values, but estimates may also be based on the functional form of the relationship (McDill and Amateis 1993).

Figure 8.1 illustrates that the average increment observed over a 13-year period (Δd, thick line) is a good predictor of the expected annual growth (∂d) for the average conditions (i.e. $(d_0 + d_{13})/2$ and $(G_{>d,0} + G_{>d,13})/2$) during that period, but not of the conditions at the beginning of the period. The instantaneous growth ($\mathrm{d}d/\mathrm{d}t$) cannot be observed, and annual increments may be obscured by measurement error and other noise, so in practice we have to make do with periodic average increments. Here we infer expected growth from the results of Exercise 6.4 (p. 132). Notice that the expected growth (thin solid line) decreases as the tree becomes larger and the basal area in larger trees ($G_{>d}$) increases, and that a different trend emerges if $G_{>d}$ is assumed to be constant (dashed lines). Thus a sensible compromise when fitting growth models may be, for example, to regress the observed increment on the mean diameter $(d_n + d_0)/2$ and the mean basal area in larger trees rather than on the initial values. A better but more complex alternative is to interpolate using the functional form of the equation being fitted (McDill and Amateis 1993).

Many modellers choose the growth rather than yield formulation for predicting diameter increment because the implications regarding serial correlation of errors are minimized, but they remain divided between predicting diameter increment or basal area increment. In practice, this choice is immaterial. Any of the four alternatives (growth or yield of diameter or basal area) can be formulated to provide reasonable estimates.

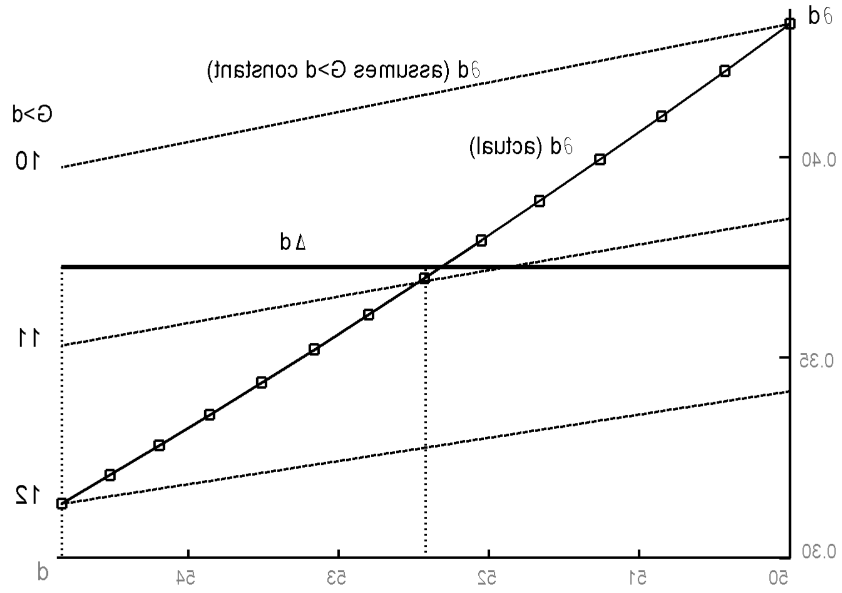


Fig. 8.1. Observed periodic increment *(Δd)* may be a better predictor of annual growth *(∂d)* at the mean diameter and competition ($G_{>d}$), than at initial conditions.

Competition Indices, Modifier Functions and Allocation Rules

Competition indices are an attempt to quantify in a simple expression, the effects of neighbouring trees (or other plants) on the growth of an individual in a forest stand. They may be expressed as an absolute value such as stand basal area, or as a relative index comparing tree growth to expected growth under ideal conditions. Many competition indices have been proposed (see Chapter 4, p. 58), but several empirical trials have suggested that general non-spatial measures of competition such as stand basal area and basal area in larger trees are as effective as other indices for many applications (e.g. Opie 1968, Lorimer 1983, Martin and Ek 1984). Many competition indices require spatial data and are thus unsuitable for non-spatial models such as size class and tree list approaches.

One robust way to model diameter increment is to predict potential growth and use a *modifier function* to estimate actual increments (e.g. Ek and Monserud 1974, Leary 1979, Arney 1985, Pretzsch 1992*a*):

$$\textit{Expected growth} = (\textit{Potential growth}) \times (\textit{Modifier function})$$

Whilst this has some attractions, it poses several difficulties. One difficulty is estimating the potential growth rate. Shifley (1987) based his equation for potential growth rate on the fastest growing five percent of trees in his data. Such approaches should be based on the average growth rate between the first and last measure, since the use of growth estimates between individual consecutive pairs of measures may select for measurement errors rather than for real growth. An alternative that avoids this problem is to base the potential growth equation on trees assessed as open-grown or free of competition.

An alternative approach is to predict the stand increment, and distribute this among the stems comprising the stand. Stand increment may be predicted as basal area increment (e.g. Opie 1972, Clutter and Allison 1974), as increment in the sum of diameters (e.g. Leary 1979, 1980), as biomass increment (e.g. Bossel and Krieger 1991), or as canopy photosynthesis (e.g. Sievänen and Burk 1993). Allocation rules range from the simple to the complex. One widely used approach is to allocate the basal area increment according to the basal area:

$$\frac{\Delta g_i}{\Delta G_i} = \left(\frac{g_i}{G_i} \right)^w$$

where the weight w may vary from 0.93 (Campbell *et al.* 1979) to 1.25 (Opie 1972) for *Eucalyptus regnans*.

Vanclay (1988*a*) found that this relationship (with $w = 1$) held for even-aged stands of *Callitris*, but that in uneven-aged stands, the smaller trees got a greater share of the increment (i.e. $w < 1$). This result may reflect the fact that competition in this semi-arid part of southern Queensland is mainly for moisture rather than for light. The allocation of stand increment to individual trees was predicted from the cumulative basal area distribution using the diameter of the largest tree, the stand mean diameter and standard deviation of diameters as explanatory variables.

Leary *et al.* (1979*b*) predicted the share of the increment in sum of diameters with a more complex relationship:

$$\ln(Y+\gamma_1) = \beta_0 + \beta_1 \ln(X+\gamma_2) + \beta_2 \left[\ln(X+\gamma_2)\right]^2 + \beta_3 \left[\ln(X+\gamma_2)\right]^3$$

where $Y = \Delta d_i / \Delta D_i$ and $X = d_i / D_i$, the βs are constants common to species and stands, and the γs depend upon species and stand condition. All these methods may provide good results in well-stocked stands with few species, but the allocation rules become complex for stands with many species.

Explanatory Variables

Many of the explanatory variables used in plantation growth and yield models are of little relevance to models for mixed forests as they cannot be determined or have no apparent meaning. These include stand-level variables such as age, site index, top height, and mean diameter. Spatial competition indices may be useful in research applications, but often cannot be used in growth models for forest management because the necessary spatial data are rarely available from operational inventory.

Variables which may be more accessible in mixed stands include species, diameter, stand basal area and estimates of site productivity (Chapter 7). The advantage of these variables is that the model can readily predict future values during long simulations.

Other variables such as crown size and position (e.g. dominant, intermediate, suppressed) may exhibit a high correlation with diameter increment, but pose difficulties in predicting how these variables themselves change over time. It may be valid to assume that crown size and position do not change during short projections, but it is unreasonable to assume this during long simulations. Some studies (e.g. Alder's 1990 Ghafosim model for high forest in Ghana) recruit the largest trees to the dominant crown classes to maintain a predetermined number in each class and replace dominant trees that died or were harvested. Many studies have observed a high correlation between crown characteristics and the growth during the preceding period (e.g. Wadsworth *et al.* 1989), but overlook the fact that growth in the succeeding period is not so well correlated with crown characters (i.e. many tree variables are better for description rather than prediction of growth).

Basal Area in Larger Trees

The basal area of trees larger than the subject tree ($G_{>d}$) has been found to be a useful predictor of diameter increment in both temperate (e.g. Meldahl *et al.* 1985, Wykoff 1990) and tropical forests (e.g. Vanclay 1991*b*). It is analogous in many respects to available light, used in many succession and process models. For example, the JABOWA model (Botkin 1993) assumes that leaf area is proportional to the square of tree diameter, and "shading leaf area" ($A_{<h}$) is the sum of leaf areas on taller trees. Tree height is assumed to increase monotonically with diameter, so shading leaf area is analogous to basal area in larger trees (depending on the specific height–diameter relationships). Available light, estimated as $l_a = \beta e^{-\alpha A_{<h}}$, is a multiplier in increment predictions. This is equivalent to including $\ln G_{>d}$ in a linear model which has the logarithm of diameter increment as the response variable.

Several researchers (e.g. Ford and Diggle 1981, Cannell *et al.* 1984, Hara 1986) have suggested that competition between plants in a monoculture is mainly for light, rather than for other environmental resources. This suggests that a component of competition is "one-sided" (i.e. that larger plants shade smaller ones, but not vice-versa; but note that this dependent on plant architecture, and that in forestry it may be more correct to say that "higher leaves shade lower leaves"). Thus basal area in larger trees should be a useful predictor, complementary to stand basal area which indicates "two-sided" competition (e.g. for resources other than light, including moisture and nutrients). The relative importance of these complementary measures may vary with site. On moist, fertile sites, light may be the limiting factor, whereas other resources may be limiting in other situations. Thus these two measures of competition (G and $G_{>d}$) should be viewed as complementary rather than as alternatives.

Wykoff (1990) found that the interaction between tree size and basal area in larger trees, $G_{>d} / \ln(d{+}1)$, gave better predictions than $G_{>d}$. He also argued that basal area in larger trees was a more appropriate explanatory variable than relative size ($G_{>d} /G$ and d /d_g), as the latter would be influenced by harvesting and could predict a counter-intuitive response to thinning from below.

One weakness with $G_{>d}$ is that it assumes that all species exert equal competitive influence. However, it is likely that some species (e.g. those with large, dense crowns) exert a greater influence than others, so some weighting according to species may be desirable. There is currently no paradigm for "strong" and "weak" competitors, and this remains a fertile area for further research.

Diameter Increment Functions

Many diameter growth and yield functions have been published and no attempt is made to review even a few these. Only those functions that do not require age are considered. Functions are grouped into broad classes (empirical, theoretical, probabilistic) to allow generalizations to be made. Little distinction is made between diameter and basal area increment models, or between growth and yield models.

Empirical Functions

Empirical equations are simply mathematical expressions with a resemblance to the observed growth, but without any associated hypotheses of cause or function of the phenomenon. Such equations may be useful for interpolation, but generally contribute little toward further understanding of stand dynamics, and are often unreliable when extrapolated beyond the limits of the data. However, empirical equations can be formulated to provide biologically realistic predictions across a wide range of values, and such equations may provide better predictions than theoretical equations (e.g. Martin and Ek 1984). They may also be easier to fit to the data. Many empirical equations are developed using stepwise regression analysis to select variables correlated with the response variable. Unfortunately, such equations often include an unnecessarily large number of variables, and may perform poorly when used near the limits of the data.

One commonly used empirical equation is the simple quadratic relationship $\Delta d = \beta_0 + \beta_1 d + \beta_2 d^2$, but it does not ensure robust results. If

β_2 is positive, the quadratic equation predicts ever-increasing increments for larger and larger trees. This is biologically untenable but can be seen in some models. Quadratic equations may provide reliable estimates of growth over a limited range of diameters, but are unsuitable for extrapolations, and should not be used in models which may be used to provide long term simulations.

Mawson (1982) recognized these limitations of the quadratic equation, and proposed a variation of Schumacher's (1939) equation: $\ln(\Delta d) = \alpha + \beta d^{-1}$. He argued that the parameter α could be expressed as a function of site, and the parameter β as a function of past stand treatment. Predictions from this equation are asymptotic to e^{α} as d becomes large, so care should be taken that sensible values are estimated for the parameter α (and all other estimated parameters).

Hilt (1983) used a two-stage analysis to establish diameter increment functions for even-aged oak forests. The first stage fitted $\Delta g = \beta d^2$ (Δg is tree basal area increment and d is dbh) to individual tree data from each plot, and the second stage fitted

$$\ln\beta = \gamma_1 + \gamma_2 \ln S + \gamma_3 d_g + \gamma_4 N_{\%}$$

to each plot, where S is site index, d_g is the quadratic mean diameter and $N_{\%}$ is percent stocking. This can be expressed as a diameter increment function:

$$\Delta d = \alpha\, d\, S^{\gamma_2} e^{\gamma_3 d_g + \gamma_4 N_{\%}}$$

Although this function performed well with the data used to fit the model, it assumes a linear relationship between Δd and d, and may predict excessive increments for large trees.

Theoretical Functions

There are no particular theoretical equations relating specifically to the growth of trees, but the Bertalanffy equation (see p. 108), originally formulated for weight gain in animals, is often used to model diameter growth in trees:

$$\Delta d = \beta_1 d^{\alpha} - \beta_2 d = \beta_2 d \left\{ \left(\frac{d_{max}}{d} \right)^{1-\alpha} - 1 \right\}$$

where d is tree diameter and d_{max} is the maximum attainable diameter.

The Bertalanffy equation overcomes many of the shortcomings of empirical equations, and provides for an asymptotic maximum diameter which cannot be exceeded. Martin and Ek (1984) considered a modification of the Bertalanffy equation for *Pinus resinosa* plantations:

$$\Delta d = \left(0.283\, d^{2/3} - 0.0492\, d\right) e^{-0.0392\, G}$$

but found that carefully formulated empirical equations could provide more accurate predictions within the range of the data. Shifley (1987) used a similar equation to model potential growth of many species (e.g. for eastern red cedar):

$$\Delta g = \left(0.0124 g^{0.515} - 0.0149 g\right)\left(0.397 + 0.00236 S_{50} + 0.749 c_r\right)$$

where g is tree basal area (m^2), S_{50} is site index (m) at age 50 and c_r is crown ratio, the ratio of crown length to total tree height. This was fitted using a two stage approach. The first stage fitted

$$\Delta g = \beta_1 g^{\alpha} - \beta_2 g$$

The estimated asymptotic maximum size

$$g_{max} = \left(\frac{\beta_1}{\beta_2}\right)^{1/(1-\alpha)}$$

was then compared with the national register of big trees. If the asymptote g_{max} seemed unreasonable, the parameter β_2 was revised to take the value $\beta_2 = \beta_1\, g_{max}^{\alpha-1}$. The second stage estimated the final three parameters in the equation. These potential increments were reduced by a modifier predicted from tree size, basal area in larger trees and stand basal area.

Another variation of the Bertalanffy equation is the equation of Hahn and Leary (1979) and Leary (1980):

$$\Delta d = \beta_0 + \beta_1 d^{\beta_2} + \beta_3 S c_r d^{\beta_4}$$

where d is diameter, S is site index and c_r is crown ratio. This equation includes an intercept β_0, which improves the prediction of increments of very small trees, but detracts from the theoretical attraction of the equation. Unlike the mass of an organism, tree biomass is not zero when dbh is zero, so the Bertalanffy equation may underestimate increments for small trees. An intercept is one solution, but it is theoretically preferable to use an allometric relationship of $d+\gamma$ (i.e. $\beta_1 (d+\gamma)^{\alpha} - \beta_2 (d+\gamma)$) rather than an intercept. The assumption that site index and crown ratio affect only photosynthesis and not respiration may not be entirely realistic, but means that the increment pattern (as well as rate) may vary with site and crown ratio.

The JABOWA model (Botkin *et al.* 1972, Botkin 1993) uses a similar equation, based on the logistic equation ($dY/dt = \beta Y(1 - Y/Y_{max})$), one of the Bertalanffy family of equations:

$$\frac{d}{dt}\left(d^2 h\right) = \alpha\, d^2 \left(1 - \frac{d}{d_{max}}\frac{h}{h_{max}}\right)$$

This equation can be re-arranged to give a diameter increment function, and a few simple assumptions (e.g. a quadratic height–diameter relationship, see p. 70) reduce the equation to one with only three parameters which can be determined subjectively (e.g. h_{max}, d_{max} and Δd_{max}). This basic growth function was modified to account for shading, climate and soil quality.

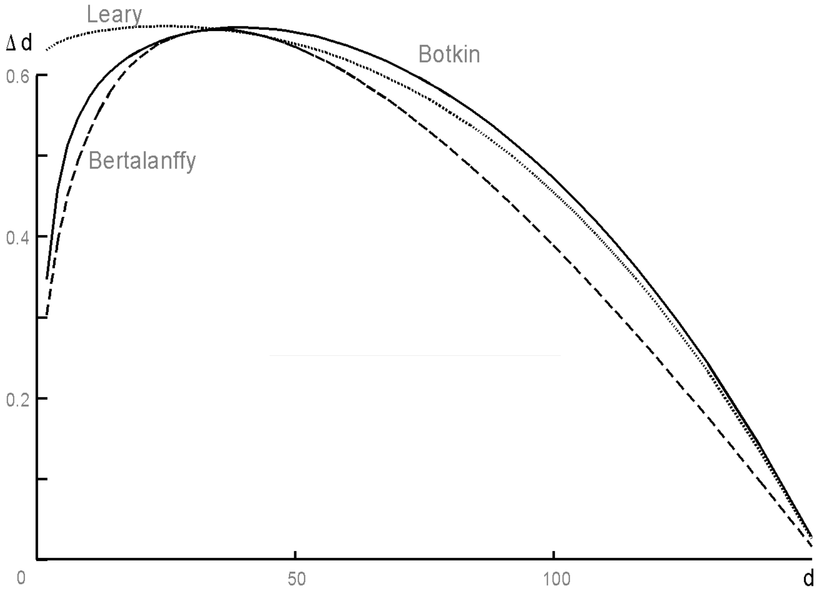


Fig. 8.2. Comparison of three "theoretical" diameter increment functions (drawn from Equations 8.1–8.3).

The Bertalanffy, Leary and Botkin equations describe rather similar diameter increment patterns. Figure 8.2 illustrates the following three equations, with parameters selected so that all three equations coincide approximately (based on sugar maple with $S_{h,t} = 21$ m and $c_r = 0.32$):

Bertalanffy: $$\Delta d = 0.245 d^{0.44} - 0.0147 d \tag{8.1}$$

Botkin: $$\Delta d = \frac{d - d^2 \dfrac{137 + 50.9d - 0.167d^2}{611677}}{2.74 + 1.527d - 0.00668d^2} \tag{8.2}$$

Hahn and Leary: $$\Delta d = 0.477 - 0.0000016 d^{2.58} + 0.145 d^{0.0838} \tag{8.3}$$

Empirical Analogues of Theoretical Functions

Although the Bertalanffy equation is flexible and frequently used to model tree growth, it has limitations that may be attributed to its origins as a particular theoretical equation for growth in weight of animals. When applied to diameter increment data of forest trees, the equation may overestimate the increment of large trees. This may be attributed to the fact that although an animal consists almost entirely of living tissue, a tree is a thin layer of living tissue enclosing a corpse of dead wood. Thus an asymptotic relationship may describe tree respiration better than the linear relationship assumed in the Bertalanffy equation. The effect of such an asymptotic relationship is to "bend" the right-hand tail of the increment curve, allowing trees to attain larger diameters (Fig. 8.3), and can be achieved using a variant of the Bertalanffy equation (e.g. Equation 6.2, p. 108, 111). One such family of equations relies on the underlying relationship:

$$\ln\Delta d = \beta_0 + \beta_1 \ln d + \beta_2 d^k$$

where d is tree diameter and typically $k=1$ or $k=2$. Zeide (1993) examined many equations for predicting tree growth (including many employing tree age as a predictor variable), and found this equation (with $k=1$) the best alternative for estimating individual tree diameter increment.

This equation is the basis of Wykoff's (1990) revision of growth functions in the Prognosis model:

$$\ln\Delta d^2 = \beta_0 + \beta_1 \ln d + \beta_2 d^2 + E + C$$

where E describes the environmental effects and is estimated from habitat type, location, elevation, slope and aspect; and where C describes competition and is estimated from crown ratio, crown competition and relative tree size. The final function was:

$$\begin{aligned}\ln\Delta d^2 = \beta_0 &+ \beta_1 \ln d + \beta_2 d^2 \\ &+ \beta_3 E_s \cos E_a + \beta_4 E_s \sin E_a + \beta_5 E_s + \beta_6 E_s^2 + \beta_7 E_e + \beta_8 E_e^2 \\ &+ \beta_9 c_r + \beta_{10} c_r^2 + \beta_{11} \frac{G_{>d}}{\ln(d+1)} + \beta_{12} C_{cf}\end{aligned}$$

where d is diameter, the Es are environmental variables (e.g. slope, aspect, elevation), $G_{>d}$ is basal area in larger trees, c_r is crown ratio, and C_{cf} is crown competition factor. The first line of this equation reflects the effect of tree size on increment, the second line is a proxy for site productivity, and the third line accounts for competition. This equation produces diameter increments similar to those produced by the theoretical equations, and is constrained to produce sensible increment predictions for any tree size and any stand density.

Another empirical linear equation (Vanclay 1991*b*) with a similar shape does not require crown characteristics, and may provide good predictions (see Fig. 7.4, p. 144):

$$\ln(\Delta d+\alpha) = \beta_0 + \beta_1 d + \beta_2 \ln d + \beta_3 S_{\Delta d} \ln d + \beta_4 \ln G + \beta_5 G_{>d}$$

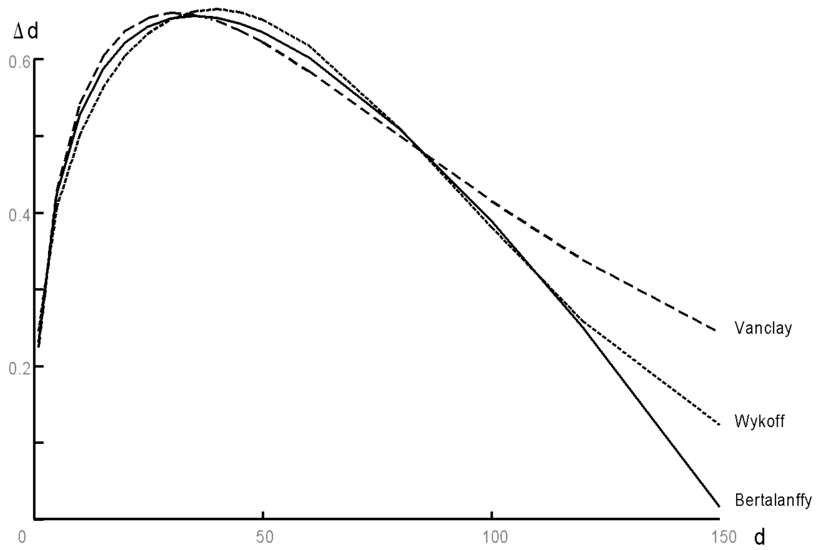


Fig. 8.3. Comparison of Bertalanffy, Vanclay and Wykoff equations for predicting diameter increment (drawn from Eqns 8.1, 8.4 and 8.5).

These two empirical analogues are contrasted with the Bertalanffy equation in Fig. 8.3. The Bertalanffy equation is the same as before (Equation 8.1), while the other equations were fitted using least squares to data generated from the Bertalanffy equation for $d = 5, 10, 15, \ldots, 90$.

Wykoff: $$\ln \Delta d = -1.41 + 0.314 \ln d - 0.000101 d^2 \qquad (8.4)$$

Vanclay: $$\ln \Delta d = -1.48 + 0.439 \ln d - 0.0142 d \qquad (8.5)$$

The equations describe a similar shape for trees less than 80 cm diameter, but the Vanclay and Wykoff equations always predict a positive increment, allowing some growth on trees of very large sizes, while the Bertalanffy equation imposes a maximum attainable size (here 150 cm dbh). Although the Bertalanffy equation provides a convenient way to constrain the upper-end of a relationship where data are few (but it may be impossible to estimate objectively when data are few), this imposition of a maximum tree size may not be very realistic, since heartwood has no respiratory cost (e.g. Prentice and Helmisaari 1991).

Probabilistic Functions

Lowell and Mitchell (1987) used a probabilistic function to predict diameter increment, an approach that enables simultaneous estimation of increment and mortality. They argued that since growth and mortality are biologically related they should be modelled simultaneously. Lowell and Mitchell (1987) illustrated the approach for 4 species groups in even-aged mixed oak forests in the USA. Their equation for white oaks was

$$p_{\Delta d} = \left(1 + \exp\left[-8.901 + 271.1\Delta d - 2.733 d - 1.594 \ln\frac{d}{D}\right]\right)^{-1}$$

where $p_{\Delta d}$ is the probability that a tree of d cm dbh will achieve a diameter increment exceeding Δd cm over a five year period (D is the sum of diameters $\sum d$ of all trees per hectare; note that d/D can also be expressed as the relative diameter $d/\overline{d}$ divided by the stem number N). The probability p_0 of achieving a zero increment is the probability of survival. Since Δd enters the equation in a linear form, the model will predict a normal distribution of diameter increments. Transformations such as $\ln\Delta d$ may be necessary to provide for a non-normal distribution of increments.

A similar probabilistic diameter increment function was used in the NORM model (Vanclay 1991*d*). The probability that a tree would complete one centimetre of growth during a given year (i.e. that a tree less than d cm dbh would attain a size of d cm or more within a one year interval, for any integer d) was predicted from tree size, competition and site factors (e.g. for *Flindersia pimenteliana*, Fig. 8.4):

$$\begin{aligned} p = (1 + \exp[&-0.738 + 0.108 d - 1.99 \ln d - 0.146 S_{\Delta d} \ln d \\ &+ 1.99 \ln G + 0.0355 G_{>d} - 0.422 S_{av}])^{-1} \end{aligned} \tag{8.6}$$

where p is the predicted probability, d is tree size (cm dbh), $S_{\Delta d}$ is the growth index, G is stand basal area ($m^2 ha^{-1}$), $G_{>d}$ is basal area in larger trees ($m^2 ha^{-1}$), and S_{av} is a binary variable which takes the value one on soils derived from recent alluvial, volcanic or granitic parent material, and zero on soils derived from sedimentary or metamorphic parent materials. This formulation offers two advantages over alternatives: it is robust in the presence of outliers and simplifies the construction of compatible deterministic/stochastic growth models.

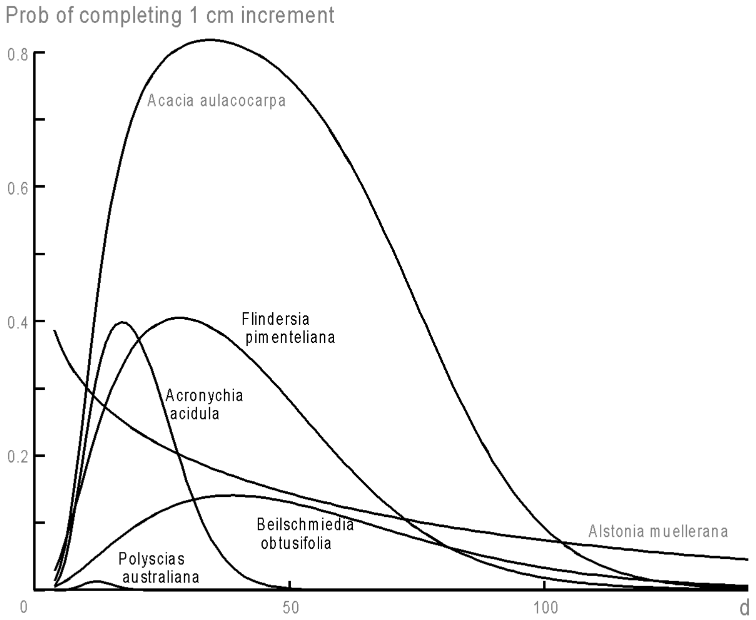


Fig. 8.4. Annual diameter increment probabilities for several species in the NORM model, under typical conditions (drawn from Equation 8.6 with $S_{\Delta d}=6$, $G=30$ and $G_{>d}=G\times(1-d/140)$).

Diameter Increment as a Stochastic Process

Stochastic variation in diameter increment may be introduced into a model by adding random variates to predicted increments, or by simulating the same effect in a deterministic way by using a "swindle" such as record tripling.

Stage (1973) observed that the introduction of stochastic variation in predicted diameter increments would have ramifications on all other components of the Prognosis model. The distribution of the logarithm of the tree basal area increments is assumed to be normally distributed, consistent with the regression model. The logarithmic transformation provides for heteroscedastic increments, and ensures that negative increments do not arise. In stands with many trees where the effect of the stochastic element is not likely to influence the total stand estimate, uniform random variables are assigned to each tree, and the corresponding normal deviate is added to the logarithm of its increment. Because of the logarithmic transform, the effect of this is multiplicative and allows a small proportion of trees to have comparatively large increments. The random variable associated with each tree is retained to preserve the appropriate serial correlation in increment estimates. Where fewer cohorts were present, the same effect is introduced in a deterministic way by "tripling" cohorts (see p. 72).

Failure to preserve serial correlation in single-tree models may, in theory, lead to underestimates of yield. Swaine *et al.* (1987) reported significant correlation between successive increment observations and found that these persisted for several years. Hilt (1983) predicted the standard deviation for a given increment prediction as $\sigma = \alpha e^{\beta d}$ and drew random numbers to assign increments to individual (simulated) trees. He observed a correlation of 0.63 between successive increments and assumed the bivariate normal distribution in calculating successive increments. Stage and Wykoff (1993) demonstrated how to construct a model for unexplained variation in estimates of periodic tree increments, including serial correlation. Their work improved the theoretical basis of the Prognosis model, but made little difference to model estimates because the reduction in variance thus attained was compensated for by modelling serial correlation (which tends to increase the range of estimates).

Synthesis

No attempt has been made to review the vast number of functions used to predict diameter increment. Instead, selected equations have been used to illustrate strengths and weaknesses of some popular relationships. The most appropriate equation in any situation may depend upon the data available and the resources available for analysis. It is irrelevant whether growth or yield, basal area or diameter is modelled. But it is important that the researcher takes proper and explicit account of the error distribution, and ensures that the function provides reliable results over the full range of possible tree sizes, sites and stand conditions.

One diameter increment function that offers promise is a generalization of the Bertalanffy equation based on

$$\ln\Delta d = \beta_0 + \beta_1 \ln d + \beta_2 d^k$$

This equation has been used in several models, can easily be fitted to data, and provides robust predictions.

Exercises

8.1. Using the data given below, fit a simple linear regression to estimate diameter increment from initial diameter ($\Delta d = \beta_0 + \beta_1 d$). Fit the equivalent models for basal area increment ($\Delta g = \beta_0 d + \beta_1 d^2$) and future diameter ($d_1 = \beta_0 + \beta_1 d_0$). Explain the differences in the parameter estimates (βs) and the R²s. Which is the best model? Why?

d	10	11	12	13	14	15	16	17	18	19	20
Δd	0.74	0.82	0.80	0.81	0.91	0.94	0.90	0.93	1.02	0.97	1.01

8.2. The data below relate to the growth of *Anisoptera thurifera* in Papua New Guinea (Oavika 1990). Fit a diameter increment function to these data. Justify your choice of equation, and explain its strengths and weaknesses.

d	G	Δd	d	G	Δd	d	G	Δd
25	5	1.04	15	15	0.40	25	25	0.24
35	5	2.16	25	15	0.60	35	25	0.68
45	5	0.90	35	15	1.02	45	25	0.64
55	35	0.40	45	15	0.68	65	25	0.26
65	35	0.30	55	15	0.60			

8.3. Howard and Valerio (1992) used the data below in their study of silvicultural prescriptions for natural forests in the Osa Peninsula of Costa Rica ($G_{r\%}$ represents the basal area harvested or destroyed in logging as a percentage of the initial basal area before harvesting). Fit a diameter increment function to these data. Justify your solution by explaining its strengths and weaknesses. Howard and Valerio (1992) used the following relationship:

$$\Delta d = \beta_0 + \beta_1 G_{r\%} + \beta_3 (d - \alpha)^2$$

Explain why your model is better than theirs. Is the growth response to harvesting realistic, particularly as it relates to tree size?

d	$G_{r\%}$	Δd	d	$G_{r\%}$	Δd	d	$G_{r\%}$	Δd
7.5	6.7	0.19	17.5	17.1	0.38	7.5	43.3	0.29
15.0	6.7	0.30	22.5	17.1	0.28	15.0	43.3	0.56
25.0	6.7	0.45	27.5	17.1	0.45	25.0	43.3	0.47
35.0	6.7	0.40	32.5	17.1	0.51	35.0	43.3	0.79
45.0	6.7	0.46	37.5	17.1	0.51	45.0	43.3	0.53
17.5	10.4	0.28	42.5	17.1	0.57	7.5	76.7	0.45
22.5	10.4	0.25	47.5	17.1	0.46	15.0	76.7	0.95
27.5	10.4	0.29	52.5	17.1	0.45	25.0	76.7	0.87
32.5	10.4	0.35	57.5	17.1	0.46	35.0	76.7	0.57
42.5	10.4	0.35	67.5	17.1	0.25	45.0	76.7	0.34
47.5	10.4	0.33						

Chapter Nine

Mortality and Merchantability

Many growth models for plantations avoid the problem of predicting mortality by assuming that no mortality occurs in well managed stands. This assumption may be reasonable for intensively managed plantations, but is inappropriate when modelling natural forests where mortality is significant and should be taken into account. Stage and Renner (1988) found that most (80%) of the variability in volume predictions for mixed conifer forests in the Rocky Mountains region of the USA was due to uncertainty in mortality estimates. Clearly, mortality estimates may have a major influence on the accuracy of growth and yield forecasts.

Criticism and comparison of the alternative methods for predicting mortality is easier if we begin by classifying mortality according to cause and pattern of occurrence. Mortality patterns may differ in scale, frequency and severity, and these should be considered when choosing a modelling strategy.

Tree age may be a contributing factor in the death of a tree, but may not be the cause of death, and many deaths may be attributed to competition, pests and diseases, and chance events. Tree size does not appear to be a cause of tree death, but is in practice a good predictor of the probability of mortality.

Much attention has been given to the processes of competition and suppression, especially in pure stands. Much of this debate (e.g. Drew and Flewelling 1977, Aikman and Watkinson 1980, White 1981, Smith and Hann 1984, Westoby 1984, Lonsdale 1990) focuses on Reineke's (1933) stand density index and Yoda *et al.*'s (1963) self-thinning line, but it appears impossible to explain density-dependent mortality in sufficient detail with these simplistic theories, even in pure, even-aged stands (e.g. Zeide 1987, Norberg 1988, Skovsgaard 1994). In mixed forests, the situation is more complex (e.g. Sterba and Monserud 1993), but the need for light, nutrients and physical space continues, and any reduction below the minimum requirements will eventually lead to death.

Pests and diseases may also kill trees. In many natural forests, these often occur in a dynamic equilibrium, but epidemics may occur. The effects of an epidemic may range from a comparatively small reduction in increment for a year, to the elimination of the host species from the region.

Weather also influences mortality patterns in forest stands. Drought or prolonged waterlogging may hasten the demise of aged, diseased or suppressed stems, but may also cause death of healthy trees. Lightning, hail and wind storms may kill or damage trees. Wind damage may be catastrophic, and may or may not be dependent upon stand condition. Wildfires may also cause catastrophic losses. Many other tree deaths can be best attributed to chance, as no satisfactory explanation can be given.

Mortality or removal of trees from the forest as a result of human interference should also be simulated in forest growth and yield studies. The most obvious aspects of human interference may be the official harvest of timber and any silvicultural treatment (felling, girdling or poisoning of unwanted stems). Pollution and unauthorized timber harvesting may be significant in some areas, but modelling these impacts may also require socio-economic indicators and will not be considered in this chapter.

Authorized timber harvesting usually follows some guidelines, and these may be included in a model for growth and yield studies. From a growth modelling viewpoint, silvicultural treatment can be simulated in the same way as harvesting. Damage arising from harvesting activities may also contribute to mortality in forest stands. Logging operations may destroy smaller stems, may damage roots and bark, and may create entry points for disease and decay. Disturbance to the canopy caused by logging may also disadvantage some tree species, or favour pests and diseases. Thus harvesting practices may need to be considered in mortality models.

The relevant amount of detail to include in the model depends on its purpose. A succession model may simulate only natural mortality, but growth and yield studies require that many of these factors are taken into account. Here we consider some aspects of anthropogenic (human-related) mortality as well as natural mortality.

Natural Mortality

Two major categories of natural mortality can be distinguished: regular and catastrophic mortality. *Regular mortality* refers principally to ageing, suppression and competition, but also to mortality arising from "chance", and from normal incidence of pests, diseases, and weather phenomena (e.g. drought, storms, etc.; events which typically occur less frequently than once every ten years). *Catastrophic mortality* includes wildfire, occasional but severe losses from "abnormal" weather conditions, and major pest and disease outbreaks.

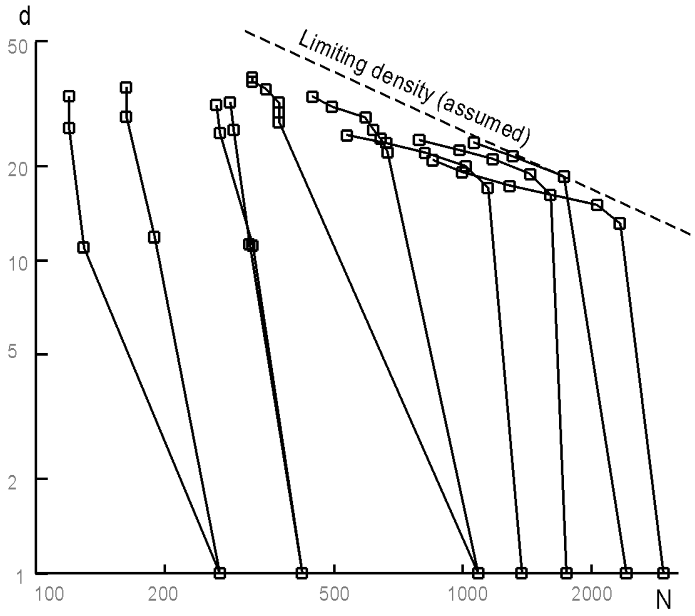


Fig. 9.1. The concept of limiting stand density. Stands grow "upwards" with little mortality until they approach the limiting density, after which deaths occur (data from T.J. Dean, Louisiana State Univ.).

Estimates of regular mortality may be based on theories of limiting stand density (e.g. Reineke 1933, Yoda *et al.* 1963), or may use empirical relationships fitted to stand- or tree-level data. Most theoretical and empirical approaches rely on estimates of stand density, competition or tree vigour, but some methods also draw on tree age.

The JABOWA model (Botkin *et al.* 1972, Botkin 1993) predicts chance mortality from the estimated potential maximum age of a tree species, assuming that only two percent of trees reach their maximum age. This is supplemented with competition-related mortality. No further mention is made of age-based methods, because of the difficulty of determining age in many natural forests.

Theoretical Approaches

Many growth models for even-aged forests predict density-dependent mortality, assuming that there is a simple relationship between maximum stand density (stems ha^{-1}) and mean tree size. Three relationships of this type have been proposed:

1. Relative spacing (Hart 1928, Wilson 1951) was originally proposed as a practical thinning guide, but has also been used as a guide to limiting stand density (e.g. Mitchell 1975). It relates the average spacing to top height (h_T), and can be expressed in terms of stem number (N):

$$\ln N + 2 \ln h_T = constant$$

2. Reineke's (1933) stand density index relates limiting stocking (N) to mean stem diameter (d_g) (Fig. 9.1):

$$\ln N + 1.6 \ln d_g = constant$$

3. Yoda *et al.*'s (1963) self-thinning line relates limiting stocking to plant mass, but in forestry applications, mean tree volume ($\bar{v}$) is often used as a surrogate for plant mass:

$$\ln N + \frac{2}{3} \ln \bar{v} = constant$$

These three alternatives are related, and some unification is possible if the stand development trajectories are viewed in three-dimensions (h_T, d_g, N; note that $\bar{v} \approx f(h_T, d_g)$) and these limits are expressed collectively as a plane in the three-dimensional space (García 1994). Here we deal mainly with Reineke's stand density index because diameter measurements are more accessible than estimates of height and mean tree volume (or mass).

Although these concepts of limiting density are elegant (at least for pure, even-aged stands), there are some practical limitations:

1. They indicate the residual stem number, but not the trees that die so that other assumptions are needed. Many modellers assume the smallest trees die, but this is not necessarily correct.

2. They do not indicate deaths from causes unrelated to competition (e.g. weather, physical injury, and some pests and diseases). These may contribute a substantial proportion of total mortality. For example, half of juvenile mortality in tropical forests may be due to physical causes (e.g. Hartshorn 1975, Osunkjoya *et al.* 1992).

3. The size–stocking relationship may not be linear, and its slope may not be constant. A constant slope assumes that (Zeide 1987):

(a) the combined action of crown growth and self-thinning maintains a closed canopy, and

(b) plants of the same species are geometrically similar in shape, irrespective of growth stage and habitat condition.

These two assumptions are unlikely to be satisfied for forest trees, as tree death creates gaps in the canopy which are not filled immediately by crown growth, and mechanical considerations require tree allometry (i.e. the relationships governing tree dimensions) to change with tree size.

4. There may be problems in identifying the onset of density-dependent mortality in data, and in obtaining reliable estimates of the coefficients describing the relationship (e.g. Ferguson and Leech 1976).

5. Unless the stand is pure, even-aged and evenly-spaced (i.e. a plantation), the relationship is likely to be highly data-dependent (e.g. dependent on plot size), as the onset of density-dependent mortality may not occur in all parts of the stand at the same time. Sterba and Monserud (1993) found that in uneven-aged mixed stands, the slope of Reineke's line depends on the skewness of the $d^{1.5}$ distribution. Reineke's line is almost flat for stands with a reverse-J diameter distribution (i.e. $\ln N + 0.0 \ln d_g = constant$).

Despite these limitations, the concept of a limiting stand density has been used as the basis for modelling mortality in several growth models for pure stands. However, many of these require subjective decisions regarding the trees that "die". For instance, Opie's (1972) model was based on Reineke's line, and "killed" every third tree commencing from the smallest, until the required reduction in stocking was obtained. Any trees smaller than one-seventh of the largest diameter in the stand were also assumed dead. Campbell *et al.* (1979) later modified the model to remove every sixth tree from the smallest, up to a maximum of 100 deaths per hectare in any year. Both of these alternatives are rather arbitrary.

Some alternatives to Reineke's line rely on tree height and crown competition. Mitchell's (1975) model used a relative spacing guide (Wilson 1951) based on the square of the site height (tallest 40% of trees), and assumed that the shortest trees would die first. Arney's (1985) approach was based on the crown competition factor. Leak (1969) used an exponential function of mean diameter increment, implying that every centimetre increase in stand mean diameter is associated with a fixed decrease in stocking: in his case, 15% for commercial species, and 19% for intolerant and intermediate species.

None of these options is entirely satisfactory, and a better alternative may be to predict limiting conditions directly from growing space, competition index or crown dynamics. Mitchell (1969) modelled the crown development of trees, and assumed that when the actual crown width fell below 17% of the potential open-grown crown width for a tree of that size, it had a 50% probability of being overtopped and dying. Arney (1972) assumed that trees would die if the crown length receded to less than 5% of the tree height. However, there is no threshold at which mortality increases markedly, and these limits are rather arbitrary.

Another alternative often used in models for uneven-aged stands is to estimate a threshold increment, and assume that all (e.g. Newnham 1964) or some (e.g. Botkin *et al.* 1972, Reed 1980) of the trees with predicted increments less than the threshold will die. Ek and Monserud (1974) used a stochastic function of diameter to predict the threshold, and assumed that all trees with predicted increments less than the threshold would die.

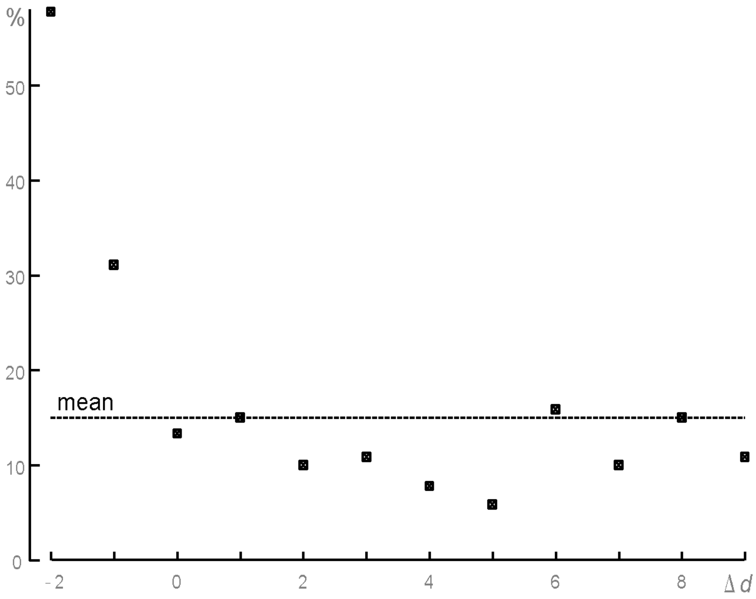


Fig. 9.2. Mortality (% of trees in class dying during 1970–82) in relation to growth rate (mm y^{-1}) at Kade, Ghana (redrawn from Swaine *et al.* 1987).

The threshold increment approach makes sense and is supported by some empirical evidence. Spurr (1962*b*) reported that any *Pinus radiata* tree achieving an annual increment less than twenty square centimetres, has on average, less than eight years to live. Bevege (1972) observed that stems dying in a *P. elliottii* plantation had existed for a considerable period (3 to 10 years, mean 6) with negligible increment, and this moribund period was independent of tree age and size. These observations show that the dead trees had been moribund for several years prior to death, but they do not prove that failure to achieve a critical increment results in death.

Swaine *et al.* (1987) observed that mortality in semi-deciduous forests in Ghana was significantly higher among trees that did not increase in size. Trees which decreased in diameter by 1 and 2 mm y^{-1} had twice and four times the average mortality rate (Fig. 9.2). Dead trees may have exhibited very little increment in the years prior to death, may have grown rapidly until shortly before death, or may have died while rapidly growing (e.g. Hartshorn 1975).

Threshold increments may provide a suitable basis for modelling some components of mortality amongst light-demanding species, but are less relevant to shade-tolerant species. For example, Douglas-fir trees may survive for more than 25 years with an average diameter increment less than 0.3 mm y^{-1} (Newnham 1964), and *Araucaria cunninghamii* seedlings may reach 50 years of age before attaining a height of 0.5 metres.

These theoretical approaches based on limiting conditions do not account for all regular mortality, and may need to be supplemented with empirical models of "chance" mortality. Another problem with predicting mortality according to causal agent (competition, etc.) is that it may be difficult to determine the agent responsible. The assumption that most mortality can be attributed to competition is not always reliable. An alternative is to predict a composite estimate of all regular mortality from stand or tree characteristics with an empirical function.

Empirical Approaches

Most size class models use an empirical method to predict mortality, often as simple linear functions of stand density and relative size (e.g. Moser 1972, Leak and Graber 1976, West 1981, Shifley *et al.* 1982). Such functions should be used with caution because they may give estimates outside the feasible range (0,1), if used beyond the range of data for which they were developed.

The logistic function (Fig. 9.3, overleaf) offers a convenient way to constrain predictions to the interval (0, 1), and provides a realistic (binomial) distribution of errors. The logistic function can be expressed in several ways:

$$P = \left(1+e^{-f(X)}\right)^{-1} = 1-\left(1+e^{f(X)}\right)^{-1} = \frac{e^{f(X)}}{1+e^{f(X)}} \tag{9.1}$$

where P is the probability of survival and $f(X)$ is a function (often linear) of several explanatory variables. Probability of mortality may be obtained from $1-P$. The advantage of predicting survival is that it, unlike mortality, is a Markov process (i.e. it is a transient rather than absorbing state), so that survival over an n year period is given by the n^{th} power of the annual probability of survival. Unequal measure intervals can be accommodated by using the interval as an exponent (i.e. $-t$ replaces -1 in Equation 9.1). This may sometimes cause difficulties, and an efficient alternative is to weight Equation 9.1 by the remeasure interval (Hamilton and Edwards 1976). Although not technically correct, this alternative seems reasonable provided that remeasurements are fairly frequent and mortality is relatively low (e.g. frequency < 8 years and mortality < 0.7%; Hamilton, pers. comm.).

Equation 9.1 may be converted to a linear equation if the data are grouped into categories with at least one death and one survival in each category:

$$Y = \ln\left(\frac{P}{1-P}\right) = \ln\left(\frac{N_{surv}}{N_{dead}}\right) = f(X)$$

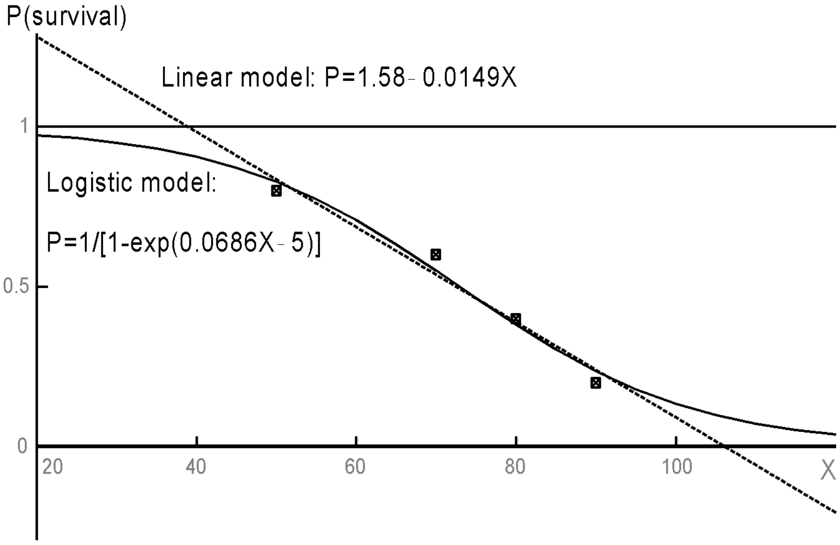


Fig. 9.3. The logistic is better suited for modelling mortality than linear equations because it is constrained between 0 and 1.

where P is the proportion of trees surviving, N_{surv} is the number of survivors and N_{dead} is the number of deaths. This is useful for initial investigation of variables and for plotting data, but is inefficient for parameter estimation and inappropriately affects the error distribution. Many statistical packages allow logistic functions to be fitted to data without this conversion to a linear form (e.g. using generalised linear models; Aitkin *et al.* 1989, Crawley 1993), and these should be used if they are available. The beauty of fitting Equation 9.1 without transformation is that it allows individual tree data to be used, allowing the maximum information to be extracted (e.g. there is no need to use the class mean for tree size and other explanatory variables).

Equation 9.1 has been used in many mortality models, often with empirical functions of diameter, height, defect, crown class and stand basal area (e.g. Hamilton and Edwards 1976). It is useful to include two transformations of tree size (e.g. $d^{0.5}$ and d^{-1}, or d and ln$d$) in the model to allow a good fit for both large and small trees (Hamilton 1986). Many tree variables are highly correlated and provide equally good predictions of mortality when used separately, but offer no further improvement when more than one variable is included (e.g. tree height and diameter).

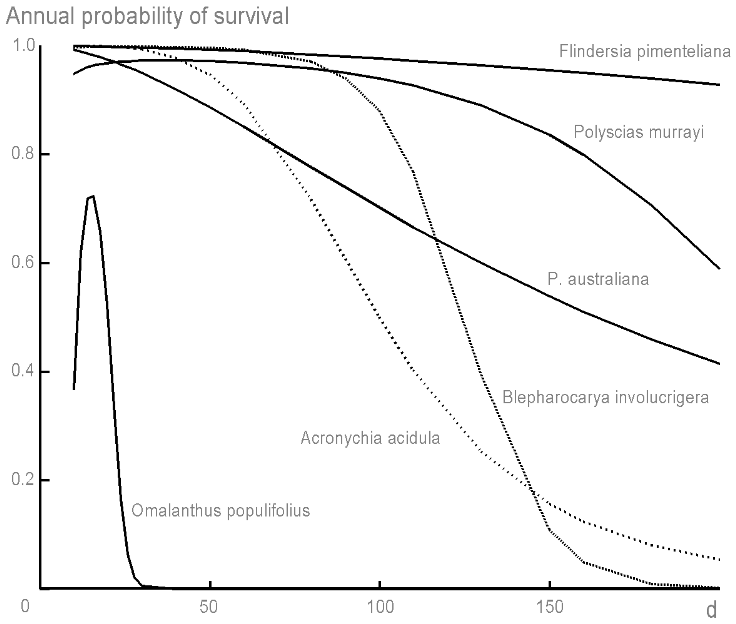


Fig. 9.4. Predicted survival patterns for several tree species in the NORM model (relative size 0.5, growth index 7, basal area 30 $m^2 ha^{-1}$; drawn from Equation 9.2).

The relative size of individual trees may also be useful, and has been included in mortality models as a relative diameter ($d/\overline{d}$) (e.g. Hamilton 1986, 1990) and as the rank in the cumulative stand basal area distribution ($G_{>d}/G$). Although analogous, these alternative measures of relative tree size imply a different response to harvesting. Relative diameter tends to increase (thinning from above) or decrease (thinning from below) by the same proportion for all trees in the stand (i.e. $\overline{d}$ changes while the ds remain unchanged), whereas changes in relative rank depend on tree size as well as harvesting direction (since both $G_{>d}$ and G may change).

Tree survival in the NORM model was predicted from tree size, site and competition. For *Flindersia pimenteliana*, survival of individual trees was predicted as (Vanclay 1991*c*, Fig. 9.4):

$$p = \left(1 + \exp\left[-11.0 + 1.67 \ln d + 0.0 d + 2.88 \left(\frac{G_{>d}}{G}\right)^3 + 0.104 S_{\Delta d} + 0.0 G + 0.0 \ln G\right]\right)^{-1} \tag{9.2}$$

where p is the annual probability of survival, d is tree diameter, $G_{>d}$ is basal area in larger trees, G is stand basal area, and $S_{\Delta d}$ is site quality. The zero parameters are non-zero for some other species groups. These probabilities are used to reduce the expansion factors of cohorts in the annual simulation cycle in the NORM model (e.g. see Fig. 4.3, p. 72).

The effect of site productivity on survival is unclear. There is empirical evidence that in plantations, density-dependent mortality expresses itself earlier on better sites, and if mortality is expressed as a function of age, it appears that mortality increases with increasing site productivity. However, if mortality is expressed with respect to top height or stand density, a different picture emerges. Better sites should be able to sustain a higher basal area, and all other things being equal, should have lower mortality. Half the species groups in north Queensland rainforests show a positive coefficient for site productivity ($S_{\Delta d}$ in Equation 9.2), while the others show no significant trend. Empirical investigations into the interaction between site productivity and mortality appear to be hampered by a lack of suitable experimental data.

Many models have employed past diameter increment to predict probability of mortality (e.g. Ek and Monserud 1979, Hamilton 1986, Wan Razali 1989, Pretzsch 1992*a*). However, predicted and observed diameter increments will provide different parameter estimates, and it is important to choose the appropriate formulation to suit intended uses of the model. Generally, observations of past diameter increments will not be available, and models fitted with predicted increments will have greater utility, despite poor goodness-of-fit statistics. Monserud (1976) predicted the survival of all species in mixed northern hardwoods stands with the equation:

$$p = \left(1 + \exp\left[-1.45 - 0.088d - 0.62\Delta\hat{d} + 0.0015c\right]\right)^{-t}$$

where p is the probability of survival over a t year period, d is diameter, $\Delta\hat{d}$ is predicted diameter increment and c is a competition index. This function correctly classified 88% of survivals and 35% of deaths. An analogue of this equation using actual rather than predicted diameter increment correctly classified 98% of deaths and 90% of survivals.

The use of predicted diameter increment in the mortality function makes it dependent on the diameter increment function, which then should not be revised without also revising the mortality model. If mortality can be estimated from predicted increments, then it should also be possible to estimate it from the explanatory variables for increment without the intermediate step of estimating diameter increment. Modelling mortality directly from tree and stand variables is a more robust approach and should be preferred.

It takes a lot of data to obtain good mortality models, and data available for model development may not represent all sites and stand conditions for which the model may be used. The fact that a model provides biologically reasonable predictions over a wide range of conditions does not necessarily ensure reliable extrapolations, but the inclusion of some specific provisions in its implementation may provide for more robust predictions. Hamilton (1990) found that he could improve predictions by specifically constraining mortality estimates in very dense stands, for very small trees, and in habitats and for species not represented in the development data set. His modifications included provisions to ensure that the site's maximum basal area would not be exceeded, and that mortality amongst new recruits (these have no previous diameter increment and may have $d < 0.5$) was bounded (i.e. same as if $d = 0.5$).

Mortality predictions may be applied in the model in a deterministic or stochastic way. The stochastic implementation draws a random number and compares it with the predicted probability of mortality. Alternatively, these probabilities can be interpreted as proportions, and the stocking in each cohort or class can be reduced by the predicted proportion. These alternative approaches should produce compatible predictions (e.g. Weber *et al.* 1986), but there are computational advantages in simulating mortality in a deterministic way unless the user is specifically interested in studies of variability.

Catastrophic Mortality

Catastrophic mortality is generally ignored in mortality functions. If it is to be taken into account, yields may be predicted in the absence of catastrophic mortality, and the final estimate may be reduced by an arbitrary allowance to account for such losses. There may be good reasons for adopting this approach, as any attempt to include catastrophic mortality data in regression analyses may severely compromise the assumptions of normality, and result in biased estimates. However, an objective estimate of the reduction to apply to the final yield is preferable to a subjective guess.

Catastrophic mortality can be modelled in two stages: first the probability of a catastrophe should be predicted, and then a conditional function should be used to predict the probability of mortality given that a catastrophe has occurred. This approach can provide a weighted estimate of annual mortality for a deterministic model including catastrophic mortality, or may be implemented directly in a stochastic model.

Several models simulate the interacting effects of pest or disease populations and stand condition. Stage (1973) explicitly modelled mortality due to mountain pine beetle using a deterministic model incorporating tree and stand characteristics (phloem thickness, bole surface area, stand density, etc.) and beetle population. Similar models exist for other pests and diseases (e.g. Valentine and Campbell 1975 for gypsy moth). Reed (1980) examined the development of a forest after the catastrophic elimination of one of its component species.

Harvesting and other Human Factors

As well as natural mortality, a growth model should be able to predict anthropogenic mortality, including planned harvesting, silvicultural treatment, and any deaths which may arise indirectly from these activities through damage, etc.

Here we are concerned with forest management-related activities including harvesting (i.e. commercial thinning, logging) and other silvicultural treatments (i.e. non-commercial thinning, timber stand improvement, climber cutting, poisoning of unwanted stems, etc.). In many forests, other human activities (e.g. unauthorized logging and cultivation, pollution) may have a significant influence on forest yields, and should be considered in yield studies. These activities cannot be predicted from forest conditions alone, but may also involve socio-economic indicators, and are outside the scope of this book. Nonetheless, they may have a significant impact on forest yields in many places, and should be taken into account in yield predictions.

Harvesting

A clearfelling operation is easy to model; all the merchantable stems are removed. Thinnings and selection harvests are more complex to model, as the size distribution of the removals or of the residual stand must be predicted. Systematic thinnings in pure even-aged stands are simple, as a constant proportion of stems in all size classes can be removed. In uneven-aged stands, there may be a greater variation in the stand condition, and selection logging tends to remove the largest stems, often according to defined size limits (*cutting limits*). Thus functions to predict the distribution of removals are likely to be cumbersome to use, and a better approach may be to specify "harvesting rules". Harvesting rules may specify the

1. number of trees in each size class to be removed at the time of harvesting,
2. percentage of trees in each size class to be removed, or the
3. number of trees in each size class to be retained after the harvest.

The first of these reflects a market-oriented approach which may be impractical to implement (i.e. cannot remove more trees than exist). The third option is ideal for plantations where the desired residual stocking can be specified; in natural forests the composition and stocking may be so variable that this option may be impractical. Thus for many natural forests, the second approach may be the only viable alternative, and is of practical relevance in computer simulation studies and as a field guide for officers marking trees for removal.

Logistic functions were used in the NORM model to predict the probability that a given tree would be harvested, given its species and size, and the time since the last harvest (Vanclay 1989*b*). The equation for some of the prime veneer species (e.g. *Flindersia pimenteliana*) was

$$p = \left(1 + \exp\left[6.09 - 0.0741\,d + 19.3\,t_l^{-1} - 1.70\,d_c\right]\right)^{-1} \qquad (9.3)$$

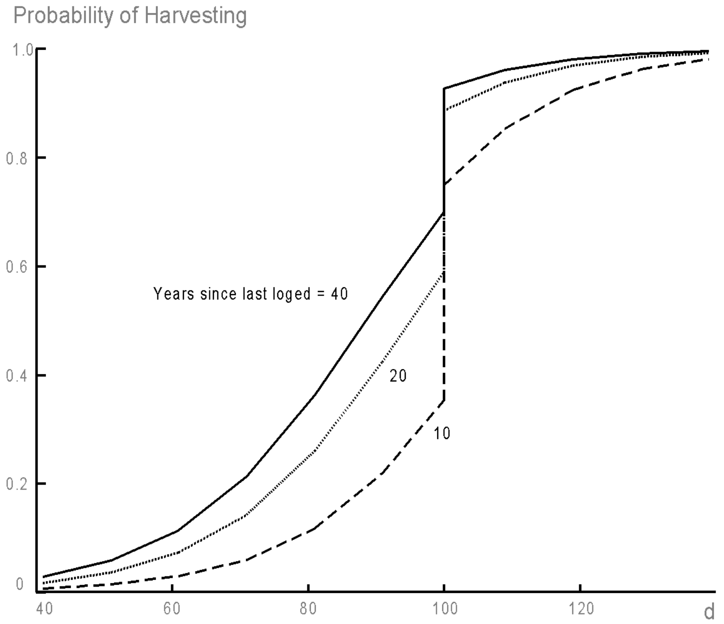


Fig. 9.5. Probability of harvesting a *Flindersia pimenteliana* tree, given its diameter and the time since last harvest (drawn from Equation 9.3).

where p is the probability of harvesting, d is tree diameter (cm dbh or above buttress), t_l is years since last harvest and d_c is a binary variable which takes the value one if the tree exceeds the cutting limit ($d > 100$) and zero otherwise (Fig. 9.5, overleaf).

Silvicultural Treatment

Silvicultural treatments such as liberation thinning, timber stand improvement, and other operations to reduce stand density can be modelled in much the same way as harvesting. Treatment prescriptions may dictate the removal of all old stags remaining after logging, a reduction in stocking in the smallest size classes to some specified amount, or poisoning of non-commercial species. These and similar treatments are essentially the same as harvesting, and can be implemented in the model in the same way.

Some treatments (e.g. climber cutting) may induce a response that is greater than can be attributed to the reduction in stand basal area, and must be modelled explicitly. Transient responses in diameter increment (additional to that attributed to the reduction in stand basal area) following silvicultural treatments have been found in many forest types, ranging from tropical rainforests to semi-arid forests (e.g. Vanclay 1988*a*). Such

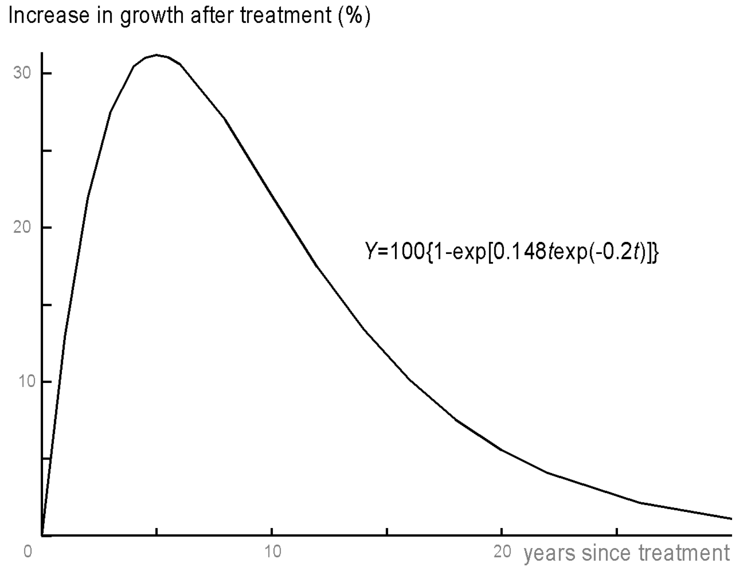


Fig. 9.6. Silvicultural treatment may stimulate tree growth for several years following treatment, and this may be modelled with a multiplier, illustrated here for *Flindersia brayleyana* in north Queensland.

responses can be accommodated directly in diameter increment functions (e.g. *Flindersia brayleyana* in north Queensland, Vanclay 1991*b*):

$$\ln(\Delta d + 0.02) = -1.08 + 1.19 \ln d - 0.060 d + 0.0964 S_{\Delta d} \ln d$$
$$- 1.02 \ln G - 0.0366 G_{>d} + 0.351 S_{soil} + 0.148 t_s e^{0.2 t_s}$$

where t_s is the time (years) since the last silvicultural treatment. The term $\beta t_s e^{0.2t_s}$ acts as a multiplier on the basic (untreated) growth rate, boosting growth predictions for several years (Fig. 9.6).

Logging Damage

Harvesting not only removes stems from the forest, but may also damage some of the trees remaining in the residual stand, and the effects of this damage should be included in the model. Logging damage should not be included with regular mortality, as its impact may be dependent upon the frequency and nature of harvesting. This distinction between logging damage and regular mortality may be particularly significant if the model is to be used to determine optimal stand condition and optimum cutting cycle, when failure to identify this component separately may bias estimates of the optimum cutting cycle.

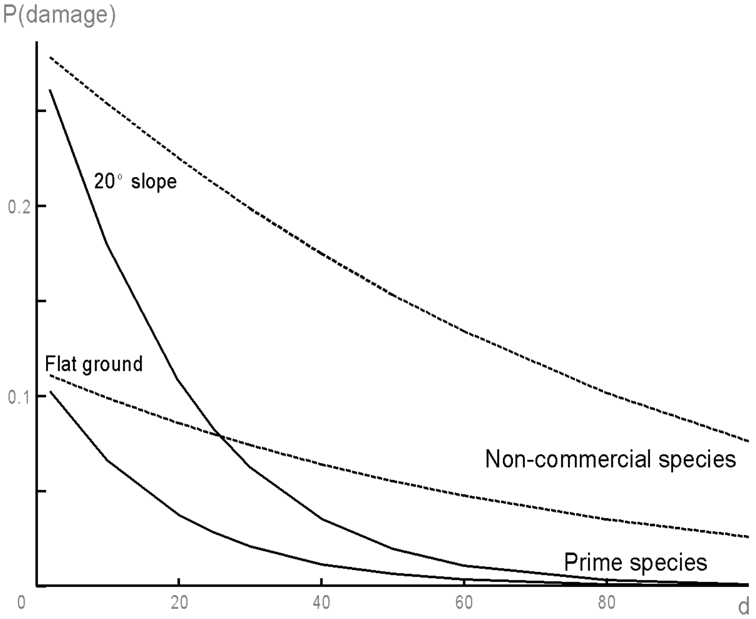


Fig. 9.7. Probability of logging damage in north Queensland rainforest. Directional felling minimizes damage to larger trees of the prime species, but slope has a big impact on the ability to control damage (drawn from Equation 9.4).

Relatively few studies of logging damage have been published as equations suitable for inclusion in growth models. The NORM model predicted the proportion of stems destroyed as a function of tree size, topographic slope and the basal area removed in harvesting (Fig. 9.7):

$$p = \left(1 + \exp\left[3.990 + \beta d - 9.689 G_r - 0.05648 E_s\right]\right)^{-1} \quad (9.4)$$

where p is the probability that a tree will be destroyed in a harvesting operation which removes a proportion G_r of the standing basal area, and where E_s is the topographic slope in degrees. The value of β depends on the species group, and takes the value 0.0596 for prime commercial species, 0.0361 for other commercial species, and 0.0157 for non-commercial species. This reflects the practice of directional felling which attempts to minimize damage to commercial trees in the residual stand.

Tree size, topographic slope and harvesting intensity explained most of the variation in this study, but canopy height may also be useful as explanatory variable in some cases.

Logging may cause mortality indirectly in several ways. The presence of logging waste or changes in the canopy conditions may favour pests or diseases, and trees with mechanical injuries may slowly deteriorate and die some years later. Walters *et al.* (1982) reported that tree mortality may treble for several years following logging in the Rocky Mountain region of the USA, apparently because of injuries inflicted during harvesting operations. A similar trend is also evident in Malaysian dipterocarp forests. There are two possible ways to model this effect. One way is to try to identify mortality arising from logging-related causes, and to fit a model to this data. An alternative and probably superior approach is to include these deaths in the composite model, and to include time since logging (or some suitable transformation) as a variable in the model to account for the increased mortality following harvesting.

Hann (1980) included time since logging in his logistic function for composite mortality, and this increased mortality estimates for a few years after a simulated harvest. However, time since logging did not improve predictions of mortality in the NORM model. Hamilton (1986) also found that time since and type of thinning had no impact on mortality rates following thinning in forests in Idaho (USA), and this did not explicitly include thinning in mortality functions for the Prognosis model. However, it remains important to discriminate between this logging-related mortality and natural mortality, or biased predictions may result, especially in studies concerning the optimal cutting cycle.

Merchantability

The assessment of merchantability may seem unrelated to mortality prediction, but for modelling purposes, they are analogous. Provided that we accept the assumption that once defective, a stem will never again become merchantable (i.e. "defect" is an absorbing state, and "merchantable", like survival, is a transient state), then we can predict deterioration of merchantable trees in the same way as we predict the death of living trees. Similarly, estimating the proportion of commercial stems among those felled is essentially the same as predicting the proportion to be felled.

Deterioration of Living Trees

Stems assessed as merchantable at time of inventory may not remain so until the next harvest. Some of these once-merchantable trees may deteriorate to the extent that they are no longer of commercial importance, and this deterioration should be taken into account in growth models.

Although small, this deterioration is cumulative and becomes sufficiently large during a cutting cycle to warrant inclusion in yield studies.

The NORM model predicted deterioration of rainforest trees using a logistic function of stand basal area, tree size, time since logging and soil type (Vanclay 1991*a*). For the more durable and valuable species (e.g.

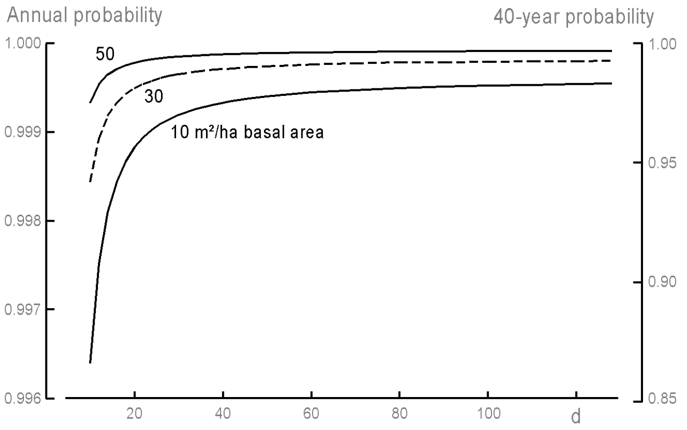


Fig. 9.8. Proportion of stems assessed as merchantable and remaining merchantable, as a function of tree size and stand basal area (drawn from Equation 9.5).

Flindersia pimenteliana) time since logging was not significant, and the prediction equation was (Fig. 9.8):

$$p = \left(1 + \exp\left[-7.450 - 0.04195G + 22.49d^{-1} + 0.4213S_{cg}\right]\right)^{-1} \tag{9.5}$$

where p is the annual probability that a tree remains merchantable, G is stand basal area (m^2ha^{-1}) and S_{cg} is a binary variable which takes the value one on soils derived from coarse granite parent material, and zero elsewhere.

Hann (1980) studied an analogous problem in *Pinus ponderosa*. Young vigorous trees have dark-coloured bark and are known as blackjack pine, while the mature slow-growing trees develop a yellow-coloured bark and are called yellow pine. Hann found that trees in these two categories displayed different growth and mortality rates, and that predictions could be improved by modelling the two categories separately. Thus it was necessary to predict the conversion from blackjack to yellow pine. One problem is that this conversion, like deterioration, is a slow but continual process, and the classification of a tree as one or the other is extremely subjective. Hann's model took the form

$$p = \left(1 + \exp\left[\beta_0 + \beta_1 d + \beta_2 d^2 + \beta_3 S_{h,t}\right]\right)^{-1}$$

where p is the proportion of blackjack (vigorous) trees converting to yellow pine, d is diameter and $S_{h,t}$ is site index.

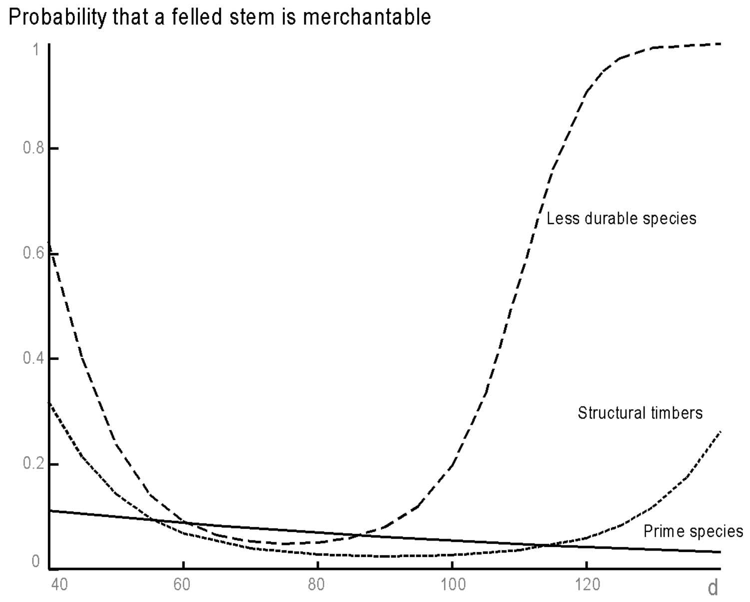


Fig. 9.9. Probability that a felled tree does not contain at least one commercial log. Species groups are prime (e.g. *Cardwellia sublimis*), structural (e.g. *Syzygium wesa*) and less-durable timbers (e.g. *Acacia mangium*) (Equation 9.6).

Merchantability of Harvested Stems

Not all trees assessed as merchantable and felled during harvesting will yield a merchantable log; some will be found, after felling, to be unmerchantable. The harvesting model could treat these stems as logging damage, and the predicted harvest could comprise only the merchantable stems. An alternative approach is to use three functions for modelling harvesting: one to predict all felled stems, one to predict the destruction of trees in the residual stand, and one to predict the proportion of felled stems which are merchantable.

The NORM model predicted the proportion of stems which were assessed as merchantable prior to felling, but which once felled, failed to realize at least one merchantable log. This is consistent with field practice: stems which are thought to be defective are not felled, and remain in the residual stand. The model used a logistic function with tree species and size as explanatory variables (e.g. for *Flindersia pimenteliana*, Vanclay 1989*b*; Fig. 9.9):

$$p = \left(1 + \exp\left[1.56 + 0.0129d + 0.0d^2\right]\right)^{-1} \tag{9.6}$$

where p is the proportion of apparently merchantable trees which fail to realize a commercial log and d is tree diameter (cm dbh). The values of the βs depended on the species group, and the resulting trends reflect the commercial value of the timber, the higher costs associated with smaller logs, and the increasing amounts of defect in the very large trees.

Strub *et al.* (1986) predicted merchantability of *Pinus taeda* stands at clearfall, and were concerned with the proportion of the total number that were merchantable (cf. previous example which predicted the proportion of stems assessed as merchantable). They used an exponential function of tree diameter and mean height of dominant and codominant trees:

$$P = \frac{1 - \exp\left[0.511(19 - d)\right]}{1 + \exp\left[9.09 - 0.751\bar{h}\right]}$$

where d and $\bar{h}$ are tree diameter (cm dbh) and mean height (m) of dominants respectively, and P is the proportion of the total stocking which is merchantable.

Synthesis

One reliable and efficient way to model mortality, merchantability and similar relationships, is to fit logistic functions to the individual tree data using maximum likelihood estimation or generalized linear modelling. Although these techniques are not always presented in elementary statistics texts, they are available in many statistical packages (e.g. Aitkin *et al.* 1989) and are not difficult to use. Such packages allow models to be fitted to the individual tree data enabling the most efficient use of these data and realistic assumptions regarding the error distribution.

Growth modellers have devoted more attention to natural stand dynamics than to the modelling of forest management (e.g. harvesting, etc.) but management models may be equally important in the context of management-oriented modelling of forest growth and yield. These models are sometimes based on few data and may not be very accurate, and this can degrade the overall accuracy of yield forecasts. Commensurate effort is required for constructing and testing all model components if a good overall result is required.

Exercises

9.1. The data opposite are a summary of tree deaths recorded on 212 permanent plots in a tropical rainforest (Vanclay 1991*c*). These 70871 observations were made on 30523 individual trees, some of which were measured more than once. The time interval varied, but averaged about 5 years. Fit a mortality function to these data. Discuss the merits of your equation.

Dbh (cm)	Total trees	Deaths	Dbh (cm)	Total trees	Deaths
10–14	27109	1306	50–54	1399	52
15–19	13106	608	55–59	928	40
20–24	8272	301	60–64	582	20
25–29	5840	200	65–69	450	22
30–34	4515	144	70–79	491	14
35–39	3325	109	80–99	298	7
40–44	2407	70	100–119	95	4
45–49	1881	62	120+	83	4

9.2. The (simulated) data below represent a harvesting study, in which a compartment was enumerated before and after logging to record the initial stocking and the number of stems felled, removed (i.e. merchantable) and damaged. Fit equations to these data so that you can predict the harvest, the proportion of merchantable stems, and the incidence of damage in the residual stand. Are your solutions generally applicable? Discuss.

d (cm)	Number of stems			
	Initial	Cut	Removed	Damaged
15	3571	0	0	307
25	1820	0	0	113
35	927	0	0	41
45	486	6	4	12
55	243	9	9	3
65	116	11	10	1
75	58	18	14	0
85	29	15	12	0
95	14	12	12	0
105	7	7	4	0
Total	7271	78	65	477

9.3. Howard and Valerio (1992) predicted logging damage as

$$P = 0.146 + 3.98G_r^2 - 0.00217d$$

where G_r is the proportion of the basal area removed in logging. Compare this equation with Equation 9.4 (p. 186), and discuss their relative merits and limitations.

Chapter Ten

Regeneration and Recruitment

Regeneration is the renewal of forest stands by natural (self-sown seed, coppice, suckers, lignotubers) or artificial means (sowing and planting). *Recruitment* refers to individuals that reach a specified size (e.g. breast height or a specified dbh). Both concepts are related, and the distinction depends on some rather arbitrary milestones in the development of an individual from a seed to a small tree (Fig. 10.1). Simulations may begin at any stage of development, but many models start with the "established seedling" or when trees reach some specified threshold size, usually based on its height or diameter. Accordingly, it is convenient to classify models as:

1. *Regeneration models,* which predict the development of trees from seed or seedlings, and
2. *Recruitment models,* which predict the number of stems reaching or exceeding some specified size limit (e.g. 1.3 m height, 10 cm dbh, etc.).

These are simply convenient categories to consider in the present context. Other alternatives have also been used in modelling forest renewal (e.g. Leak 1968, who started with flower development).

In theory, there are some merits in simulating as much of this process as possible, but in practice, this adds considerably to the complexity of a model and may provide no measurable improvement in model predictions. Efforts to build "seeding models" (Fig. 10.1) may be hampered by the large part that chance events seem to play in the development and germination of seed, the lack of empirical data to fit the model, and insufficient understanding of the physiological processes involved. So in practice, many modellers resort to regeneration and recruitment models. As suitable data for modelling regeneration are often difficult to obtain, recruitment rather than regeneration models are often used to model forest renewal.

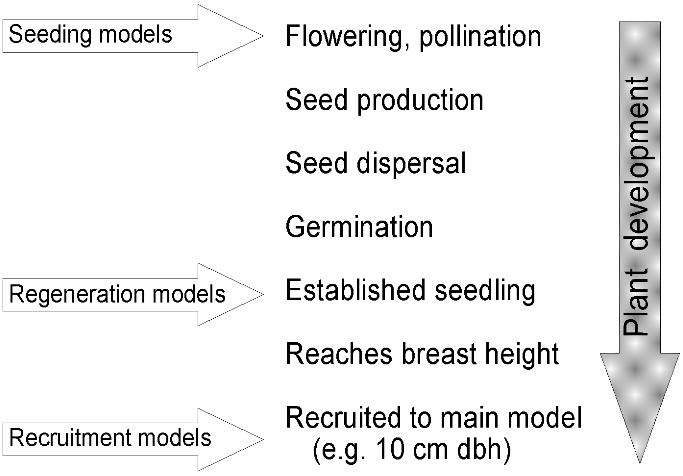


Fig. 10.1. Modelling of regeneration may commence at any of several stages during the development of an individual from the flower to the seedling and the small tree. Arrows indicate stages and convenient names for popular modelling approaches.

This chapter is concerned primarily with natural regeneration and recruitment in uneven-aged stands. Planting (and to a lesser extent, sowing) after clearfelling is relatively easy to model, and is not considered here. Regeneration within established plantations may be negligible, and many plantation growth models disregard any recruitment. Many models for uneven-aged forests also assume that recruitment is negligible or will not influence short term estimates of yield. This assumption is unsatisfactory for longer simulations of natural forests, as recruitment may contribute substantially to future stand basal area and thus influence growth and yield forecasts.

Recruitment Models

Recruitment models predict trees reaching a specified threshold size, usually based on height (e.g. breast height) or diameter (e.g. 10 cm dbh). It is convenient to distinguish two approaches:

1. *Static* approaches which take relatively little account of stand condition and thus predict a fairly constant amount of recruitment indicating the long-term average expectation under "typical" conditions, and
2. *Dynamic* approaches which respond to stand condition, predicting recruitment as a function of stand density, composition and other parameters.

Static Approaches

Some recruitment models assume that the amount of recruitment observed during the period of data collection reflects the long term average, and that this amount will not vary greatly during simulations. This assumption is common in many stand table and matrix approaches.

One common assumption is that the number of trees in the smallest class remains the same, irrespective of upgrowth to the next class, and of the total stand density. Size class models which smooth the stand table (p. 38) may use the smoothing function to estimate potential recruitment (e.g. by extrapolating the function to estimate stocking below the threshold size for the model). Other models assume a constant amount of recruitment each cycle, and assume that the mortality function will compensate for the errors that this over-simplification introduces. These assumptions are not realistic, and better approaches are available.

Many matrix models also make unsatisfactory assumptions. Positive fecundity coefficients in transition matrices (p. 46) predict an exponential increase in tree numbers, with the number of recruits increasing proportionally with the number of trees in the stand table. This is easy to implement in a matrix model, but is unrealistic and is unlikely to provide acceptable results. More satisfactory approaches are illustrated in Buongiorno and Michie's (1980) generalized matrix where recruitment was negatively correlated with stand density, and in Bosch's (1971) Leslie matrix which allowed regeneration to occur only after the death of another tree.

The special case of even-aged regeneration (natural or artificial) following clearfelling is often modelled by predicting the future stand structure at a nominated point in time. Alder's (1979) cohort model commenced with a prediction of stand structure when dominant height was seven metres. Belcher *et al.* (1982) predicted regeneration entering the STEMS projection system 15 years after clearfelling, using subjective decision trees based on site characteristics and details of the original stand.

Although these approaches are rather empirical, they may provide useful estimates of recruitment for stands that do not differ greatly from the source stands used for model development.

Dynamic Models

Where data permit, a better approach is to predict recruitment as a function of site and stand condition. Several such models have been published, and range from the highly empirical to those with a biological basis. Variables commonly used include site productivity, stand density (e.g. stand basal area, stem number, competition indices), and time since and nature of harvesting. Including details of the last harvest in the model may help to obtain a good fit to the data, but this may limit applications of the model, as these details may not be known for all stands. Where several species are involved, models may predict recruitment for each species group

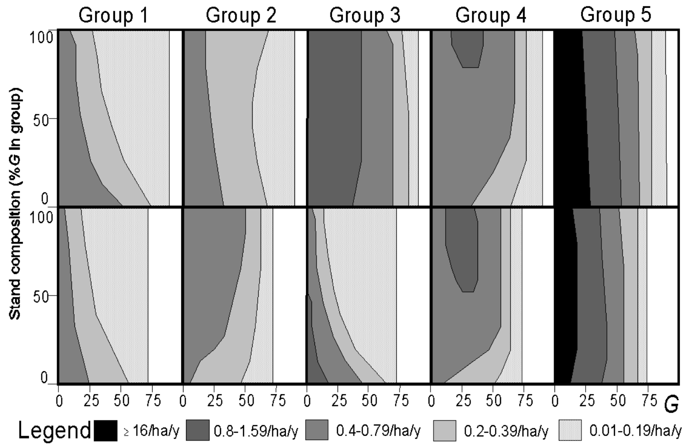


Fig. 10.2. Recruitment at 20 cm dbh predicted from site quality, stand composition and basal area (redrawn from Vanclay 1989*a*).

independently, or may predict the total and assign it to the candidate species.

Vanclay (1989*a*) predicted the total recruitment at 20 cm diameter in Queensland rainforests as a linear function of stand basal area and site quality:

$$N_r = 5.466 - 0.06469G + 1.013S$$

where N_r is number of recruits (trees $\text{ha}^{-1}\text{y}^{-1}$), G is stand basal area (m^2ha^{-1}) and S is site quality. The composition of this recruitment was determined by predicting the proportion in each of five species groups, and standardizing the proportions. The proportion for each species group was predicted from the stand basal area, the site quality and the basal area of that species group (e.g. for the large, fast-growing species):

$$P_1 = \left(1 + \exp\left[2.41 + 0.00561G - 0.0110G_1 - 0.00464G_1S\right]\right)^{-1}$$

where G_1 is the basal area of group 1 species ($\text{m}^2\,\text{ha}^{-1}$). These proportions were then standardized to ensure they summed to unity:

$$P_i' = \frac{P_i}{\sum P}$$

Although simplistic, this model allowed sufficient flexibility for species group dynamics to emerge (Fig. 10.2), and provided reasonable results.

The JABOWA model simulates stand dynamics on 10×10 metre plots. The original version of JABOWA (Botkin *et al.* 1972) predicted recruitment at 2 cm dbh, but the model was later revised to predict recruitment at breast height (Botkin 1993). A seed source is assumed to be available for each of the major species considered in JABOWA, and candidate species for the plot being modelled are selected from a master list according to the requirements of each species for light, warmth and moisture. For a given site where the temperature and soil moisture requirements are met, the composition and amount of recruitment is determined mainly by the light available at the forest floor. For example, if the plot's leaf area index (LAI) is less than a specified threshold, 60 to 75 cherry trees are recruited on the plot. If the LAI is between the first and second threshold, some (0–13) birches are recruited. If LAI exceeds both thresholds, a random choice of the remaining shade tolerant species is made, and a random number (0, 1 or 2) of each is recruited. In the current version of JABOWA, the nominal threshold for recruitment is breast height, but saplings are assigned a height selected randomly from the interval 137–167 cm.

Shugart and West (1977) followed a similar approach, but identified specific requirements for mineral soil or leaf litter, simulated weather and browsing stochastically, and allowed sprouting from dead trees. Trees were recruited at breast height. Similar succession models exist for subtropical rainforest in Australia (Shugart *et al.* 1980) and a tropical montane forest in central America (Doyle 1981).

Recruitment can only be defined relative to a threshold size, and the one threshold may not suit all applications equally. The smallest threshold consistent with permanent plot procedures may provide the best model, but operational inventories may adopt different standards and may use a larger threshold size for measurement of trees. Data "censorship" occurs when the threshold in the model is less than the minimum diameter recorded in operational inventory, and degrades model performance. One solution is to augment the censored data with an "average" stem distribution for the forest type (Randall *et al.* 1988); this is preferable to using the unadjusted data. An alternative is to make the recruitment model more flexible, so that it can predict recruitment at any threshold.

Shifley *et al.* (1993) devised a model to predict recruitment of six tree species in the North Central USA, at any threshold in the range 1–13 inches (*c*. 3–33 cm). They argued the following:

1. Stands tend to move toward full site occupancy.
2. Stands that are understocked have potential space for new recruits.
3. Resources available to support ingrowth decrease as stocking increases, so recruitment should decrease as stand density increases.

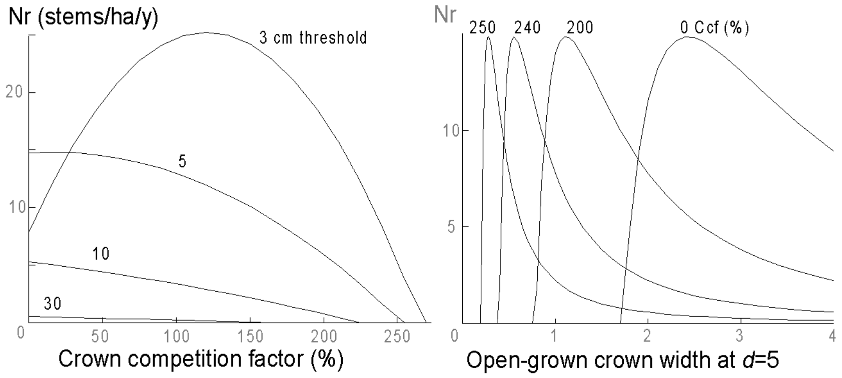


Fig. 10.3. Recruitment estimated for different threshold diameters (left, drawn from Equation 10.1). The model is very sensitive to estimates of open-grown crown width, and a 10% perturbation may alter recruitment estimates by 100% (right).

They assumed that site occupancy is revealed by crown competition factor:

$$C_{cf} = \frac{\pi}{400}\sum c_{wo}^2 = \frac{\pi}{400}\sum \left(\beta_0 + \beta_1 d^{\beta_2}\right)^2$$

where c_{wo} is open-grown crown width and the βs are species-specific, and that recruitment could potentially be equal to the number of trees of threshold size need to lift the crown competition factor to the maximum for the site:

$$N_{max} = \frac{C_{max} - C_{cf}}{c_{wr}^2 \frac{\pi}{400}}$$

where c_{wr} is the open-grown crown width of a tree at the recruitment threshold. The maximum crown competition factor appeared to be constant across several sites (except for elm-ash-cottonwood forests), but is dependent on the threshold diameter d_r (cm):

$$C_{max} = 287 - 7.28\, d_r + 0.109\, d_r^2$$

The actual recruitment was then predicted from the potential (e.g. for shortleaf pine, as shown in Fig. 10.3):

$$N_r = 5.42\times10^{-3} N_{max} - 9.4\times10^{-8} d_r N_{max}^2 - 1.0\times10^{-9} d_r^2 N_{max}^2 \tag{10.1}$$

The low precision of the model reflects the highly variable nature of recruitment, but Shifley *et al.* (1993) felt that the model was well-behaved from a biological perspective, and saw few options for further improvement without substantially increasing model complexity or reducing its geographic applicability. However, the model is very sensitive to estimates of open-grown crown width at the threshold diameter (Fig. 10.3), and this may limit the utility of the model, as it may be difficult to obtain reliable estimates of open-grown diameter for many natural forest trees. The analogue based on stand basal area (rather than crown competition factor) may be more robust.

Two-stage Approaches

One of the difficulties in modelling recruitment is the great variability in regeneration. Stand condition accounts for some of this variation, periodicity of mast years and prevailing climate accounts for some, but regeneration remains a stochastic process, providing difficulties for efficient model estimation.

Much of the variability seen in recruitment data reflects the fact that during any period regeneration may or may not occur. This characteristic (no/some recruitment) becomes especially apparent when individual species or species groups are modelled. Such data may be modelled using a two-state approach, which helps to satisfy the usual statistical assumptions (p. 119). With this approach, we first predict the probability that some recruitment will occur, usually with a logistic function with presence (or absence) of recruitment as the response variable (cf. mortality functions, p. 178). Then a conditional function can be used to predict the amount of recruitment, given that some is known to occur. This conditional function can be estimated using ordinary linear regression. Hamilton and Brickell (1983) used this approach to predict defective volume in standing trees, and the method applies equally well to modelling recruitment.

The Prognosis model uses a two-stage approach to predict recruitment (Ferguson *et al.* 1986, Ferguson and Carlson 1993). It predicts regeneration stochastically, assuming 50 subplots each 1/300 acre (about 0.001 ha), and add these stems into the main Prognosis model at 10 and 20 years after disturbance (the Prognosis model uses a 10-year simulation cycle). The probability that some regeneration occurs is predicted from environmental variables (habitat, slope, aspect, elevation), distance to seed source, residual basal area and time since disturbance. Given that regeneration is known to occur, the expected number of trees is chosen with random numbers, and determines the number of cohorts for that subplot. The number of species present (1–6 species), and the identity of these species is also stochastically determined. Stochastic perturbations are also added to predicted heights of recruits. Consistent with field procedures, they discriminated between "best" trees and "excess recruitment", using smaller expansion factors to model the former category.

The NORM model uses a two-stage approach (Fig. 10.4) to predict regeneration occurring for each of the 100 species which contributed 97% of all recruitment (at 10 cm dbh) recorded in 217 permanent plots. The remaining 3% of recruitment comprised 113 species, and contributed insufficient data for meaningful analyses of regeneration characteristics.

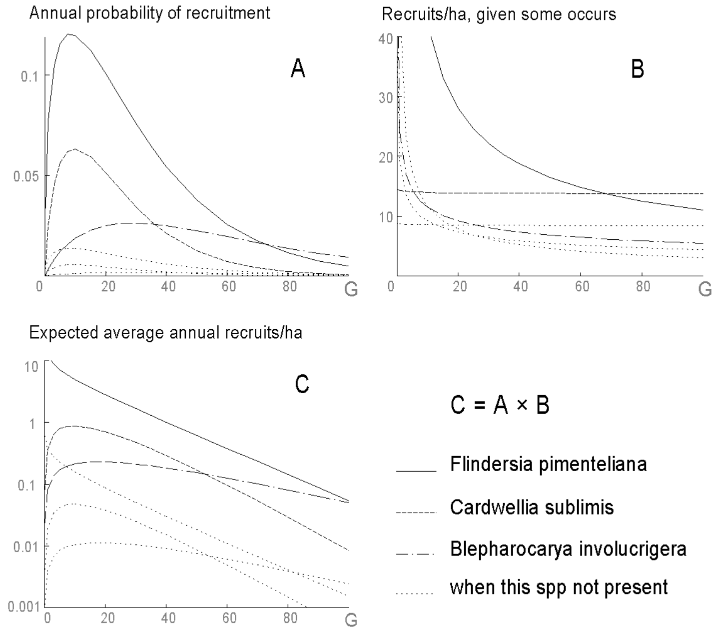


Fig. 10.4. Recruitment predicted from the NORM model for typical conditions ($N_r/N=0.2$, $S_{\Delta d}=7$, $S_v=0$, $R_t=0$; drawn from Eqns 10.2–10.3).

There is no easy way to decide how many species to model recruitment for; the law of diminishing returns seems to apply. In north Queensland rainforests, 60 species contributed 90% of all recruitment, 80 species accounted for 95%, and 100 species brought the total to 97% of all recruitment observed (Vanclay 1992). The optimal number of species to model depends very much on the data and resources available, and on the intended uses of the model.

The probability that any recruitment occurred is predicted separately for each species, in contrast to the approach used in the Prognosis model (Ferguson *et al.* 1986) where the collective probability of regeneration is modelled. Five species groups are used to estimate annual probabilities of recruitment from stand basal area, the presence of the species in the existing stand, and years since the last disturbance (e.g. for *Flindersia pimenteliana*; Fig. 10.4, A):

$$P_{r,i} = \left(1 + \exp\left[4.73 - 2.29 Z_i + 0.0470 G - 0.396 \ln G - 0.207 S_v - 0.219\, t_s e^{0.11 t_s}\right]\right)^{-1} \tag{10.2}$$

where Z_i is a binary variable indicating the presence of species *i* in the stand (1 = present, 0 = absent), G is stand basal area ($m^2\ ha^{-1}$), t is years since silvicultural treatment, and S_v is a binary variable which takes the value 1 on soils derived from basic volcanic and coarse granite parent materials, and 0 elsewhere. This function can be employed in a stochastic or deterministic manner. The deterministic implementation sums the annual probabilities for each species until the cumulative probability exceeds unity.

Species were grouped in another fashion to analyze the amount of regeneration, given that it is known to occur, and eight species groups were used to estimate the final function. Recruitment is predicted from stand basal area, site productivity and the relative abundance of the species in the stand (e.g. for *Flindersia pimenteliana*; Fig. 10.4, B):

$$\ln N_{r,i} = 6.06 - 0.580 \ln G + 1.84 \ln\left(\frac{N_i}{N} + 0.2\right) + 0.0997 S_{\Delta d} - 0.317 S_a \qquad (10.3)$$

where N_i is the number of stems of species *i* in the stand, $S_{\Delta d}$ is an index of site productivity, and S_a is a binary variable that takes the value 1 on alluvial soils and fine-grained granite soils, and 0 elsewhere. Figure 10.4 illustrates these two functions, and shows the expected annual recruitment for selected species. There is no biological significance in the apparent asymptote in Fig. 10.4, B; it arises because many of the plots were 0.2 ha (range 0.04–0.5 ha), and thus a single observed recruit would contribute 5 stems ha^{-1}. The predicted number of recruits may seem high, especially considering that 100 species are modelled in this way, but the expected average annual recruitment is low; so low for some species that it is useful to display the expected recruitment on a logarithmic scale (Fig. 10.4, C).

The two-stage approach introduces some complications when the data represent different time intervals. Longer intervals will have a greater probability of recruitment, and there may be more recruitment if it occurs (the observed recruitment may represent more than occurrence). Thus it may be necessary to adjust for the time interval in both functions, to obtain a satisfactory fit to the data. However, this double correction may bias predictions, so it is critical to compare predictions from the two equations with the raw data, and make adjust predictions if necessary. It is a good idea to exclude data drawn from plots with very long measurement intervals (e.g. choose the maximum remeasure interval so that there will be relatively few data with two recruitment events within a single datum).

The use of variables such as presence and relative number (N_i/N) of a tree species raises some complications. In medium- to long-term simulations, the predicted composition of the stand may change. The question then is, should these equations employ the initial presence or the simulated presence? There is no easy answer. The composition of natural stands may vary considerably over time. Flushes of pioneer species appearing after disturbances may be short-lived, and the turnover rate of species may be high even in undisturbed stands. However, recruitment predictions cannot always be correct, and using the simulated composition may accelerate bias. The NORM model uses a compromise: it uses the original relative abundance, but only if the species remains in the simulated stand. However, all three options (original, simulated, compromise) are available in the model, and users are encouraged to try all three and explore the implications. Similarly, users may elect to invoke deterministic or stochastic simulation, and are encouraged to investigate both options in critical applications.

Regeneration Models

Recruitment models begin rather late in the process of stand renewal (e.g. Fig. 10.1), and it is not always clear what assumptions regarding earlier development are implicit in predictions. This limitation can be partially overcome by commencing simulation at an earlier stage of seedling development. *Regeneration models* are those that simulate the development of trees from seed or seedlings, and thus attempt to take into account more of the factors influencing the process of stand renewal. Some regeneration models recruit seedlings directly into the main model, but most simulate the growth of trees from seedlings to breast height within a separate regeneration model (e.g. Ek and Monserud 1974). This approach is sufficiently flexible that almost any size may be used as the criterion for recruitment into the main stand.

An advantage of the approach is that it more realistically models the time taken for regeneration to be recruited following a reduction in stand density due to harvesting. Recruitment models which employ an expression of stand density may lead to overestimates of recruitment in the projection period immediately following the harvest, unless there are large reserves of advance growth. Modelling may start at any of several stages. Leak (1968) modelled regeneration from the stage of flower development, and Ek and Monserud (1974) began with seed fall. Germination could provide a suitable starting point, and many models started with "establishment" when the seedling has survived its first year after germination.

Growth and development of the regeneration can be modelled in several ways. Leslie matrices (p. 46) or "life tables" have been used (e.g. Hett and Loucks 1968), but one problem is that age is not a good predictor of growth in the understorey. Tree height is a better predictor, and a better approach may be to use height cohorts.

Vanclay (1988*a*) predicted the amount of established one-year-old regeneration in *Callitris* forest from stand basal area and site productivity:

$$\ln(N_r+1) = 0.497\,S_{h,d} - 0.01723\,S_{h,d}\,G - 4.45$$

where N_r is the number (ha^{-1}) of one year old seedlings established, $S_{h,d}$ is site quality and G is stand basal area (m^2 ha^{-1}). Seedlings were modelled using cohorts representing height classes until they reached breast height, when they were recruited to the main model. A maximum of ten cohorts was imposed. Under ideal conditions (good sites with low stocking), these cohorts represented annual flushes of regeneration. Where growth of regeneration was slower and took more than ten years to reach breast height, the most similar cohorts were amalgamated to ensure that the limit of ten cohorts was not exceeded.

Ek and Monserud (1974) adopted a more sophisticated approach to predict recruitment into their stochastic spatial single-tree model. The regeneration model used cohorts representing the number of stems for each species and age in each subplot within the main plot being simulated. Despite having seed-crop records over a 26-year period, they were unable to fit a model to describe seed production, and used random numbers to select good, moderate and poor seed years according to the observed frequency for each species. Seed and sprout production were estimated for each overstorey tree as a function of its size and the threshold age, and were distributed across the subplots according to the parent tree's position, height and crown width. Germination was predicted as a stochastic function of microsite and canopy cover conditions. Each year, a germinant or tree in the understorey may die, or survive and grow in height by an amount predicted from cover, species and age. When trees reached breast height they were recruited into the main model. If trees did not attain this height within a specified time (e.g. 25 years for black spruce), they "died".

Monserud and Ek (1977) refined this model, improving the efficiency by reducing the number of cohorts to be modelled. They assumed that understorey tree size was more relevant than tree age, and modelled the development of trees to 7.6 metres height using five height cohorts of varying size and movement ratios (p. 36). The height increment of the mean tree was predicted from the potential height increment (a function of height and site), overstorey competition index (a relative size-distance index), shade tolerance (a function of species and height), and stand density. The model gave good predictions, even when extrapolated to clearfall conditions which were not represented in the database (Ek and Monserud 1979). Detailed approaches such as this may not be warranted in all yield studies, but may be relevant in models used to analyze silvicultural alternatives for intensively managed stands.

Synthesis

Although regeneration models offer several desirable features, they may be impractical in many natural forests because of difficulties with species identification in small trees and an absence of suitable data. Many shade tolerant species may exist as advance growth for decades, until a gap appears in the canopy and provides an opportunity for these stems to grow into the overstorey. If such advance growth contributes substantially to recruitment in the forest and with the management system under consideration, a regeneration model may confer any advantage over simpler recruitment models. As in all model components, there is no single best strategy for all occasions, and the most suitable approach to use depends on the situation.

In the next chapter we will consider procedures and tests to evaluate the performance and suitability of models for various applications. This is an important aspect of model construction, and should not be a mere afterthought.

Exercises

10.1. Contrast the various methods of predicting regeneration and recruitment and discuss their strengths and weaknesses. What method would be most suitable for your application? Why? What data would you need to fit the model and how would you implement it within an overall growth modelling framework (e.g. your solution to Exercise 4.4)?

10.2. Construct an Usher matrix model (p. 46) which predicts recruitment only on the death of another tree. State the assumptions you make regarding growth and mortality rates, as well as any implications regarding regeneration. How would you test this model?

10.3. The (synthesized) data below concern recruitment at 10 cm dbh in a series of plots remeasured at 5-year intervals in a tropical forest. Build a model to predict recruitment for this forest. Discuss the strengths and weaknesses of your method. How could you test this model, and what data would you need?

Plots without recruits				Plots with recruits			
Basal area (G, m^2 ha^{-1})				G	N_r	G	N_r
2.2	10.2	27.5	40.7	2.5	81	9.1	63
2.6	15.2	28.1	41.8	5.0	87	22.2	22
4.3	16.4	30.2	45.8	5.2	72	22.4	20
5.3	17.2	30.9	56.9	6.3	91	26.1	23
6.1	18.3	33.4	65.0	6.3	62	28.5	23
6.6	22.7	36.5	65.6	7.1	41	31.3	30
9.7	22.8	40.4	72.9	8.7	44	44.7	25

Chapter Eleven

Model Evaluation and Re-calibration

Model evaluation is an important part of model building, and some examination of the model should be made at all stages of model design, fitting and implementation. Evaluation should not merely be an afterthought or an acceptance trial. A thorough evaluation of a model involves several steps, including two which are often called *verification* and *validation*. In forest growth modelling, these usually denote *qualitative* and *quantitative* tests of the model, respectively. There are several objections to these terms:

1. They are value-loaded, and it is preferable to use neutral language to assess model performance (see e.g. Oreskes *et al.* 1994).
2. The same terms are used in other branches of mathematics and logic to denote other meanings: a model is valid if it is logically correct, and verified if it is true.
3. Verity implies truth, but it is impossible to prove a model "true" (except in the special case of a closed system; e.g. Oreskes *et al.* 1994). The only truth that can be established in a growth model is in the context of Goulding (1979), namely that the model is a faithful representation of what the modeller intended.

Thus it is appropriate to avoid these terms, and to use alternatives such as *model criticism* and *benchmarking* respectively. The former should involve examination of the structure and properties of model properties, with or without supplementary data, to confirm that it has no internal inconsistencies and is biologically realistic. Benchmarking requires comparisons with data to quantify the performance of the model. Thorough benchmark tests should involve data which are in some sense unlike the data used to fit the model (Snee 1977), but useful insights can also be obtained with the data used to fit the model. These tests cannot prove a model to be "correct", but may be used to *falsify* inferences made from the model. The quality of a model can only be evaluated in relative terms, and its predictive ability always remains open to question. However, the failure of several attempts to falsify a model should increase its credibility and build user confidence. This is the role of model evaluation.

Re-calibration refers to the search for adjustments to improve model predictions for a specific locality. It relates to model evaluation, because if benchmark tests reveal deficiencies in a model, the question is if the same data may be used to re-calibrate the model so that predictions are improved for that locality, and if so, how. Both these activities require an operational growth model, but should not be seen as "optional extras". They remain an integral part of model development.

Model Criticism

Decision makers and managers may rely on predictions from growth models to examine implications and consequences of forest policy options and management strategies. Unreliable predictions may lead to suboptimal decisions, so models should be examined and their limitations ascertained before they are used. Model evaluation should be convincing enough to boost user confidence so that the model will be used as a basis for action. If the evaluation reveals faults, the question then is where and how the model fails, and what can be done to improve it. The quality of a model thus depends on the application and objectives.

Model evaluation should reveal any errors and deficiencies in the model, and should establish:

1. whether the equations used adequately represent the processes involved,
2. if the equations have been combined in the model correctly,
3. that the numerical constants obtained in fitting the model are the best estimates (unbiased minimum-variance estimators),
4. the range of site and stand conditions over which the model applies,
5. if the model satisfies specified accuracy requirements,
6. whether the model provides realistic predictions throughout this range, and
7. how sensitive model predictions are to errors in estimated coefficients and input variables.

Some of these criteria can be examined at several stages during model design and construction, but some aspects require additional data for benchmark tests. In this section on model criticism, we consider tests of a model that can be completed without additional data, while aspects of benchmarking will be considered later in this chapter.

Evaluation Procedures[1]

"Good science" or strong inference requires that scientists make imaginative conjectures (such as models), deduce consequences amenable to testing, and attempt to disprove the conjecture experimentally (e.g. Popper 1958). A conjecture which survives repeated attempts to falsify it is corroborated, or not yet falsified, and may begin to assume the status of an axiom. Tests of model implications and consequences may involve analytical investigation of component equations, or comparisons with empirical data.

Rigorous testing of a growth model may require data drawn from permanent plots not used in the development of the model, and may be more convincing if these are derived from a different population (e.g. another forest of the same type, a different series of experiments, or data collected by another agency). Repeated testing until a growth model fails does not mean that the growth model will be rejected. Rather, it helps to define the region within which the model behaves satisfactorily, and indicates the areas where further research is warranted.

Forest growth models may comprise many separate but interrelated components, each of which may influence, and be influenced by other components and assumptions in the model. Model evaluation should extend to all model components and assumptions, and this requires a thorough understanding of the structure of the model and the interrelationships between components. A thorough evaluation should include the following aspects (Soares *et al.* 1994):

1. Examine the model and its components for logical consistency and biological realism (e.g. Liu *et al.* 1989, Oderwald and Hans 1993) to ensure that they
   (a) are parsimonious and are biologically realistic,
   (b) agree with existing theories of forest growth, and
   (c) predict sensible responses to management actions.

2. Ascertain the statistical properties of the model in relation to data (e.g. Ratkowsky 1983, Weisberg 1985), including the:
   (a) nature of the error term (i.e. additive or multiplicative, independence, etc.), and the
   (b) estimation properties of parameters in model functions.

3. Characterize errors (e.g. Reynolds and Chung 1986, Gregoire and Reynolds 1988) in terms of:
   (a) magnitude (including confidence intervals and critical errors),
   (b) residuals (distribution, dependencies on initial conditions and projection length), and

---

[1] This section on model criticism draws heavily from the paper by Soares *et al.* (1994), and owes much to fertile discussions with J.P. Skovsgaard of the Danish Forest and Landscape Research Institute.

  - **(c)** contributions by each model component to total error.

4. Test, using statistical approaches (e.g. Reynolds *et al.* 1988, Mayer and Butler 1993) for:
  - **(a)** bias and precision of the model and its components,
  - **(b)** goodness-of-fit of predicted size distributions,
  - **(c)** patterns in, and distribution of residuals, and
  - **(d)** correlations over time and between components.

5. Identify model components with the greatest influence on predictions, by
  - **(a)** analytical and simulation studies of model sensitivity to perturbations in parameter estimates (e.g. Liu *et al.* 1989, Elston 1992), and
  - **(b)** analytical and simulation studies of the propagation of errors in model inputs (e.g. Gertner 1987*a*, Mowrer 1991).

These analyses need not be sequential, but all relevant aspects should be examined in each model component and in the assembled model. Each of these steps should involve both graphical analyses as well as statistical indices. The statistical properties of models have been dealt with in Chapter 6 (p. 119), and are not considered any further in this chapter, but this does not diminish their importance.

Logical and Biological Consistency

Each model component and the model as a whole should be logically consistent and biologically realistic. Many model properties can be examined for consistency. Some aspects include (Oderwald and Hans 1993):

1. variables included in, and omitted from the model should agree with expectations,
2. signs and values of coefficients should agree with expectations,
3. extrapolations outside the range of the development data should be reasonable,
4. transformations of model predictions should be reasonable (e.g. model forecasts of future diameters should also provide reasonable estimates of diameter increments, future volumes, mean increment curves, etc.),
5. contradictions should not be evident in a model, and
6. derivatives, limits, maxima, minima, inflections, etc. should agree with expectations.

Matrix plots of simulated stand development trajectories showing a range of property-time and property-property relationships may offer useful insights into model behaviour, and may provide an efficient way to reveal discrepancies in model predictions (Leary 1988).

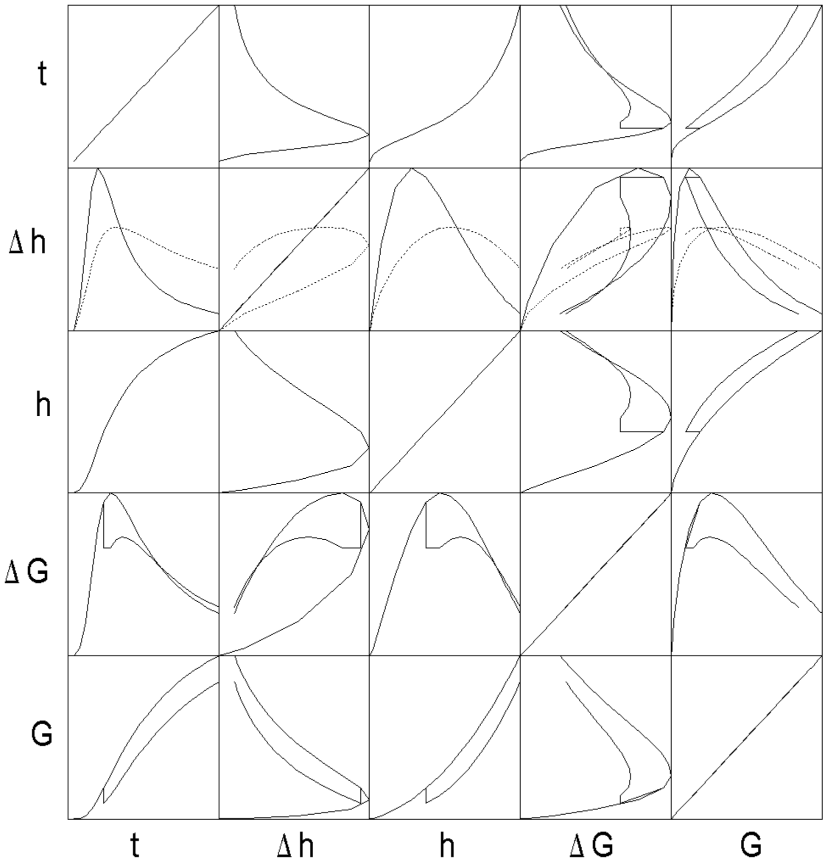


Fig. 11.1. Matrix plot of model defined by Equations 2.13 and 2.14, showing age, height increment, height, basal area increment and stand basal area. Unthinned and thinned (age 12) trajectories are shown. The dotted line shows mean annual height growth.

Figure 11.1 illustrates such a matrix plot for the simple model defined by Equations 2.13 and 2.14 (p. 28), showing two simulated stand development trajectories, one for an unthinned stand, the other for the same stand thinned when top height reached 10 m (age 12). The dotted line in the second row of the matrix indicates the mean annual height increment (it is omitted from other cells for clarity). The matrix plot reveals a lot of information and gives a good insight into model behaviour, but may become very large. Here, a simple model for a pure, even-aged plantatio, with just two equations and three variables produces a 5×5 matrix with a lot of information to digest (and it could be expanded to 7×7 by including mean annual increments in height and basal area). Because of the high information content of these matrices, it is a good idea to examine critical parts of the model in more detail, by creating subsets of the matrix with inter-relationships of particular interest. In this particular illustration, one might choose to examine the thinning response more closely by experimenting with different thinning intensities and timings. Further criticism of this model is left to the reader as an exercise (11.1).

Figure 11.1 illustrates a very simple model of the most simple forest, the pure even-aged stand. Figures like this can also be compiled for uneven-aged mixed forests, but the number of cells that could potentially be drawn in the matrix becomes very large (e.g. tree growth characteristics for different species-site-competition permutations, and stand growth characteristics for stands comprising different species and size structures), and it is expedient to choose a few graphs for close scrutiny. There are a few principles that may provide a good basis for criticising models for pure, even-aged stands (e.g. Eichhorn's rule, Langsaeter's hypothesis, etc.; see Assmann 1961, Wenk *et al.* 1990), but few such guides exist specifically for mixed, uneven-aged forests.

Parameter estimates and model forecasts should agree with both empirical data and current understanding of growth processes. Experienced foresters and other experts may indicate areas where model predictions are deficient. Simulations at extremes stand conditions may be particularly revealing. Monserud (1989) suggested that optimization studies provided a discriminating test of a model, as his optimizer seemed to be remarkably efficient at exploiting seemingly minor quirks in the Prognosis model to arrive at unrealistic solutions. Thus optimization studies coupled with expert insights may provide a good basis for model criticism. However, a model should not be rejected simply because it behaves in a counter-intuitive fashion; it may be our preconceptions that are wrong. Thus discrepancies should cause a critical reappraisal of the model, the data, and of preconceptions.

Sensitivity Analyses

A sensitivity analysis attempts to reveal model parameters and submodels which, when perturbed, cause the greatest fluctuations in model predictions. These studies may reveal model components with low and high sensitivity, both of which are of interest. Insensitive components may contribute little toward model predictions and could be targets for omission from the model during model revisions. Conversely, it is useful to know about model components with high sensitivity, because these may have the greatest impact on model predictions. All model parameters and inputs should be estimated accurately, but particular care is required with the most sensitive variables.

In theory, sensitivity studies can be done analytically (e.g. by taking derivatives), but in practice this may be complicated by the interaction of various model components and feedback loops. Thus sensitivity analyses are often carried out as simulation studies in which the parameters or components are changed to observe corresponding effect on predicted outputs ($\hat{y}$):

$$S_\beta = \frac{\partial\hat{y}/\hat{y}}{\partial\hat{\beta}/\hat{\beta}} = \frac{\partial\hat{y}}{\partial\hat{\beta}}\frac{\hat{\beta}}{\hat{y}}$$

where β is a parameter estimate (e.g. if a 10% change in a parameter β results in a 50% change in the predicted outcome ŷ, the sensitivity S_β is 5). In practice, meaningful sensitivity studies are difficult, as the estimate of sensitivity S_β may depend on the values of both ŷ and β, so that many simulations may be necessary to complete the picture. This may be a tedious undertaking, especially where there are many parameters.

The parameter estimates examined in sensitivity analyses are usually those estimated by regression (i.e. the βs), but other aspects of the model may be examined analytically or by simulation. Liu *et al.* (1989) examined several other aspects of a yield function for loblolly pine, including

(a) factor interdependence $\partial^2 Y/\partial X_i \partial X_j$, which indicates if the explanatory variables are complementary (> 0) or competitive (< 0),

(b) elasticity of output, $(\partial Y/Y)/(\partial X_i/X_i)$, which reveals how Y changes with respect to X_i when all other Xs are held constant, and

(c) elasticity of substitution $(\partial X_i/X_i)/(\partial X_j/X_j)$, which indicates how readily one input may be substituted for another.

These criteria provide useful insights into model behaviour, but may be complex to explore fully.

Studies of error propagation (Gertner 1987*a*, Mowrer 1991) may reveal model limitations, and are particularly useful in offering insights into the interaction of errors in the input data and in the simulation. One application of stochastic simulation studies is to investigate the "quality" of predictions. Variance approximation provides an efficient alternative to such studies, and enables the variance of predictions to be estimated deterministically. It also enables the variance of the input data to be incorporated into the analysis. Mowrer and Frayer (1986) and Gertner (1987*a*) used a simple first-order Taylor series to estimate the errors propagated through growth and yield projections. The general formula for error propagation is (Mowrer 1989):

$$\hat{Var}(\hat{Y}) = \sum_i^n \sum_j^n \left\{ \frac{d\hat{Y}}{dX_i}\frac{d\hat{Y}}{dX_j} S_{X_i} S_{X_j} R_{X_i X_j} + \frac{d\hat{Y}}{d\beta_i}\frac{d\hat{Y}}{d\beta_j} S_{\beta_i} S_{\beta_j} R_{\beta_i \beta_j} \right\}$$

Terms involving the variable X estimate the variance in the estimated response variable (Ŷ) from propagated errors (S_X) in the predictor variables (X) and their cross-products, with correlation coefficients R_X. Terms involving the variable β approximate fixed variance contributions (S_β) from the estimated regressor coefficients (β) with correlation coefficients R_β. These contributions may be obtained from the covariance matrix of the estimated regression coefficients, and reflect the quality of the various estimators used in the model. Equations such as these can readily be incorporated into computer implementation of models and can provide concurrent estimates of the variances associated with each variable at any time during the simulation.

Growth models may contain many assumptions which are not supported by data (e.g. simulated plot size, maximum number of tree records, record doubling or tripling, attenuation of re-calibration factor, etc.), which may influence predicted outcomes, and these assumptions may not be easily confirmed. If benchmark tests reveal that the model gives acceptable predictions, then these assumptions are probably reasonable. However, it remains important for the modeller to know if any of these assumptions are critical to model preformance, and to be aware of the consequences of modifying any of these parameters. Such sensitivity tests are especially critical if the parameters can be modified by users. Users of a model may wish to vary some parameters (e.g. to improve computational efficiency during complex calculations), so model testing should include a demonstration that predictions are not unduly influenced by these assumptions.

Results of sensitivity tests may reveal parameters critical to model predictions, and parameters which may be redundant. Knowledge of sensitive parameters may guide applications (especially extrapolations) and the planning of model enhancements. Clearly, sensitivity analysis should be an integral part of model design and evaluation.

Benchmark Tests

Model criticism and benchmarking are two aspects of a range of tests that should be undertaken to demonstrate the utility and limitations of a model. The distinction between these phases is based, in part, on the data required to make the tests. Model criticism can be completed without additional data, whereas benchmark tests require data (preferably independent data) to compare model predictions with observed values. This raises some additional questions about the nature and amount of data used for such comparisons.

Characterizing Model Error

One of the most efficient ways to examine model performance is to plot residuals for all possible combinations of tree and stand variables, and to look for patterns which may indicate serial correlation, dependencies on initial conditions or on projection length, or other systematic patterns (e.g. Soares *et al*. 1994). Such plots may be interpreted visually, but formal tests are also available (e.g. Draper and Smith 1981, Weisberg 1985). The common practice of plotting both the observed and predicted values against time is not recommended as it is rarely particularly revealing (e.g. Mayer and Butler 1993). It is common to plot observed values (y) against predicted ($\hat{y}$) values, but in many cases it is more revealing to plot residuals ($e = y - \hat{y}$) versus observed values (Fig. 11.2). Notice that plots of e versus y and e versus $\hat{y}$ are analogous, but the latter involves a re-arrangement of the

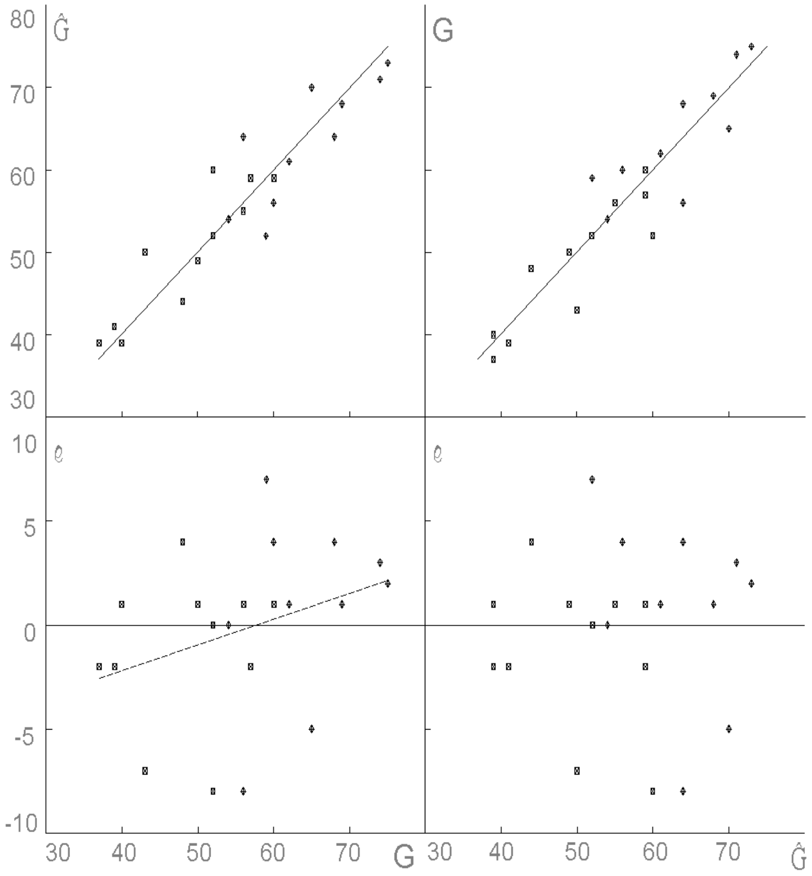


Fig. 11.2. Observed (G) and predicted ($\hat{G}$) values plotted directly and as differences ($e=G-\hat{G}$). Symbols indicate $G_0<50$ (#) and $G_0>50$ (♦) (see Ex. 11.3).

residuals (e) which may conceal some trends. In plots like this, it is often informative to stratify the data and use a different symbol for each stratum.

Errors may depend on projection length or initial forest condition, and such dependencies can be shown graphically (Fig. 11.3). Many graphs may be required to screen all candidate dependencies, and preliminary screening can be done efficiently using stepwise linear regression to reveal variables for which scatterplots should be prepared. Time should be examined in three contexts: (i) stand age or time since disturbance (e.g. logging; if known), (ii) length of projection, and (iii) year of measure. These three aspects of time may reveal different properties of the model.

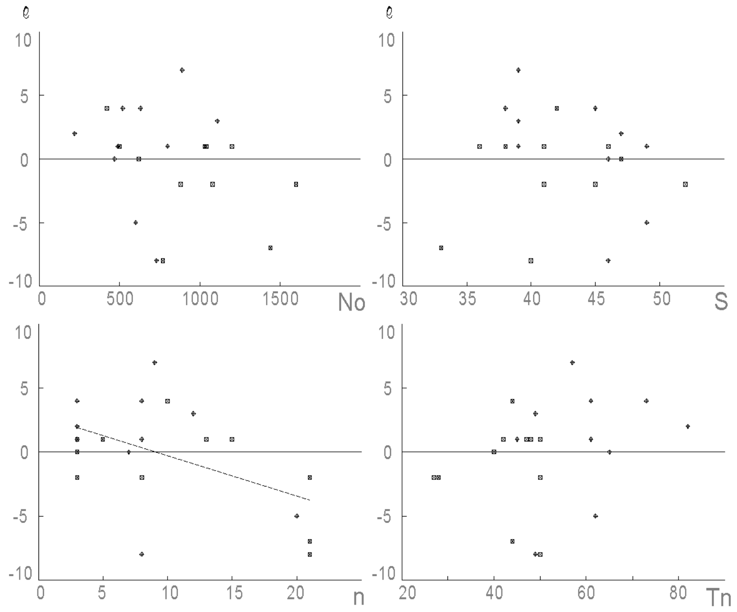


Fig. 11.3. Residuals vs initial values (stem number N_0 & site index S), length of (n), and age after simulation (T_n). Symbols indicate initial basal areas and dotted line shows trend.

Two simple criteria, when used in conjunction, provide a useful summary of the overall model performance: average model bias ($\sum(\hat{y}_i - y_i)/n$) and mean absolute difference ($\sum|\hat{y}_i - y_i|/n$). These formulae apply at the stand level and for size or product breakdowns. Average model bias is a measure of the expected error when several observations are to be combined by totalling or averaging, while the mean absolute difference indicates the average error associated with a single prediction. These statistics may be weighted to reflect the discrepancy in terms of stand basal area or volume. It may also be useful to express these as percentages (e.g. $100[\sum|\hat{y}_i - y_i|/|y_i|]/n$), but if some observed values are very small ($y_i \rightarrow 0$), it is preferable to use the alternative $100[\sum|\hat{y}_i - y_i|/n]/\bar{y}$.

Another useful technique is to compare predictions directly with observed data using a statistic analogous to R^2, and sometimes called modelling efficiency:

$$E_m = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

This statistic provides a simple index of performance on a relative scale, where one indicates a "perfect" fit, zero reveals that the model is no better than a simple average, and negative values indicate a poor model indeed. However, this statistic suffers all the weaknesses of R^2 (see p. 125), and a high "efficiency" does not necessarily mean an adequate model.

Reynolds (1984) pointed out that two different philosophies of model evaluation can be identified. One approach uses statistical hypothesis testing and the other uses statistical estimation. Hypothesis tests are appropriate when the question is whether a model behaves like the real system or whether a model meets a specified accuracy requirement of the user. In other situations, however, there is no particular standard for comparison and the objective is simply to give the user of the model some estimate of how far predictions from the model will be from actual values. Reynolds (1984) presented formulae to calculate critical errors which users could use to judge if a model prediction would meet their requirements. Unfortunately, these are sensitive to the assumption of normality, and are open to misinterpretation.

However, prediction intervals can be calculated even when the assumption of normality is not reasonable. The usual formula for the $100(1-\alpha)$ percent prediction interval for the mean of k predictions is:

$$\bar{e} \pm t_{1-\alpha/2,\, n-1} \sqrt{\left(\frac{1}{n} + \frac{1}{k}\right) \frac{\sum (e_i - \bar{e})^2}{n-1}}$$

where e is the difference between predictions and observed values in the benchmark data, and n is the number of benchmark data. Thus a model user who makes forecasts for k stands could expect that the mean of these forecasts would be within this range of the true value with probability $100(1-\alpha)\%$.

The non-parametric equivalent should be used if the assumption of normality is not reasonable. It indicates the probability that all of the next k predictions will lie between the smallest and largest values in the benchmark data:

$$1 - \alpha = \frac{n(n-1)}{(n+k)(n+k+1)}$$

These formulae are somewhat limited because they only apply to the projection lengths available in the benchmarking data. Reynolds and Chung (1986) illustrate how regression analyses may be used to formulate equations for expected bias and prediction intervals for predictions of different lengths.

The error structure and the contribution of each model component to total error may be more revealing than a mere evaluation of total model performance. Thus an error budget of variance components of the model may help to identify weaknesses and define priorities for future research. If suitable data are available, error budgets may be compiled by successive simulations using predicted values for only one component and observed data for all other model components (e.g. Hann 1980, Soares *et al.* 1994). Hann (1980) used several simulation cycles to identify faulty components in his model. His first cycle simulated only increments, and used actual mortality, harvesting and recruitment and thus resolved that increment prediction was satisfactory. The second cycle predicted increments and mortality, and employed actual harvesting and recruitment, and so forth, until the final test included all the predicted values.

Several researchers have advocated Turing tests in which experts are asked to discriminate between simulated and real world data, but this is not a good basis for comparison. If the real and simulated data are sufficiently alike to offer a realistic test, they should be amenable to statistical testing which avoids the difficulties with personal bias. Conversely, if the data are unsuited to statistical testing, it is likely that they will contain certain identifiable features which may make the distinction easy. Turing-type tests are not recommended.

Statistical tests

Many statistical tests of model performance have been suggested, but no single criterion can incorporate all aspects of model evaluation, and it is desirable to use several simple tests to examine different facets of model behaviour.

One simple but efficient technique is based on linear regression of residuals on observed data. Some useful insights into the quality of predictions may be given by the slope of the fitted line, and a good test for bias is the simultaneous F-test for zero intercept and slope:

$$F_{2,n} = \frac{\frac{1}{2}\left[\sum e^2 - \sum \varepsilon^2\right]}{\frac{1}{n}\sum e^2}$$

where $e = y - \hat{y}$ and the ϵs are the residuals about the regression of the observed values on the residuals, $\epsilon = (\beta_0 + \beta_1 y) + e$.

A simultaneous F-test for $\beta_0 = 0$ and $\beta_1 = 1$ in the regression of observed on predicted data ($y = \beta_0 + \beta_1 \hat{y}$) is often recommended, and is effectively the same as the former F-test, but substitutes $\hat{y}$ for y. Regressing observations against predictions is consistent with the statistical notation $y = \hat{y} + e$, and recognizes that the observations contain natural variation, but assumes that model predictions are accurate. An alternative view is to recognize that the model is an incomplete abstraction of reality, and that its predictions are only approximations. The second view is preferable, and thus residuals should be regressed on observed values. These alternatives may give different results (e.g. for the F-test), and the latter approach is generally more revealing.

In addition to overall appraisals, it is desirable to partition data (e.g. by age, site index or stand density), and examine model performance in each of several strata (e.g. Mayer and Butler 1993). The most revealing insights may be obtained by devising strata based on a knowledge of the biological system, the model and the characteristics of the data. However, the absence of any visible inadequacies in any particular stratification does not imply that weaknesses cannot be found in an alternative stratification.

Partitioning Data

Benchmarking in its purest form requires that some data are set aside, or that new data are obtained for benchmark tests. The most convincing test would use a set of data drawn from an independent population measured over a long period, but such data are rarely available. Growth modellers frequently are faced with the decision of having to partition a data set from a single population into two subsets, one for development, and the other for the testing the model. Where ample data are available, this partitioning causes few problems. However, when data are scarce, there is a temptation to use all the available data for development, in an attempt to improve the model. Unfortunately, this diminishes the ability of the modeller to demonstrate the quality of his model, and ignores the role of falsification of hypotheses in science.

Statistical principles suggest several guidelines to use when partitioning the data set, but in practice the choice seems to be largely arbitrary. When testing a model, the modeller generally hopes to find the model acceptable, and must avoid the temptation to weaken the tests, for example, by reducing the number of data available for benchmarking. A half and half split is popular in many disciplines (Snee 1977), but fewer data are often used to benchmark forest growth models (e.g. West 1981 and Shifley 1987 reserved one quarter of their data). Plots established for long periods with regular remeasurement, particularly those remaining unlogged and without other disturbance, may prove useful as a discriminating test. If data from a geographically independent area (but preferably not an outlying extreme) can be reserved without compromising the range of site and stand conditions represented in the model, they could provide a convincing test of the model. Alternatively, data collected by an independent agency could provide a good basis for benchmarking.

The outcome of benchmark tests can be influenced by the selection of data: "like" data will provide a more optimistic result than comparisons with "unlike" data from another population. Thus the most convincing demonstration of model quality can be made only if the test data are in some sense unlike the development data. Since growth models are used to forecast future forest conditions, one way to split the data is on time, and for

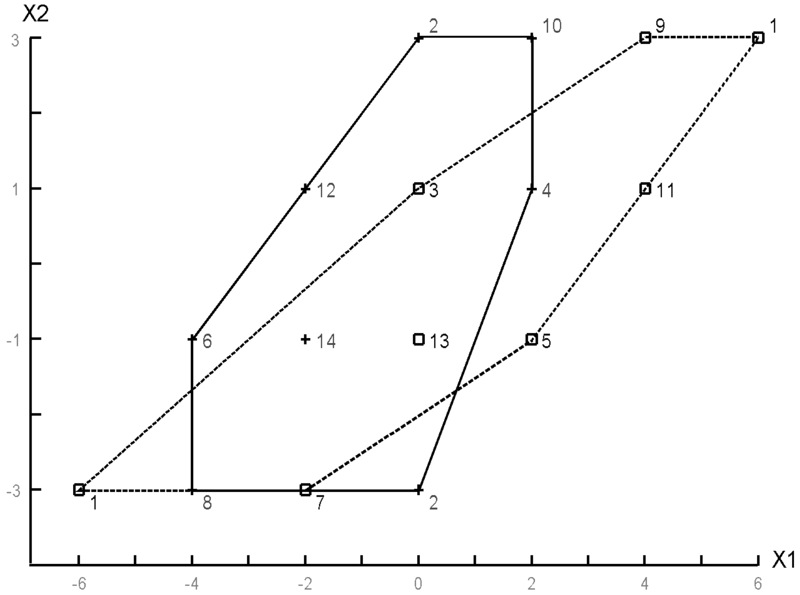


Fig. 11.4. Example of data partitioned using the duplex algorithm (Snee 1977) into estimation (G) and benchmark (+) subsets. Numbers indicate sequence in which data were assigned to subsets (see text).

example, to use data collected before 1980 for development, and since 1980 for testing. It is desirable to have test data available to test the model over the widest possible range of site and stand conditions. Objective procedures should be used to select these data to minimize the dangers of bias.

Snee (1977) discussed the duplex data-splitting algorithm. This algorithm splits the data so that the two sets cover approximately the same region and have similar statistical properties. The duplex algorithm takes a list of candidate points, standardizes them and computes the Euclidean distance between all possible pairs. The two points furthest apart are assigned to the estimation set, the pair with the next greatest separation are assigned to the benchmark set, and the remaining points are assigned in turn to the estimation and benchmark set, selected in order of the greatest distance from the data in the set (see sequence numbers in Fig. 11.4). This has the effect of creating two overlapping sets which cover different parts of the data space, thus providing an "unlike" set for rigorous testing. If the data contain replications or pseudo-replications, the candidate list should comprise clusters of data rather than individual points, otherwise the two tests will cover the same space.

The benchmark data should contain sufficient replications to enable the natural variability to be expressed. There are really two sources of error to be found in benchmarking a model: model error and observed data variability. Nature is not deterministic: two stands apparently with the same initial conditions may develop differently. Evaluations should employ sufficient data to provide a reasonable estimate of the expected actual stand condition.

Resampling Procedures

Suitable data for modelling are scarce, and researchers may be reluctant to set aside a sufficiently large amount of data for benchmarking. An alternative is to mimic the use of independent data using resampling techniques such as cross-validation, boot-strapping and jackknifing (e.g. Weisberg 1985).

Regression models may take the form: $\boldsymbol{Y} = f(\boldsymbol{X}, \beta) + \epsilon$ where the error ϵ is unknown, f may be non-linear and β are parameters to be estimated. *Apparent error* is computed by applying the fitted equation to the data used in fitting (e.g. R^2) and will normally give an optimistic view of the quality of a model. *"True" error* is best estimated by fitting the model to independent data. However this is often not possible, especially when one considers how many additional data should be used. Resampling techniques provide an alternative.

Consider partitioning the data, with half for development and half for benchmarking. If the resulting model failed to satisfy expectation, would it be improved by using all the data for development? But then there would be no independent test. Several re-sampling methods attempt to mimic independent testing whilst enabling the full data set to be used for model development.

Half-splitting provides a solution, by fitting with half the data and benchmarking with the remainder; the roles are then reversed. Each evaluation provides an estimate of true error which may be averaged. The entire data set may be used to obtain the final model and estimates of apparent error. This method provides pessimistic estimates of model performance but is computationally efficient (Burk 1990).

Cross-validation is the logical generalization of half-splitting. Rather than deleting half the data, each datum is deleted in turn and the model is fitted to the remaining $n-1$ data. Benchmark tests are averaged from the individual deleted data. If the test statistic is squared error and the model is linear, the cross-validation estimate of true error is n times the PRESS statistic computed by many regression packages. The boot-strap and jackknife are similar to cross-validation, especially as sample size increases, but are computationally more complex (Efron and Gong 1983).

If liberal amounts of data are available to the modeller, some data should be set aside to allow an independent check of the model. In data-poor situations, it may be wise to use a re-sampling procedure. Burk (1990) found that half-splitting provided a pessimistic evaluation, and suggested the use of cross-validation and jackknifing, coupled with tests based on limited additional data.

One shortcoming of any resampling procedure lies in its dependence on the data. The sample should adequately represent the variability and other characteristics of the population of interest, or the resampling procedure will not provide an adequate test of the model. Unfortunately, these are the very circumstances under which the model itself should come under heaviest criticism.

Re-calibration

Re-calibration implies adjusting a growth model so that it provides good predictions for a new population. This may entail estimating new parameters for some or all of the equations in the model, or may use a scaling factor to adjust predictions.

STEMS (Belcher *et al.* 1982) is one growth model which has been "transplanted" to several other regions. Most copies have retained all the computer code and retained the form of all equations. Some have estimated new coefficients for all or most equations (e.g. Shifley 1987), whilst others have developed scaling factors to adjust existing equations (e.g. Holdaway 1985). Such scaling factors may comprise a single adjustment for each species, or may be correlated with some tree or stand variables (e.g. tree diameter or stand basal area). Smith (1983) applied three different scaling factors for each species, for each of three diameter classes. This means that the diameter increment function effectively becomes a discontinuous step-function and may have undesirable properties for some applications.

Neither single nor multiple scaling factors guarantee a general improvement in the model, and some tests are necessary to ensure that the adjustment is adequate for the region over which it will be applied. All equations used should be appropriate for the new location. Residuals should be inspected to ensure that the models are sufficient and adequate.

Re-calibration is not a "cure-all", even for existing "good" models. Attempts to re-calibrate STEMS to Australian forests using a single scaling factor (e.g. Swain and Turner 1988 for *Eucalyptus marginata* forest) or by re-estimating coefficients in component equations (e.g. Goodwin 1988 for mixed eucalypt forest) have been fraught with difficulty and results to date have been poor.

The creation of a variant of a growth model for a new locality, may involve several steps. Firstly, the model should be benchmarked using data from the new locality to determine if any re-calibration is needed. Given that some adjustment is necessary, the residuals about predictions should be examined to see if a single scaling factor would be adequate, or if a more sophisticated adjustment is necessary. If inspection of residuals indicates that a simple adjustment to increment rates would provide satisfactory predictions (e.g. analogous to a better site productivity), then such re-calibration may be attempted. However, if a more complex adjustment to growth patterns is indicated, it may be preferable to abandon re-calibration attempts and to estimate new parameters for all coefficients in the model.

Self-calibration

The concept of self-calibration appears to have been first applied by Stage (1973), who used a single, simple correction factor based on increment cores taken on temporary inventory plots to adjust estimates of diameter increment. The method was formalized statistically by Meng *et al.* (1990). Stage (1973) calculated the deviation (observed − predicted) in the logarithm of the tree basal area increment, and added this to the logarithm of his diameter increment function. He, in effect, assumed that his function has the correct "shape", and that only the rate of growth needs to be adjusted using a single parameter. This robust approach may be advisable if estimates of site productivity are unattainable, or if the climate varies greatly between the areas from which the development data were obtained and where it is intended to apply the model. Stage (1973) argued that the approach accommodates the model to "local peculiarities of site quality, genetic character and tree vigour", but cautions that "growth functions should be based on data derived from the area to which the model is to be applied; the self-calibration feature . . . only partially mitigates that admonition".

The self-calibration feature in Prognosis has been implemented so that local adjustments are attenuated over time, and predictions gradually revert towards that of the standard uncalibrated model. Attenuation is modelled using an exponential decay function (Wykoff 1986) with the asymptote midway between 1.0 and the calculated multiplier. The decay rate is such that it takes 25 years to reach midway between the calculated multiplier and the asymptote.

A contrasting application is the GROPE model of Alder *et al.* (1977) which was to have all its parameters estimated "automatically" by self-calibration from site-specific data. Such an empirical approach must ensure that all equations used are inherently robust, and that the data available for each site are extensive and sound. Although this is intuitively possible, effective implementation would require an expert system embodying many concepts of statistics, ecology and silviculture.

Benchmarking after Re-calibration

Benchmark testing of a re-calibrated model poses a special problem. In Stage's (1973) approach, the model can be thoroughly tested prior to re-calibration. After re-calibration for a site, two factors should be benchmarked: (i) the assumption that the "shape" of the basic functions remain unchanged, and (ii) that the re-calibration process has in fact removed the bias evident in the model prior to re-calibration. Whether or not the basic functions hold for a large range of sites may depend on the specific functions, on the linkages between functions in the particular model involved, and on many other factors. Stage (1973) re-calibrated only the diameter increment function, and assumed that the other functions (height increment, bark ratio, crown ratio and mortality) would remain unaffected by "site quality, genetic factors and tree vigour".

The GROPE model (Alder *et al.* 1977) provides an option for data to be reserved for benchmark tests of the model, and this is clearly required in an empirical approach of this nature. The "model fitting cycle" of their model may be viewed as the normal development phase of model construction, and predictions should be compared with additional data before the model is used for forest management applications.

Synthesis

If growth modelling pretends to be a science rather than an art, the models proposed should be falsifiable. Thus, they must be able to be rejected through the normal process of experimental testing. Model evaluation is an important part of model construction, indicating the nature of the forests for which the model may be expected to yield reliable results, as well as areas in which further research is required.

The temptation to use all the available data for the development of the model must be avoided, as it is equally important to have an independent set of data available for benchmark testing. The need for such testing is not diminished through the use of "self-calibrating" models.

Exercises

11.1 Criticize the model illustrated in Fig. 11.1 and given by Equations 2.13 and 2.14 (p. 28).

11.2. West (1981) published a growth model for natural regrowth stands of three eucalypt species in Tasmania. For *Eucalyptus obliqua*, the diameter increment is predicted as:

$$\Delta d = \left(-792 + 7400t^{-3} + 0.679N\right)\times 10^{-4} + \left(1.8 + 1750e^{-t} + 0.029N\right)\times 10^{-5}\left(d - 10\right)^2$$

and mortality is predicted as

$$p = \left(10.0 - 4.77t + 1.40t^2 - 0.113t^3\right)\times 10^{-3}\left(\frac{\bar{g}}{g}\right)^{1.75}$$

Criticize the model.

11.3. West (1981) reserved 21 plots to allow testing of his model. A summary of these data is given below (n is the number of years simulated, and $\hat{G}$ and $\hat{N}$ are predictions of G and N respectively). Benchmark his model. Is it an adequate model?

Plot	G_0 m^2ha^{-1}	N_0 ha^{-1} ×10	$S_{h,t}$ m	n y	t_n y	G_n m^2ha^{-1}	$\hat{G}_n$ m^2ha^{-1}	N_n m^2ha^{-1} ×10	$\hat{N}_n$ m^2ha^{-1} ×10
E3	60	80	39	3	45	62	61	77	74
O12	38	50	38	3	48	40	39	46	45
O25	56	63	38	3	61	60	56	60	56
R2	49	62	47	3	40	52	52	53	59
R7	71	22	47	3	82	75	73	21	21
O68	37	108	52	3	27	39	41	108	104
O15	46	120	36	5	42	50	49	110	110
R13	52	47	46	7	65	54	54	39	38
G4	63	52	45	8	73	68	64	46	41
R20	32	160	45	8	28	37	39	157	151
G10	56	73	46	8	49	56	64	57	62
R11	65	49	49	8	61	69	68	39	33
E9	51	89	39	9	57	59	52	80	67
O19	35	42	42	10	44	48	44	42	29
O21	58	111	39	12	49	74	71	99	78
O61	48	103	46	13	47	60	59	65	77
E10	45	104	41	15	50	56	55	82	74
O58	57	60	49	20	62	65	70	41	38
O6	35	144	33	21	44	43	50	63	114
O2	39	77	40	21	50	52	60	42	46
O4	42	88	41	21	50	57	59	56	55

11.4. Equations for growth, death, recruitment and harvesting have been given for *Flindersia pimenteliana* (i.e. increment Eqn 8.6 p. 168; survival 9.2 p. 180; harvesting 9.3 p. 183; logging damage 9.4 p. 186; deterioration 9.5 p. 188; defect 9.6 p. 189; recruitment 10.2 and 10.3 p. 199). Condense these into a transition matrix model and criticize the model. What assumptions do you need to make? Is the model generally applicable? Could you use it for both pure and mixed stands? For both even-aged and uneven-aged stands?

11.5. Formulate the *Flindersia pimenteliana* equations as a cohort or tree list model, and discuss assumptions, implications and the general applicability of the model. You may need the following volume equation: $v = -0.661 + 8.97g$. Is the cohort implementation a better alternative than the matrix model?

Chapter Twelve

Implementation and Use

There is little point in developing a growth model unless it is to be used. The process of model construction may reveal some implications for forest management, but the greatest benefit will accrue if forest managers use the model to investigate forest management alternatives. Thus the model should provide information required for forest management, in a form useful to forest managers. The information provided should be

1. *timely*: available when needed, and not outdated when made available,
2. *accurate*: reasonable estimate of reality,
3. *complete*: include all the details the user needs to know about the situation,
4. *concise*: exclude elements not required by the user,
5. *relevant*: directly related to the issues under consideration, and
6. *appropriate*: in presentation for the particular audience.

This requires that the model is implemented in a flexible information system linked to other resource databases.

To get maximum benefit from a growth model, it should be made available to all potential users, including forest managers, planners, and policy-makers in head office, research and in field management. It should also be made available for education and training at a variety of levels and times (i.e. pre-vocational and in-service). Field managers should be able to use the models themselves, preferably in a "hands-on" way. The growth model should be viewed as a tool to provide better information for forest management, but must not threaten the experience and judgement of the manager. To achieve these lofty goals, a growth model must be easy to use, well documented and integrated with other information systems. The question is then how the model can be used most efficiently to provide the information needed for forest management and planning.

Implementing the Model

To make it accessible to users, a model is usually implemented on a computer, and the task of implementing and maintaining the computer program may be simplified if a few simple guidelines are followed. In essence, the model should be as simple, logical, modular and portable as possible. Unnecessary complexity invites errors and may offer few benefits.

Model implementation on a computer involves two aspects: logical algorithms and appropriate data structures. The computer language doesn't matter, but it is important that the program is clear and logical, so that it is easy to find and fix errors, and easy for others to see what has been done. Many pioneering modelling approaches resorted to a variety of "tricks" to reduce computer storage and execution time. However, the relativities of computer and staff costs have changed greatly, and it is desirable to build a model that is easy to maintain, even if it is necessary to sacrifice some computational efficiency. It is likely that a model may be implemented on several different computers during the course of its useful life, and this means that the model should be written in a standard high-level language without using vendor-specific features.

Models may involve interim components because of limitations in data, techniques or knowledge. The model should be constructed in a modular fashion so that each module is, as far as possible, independent of other components in the system. This makes it possible to refine one component equation using new data or analyses without altering all other components.

The application dictates the nature of the output required from a growth model. Initial testing may require detailed output, whereas an analysis of optimal management regimes may involve many long simulations but produce only brief output. Thus input-output routines should be placed in a separate module and should not be embedded in the growth simulation module. The model should be constructed so that it can be used as a "black box" which grows the stand for one growth period (generally one year but sometimes as much as 10 years, e.g. the Prognosis model). This approach has been used in many growth models and has proved to be a useful and flexible approach.

In summary, a model should be:

1. General and applicable to a wide range of site and stand conditions.
2. Able to simulate effects of the major management options.
3. Realistic, with biologically sound functions rather than empirical surrogates.
4. Modular with growth modules separate from the management simulator and input-output routines.
5. Driven by operational inventory data rather than by data that require prolonged scientific measurement.
6. Diagnostic, allowing users to identify data entry errors and unacceptable model performance.

7. Flexible with options to alter simulation of individual processes and format of reports (e.g. ability to write reports to file or to pass them to graphics packages).
8. Controlled by the user, with subjective parameters kept to a minimum and amenable to sensitivity testing by the user.
9. Well-documented and easy to use with clear reports that simplify the interpretation of results.

It is useful to integrate growth models into inventory reporting systems, so that plots can be reported as at the date of measure or after projecting to any future date. Models should also be offered as interactive packages inviting investigation of "what if . . ." questions. Interactive simulation options that may be provided could include:

1. Grow the stand a specified number of years, or until a given stand parameter (e.g. stand basal area, standing or merchantable volume or average stem size) reaches a specified level.
2. Log the stand, using a harvest simulator, removing all merchantable stems over a given size limit, or soliciting the user's intention for each species and size class specified by the user.
3. Treat the stand, applying standard treatment rules, reducing the stocking of specified species and sizes to a specified residual, or soliciting user's intention for each species and size class.
4. Undo the last (grow, log, or treat) command.
5. Save the present stand for future reference. This is useful for examining different harvesting or treatment options for a stand which can be logged, treated and grown several times, using different strategies.
6. Restore the stand as at last save.
7. Display the current stand as a stand table, using size classes specified by the user.

A macro facility should also be provided, so that silvicultural regimes can be defined in a series of instructions, and repeatedly invoked so that long term consequences can be investigated. Such studies can be done individually for single stands, or simultaneously for a series of stands. Whilst such studies may provide useful results, they should always be investigated further in field trials.

Resource Data for Simulation Studies

Considerable effort may be expended in developing reliable growth models, but good predictions can only be made if user inputs are also reliable. Thus users of growth models should take commensurate care in collecting the necessary input data. Mowrer and Frayer (1986) examined some effects of errors in initial conditions, and suggested that errors in user-supplied data could have a bigger effect on the overall variability of predictions than the contributions from the growth model.

It is important that sampling should be efficient and unbiased, and this requires decisions on stratification, plot size and trees to be measured. These topics are discussed in many inventory texts (e.g. Loetsch *et al.* 1973, de Vries 1986, Schreuder *et al.* 1993, Avery and Burkhart 1994), and the following is only a brief overview.

The general precaution of data collection applies equally to the gathering of data for use in growth modelling and resource forecasting: know your information requirements, and ensure that field and office procedures contribute towards meeting those needs. This caveat applies to inventory design, and especially to details such as stratification, plot size and measurement limits.

Stratification

Improved stratification may be the most efficient way to improve the precision of resource estimates and forecasts. Precision can be gained by dividing the population into as many blocks as expedient, even though the number of random sampling units taken from each may be the minimum of two (Schumacher and Chapman 1954). Gains in precision are greatest when the within-stratum variation is small compared with the between-strata variation, but even geometric blocking (i.e. dividing the resource into rectangular blocks) may help to improve estimates and forecasts.

Table 12.1 summarizes some options for resource inventory. It takes the form of a binary key to help users select an appropriate method. Start at step 1 in the first row of the table and consider the criteria in the left column; if you think that the central column best describes your situation, proceed to the next row (step 2). Alternatively, the right column suggests one possible design that may fit your circumstances and highlights some possible consequences (in *italics*). For example, Table 12.1 suggests that if the requirement is for a reliable estimate suitable for spatial interpolation (e.g. for site quality or forest type maps), then systematic sampling may be a good approach to use. If estimates are critical and confidence intervals are required, it may be better to use stratified random sampling, so that unbiased estimates of the precision can be obtained. The decision to use geometric blocking or statistical strata may be based on the availability and reliability of existing information; complex strata are beneficial only if they reduce the within-stratum variation relative to the between-strata variability.

Notice that this key encapsulates some dilemmas, reflecting situations where satisfactory samples may be unattainable. For instance, if resources are insufficient, it may be appropriate not to attempt any sampling because objective samples may lack the precision desired, and subjective samples may be biased. Despite the compact nature of Table 12.1, it summarizes many important factors influencing sampling design, and readers should

Table 12.1. Key to some alternative sampling designs showing selected selection criteria and some possible consequences (adapted from H.C. Dawkins, pers. comm.).

Criteria	Inventory alternatives & possible consequences	
Step 1		
Nature of estimate	Critical	Unimportant/personal
Forest characteristics	Unknown/diverse	Familiar or uniform
Representative selection	Unreliable	Reliable
Time and resources	Sufficient	Very limited
	Objective → Go to 2	**Subjective sampling**
Bias	*Absent*	*Unavoidable*
Precision	*Can be estimated*	*Unknown*
Step 2		
Periodicity	Possible/unknown	Unlikely or known
Interpolation	Not required	Necessary
Estimate of precision	Required	Unimportant
	Random → Go to 3	**Systematic sampling**
Sampling error	*Correct estimate*	*Probably inflated*
Periodic bias	*Unlikely*	*Possible*
Step 3		
Pattern in population	Clear or likely	Absent or unlikely
Sampling intensity	Relatively low	High
	Stratified random → 4	**Unrestricted random**
Inherent risks	*Misjudge pattern*	*Sample clustering*
Step 4		
Pattern in population	Obscure/unknown	Visible or well known
	Geometrical blocks	**Statistical blocking**
Calculations	*Simple*	*Possibly complicated*

consider all the issues raised before commencing an inventory. In many cases, some form of stratified sampling may be optimal.

Smith and Burkhart (1984) found that stratifying by both site index and stocking improved the precision of volume estimates by two-thirds over simple random samples. Site index was the more useful variable, and sample size had no effect on the relative gains achieved through stratified sampling.

Plot Size

Plots should be of an appropriate size to ensure that results obtained are realistic, and are not an artifact of sampling. There are two issues: the size of the plot measured in the field, and the plot size assumed in the model simulations. Generally, it seems preferable that the three "plot" types used in modelling should be similar: (i) permanent plots used to fit the model, (ii) the "plot" simulated within the model, and (iii) the sample unit (e.g. temporary plot or point sample) used to gather additional resource data for simulation studies. However, in practice, each of these serves different purposes, and they need not be the same size or dimensions.

The size of plots used to sample stands for growth forecasts may be dictated by the need for plots to be *homogeneous* and representative of the forest stand, and by practical and financial aspects of inventory. Notice that data for modelling studies, like model development data, should be drawn from plots which are relatively homogeneous and in which edge-effects are minimized. Other inventory applications may attempt to maximize within-plot variation to reduce between-plot variation and the sampling error, but this may not provide the best data for modelling studies.

In theory, optimal plot size may be determined from statistical requirements and the time taken to locate and measure a plot (e.g. Loetsch *et al.* 1973, Schreuder *et al.* 1993, Zeide 1980). However, the optimal size depends on the purpose, and the size that is optimal for reporting a static inventory may not be optimal for the provision of input data for growth and yield studies. In any case, practical aspects may override theoretical considerations. The need for relatively homogeneous plots may provide an upper limit on plot size. West *et al.* (1988) suggested that 0.5 hectares was about the practical limit in north Queensland rainforests. Even if the forest exhibits few discontinuities, excessively large plots tend to average out variations, and may provide less reliable and overly-optimistic yield predictions. Hann (1980) compared yield forecasts from large (32 ha) plots with those from one-hectare subplots, and found that the smaller (1 ha) plots gave more reliable forecasts. The lower limit may be influenced by the need for representative plots, and the large perturbations that may be caused by borderline trees in small plots.

The size of physical plot on which data were gathered may have a considerable influence on resource estimates and forecasts, but predictions may be less sensitive to the plot size simulated within the model. In many models, the simulated plot size is fixed (e.g. at 1 ha) and is not accessible to the user. In other models, the simulated plot size is accessible to the user and may be varied to maintain consistency with input data, and this may have a major influence on the variance of stochastic predictions.

Censored Data

It is important that inventory data be compatible with the data requirements of the model. Incompatibilities may arise most frequently in the measurement of small trees. For example, the growth model may predict recruitment at 10 cm dbh and may require that data inputs include all trees in the stand exceeding 10 cm diameter. However, some inventory data may only record trees exceeding 20 cm diameter, and such data are known as censored data (i.e. trees between 10 and 20 cm dbh have been "censored"). Another form of censorship is the failure to record non-commercial species during inventory. Censored data may lead to biased estimates of growth and yield. Randall *et al.* (1988) investigated the impact of censored data (omission of trees < 13 cm dbh) on yield predictions, and recommended that such data be augmented by "average" small tree distributions for that forest type.

Obviously, it is preferable to co-ordinate inventory design and model design to ensure that no data censorship occurs, either by ensuring that measurement limits used in inventory are compatible with the recruitment threshold used in the model, or by providing a recruitment function with a variable threshold (e.g. Shifley *et al.* 1993; but note that some censorship may remain, since estimates of stand basal area may utilize different lower limits).

Optimization Studies

Forest managers may be interested in details of the theoretically optimal management regime, and, for uneven-aged stands, may seek information concerning:

1. sustainable diameter distributions, maximum tree sizes, and optimal stocking by species and size classes,
2. the optimal cutting cycle and the best strategy to convert present stands to that regime, and
3. stand-specific treatment schedules to meet forest-wide objectives and constraints.

Such information may be obtained from analyses at the stand level, or at the forest level. Stand level analyses (e.g. of optimal sustainable diameter distribution, cutting cycle or conversion strategy) are relatively easy to undertake, but do not indicate the best strategy for a forest comprising many stands. The optimal single-stand strategy is not necessarily the best approach to adopt on all stands in a forest. For instance, the need to maintain a steady supply of timber may preclude the immediate adoption of an optimal cutting cycle in all stands in the forest estate. Stand level analyses may indicate the optimal strategy for a single stand in isolation, but an optimal schedule for all stands requires forest-level optimization.

A very brief overview of some techniques relevant to mixed forests is presented here. Comprehensive reviews of optimization studies and methodologies for uneven-aged forests were given by Bare and Opalach (1987), Haight and Monserud (1990*a*,*b*) and Hof (1993), and readers should consult these for more details. The purpose of this overview is not to teach optimization techniques, but to consider some aspects of optimization that may have a bearing on model design and construction.

It is convenient to distinguish two classes of optimization study. *Static* studies may often be inferred from a single equation in a model, and may assume that many factors not under immediate consideration (e.g. competitive status of trees) remain constant. In contrast, *dynamic* studies are simulation studies that take all factors (included in the model) into account, and may thus provide a more detailed picture. Haight (1985) demonstrated that static and dynamic studies may not provide the same results; careful consideration of the assumptions implicit in each approach is necessary to determine the appropriate result.

Static Studies

Useful information regarding the stand structure and cutting limits (i.e. size for harvesting) required to achieve near-optimal production can be obtained from simple analyses of the component equations comprising the growth model. Such studies may use a diameter increment function and a volume equation to determine the current and mean annual volume increments for a single tree, and thus determine the optimal size for harvesting (i.e. the size at which the mean annual increment is maximized, if volume is to be optimized). Such studies generally assume that stand basal area and the relative competitive status of the tree remain constant, and are thus termed static studies. Dynamic studies simulate the stand dynamics and provide a more accurate picture, but require additional resources.

Vanclay (1989*a*) reported some simple static studies of optimal tree size for harvesting, and illustrated how the optimum changes with site productivity and stand basal area. His findings tended to support the existing harvesting guidelines, but some of the guides were below the apparent optima (Table 12.2). The guidelines specified that trees exceeding the cutting limit should be harvested unless they are of outstanding vigour and form, and that all trees exceeding the "retention" limit should be harvested, irrespective of vigour and form, unless required as seed trees. The optimal values for these cutting and retention limits were estimated by calculating the diameter at which the mean annual volume increment of individual trees reached its maximum, assuming no mortality for the retention limit and average mortality losses for the cutting limit (Table 12.2). The incidence of internal stem defect was not considered, and may slightly reduce the optimal diameters for some species.

Table 12.2. Diameter (cm) at which tree volume growth is maximized, assuming soils derived from coarse granite parent material, compared with Queensland treemarking guidelines (after Vanclay 1989a, and Preston and Vanclay 1988).

Species group	Static optimization study				Treemarking guidelines	
	Retention limit (cm dbh)	Cutting limit when basal area is (m^2/ha > 20 cm dbh)			Retention limit (cm dbh)	Cutting limit (cm dbh)
		20	30	40		
Large, fast	128	105	101	96	80–100	60–100
Large, slow	143	108	105	101	50–100	50–80
Small, fast	109	71	64	54	50–90	50–70
Small, slow	101	66	60	53	50–90	50–70

In all optimization studies, it is important to consider if the apparent optimum is "true" or is an artifact of the model. This caveat applies particularly to size class models, where the class boundaries may appear to attain special significance. Matrix models deserve special attention in this regard. Their mathematical elegance simplifies optimization studies, and allows "dynamic-like" studies (e.g. steady state, eigenvalues, p. 47) to be performed easily, but the stationary and other assumptions (p. 44) cast some doubts about the validity of conclusions drawn in this way.

Dynamic Studies

Dynamic optimization studies are more complex, but provide a more realistic analysis of stand dynamics. The growth model's ability to predict yields corresponding to various stand conditions can be exploited to compute the stand condition that maximizes some benefit such as log volume, sawn volume, discounted net revenue or some other criterion, subject to several constraints. Although many constraints can be accommodated, a single objective function is required and may prove something of a limitation. If it is desired to maximize several products, these must be converted to common units (e.g. dollars or cubic metre equivalents) and combined into a single objective function. Some agencies may not be profit motivated, and non-timber benefits may be difficult to quantify, but this does not eliminate the need for potential outputs to be appraised in common units. The use of any criteria other than revenue will lead to lower returns (Adams 1976, Hof 1993). Most optimization studies optimize the expected outcome, but some applications may be better served by maximizing the lower confidence limit. Most research on optimal stand structures has focused on diameter distributions, but some studies have examined the question of species composition (e.g. Valsta 1988).

The problem of finding the optimum stand condition may be expressed as: find the initial stand condition X_0 which gives rise to the stand X_t in t years time, such that the increase $X_t - X_0$ is maximized. The initial state X_0 may be zero for an even-aged stand which is ultimately clearfelled. Meyer (1952) expressed this as the stand structure in which "current growth can be removed periodically while maintaining the diameter distribution and the initial volume of the forest". Many simple optimization studies rely on the existence of this "sustainable distribution" (Adams and Ek 1974): $N_{d,t+1} \geq N_{d,t}$ for all d, where $N_{d,t}$ is the number of trees in a diameter class d at time t. This ensures that at any future time, some trees can be removed from each class to return to the original distribution. Michie (1985) used a matrix model to develop a more general case which converts any initial stand structure to an investment-efficient sustainable distribution in a predetermined number of harvests.

Adams and Ek (1974) maximized the stumpage value of stems removed to restore the initial diameter distribution with the constraints that the number of stems removed from any size class must always be non-negative, the specified initial basal area must be maintained, and that the number of trees in each diameter class must never be less than zero. This simple analysis led to a considerable increase in predicted value of cut, of both stumpage value and in volume, over the existing management guides. They considered five and ten year cutting cycles in their analysis, but did not consider a variable cutting cycle. The problem of the optimal conversion of some stand to the optimum stand condition was formulated so as to maximize the revenue from a specified number of cuts which result in the specified optimum stand condition after the final cut. They gave an example where a stand could have been converted immediately to the optimal condition, but it would have been uneconomic to do so. Their analysis led to an optimal conversion schedule which maximized the perceived benefits. Adams and Ek (1974) did not consider the holding cost of the residual growing stock, so their analysis yielded different results to those obtained from dynamic present net value studies (e.g. Haight 1985).

Bare and Opalach (1987) assumed that the sustainable diameter distribution in spruce-fir forests could be modelled by a Weibull distribution. This assumption enabled computational efficiencies, but such a uni-modal distribution may be inappropriate for uneven-aged forests. They found that the "investment efficient" sustainable equilibrium diameter distribution associated with maximum land expectation value differs from the maximum managed forest value. These and other studies demonstrate the interrelationship between management objectives and optimal stand structures (e.g. Hof 1993).

Many optimization studies simplify the growth model to provide a tractable model for their analyses, and this may influence the results. Haight and Monserud (1990*a,b*) demonstrated a method for optimizing any-aged management of mixed species stands using a tree list growth model (Prognosis) without modification. They did not assume a sustainable diameter distribution, but allowed any silviculture ranging from selection logging and shelterwood systems to clearfelling with natural regeneration or planting. Their formulation of the problem was to define the *n*-dimensional vector $\boldsymbol{x}(t)$ as the initial state (i.e. the present stand table) at the beginning of time period *t*, and let $\boldsymbol{u}(t)$ be an *m*-dimensional vector of controls where *n* and *m* are the number of state and control variables respectively. Typically, $\boldsymbol{x}(t)$ represents the distribution of trees by diameter and species classes, and $\boldsymbol{u}(t)$ represents the harvest levels for various classes. Each element $u_j(t), j = 2, \ldots, m,$ of $\boldsymbol{u}(t)$ is defined as the proportion of the trees harvested from diameter and species class *j* and is constrained between 0.0 and 1.0. The smallest size class $u_1(t)$ is unconstrained to allow for planting, and when $u_1(t) < 0$ it represents the number of seedlings planted, and when $u_1(t) > 0$ it represents the number of seedlings thinned. Let $R[\boldsymbol{x}(t), \boldsymbol{u}(t)]$ denote the revenue obtained in period *t*, where the resource is in state $\boldsymbol{x}(t)$ at the beginning of this period before harvesting, and the harvest and planting control $\boldsymbol{u}(t)$ takes place at the beginning of the period. Then letting δ denote a discount factor ($\delta = (1+r)^{-1}$, where *r* is discount rate), and *T* the planning horizon, the optimization problem can be expressed as

$$\underset{\{\boldsymbol{u}(t),t=0,1,\ldots,T-1\}}{Max} \quad J_T\left[\boldsymbol{x}_{(0)}\right] = \sum_{t=0}^{T-1} \delta^t R\left[\boldsymbol{x}_{(t)}, \boldsymbol{u}_{(t)}\right] + \delta^T G\left[\boldsymbol{x}_{(T)}\right]$$

where $\boldsymbol{x}(0)$ is the initial state and $G[\boldsymbol{x}(T)]$ is the terminal payoff function for the stand in state $\boldsymbol{x}(T)$. The maximization was solved subject to an *n*-dimensional difference equation for the stand dynamics which implicitly represents natural regeneration, tree growth and survival:

$$\boldsymbol{x}_{(t+1)} = \boldsymbol{f}\left[\boldsymbol{x}_{(t)}, \boldsymbol{u}_{(t)}\right], \; t = 0, 1, \ldots, T-1.$$

The solution to these equations is the control variable set $\{\boldsymbol{u}(t), t = 0, 1, \ldots, T-1\}$ that maximizes $J_T[\boldsymbol{x}(0)]$, the present value of the existing stand over the *T*-year planning horizon. Haight and Monserud (1990*a,b*) found that even-aged plantation management and uneven-aged shelterwood systems were both capable of producing the same high level of yield indefinitely.

Haight and Monserud (1990*a*) cautioned that optimal solutions to any-aged management problems may vary considerably depending on initial estimates for parameters to be estimated. Wide diameter classes and broad species groups provided more robust results. They stressed that sensitivity analyses were essential to investigate the sensitivity to starting condition, and to ensure that a global rather than a local optimum had been reached. Optima may also be sensitive to the planning horizon, and long term simulations (e.g. >100 years) may be necessary for uneven-aged forests.

Information concerning the sensitivity of the optimum is as important to the manager as the optimum itself. In many cases it may be difficult or even impossible to achieve the optimum, and the manager will want to know how close to the optimum stand condition his forest needs to be to achieve a near optimal return. The optimum may well be attained over a wide range of stand conditions, and sensitivity analyses should be carried out to indicate, for example, the range of stand conditions which return a yield within five percent of the maximum. The forest manager may then strive to attain a stand condition somewhere within this range, and may well exert his own beliefs, based on personal judgement and experience, as to the best condition for stands under his control.

This brief overview is not intended to be comprehensive, and merely attempts to give the growth modeller a feel for some requirements that potential model users may have. Readers attempting optimization studies should consult standard references (e.g. Leuschner 1990, Hof 1993).

Yield Prediction

Yield forecasts are amongst the most important applications of many growth models. Several factors need to be considered, to ensure that yield estimates are not biased. Yield estimates prepared by multiplying the productive area of the forest estate by the theoretical per-hectare yield usually overestimate attainable yields. More detailed simulation studies are required to provide reliable estimates of sustainable yields for a forest estate.

Determining the yield from a single stand with the aid of a growth model is a relatively simple matter, but extending the concept to the whole forest estate requires some account to be taken of the distribution of yield over time. Agency objectives may stipulate a non-declining even flow of timber onto the market, even though the market may be fickle and the forest estate far removed from the "normal forest". This generally requires that some management units may need to be cut before, after, heavier or lighter than the optimal cutting rules would suggest.

Managers may also wish to capitalize on buoyant market opportunities by temporarily increasing the harvest. This is possible, as the sustained yield is only a guide, and variations in cutting cycle length and intensity are inevitable. Haight (1990) investigated thinning strategies in situations where prices vary stochastically. When the sustained yield is temporarily exceeded, the standing crop and its increment is reduced, and production and future yields may be depressed for some time. Similarly, if the stand becomes overstocked through "undercutting", production may also be depressed, although the standing volume will continue to accumulate to the maximum for the site.

Simulation studies can assist the manager to find the best path through the sometimes contradictory requirements of maximizing sustained yield and achieving a non-declining even flow. Such studies require basic resource data (i.e. stand table and nett area) for each management unit within the forest estate and a good growth model.

Averaging Yields from Individual Plots

Hann (1980) observed that yields estimated from individual one hectare plots were more accurate than an estimate from the mean of these plots. This suggests that plots should be projected individually before results are averaged or summed. The easy alternative, projecting the average of all the plots, may introduce a bias. Suppose that the forest under consideration is overstocked on about half its area, and understocked elsewhere. Then the average of the plots will indicate a near optimal stocking, and will forecast yields higher than the average of the individual projections. However, if the forest is stratified into a number of units, each homogeneous with respect to site and stand composition, it should be possible to project the plot averages for any stratum. Hägglund (1981) also observed that for uniform, even-aged stands, projections of the stand mean provided the same estimate as averaged projections of each individual plot. However, these results apply only to homogeneous, well-managed stands inventoried using objective and accurate methods, and need not apply under other circumstances.

Moeur and Ek (1981) compared predictions from individual plots, averages for homogeneous stands, and averages for forest types. Their study involved 134 permanent plots each 0.058 hectares, established according to a systematic design with random starts. The homogeneous stands were defined from inspection of 1:16 000 scale aerial photographs, and forest types were classified according to the major overstorey species, as aspen, red pine or jack pine. Although all their projections overestimated the actual yield, the best predictions were obtained from the average of the individual plot predictions, and the greatest overestimates resulted from projecting the average for the forest type. These findings are consistent with other studies (e.g. Smith and Burkhart 1984).

In view of the expense involved in developing a growth model and obtaining resource data, it seems unwise to economize on computing costs by projecting plot averages rather than individual plots. If time and computing resources are limited, the average of homogeneous strata may be projected, but it is preferable to project the individual plots and to determine the average yield after the projections.

Cutting Cycle Analysis

Cutting cycle analysis has been extensively used for yield forecasting in uneven-aged forests, partly because of its ease of application and because it is the most reliable technique that can be performed without computers. The basic method (e.g. Davis and Johnson 1987) is to nominate a cutting cycle length, construct a typical stand table, project this stand to the mid-point of the cutting cycle, and apply a harvesting rule to determine the loggable volume. As the actual time of harvesting is not known, the mid-point of the cutting cycle is used as a compromise. The annual yield is determined by dividing the loggable volume by the cutting cycle length.

Cutting cycle analysis offers several advantages over alternatives such as area control (i.e. harvesting an equal area each year on a nominal rotation). As the method employs the current stand table and a harvesting rule that approximates field practice, it is able to predict yields available for harvesting. Continuing the analysis for several cycles indicates the long term yield, and the viability of the nominal cutting cycle length and harvesting rule (Fig. 12.1).

The method poses some questions which require subjective decisions and which may have a substantial impact on forecasts: What if the yields derived from successive cycles differ? Should the yield be set at the average of these, should the nominal cutting cycle length be altered, or should the harvesting rule be changed? Is the "typical" stand employed representative? The method can be improved by stratifying on site productivity and standing volume, and by simulating individual plots rather than stratum averages. Other deficiencies include the assumption of a fixed cutting cycle for all stands in the stratum, the assumption of harvesting at mid-cycle, and the implicit assumption that all stands will be cut in the same sequence in subsequent cutting cycles.

Yield Scheduling

Yield scheduling by heuristic (trial and error) simulation can overcome many of the deficiencies of traditional cutting cycle analysis. It attempts to emulate the sequence of harvesting across the resource, and may determine the most appropriate "cutting cycle" for each individual plot, rather than using the nominal cycle in cutting cycle analysis. Yield scheduling enables the user to specify any allowable cut and examine the outcome. In this way, it reveals for how long a given harvest can be sustained. Heuristic simulation does not provide the maximum sustainable yield, but takes an initial estimate by the user and provides information to enable the user to make a better estimate for a subsequent iteration (Davis and Johnson 1987, Leuschner 1990).

Grosenbaugh (1955) recognized the deficiencies in the horizontal cut and cutting cycle analysis approaches, and advocated the recognition of homogeneous "record-units" which were to be the sole area unit for all

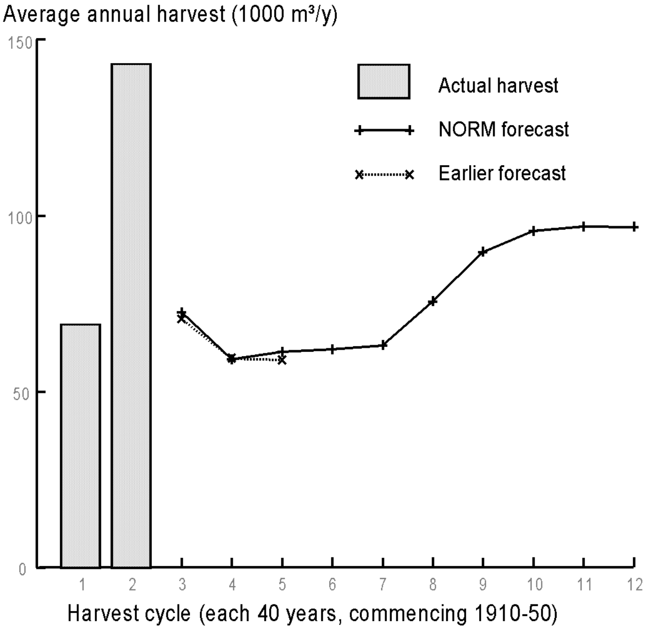


Fig. 12.1. Harvesting history (1910–89) in Queensland rainforests (Vanclay 1993), and forecasts (1990–2389) based on cutting cycle analysis (Preston and Vanclay 1988, Vanclay 1994).

mapping, sampling, forecasting and operational work. He also insisted that yield forecasting should recognize the actual order of working over the resource. Phillis (1971) reported a trial of Grosenbaugh's "diagnostic survey technique" in an irregular eucalypt forest, and reported that it was efficient in producing detailed yield estimates and other information useful for operational management at a cost comparable to the established continuous forest inventory system.

Vanclay (1994) gave an example of heuristic simulation with the NORM model to provide yield schedules for timber harvesting. The forest resource was partitioned into management units which formed the basis for management and prediction, and which were further stratified into homogeneous subunits for efficient sampling. The simulation system enabled several constraints to be specified, and included the specification of minimum yields per hectare, species mixes and other criteria, and this ensured that the predicted harvesting schedule made sense from a silvicultural and management viewpoint. Some discrepancies between yields predicted from cutting cycle analysis and yield scheduling were revealed, and can be attributed to the "half-cycle moratorium" implicit in cutting cycle analysis, which grows all stands to the mid-point of the cycle without simulating any harvests. This means that short-term yields may be overestimated, even though longer term estimates are reasonable. These predictions were rather similar to earlier medium-term predictions made with a much simpler growth model, less precise area data, and fewer static inventory data (Fig. 12.1). Provided that they are unbiased, model differences may be minor when combined with area and static inventory data in a long-term analysis of yields from a large forest area.

Heuristic simulation offers some advantages over the widely-used alternatives of linear programming and other mathematical programming techniques. It is conceptually very simple, and easy for forest managers to comprehend. It avoids many of the problems of mathematical programming methods, including the need to determine a suitable objective function (Haight 1987) and end condition (Davis and Johnson 1987). Johnson and Tedder (1983) discussed several advantages of heuristic simulation over linear programming, including the ability to portray inventory in greater detail, shift areas in and out of production more easily, to produce analyses at lower cost, and to find feasible solutions more easily.

Linear Programming

Linear programming and other mathematical programming techniques have been very useful and widely used for estimating and regulating timber harvests from plantations (e.g. Davis and Johnson 1987, Leuschner 1990, Hof 1993), but have been used to a lesser extent in natural forests. Basically, linear programming (LP) will sort through a list of choices and select one that satisfies the specified constraints and maximizes the objective (e.g. present net value). Thus an LP model requires three components: a set of decision variables, a set of constraints, and an objective function. In addition, it requires a matrix generator to formulate the list of choices, an LP package to solve the matrix, and a report writer to summarize the results in an intelligible form. Davis and Johnson (1987) give a good introduction to the topic. Advantages of linear programming include the ability to examine several alternatives simultaneously, to portray unusual yield trajectories, to accommodate more constraints, and to find the optimum.

FORPLAN is one of the most widely used LP models in forestry, and the USA Forest Service has a legal obligation to use it in forest management planning (Leuschner 1990). Yet despite this prominence in the field, it is complex, and is not well documented or used. Perhaps the major criticism of FORPLAN is that it has become an end in itself, rather than a means to investigate better forest management (Sedjo 1987). These limitations are not unique, but are common to many forestry planning systems. FORPLAN does more than timber harvest scheduling, and its major uses are in multiple use planning and in resolving conflict (e.g. McKenney 1990). Whilst FORPLAN and other mathematical programming systems are undoubtedly useful, they do not provide an easy option, and considerable skill and resource information are necessary to use these techniques.

Precision and Sensitivity

It is difficult to estimate the likely accuracy of a yield prediction. As Leary *et al*. (1979*a*) pointed out, there are two sources of error involved:

(a) error in assessing the initial state, and

(b) error in the growth prediction for the plot.

The former is a problem of resource inventory, and is dealt with in standard texts (e.g. Loetsch *et al*. 1973, de Vries 1986, Schreuder *et al*. 1993). It is significant however, and may contribute the majority of error associated with predictions (Mowrer and Frayer 1986).

Error propagation studies offer insights into the interaction of these factors and reveal the overall magnitude of errors in yield forecasts (Gertner 1987*a*, Mowrer 1991; see p. 210).

Risk Analyses

The use of stochastic models to study future yields raises three questions:

1. How many simulation runs are required in order to give a good indication of the most likely outcome and its distribution?
2. What is the best measure of the most likely outcome?
3. How should the confidence limits be estimated?

The number of simulations used to derive results published in the literature varies greatly, ranging from four (e.g. Ek and Monserud 1979) to 5000 (e.g. Gertner 1987*a*). If the distribution of yield predictions from the model is normal, the usual formulae may be used to calculate the mean, the variance and the required number of simulations. However, if the distribution of predictions is not normal, non-parametric estimates may be necessary to determine the most likely outcome and its variability.

In determining whether the distribution is likely to be normal, two approaches may be used. If, during the evaluation process, rigorous testing reveals that the distribution is always close to normal, it may be reasonable to assume that a normal distribution will generally result, and that the conventional statistics may be used. An alternative is to test for normality within the computer program during the execution of a series of simulations. This latter approach is likely to be unnecessarily intensive of computing resources, and it may be more efficient to assume a non-normal distribution in all cases. It is better to err on the conservative side, and assume that the distribution is not normal, particularly where one-sided long-tailed distributions may arise.

If the distribution is significantly different from the normal distribution, other approaches need to be used to determine the number of simulations, the most likely outcome and its variance. The median provides an unbiased estimate of the expected outcome, and is given by the $(n+1)/2^{\text{th}}$ observation of the sample, where n is an odd number and samples are ranked by expected outcome. Do say nine runs, then keep adding runs (using a power function such as 2^n+1) until no change in the median is observed. For a more precise result, keep adding more runs until the difference between the two observations either side of the median is small (i.e. $Y_{(n+3)/2} - Y_{(n-1)/2} < \delta$, where n is odd and Y_i are ranked in increasing order).

The variance of the population may be estimated as quarter of the range, but this test is not robust and performs poorly for non-normal distributions with long tails (Snedecor and Cochran 1980). A better approach may be to observe the confidence interval about the predictions directly from the simulated outcomes rather than from some formula. If, for example, the ninety percent confidence limits are required, the largest and smallest observations in 19 samples will give a rough estimate, the second largest and second smallest of 39 samples a better indication, as will the fifth largest and fifth smallest of 99 samples. For any continuous distribution, the probability that all of the next k observations are between the smallest and largest values in the current sample of size n is (Hahn and Nelson 1973):

$$1 - \alpha = \frac{n\,(n-1)}{(n+k)(n+k-1)}$$

The number of samples required may be dictated by two requirements: the need to gain a reasonable estimate of confidence limits for yield, and the need to estimate the median prediction with some precision. If reasonable estimates of confidence limits on the predicted yield are required, 30 or more simulations may be dictated. However, if concern is for the median, another approach may be used. Snedecor and Cochran (1980) gave a formula for confidence limits on the population median which is valid for any continuous distribution: $i, j = \{(n+1) \pm z\sqrt{n}\}/2$, where z is the normal deviate corresponding to the desired confidence probability, and the i^{th} and j^{th} observations indicate conservative confidence limits about the median. Other nonparametric tests are given in standard texts (e.g. Hettmansperger 1984).

Synthesis

Growth modellers cannot guess, at the time of model construction, all the possible uses to which a growth model may be put. This is why it is important to make sure that the model behaves in a realistic way for a wide range of site and stand conditions, and extrapolates safely to conditions not included in the development data. Some obvious applications of growth models have been discussed, and these techniques are likely to become increasingly important in forest management.

To fulfil this potential to assist forest managers by providing better information, growth models must not be made to appear sophisticated and complex, but must be made available for use on a regular basis as an everyday tool to improve forest management. In short, that means that the growth model should be integrated with other information systems, should provide information in a useful and flexible format, and must be easy to use, well documented and readily available.

Exercises

12.1. Discuss how you would integrate a growth model into a forest management information system. What data would you need to set up the system initially, and what on-going data collection procedures would you recommend to keep the system up-to-date?

12.2. Using the model defined by Equations 2.13 and 2.14 (p. 28), devise an optimal silvicultural regime for a loblolly pine plantation. What assumptions do you need to make to arrive at your conclusions? Are they realistic? Can you draw the same conclusions from the analogous matrix model in Fig. 2.5 (p. 29)?

12.3. Use the model in Fig. 3.4 (p. 49) to devise an optimal harvesting strategy for North American hardwoods. Many matrix models suggest optima that involve harvesting all stems in each of several size classes. It is less common for a matrix model to indicate harvesting of a proportion of the stems in a size class. Is this realistic, or is it an artifact of the model, or of the linear programming method sometimes used? Discuss.

12.4. Calculate the optimal diameter for harvesting suggested by the five different growth equations for sugar maple (Eqns 8.1–8.5 and Figs 8.2–8.3, p. 165–167). You may need the following height–diameter relationship (Botkin 1994): $h = 1.37 + 0.378\,d + 0.00111\,d^2$. Illustrate what happens to the optimum as the interval between harvests increases. How would you take mortality into account? Discuss the impact of assuming (i) constant mortality, and (ii) size-dependent mortality. What other factors should you consider to make your calculations more realistic? Discuss any assumptions that you make.

12.5. Use West's (1981) model (Exercise 11.2, p. 221) to examine silvicultural options for *Eucalyptus obliqua*, both in even-aged and uneven-aged stands. Are your solutions within the realistic bounds you defined in Exercises 11.2 and 11.3? Discuss any assumptions you make.

12.6 Use the models constructed in Exercises 11.4 and 11.5 (p. 222) to examine silvicultural options for *Flindersia pimenteliana*. Do the optima differ for the different modelling strategies? Explain. Is the "optimum silviculture" for this species biologically realistic, or is it an artifact of the model?

Chapter Thirteen

Future Directions

We conclude this book by examining what research is needed to enable models for mixed forests to meet the needs of forest planners and managers worldwide, and look toward to the future to suggest research avenues that may provide for the needs of future generations. The lead time needed to provide data and to develop techniques for modelling may be very long, and we have a responsibility to future generations to accommodate their needs. Existing models rely heavily on empirical data gathered in permanent plots established by our forebears, and we should consider what procedures we should put into place now to assist our successors.

Some of the most pressing challenges facing growth modellers are to devise ways to:

1. extrapolate existing models and techniques to sites and species for which no data are currently available;
2. devise models that can satisfy information required for different management regimes, including
  (a) different objectives (e.g. maintaining biodiversity),
  (b) multiple harvests (e.g. timber, fuel, fodder, fruits, resins, etc.), and
  (c) alternative silvicultural situations (e.g. agroforestry);
3. enhance the capability of models to provide reliable predictions under different environmental scenarios (e.g. nitrogen deposition, climate change);
4. integrate forest growth models more completely with other models at different scales (e.g. global models), with feedback in both directions;
5. develop new statistical procedures to model multi-variate systems more appropriately;
6. present model output and other resource data in more innovative ways.

Models for all Forests

It is ironic that the most sophisticated models exist only for the most simple forest ecosystems (pure even-aged plantations), and that models for the most complex of forests, the tropical moist forest, tend to be rather primitive or lacking. The technical difficulties of implementing growth models for rainforests can be overcome with sufficient time and resources, but practical difficulties may be less tractable. Limited resources and facilities hamper the efforts of researchers in the tropics to gather reliable data and build robust models, and may restrict their work to the simplest approaches. Complexity does not confer utility, and simple but reliable growth models may play an important role in demonstrating the consequences of various harvesting alternatives and the need for effective forest management. Sophisticated growth models are not always necessary, and some existing models, if re-calibrated, would be well suited to the task and could be instrumental in helping to reform tropical forest management. However, there is ample scope for improving models and extending existing methods to accommodate changing needs and conditions.

The greatest obstacle is that the present generation of forest-management-oriented growth models relies heavily on empirical data for model fitting, and for most of the world's forests, suitable data is simply not available, and forecasts for these forests can only be made by extrapolating models developed elsewhere. Careful consideration of functional relationships, and the use of eco-physiologically-based relationships where possible, may do much to provide reasonable extrapolations, but the reliability always remains in doubt. There are two immediate problems hindering model construction for forests in which no growth data are available:

1. estimating site productivity, and
2. estimating growth rates of tree species for which no data are available elsewhere.

Reliable methods for estimating site productivity in uneven-aged mixed forests rely on calibration to permanent plot data. Environmental variables and indicator species enable some extrapolation of these estimators, but it is unwise to extrapolate them long distances geographically without further calibration, as the chosen indicator species may be replaced by other species within relatively short distances, especially in moist tropical forests. There are some suggestions that general characteristics of plant groups (not necessarily taxonomic groups) known as *plant functional attributes* (Gillison 1988) may be useful in characterizing sites, and these may offer a way to extend the applicability of existing site classification methods. The application of emerging techniques for direct geocentric appraisal of site productivity (see Table 7.1, p. 135) is hampered by the shortage of environmental data, but new remote-sensing sensors may overcome these problems and research should continue in this area.

Moist tropical forests may have many species, but few for which empirical growth data exist, and general principles to extrapolate empirical growth functions to other species are urgently required. Taxonomy is no guide, and subjective appraisals of successional status may not provide a useful insight into growth patterns (see p. 127). Plant functional attributes may also be helpful in classifying species according to growth patterns, and may thus provide a basis for extrapolating empirical growth relationships. Further research on this aspect is warranted.

Different Management Needs

Most existing forest growth models focus on wood production, and are not concerned with other products that could be harvested or with other aspects of the forest environment. Notable exceptions include the JABOWA family of models (e.g. Botkin 1993) and its offspring, which address species succession, cycling of some nutrients and other aspects.

Increasing and different pressures on forests previously zoned for wood production will require greater consideration of non-wood products in management, silviculture and in modelling. The conservation aspects of all forest areas are becoming increasingly important, and models should provide details of stand structure and composition in a way that allows inferences about conservation to be assessed. Some useful inferences regarding diversity can be made from existing growth models (e.g. Hof and Joyce 1993, Buongiorno *et al*. 1994), but greater insights could be obtained with models specially adapted for this purpose. For instance, models could provide more details of "habitat" trees (previously known as "defective stags"; trees with hollows or other suitable nest-sites for birds and animals), and of *keystone* or *pivotal* species that may be critical food sources, pollinators or other vectors (e.g. Terborgh 1986). It is important to make such studies, because present objectives may be well-intentioned but infeasible (e.g. wildlife management constraints on mixed-conifer silviculture in USA, Haight *et al*. 1992).

Increasing pressure on existing forests will stimulate greater utilization and the harvesting of more products including fuel (branches), fodder (leaves), fruits and seeds, resins and other exudates, etc. This intensive level of harvesting is already visible in many tropical countries, but is likely to become even more common. This will place new demands on models, which may be required to resolve, for example,

(a) what will happen to nutrient dynamics with intense biomass harvesting?

(b) how many seeds can be harvested before regeneration dynamics are affected and the species composition of the forest is altered in an undesired way?

(c) what is the trade-off between foliage (or exudate) and wood production
(d) how much fodder can be harvested before lost wood (fuel or timber) production outweighs the value of the fodder?
(e) how are mortality trends affected by foliage harvesting?
(f) what impacts will high levels of harvesting have on fauna and non-target flora in the forest?

Traditional empirical models are not intended to resolve these questions and offer few insights into stand dynamics at this level. Some steps in this direction have been made in models such as FORCYTE (Kimmins *et al.* 1990) and LINKAGES (Pastor and Post 1986), which focus on carbon and nitrogen cycles, but further research is required. It is inevitable that models functioning at this resolution must become more physiologically-based, but unlike many of the current generation of process-based models, accurate estimates must remain a priority. Landsberg (1986) issued a challenge to forest growth modellers by defining a model as "a formal and precise statement or set of statements embodying our current knowledge or hypotheses about the working of a particular system and its responses to stimuli", and arguing that

> When such statements are made in mathematical terms it usually becomes clear that our knowledge is incomplete and assumptions have to be made about how parts of the system work. The consequences of these assumptions can be explored, either algebraically or numerically, and it must be possible to test them, and the model as a whole, experimentally. (By these criteria, conventional forestry models scarcely qualify as models; they are not hypotheses but descriptions of observations.) . . ..

Some forest growth modellers may find this assertion (regarding conventional forestry models) offensive, but the criticism is valid. Many forest growth models are excessively empirical, and a stronger basis for the functional relationships used in our models will help to (i) provide a better understanding of the processes involved, (ii) identify deficiencies in current knowledge and gaps in our empirical data, and should (iii) lead to models that extrapolate more safely to new situations. This does not imply that all model components must be physiologically-based, but rather that we should include eco-physiological (and other biological) elements when and where possible, but especially in the *key* components of models (Bossel 1991).

Increasing pressures will also mean that agroforestry and unconventional silvicultural regimes will provide an increasing share of traditional forest products and should also contribute towards conservation objectives. Growth models exist for some agroforestry situations, but the emerging need is for a model that performs adequately across much on the forest-agroforestry-agriculture continuum (e.g. Vandermeer 1989).

Increasing sophistication in temperate forest modelling has tended toward reductionist models of nutrient cycling and partitioning, but a better understanding of the tropical forest may require a more holistic approach embracing interspecific relationships, including fauna-flora interactions (e.g. Pannell, 1989, Prance 1992). Conventionally, a spatial model would account for the placement of trees within a small plot, but a spatial model for a tropical forest might consider the distance (i.e. kilometres) to the nearest individual of the same species and the implications for pollination, pests and disease. The term "ecosystem model" is sometimes used to model a small plot in great detail, but such a model for the tropical forest may require a much broader basis, including many fauna-flora interactions, intra-specific distances, and spatial data on intact forest fragments (What are the implications of size and distance of undisturbed refuges on pollinators, predators, pests, etc.?). Nutrient cycles may assume a greater importance in the moist tropics, where most of the nutrients may be held in the biomass and where cycling (especially decay) may be much more rapid. This is fertile ground for more research, and may provide important insights for conservation planning as well as for timber production.

Environmental Changes

Most empirical models based on permanent plot data implicitly assume that the future will be like the past, in terms of most environmental factors. While this assumption may be reasonable for short-term forecasts, the evidence against it increases steadily. There is strong evidence that levels of carbon dioxide and other "greenhouse" gasses in the atmosphere are increasing as a result of human activities, and it seems likely that this will change climate patterns (e.g. Wigley 1993). There are also substantial areas of forest affected by acid rain, nitrogen deposition, and other forms of pollution. These and other changes in the environment are likely to have a measurable effect on tree growth (e.g. Pretzsch 1992b). Nitrogen deposition affects tree growth in many parts of Europe, and there are some suggestions that human factors (perhaps CO_2 or other pollutants) are influencing survival and recruitment in tropical forests worldwide (Phillips and Gentry 1994).

Many of these effects are too large to be dismissed and should be accommodated in growth models. Some empirical models can be modified with scaling factors to account for these changes, but a more mechanistic approach seems necessary to account for interactions (e.g. CO_2, temperature, rainfall, etc.) in a reliable way. Some possible effects of carbon dioxide and other pollutants have been investigated with the JABOWA model (Botkin 1993), and although results have been expressed mainly in terms of species dynamics rather than production of wood and other forest products, it indicates one possible way to deal with environmental change. The most reliable way to deal with these and other aspects of environmental change is to incorporate more biological detail. Thus empirical and process models are likely to develop along convergent paths, with management-oriented models including more biological data, and process-based models employing more empirical data for model calibration to provide more accurate predictions. The current generation of process-models seems over-parameterized, creating problems for empirical estimation, and model simplification or additional process-level data will be necessary to progress with empirical calibration of process-based models. Conversely, many empirical models are excessively empirical, and would benefit from revisions which incorporate some of the knowledge obtained through experiments with process-based models.

Predictions regarding the impacts of potential climate change on forest production should be interpreted with caution, as it may be easier to predict the vegetation response to a specific regime than it is to predict the climate at a given time. Some studies have focused on plant responses to anticipated climatic means, but it may be the less predictable extremes (e.g., infrequent frosts, fires, floods and storms) that shape the vegetation. Climatic change may act indirectly via subtle changes in nutrient cycling and availability (e.g. Pastor and Post 1986), or via other feedback mechanisms (e.g. Myers 1991), and more sophisticated models are needed to examine these aspects.

Integration with other Models

Research on climate change relies heavily on general circulation models (CGMs) which simulate global atmospheric dynamics (energy, momentum and moisture) in three-dimensions with a resolution of a few hundred kilometres. The land surface may influence GCM simulations through (i) carbon uptake and release (photosynthesis and respiration), (ii) surface roughness which affects energy exchange by friction, (iii) absorption and reflection of energy, and (iv) evaporation (an energy transfer). With these few details, climate modellers attempt to examine the implications of various levels of carbon emissions, on the ecosystem and on mankind. In practice, there are many more interactions and feedbacks in the biosphere (e.g. Rambler *et al*. 1989), and a better understanding of the implications of climate requires, in part, more feedback loops between models of the biota and the CGMs. This is unlikely to be achieved in a single model, and the most fertile approach may be a hierarchical system of semi-autonomous models (e.g. O'Neill *et al*. 1986), within which forest growth models will be an important part. The challenge is to design and build growth models that can "handshake" with CGMs (probably via an intermediary such as a global vegetation model) and with models of other scales, providing them with suitable data and accepting, and responding in an appropriate way to feedback from them.

Model Construction and Use

Most models comprise several components which are fitted independently to data. However, these components and data are related, forming a multivariate system, and this should be taken into account during model fitting. Techniques are available to do this in some situations, but they need to be further developed and made more accessible to forest growth modellers. Special difficulties arise with recruitment models where it would be desirable to estimate logistic (i.e. probability of recruitment) and linear models (i.e. amount of recruitment, given that some occurs) simultaneously.

Model construction and use draws on many skills, but experts are not always available to assist model builders and users. Many of the skills and insights of experienced modellers could be incorporated into expert systems and made more widely available, and this could be an effective way to improve models and stimulate more effective use.

Presentation of Information

To have the greatest possible impact, information must be presented in a suitable format. It should be timely, concise and suitably detailed (neither too much nor too little detail). This requires considerable flexibility in the information systems to which the growth model is linked. It also requires imagination and innovation on the part of growth modellers and information technologists.

Most decision makers in forestry and other disciplines are inundated with paper, and much information is destined to be discarded unless it is presented in the most suitable way. Lobbyists and advertisers are aware of this, and take great care in presenting material, but many modellers and resource managers overlook this fact.

The technology exists to link growth models with other resource information (e.g. geographic information systems, inventory databases) interactively and to display outputs on a virtual reality (VR) interface. Model users could wear a VR headset, and take a magic carpet ride over their forest, stopping to study sites of interest, watching alternative silvicultural experiments unfold before their eyes, turning back the clock to try different alternatives, and observing several crop rotations on different sites to gauge the long-term effects. A VR interface could include non-visual outputs, allowing users to hear the birds, to smell the flowers, and feel the microclimate in the forest. Thus users could devise a management regime through their own experience, in a relatively short time. This may sound like science fiction, but the technology exists, and we could devise and implement models in this way if we thought it was appropriate (subject to sufficient financial and other resources).

Fancy technology will not improve an inadequate model, but it may make an adequate model more accessible to others. In particular, effective presentations may help foresters to give politicians and planners a better understanding of a situation, of the options for intervening, and of the implications of any action, and in this way may help to convince them to act on this information. The point of this example is to promote innovative ways to report inventory results and yield forecasts. Modellers should work with information technologists to devise better presentations, so that end-users can *visualize* information more easily (see e.g. Tufte 1983, 1990).

Synthesis

The challenge is to develop and enhance models to meet diverse needs, and to implement them in a flexible way so they can be used within an integrated system to examine the "big picture". The need to provide for "safe" extrapolations may be satisfied, in part, by greater use of process-based (and other biologically-based) relationships within the model. However, mechanistic models are not a panacea, and we should not expect them to magically solve the difficulty of accurate forecasts. A recent meeting of the GCTE (Global Change and Terrestrial Ecology project of the International Geosphere-Biosphere Programme) Wheat Network compared predictions from ten different mechanistic models of wheat growth, and found that their predictions varied three-fold (2.5–8.0 t ha^{-1}), even though dates of emergence, anthesis and maturity were prescribed (Steffen 1994). And this despite a short-rotation species amenable to experimentation !

Levins (1966) cautioned that biological models could not be general, realistic and accurate, but could only attain two of these goals. Hybrid models based on the best parts of mechanistic models and re-calibrated to empirical data (e.g. Sievänen and Burk 1993) may offer one way to approach these three goals. However, tree growth models will never be "perfect", as they are abstractions of open systems which are influenced by factors that cannot be accommodated in our models.

It is unlikely that there will be a single super-model satisfying all these diverse demands simultaneously, even though it is possible to provide expert systems offering the user problem-specific access to many different models within a single software system (Bossel 1991). Dynamic processes are shaped by the characteristic time-scales of their components, and may range from minutes (stomatal processes), to years (tree growth and senescence), to centuries (vegetation response to climate change). Diverse approaches will always be required, but the ability to integrate (or at least "handshake" in some way) the alternatives within a larger framework is desirable. In attempting to accommodate these many demands, we should not make models too complex. The need is for compact models which provide a valid structural representation with a few state variables, and which provide reasonable predictions over a wide range of conditions (Bossel 1991).

One of the challenges for modellers is to strengthen the biological and statistical basis of models and make them more widely applicable. Another is link them more efficiently with other information systems and provide information in more effective ways.

Clearly these goals will not be achieved overnight. It is not necessary or even desirable to try to do everything all at once. On the contrary, it is better to do a few things well than to do a lot inadequately. But we should remain aware of the possibilities and benefits of cooperation and integration with other modelling efforts. Start small, but think big and plan for the future. Begin with what can be implemented effectively now, but set up the framework for what should come later, working to an implementation plan, and consolidating each stage before progressing to the next.

An information broker has an important and influential role. A growth modeller deals with the future and thus has additional responsibilities, and should take special care to ensure that model forecasts provide the best possible information, presented in the best possible way.

References

Adams, D.M., 1976. A note on the interdependence of stand structure and best stocking in a selection forest. *For. Sci.* 22:180–184.

Adams, D.M. and Ek, A.R., 1974. Optimizing the management of uneven-aged forest stands. *Can. J. For. Res.* 4:274–287.

Adlard, P.G., 1990. Procedures for monitoring tree growth and site change. Oxford Forestry Institute, Tropical Forestry Papers 23. 188 p.

Aikman, D.P. and Watkinson, A.R., 1980. A model for growth and self thinning in even-aged monocultures of plants. *Annals of Botany* 45:419–427.

Aitkin, M., Anderson, D., Francis, B. and Hinde, J., 1989. *Statistical Modelling in GLIM*. Oxford Statistical Science Series 4, Oxford Science Publications, Clarendon Press, Oxford. 374 p.

Alder, D., 1979. A distance-independent tree model for exotic conifer plantations in east Africa. *For. Sci.* 25:59–71.

Alder, D., 1992. Simple methods for calculating minimum diameter and sustainable yield in mixed tropical forest. In F.R. Miller and K.L. Adam (eds) Wise Management of Tropical Forests. Proc. Oxford Conf. on Tropical Forests 1992, Oxford Forestry Institute, 288 p.

Alder, D., 1990. Ghafosim: A projection system for natural forest growth and yield in Ghana. Forestry Department, Ghana and Overseas Development Administration, U.K., July 1990, 114 p.

Alder, D. and Synnott, T.J., 1992. Permanent sample plot techniques for mixed tropical forest. Oxford Forestry Institute, Tropical Forestry Papers No 25. 124 p.

Alder, D., Synnott, T.J. and Smith J.P., 1977. GROPE: a standardized growth projection method for tropical rain forest. In E.F. Brünig (ed.) Transactions of the International MAB-IUFRO Workshop on Tropical Rainforest Ecosystems Research. Hamburg-Reinbeck 12–17 May 1977. Chair of World Forestry, Special Report No. 1, pp. 170–175.

Anscombe, F.J., 1973. Graphs in statistical analysis. *Am. Statist.* 27:17–21.

Arney, J.D., 1972. Computer simulation of Douglas-fir tree and stand growth. Ph.D. Thesis, Oregon State Univ., Corvallis. 88 p.

Arney, J.D., 1985. A modeling strategy for the growth projection of managed stands. *Can. J. For. Res.* 15:511–518.

Assmann, E., 1961. *Waldertragskunde*. BLV, München. (*The Principles of Forest Yield Study*, trans. S.H. Gardiner, 1970, Pergamon, Oxford, 506 p.)

Avery, T.E. and Burkhart, H.E., 1994. *Forest Measurements*, 4th ed. McGraw Hill, N.Y.

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 Dell, T.R., 1973. Quantifying diameter distributions with the Weibull function. *For. Sci.* 19:97–104.

Barclay, H.J. and Layton, C.R., 1990. Growth and mortality in managed Douglas fir: Relation to a competition index. *For. Ecol. Manage.* 36:187–204.

Bare, B.B. and Opalach, D., 1987. Optimizing species composition in uneven-aged forest stands. *For. Sci.* 33:958–970.

Baur, F.v., 1877. Die Fichte in Bezug auf Ertrag, Zuwachs und Form.

Beck, D.E. and Trousdell, K.B., 1973. Site index: accuracy of prediction. USDA For. Serv., Res. Pap. SE-108. 7 p.

Beetson, T., Nester, M. and Vanclay, J.K., 1992. Enhancing a permanent sample plot system in natural forests. *The Statistician* 41:525–538.

Belcher, D.W., Holdaway, M.R. and Brand, G.J., 1982. A description of STEMS: The stand and tree evaluation and modelling system. USDA For. Serv., Gen. Tech. Rep. NC-79. 18 p.

Bella, I.E., 1971. A new competition model for individual trees. *For. Sci.* 17:364–372.

Bertalanffy, L.v., 1941. Unterschungen uber die Gesetzlich beit des Wachstums. VII. Stoffwechseltypen und Wachstumstypen. *Biol. Zbl.* 61:510–532.

Bertalanffy, L.v., 1942. *Theoretische Biologie*, Band II. Borntraeger, Berlin. (2nd ed., Franke, Bern, 1951).

Bertalanffy, L.v., 1949. Problems of organic growth. *Nature* 163:156–158.

Bertalanffy, L.v., 1968. *General System Theory: Foundations, Development, Applications*. Braziller, N.Y.

Bevege, D.I., 1972. Girth and basal area growth of slash pine (*P. elliottii* var *elliottii* Engl.) and their significance to plantation management in Queensland. Qld. Dep. For. Res. Pap. 2. 69 p.

Biging, G.S.and Dobbertin, M., 1992. A comparison of distance-dependent competition measures for height and basal area growth of individual conifer trees. *For. Sci.* 38:695–720.

Binkley, C.S., 1980. Is succession in hardwood forests a stationary Markov process? *For. Sci.* 26:566–570.

Bonnor, G.M. and Morrier, K.C., 1981. Site Classification from air photos in a forest inventory. *For. Chron.* 57:265–266.

Borders, B.E., 1989. Systems of equations in forest stand modelling. *For. Sci.* 35:548–556.

Borders, B.E., Bailey, R.L. and Clutter, M.L., 1988. Forest growth models: Parameter estimation using real growth series. In A.R. Ek, S.R. Shifley and T.E. Burk (eds) Forest Growth Modelling and Prediction. Proc. IUFRO Conf., 23–27 Aug. 1987, Minneapolis, MN. USDA For. Serv., Gen. Tech. Rep. NC-120, pp. 660–667.

Bosch, C.A., 1971. Redwoods: a population model. *Science* 172:345–349.

Bossel, H., 1991. Modelling forest dynamics: moving from description to explanation. *For. Ecol. Manage.* 42:129–142.

Bossel, H. and Krieger, H., 1991. Simulation model of natural tropical forest dynamics. *Ecol. Modelling* 59:37–71.

Botkin, D.B., 1993. *Forest Dynamics: An ecological model*. Oxford University Press. 309 p.

Botkin, D.B., Janak, J.F. and Wallis, J.R., 1972. Some ecological consequences of a computer model of forest growth. *J. Ecol.* 60:849–872.

Bowling, E.H., Burkhart, H.E., Burk, T.E. and Beck, D.E., 1989. A stand-level multispecies growth model for Appalachian hardwoods. *Can. J. For. Res.* 19:405–412.

Box, G.E.P., 1966. Use and abuse of regression. *Technometrics* 8:625–629.

Bradley, R.T., Christie, J.M. and Johnston, D.R., 1966. *Forest Management Tables.* HMSO, London.

Bragg, C.T., 1988. Growth and conversion modelling: An integrated approach based on a crown development oriented tree growth model. Qld. Dep. For., Unpublished report.

Brand, G.J., 1981. GROW – A computer subroutine that projects the growth of trees in Lake State's forests. USDA For. Serv., Res. Pap. NC-207. 11 p.

Brewer, R., 1954. Soil parent material. *Aust. J. Sci.* 16:14.

Bruner, H.D. and Moser, J.W., 1973. A programmed Markov chain approach to prediction of diameter distribution in uneven-aged forest stands. *Can. J. For. Res.* 3:409–417.

Buchman, R.G. and Shifley, S.R., 1983. Guide to evaluating forest growth projection systems. *J. For.* 81:232–234.

Buckman, R.E., 1962. Growth and yield of red pine in MN. USDA For. Serv., Tech. Bull. 1272.

Buell, J.H., 1945. The prediction of growth in uneven-aged timber stands on the basis of diameter distributions. Duke Univ. Sch. For. Bull. 11. 70 p.

Bunnell, F.L., 1989. Alchemy and uncertainty: What good are models? USDA For. Serv., Gen. Tech. Rep. PNW-GTR-232.

Buongiorno, J. and Michie, B.R., 1980. A matrix model of uneven-aged forest management. *For. Sci.* 26:609–625.

Buongiorno, J., Dahir, S., Lu, H.-C. and Lin, C.-R., 1994. Tree size diversity and economic returns in uneven-aged forest stands. *For. Sci.* 40:83-103.

Burk, T.E., 1990. Prediction error evaluation: preliminary results. In L.C. Wensel and G.S. Biging (eds) Forest Simulation Systems. Proc. IUFRO Conf., Nov. 25 1988 Berkeley Calif. Univ. of Calif. Div. of Ag. and Nat. Res., Bull. 1927, pp. 81–88.

Burkhart, H.E. and Sprinz, P.T., 1984. Compatible cubic volume and basal area projection equations for thinned old-field loblolly pine plantations. *For. Sci.* 30:86–93.

Byrne, J.C. and Sweet, M.D., 1992. Managing data from remeasured plots: an evaluation of existing systems. USDA For. Serv., Res. Pap. INT-451. 26 p.

Cajander, A.K., 1909. Über Waldtypen. *Acta For. Fenn.* 1.

Cajander, A.K., 1949. Forest types and their significance. *Acta For. Fenn.* 56.

Campbell, D.G., 1989. Quantitative inventory of tropical forests. In D.G. Campbell and H.D. Hammond (eds) *Floristic Inventory of Tropical Countries.* New York Botanical Garden, pp. 524–533.

Campbell, D.G., Daly, D.C., Prance, G.T., and Maciel, U.N., 1986. Quantitative ecological inventory of terra firme and várzea tropical forest on the Rio Xingu, Brazilian Amazon. *Brittonia* 38:369–393.

Campbell, D.T. and Fiske, D.W., 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* 56:81–105.

Campbell, H., 1981. Pine Operations Thinnings Scheduler Technical Documentation: The Growth Model. W.A. For. Dep., Unpub. Manuscript. 29 p.

Campbell, R.G., Ferguson, I.S. and Opie, J.E., 1979. Simulating growth and yield of mountain ash stands: a deterministic model. *Aust. For. Res.* 9:189–202.

Cannell, M.G.R., Rothery, P. and Ford, E.D., 1984. Competition within stands of *Picea sitchensis* and *Pinus contorta*. *Annals of Botany* 53:349–362.

Canonizado, J.A., 1978. Simulation of selective forest management regimes. *Malay. Forester* 41:128–142.

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

Carmean, W.H., 1967. Soil survey refinements for predicting black oak site quality in southeastern Ohio. *Soil Sci. Soc. Am. Proc.* 31:805–810.

Carmean, W.H., 1973. Forest soils bibliography for the North Central Region. USDA For. Serv., Gen. Tech. Rep. NC-5.

Carmean, W.H., 1979. Soil-site factors affecting hardwood regeneration and growth. In H.A. Holt and B.C. Fischer (eds) Regenerating Oaks in Upland Hardwood Forests, Proc. 1979 John S. Wright Forestry Conf. Purdue Univ., pp. 61–74.

Carron, L.T., 1968. *An Outline of Forest Mensuration with Special Reference to Australia.* A.N.U. Press, Canberra. 224 p.

Cassell, R.F. and Moser, J.W., 1974. A programmed Markov model for predicting diameter distribution and species composition in uneven-aged forests. Purdue Univ. Agric. Exp. Sta. Res. Bull. 915. 43 p.

Chai, F.Y.C. and Sia, P.C., 1989. Standtable projections for a mixed swamp forest of Sarawak. In Wan Razali Mohd., H.T. Chan and S. Appanah (eds) Proceedings of the Seminar on Growth and Yield in Tropical Mixed/Moist Forests, 20–24 June 1988, Kuala Lumpur. Forest Research Institute Malaysia, pp. 60–77.

Chapman, D.G., 1961. Statistical problems in dynamics of exploited fisheries populations. In Proc. 4th Berkeley Symp. Math. Stat. and Prob., Univ. Calif. Press, Berkeley, pp. 153–168.

Chapman, H.H., 1942. The problem of forecasting growth in irregular stands. *J. For.* 40:805–806.

Clutter, J.L., 1963. Compatible growth and yield models for loblolly pine. *For. Sci.* 9:354–371.

Clutter, J.L., 1980. Forest Management opportunities for the future. In K.M. Brown and F.R. Clarke (eds) Forecasting Forest Stand Dynamics. Proc. Workshop, Thunder Bay, Ont., 24–25 June 1980. Sch. For., Lakehead Univ., pp. 246–255.

Clutter, J.L. and Allison B.J., 1974. A growth and yield model for *Pinus radiata* in New Zealand. In J. Fries (ed.) Growth Models for Tree and Stand Simulation. Proc. IUFRO Working Party S4.01-4 meetings, 1973. Dep. For. Yield Res., Royal Coll. For., Stockholm. Res. Note 30. 379 p.

Clutter, J.L. and Lenhart, J.D., 1968. Site index curves for old-field loblolly pine plantations in the Georgia Piedmont. Georgia Forestry Research Council, Report 22.

Clutter, J.L., Fortson, J.C., Pienaar, L.V., Brister, G.H. and Bailey, R.L., 1983. *Timber Management: A quantitative approach.* Wiley, N.Y. 333 p.

Corns, I.G.W. and Pluth, D.J., 1984. Vegetational indicators as independent variables in forest growth prediction in west-central Alberta, Canada. *For. Ecol. Manage.* 9:13–25.

Crawley, M.J., 1993. *GLIM for Ecologists.* Blackwell, Oxford, 379 p.

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

Curtis, R.O., 1983. Procedures for establishing and maintaining permanent plots for silvicultural and yield research. USDA For. Serv. Gen. Tech. Rep. PNW-155. 56 p.

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

Czarnowski, M.S., 1964. *Productive capability of locality as a function of soil and climate with particular reference to forest land.* Lousiana State Univ. Press, Baton Rouge. Biological Science Series No. 5. 174 p.

Czarnowski, M.S., Humphreys, F.R. and Gentle, S.W., 1976. Quantitative expression of productivity in Monterey pine plantations in Australia and New Zealand in terms of some soil and climate characteristics. IUFRO XV Congr., Florida, 1971. Sec. 21. p. 190–199.

Daniels, R.F. and Burkhart, H.E., 1988. An integrated system of forest stand models. *For. Ecol. Manage.* 23:159–177.

Daniels, R.F., Burkhart, H.E. and Strub, M.R., 1979. Yield estimates for loblolly pine plantations. *J. For.* 77:581–583.

Daniels, R.F., Burkhart, H.E. and Clason, T.R., 1986. A comparison of competition measures for predicting growth of loblolly pine trees. *Can. J. For. Res.* 16:1230–1237.

Daubenmire, R., 1961. Vegetative indications of rate of height growth in ponderosa pine. *For. Sci.* 7:24–34.

Daubenmire, R., 1976. The use of vegetation in assessing the productivity of forest lands. *Botanical Review* 42:115–143.

Davis, L.S. and Johnson, K.N., 1987. *Forest Management*, 3rd ed. McGraw-Hill, N.Y. 790 p.

Dawkins, H.C., 1958. The management of the natural tropical high forest with special reference to Uganda. Imperial For. Inst., Univ. of Oxford. Inst. Pap. 34. 155 p.

Dawkins, H.C. and Field, D.R.B., 1978. A long term surveillance system for British woodland vegetation. Commonw. For. Instit., Occas. Pap. 1.

Doyle, T.W., 1981. The role of disturbance in the gap dynamics of a montane rain forest: An application of a tropical forest succession model. In D.C. West, H.H. Shugart and D.B. Botkin (eds) *Forest Succession: Concepts and applications.* Springer Verlag, N.Y., pp. 56–73.

Draper, N.R. and Smith, H., 1981. *Applied Regression Analysis*, 2nd ed. Wiley, N.Y.

Drew, T.J. and Flewelling, J.W., 1977. Some recent Japanese theories of yield-density relationships and their application to Monterey pine plantations. *For. Sci.* 23:517–534.

Duerr, W.A. and Gevorkiantz, S.R., 1938. Growth prediction and site determination in uneven-aged timber stands. *J. Agric. Res.* 56:81–98.

Edwards, P.N. and Christie, J.M., 1981. Yield models for forest management. For. Comm. Booklet 48. 32 p.

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

Ek, A.R. and Monserud, R.A., 1974. Trials with program FOREST: growth and reproduction simulation for mixed species even- or uneven-aged forest stands. In J. Fries (ed.) Growth Models for Tree and Stand Simulation. Proc. IUFRO Working Party S4.01-4 meetings, 1973. Dep. For. Yield Res., Royal Coll. For., Stockholm. Res. Note 30, pp. 56–69.

Ek, A.R. and Monserud, R.A., 1979. Performance and comparison of stand growth models based on individual tree and diameter class growth. *Can. J. For. Res.* 9:231–244.

Ek, A.R., Shifley, S.R. and Burk, T.E. (eds), 1988. Forest Growth Modelling and Prediction. Proc. IUFRO Conf., 23–27 Aug. 1987, Minneapolis, MN. USDA For. Serv., Gen. Tech. Rep. NC-120. 1149 p.

Elston, D.A., 1982. Sensitivity analysis in the presence of correlated parameter estimates. *Ecol. Modelling* 64:11–22.

Enright, N.J. and Ogden, J., 1979. Applications of transition matrix models in forest dynamics: *Araucaria* in Papua New Guinea and *Nothofagus* in New Zealand. *Aust. J. Ecol.* 4:3–23.

Esser, G., 1984. The significance of biospheric carbon pools and fluxes for the atmospheric CO_2: a proposed model structure. *Prog. Biometeorol.* 3:253–294.

Evans, J., 1974. Some aspects of the growth of *Pinus patula* in Swaziland. *Commonw. For. Rev.* 53:57–62.

Farr, W.A. and Harris, A.S., 1979. Site index of sitka spruce along the Pacific coast related to latitude and temperatures. *For. Sci.* 25:145–153.

Ferguson, D.E. and Carlson, C.E., 1993. Predicting regeneration establishment with the Prognosis model. USDA For. Serv., Res. Pap. INT-476. 54 p.

Ferguson, D.E., Stage, A.R. and Boyd, R.J., 1986. Predicting regeneration in the grand fir-cedar-hemlock ecosystem of the northern rocky mountains. *For. Sci. Monog.* 26. 41 p.

Ferguson, I.S. and Leech, J.W., 1976. Stand dynamics and density in radiata pine plantations. *N.Z. J. For. Sci.* 6:443–454.

Fins, L. and Rust, M., 1987. Comparative costs of using an electronic data recorder and field forms. *West. J. Appl. For.* 2:28–30.

Ford, E.D. and Diggle, P.J., 1981. Competition for light in a plant monoculture modelled as a spatial stochastic process. *Annals of Botany* 48:481–500.

Ford, E.D. and Sorrensen, K.A., 1992. Theory and models of inter-plant competition as a spatial process. In D.L. DeAngelis and L.J. Gross (eds) *Individual-based Models and Approaches in Ecology*. Chapman & Hall, N.Y., p. 363–407.

Ford-Robertson, F.C. (ed.), 1977. *Terminology of Forest Science Technology Practice and Products* (with Addendum Number One by R.K. Winters, ed.). Society of American Foresters, Washington DC. 370 p.

Foster, R.W., 1959. Relation between site indices of eastern white pine and red maple. *For. Sci.* 5:279–291.

Fowler, G.W. and Arvanitis, L.G., 1979. Aspects of statistical bias due to the forest edge: fixed area circular plots. *Can. J. For. Res.* 9:383–389.

Fox, L., Brockhaus, J.A. and Tosta, N.D., 1985. Classification of timberland productivity in northwestern California using Landsat, topographic and ecological data. *Photogrammetric Engineering and Remote Sensing* 51:1745–1752.

Fralish, J.S. and Loucks, O.L., 1975. Site quality evaluation models for aspen (*Populus tremuloides* Mich.) in Wisconsin. *Can. J. For. Res.* 5:523–528.

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.

Furnival, G.M. and Wilson, R.W., 1971. Systems of equations for predicting forest growth and yield. In G.P. Patil, E.C. Pielou and W.E. Walters (eds) *Statistical Ecology*. Penn. State Univ. Press. Volume 3, p. 43–57.

Furnival, G.M. and Wilson, J.W., 1974. Regression by leaps and bounds. *Technometrics* 16:499–511.

Gadow, K.v., 1984. Die Erfassung von Durchmesserverteilungen in gleichaltrigen Kiefernbeständen. *Forstw. Cbl.* 103:360–374.

García, O., 1974. Sobre modelos matematicos de rodal. Instituto Forestal, Santiago, Chile, Informe Tecnico No. 48.

García, O., 1984. New class of growth models for even-aged stands: *Pinus radiata* in Golden Downs Forest. *N.Z. J. For. Sci.* 14:65–88.

García, O., 1994. The state-space approach in growth modelling. *Can. J. For. Res.*, in press.

Gertner, G., 1987*a*. Approximating precision in simulation projections: an efficient alternative to Monte Carlo methods. *For. Sci.* 33:230–239.

Gertner, G., 1987*b*. A procedure to evaluate sampling schemes for improving already calibrated models. *For. Sci.* 33:632–643.

Gertner, G.Z. and Dzialowy, P.J., 1984. Effects of measurement errors on an individual tree-based growth projection system. *Can. J. For. Res.* 14:311–316.

Gillison, A.N., 1988. A plant functional proforma for dynamic vegetation studies and natural resource surveys. CSIRO Instit. Nat. Res. and Environ., Tech. Memo. 88/3, 26 p.

Goelz, J.C.G. and Burk, T.E., 1992. Development of a well-behaved site index equation: jack pine in north central Ontario. *Can. J. For. Res.* 22:776–784.

Goodwin, A.N., 1988. Adaption of the STEMS growth model to Eucalypt forest. In A.R. Ek, S.R. Shifley and T.E. Burk (eds) Forest Growth Modelling and Prediction. Proc. IUFRO Conf., 23–27 Aug. 1987, Minneapolis, MN. USDA For. Serv., Gen. Tech. Rep. NC-120, pp. 76–85.

Goulding, C.J., 1979. Validation of growth models used in forest management. *N.Z. J. For.* 24:108–124.

Gower, J.C., 1967. A comparison of some methods of cluster analysis. *Biometrics* 23:626–637.

Grey, D.C., 1979. Site quality prediction for *Pinus patula* in the Glengarry area, Transkei. *S. Afr. For. J.* 111:44–48.

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

Grimes, R.F. and Pegg, R.E., 1979. Growth data for a spotted gum-ironbark forest in south-east Queensland. Qld Dep. For. Tech. Pap. 17. 30 p.

Grosenbaugh, L.R., 1955. Better diagnosis and prescription in southern forest management. USDA For. Serv., Occasional Paper 145. 27 p.

Grosenbaugh, L.R., 1967. REX – Fortran-4 system for combinatorial screening or conventional analysis of multivariate regressions. USDA For. Serv., Pac. S.W. Forest and Range Expt Sta.

Hafley, W.L. and Schreuder, H.T., 1977. Statistical distributions for fitting diameter and height data in even-aged stands. *Can. J. For. Res.* 7:481–487.

Hägglund, B., 1981. Forecasting growth and yield in established forests: an outline and analysis of the outcome of a subprogram within the HUGIN project. Swedish Univ. of Ag. Sc., Dep. For. Survey, Rep. 31. 145 p.

Hahn, G.J. and Nelson, W., 1973. A survey of prediction intervals and their application. *J. Qual. Technol.* 5:178–188.

Hahn, J.T. and Leary, R.A., 1979. Potential diameter growth functions. In: A generalized forest growth projection system applied to the lake states region. USDA For. Serv., Gen. Tech. Rep. NC-49, pp. 22–26.

Haight, R.G., 1985. A comparison of dynamic and static economic models of uneven-aged stand management. *For. Sci.* 31:957–974.

Haight, R.G., 1987. Evaluating the efficiency of even-aged and uneven-aged stand management. *For. Sci.* 33:116–134.

Haight, R.G., 1990. Feedback thinning policies for uneven-aged stand management with stochastic prices. *For. Sci.* 36:1015–1031.

Haight, R.G. and Monserud, R.A., 1990*a*. Optimizing any-aged management of mixed-species stands. I. Performance of a coordinate-search process. *Can. J. For. Res.* 20:15–25.

Haight, R.G. and Monserud, R.A., 1990*b*. Optimizing any-aged management of mixed-species stands: II. Effects of decision criteria. *For. Sci.* 36:125–144.

Haight, R.G., Monserud, R.A. and Chew, J.D., 1992. Optimal harvesting with stand density targets: managing Rocky Mountain conifer stands for multiple forest outputs. *For. Sci.* 38:554–574.

Hamilton, D.A., 1986. A logistic model of mortality in thinned and unthinned mixed conifer stands of northern Idaho. *For. Sci.* 32:989–1000.

Hamilton, D.A., 1990. Extending the range of applicability of an individual tree mortality model. *Can. J. For. Res.* 20:1212–1218.

Hamilton, D.A. and Brickell, J.E., 1983. Modeling methods for a two-state system with continuous responses. *Can. J. For. Res.* 13:1117–1121.

Hamilton, D.A. and Edwards, B.M., 1976. Modelling the probability of individual tree mortality. USDA For. Serv., Res. Pap. INT-185.

Hamilton, G.J., 1988. *Forest Mensuration Handbook*. For. Comm. Booklet No 39, HMSO, London, 274 p.

Hann, D.W., 1980. Development and evaluation of an even- and uneven-aged ponderosa pine/Arizone fescue stand simulator. USDA For. Serv., Res. Pap. INT-267. 95 p.

Hara, T., 1986. Effects of density and extinction coefficient on size variability in plant populations. *Annals of Botany* 57:885–892.

Harrison, T.P. and Michie, B.R., 1985. A generalized approach to the use of matrix growth models. *For. Sci.* 31:850–856.

Hart, H.M.J., 1928. Stamtal en dunning: Een oriënteerend onderzoek naar de beste plantwijdte en dunningswijze voor den djati. Departement van Landbouw, Nijverheid en Handel in Nederlandsch-Indië, Mededeelingen van het Proefstation voor het Boschwezen, No. 21, 219 p.

Hartshorn, G.S., 1975. A matrix model of tree population dynamics. In F.B. Golley and E. Medina (eds) *Tropical Ecological Systems*. Springer-Verlag, N.Y., pp. 41–51.

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

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

Herrick, A.M., 1938. A simple formula for making diameter predictions. *J. For.* 36:74–75.

Hett, M.M. and Loucks, O.L., 1968. Application of life table analysis to tree seedlings in Quetico Provincial Park, Ontario. *For. Chron.* 44(2):29–32.

Hettmansperger, T.P., 1984. *Statistical Inference Based on Ranks.* Wiley, N.Y. 323 p.

Higgins, M.D., 1977. A sustained yield study of north Queensland rainforests. Unpublished report, Qld Dept For., Brisbane, 182 p.

Hilt, D.E., 1983. Individual-tree diameter growth model for managed, even-aged, upland oak stands. USDA For. Serv., Res. Pap. NE-533. 15 p.

Hilt, D.E. and Teck, R.M., 1988. Individual-tree diameter growth model for northern New England. In A.R. Ek, S.R. Shifley and T.E. Burk (eds) Forest Growth Modelling and Prediction. Proc. IUFRO Conf., 23–27 Aug. 1987, Minneapolis, MN. USDA For. Serv., Gen. Tech. Rep. NC-120, pp. 86–93.

Hof, J.G., 1993. *Coactive Forest Management.* Academic, San Diego, CA, 189 p.

Hof, J.G. and Joyce, L.A., 1993. A mixed integer programming approach for spatially optimizing wildlife and timber in managed forest ecosystems. *For. Sci.* 39:816-834.

Holdaway, M.R., 1985. Adjusting STEMS growth model for Wisconsin forests. USDA For. Serv., Res. Pap. NC-267. 8 p.

Hool, J.N., 1966. A dynamic programming Markov chain approach to forest production control. *For. Sci. Monog.* 12. 26 p.

Howard, A.F. and Valerio, J., 1992. A diameter class growth model for assessing the sustainability of silvicultural prescriptions in natural tropical forests. *Commonw. For. Rev.* 71:171–177.

Hulst, R.van, 1979. On the dynamics of vegetation: Markov chains as models of succession. *Vegetatio* 40:3–14.

Husch, B., Miller, C.I. and Beers, T.W., 1982. *Forest Mensuration*, 3rd ed. Wiley, N.Y. 402 p.

Hutchinson, I.D., 1982. Field enumeration of permanent sample plots in the mixed dipterocarp forest of Sarawak. Forest Department, Kuching, FO:MAL/76/008 Field Document No. 16. 136 p.

Hyink, D.M. and Moser, J.W., 1979. Application of diameter distributions for yield projection in uneven-aged forests. In W.E. Frayer (ed.) Forest Resource Inventories. Proc. SAF/IUFRO Workshop, Colorado State Univ., Fort Collins, 23–26 July 1979. Dep. For. and Wood Sciences, Colorado State Univ., pp. 906–916.

Jackson, D.S., Gifford, H.H. and Chittenden, J., 1975. Environmental variables influencing the increment of *Pinus radiata* (2) Effects of seasonal drought on height and diameter increment. *N.Z. J. For. Sci.* 5:265–286.

Jardine, N. and Sibson, R., 1971. *Mathematical Taxonomy.* Wiley, N.Y.

Jeffers, J.N.R., 1978. *An Introduction to Systems Analysis: with ecological applications.* Edward Arnold, London, 198 p.

Johnson, E.W., 1973. Relationship between point density measurements and subsequent growth of southern pines. Ag. Exp. Sta., Auburn Univ., AL, Bull. 447. 109 p.

Johnson, K.N. and Tedder, P.L., 1983. Linear programming vs. binary search in periodic harvest level calculation. *For. Sci.* 29:569–582.

Johnston, D.R., Grayson, A.J. and Bradley, R.T., 1967. *Forest Planning.* Faber and Faber, London. 541 p.

Keenan, R.J. and Candy, S., 1983. Growth of young *Eucalyptus delegatensis* in relation to variation in site factors. *Aust. For. Res.* 13:197–205.

Kennard, R.W. and Stone, L., 1969. Computer aided design of experiments. *Technometrics* 11:137–148.

Kimmins, J.P., 1988. Community organization: methods of study and prediction of the productivity and yield of forest ecosystems. *Can. J. Bot.* 66:2654–2672.

Kimmins, J.P., Scoullar, K.A., Apps, M.J. and Kurz, W.A., 1990. The FORCYTE experience: a decade of model development. In B.J. Boughton and J.K. Samoil (eds) Forest Modelling Symposium, Proc. Symp. 13–15 March 1989, Saskatoon, Saskatchewan. Forestry Canada, Inf. Rep. NOR-X-308, pp. 60–67.

Korsgaard, S., 1989*a*. The standtable projection simulation model. In Wan Razali Mohd., H.T. Chan and S. Appanah (eds) Proceedings of the Seminar on Growth and Yield in Tropical Mixed/Moist Forests, 20–24 June 1988, Kuala Lumpur. Forest Research Institute Malaysia, Kepong, pp. 85–95.

Korsgaard, S., 1989*b*. The standtable projection simulation model. In H.E. Burkhart, M. Rauscher and K. Johann (eds) Artificial intelligence and growth models for forest management decisions. Proc. IUFRO Meeting, Vienna, 18–22 Sept 1989. VPI&SU, Blacksburg VA, Publ. FWS-1-89, pp. 43–59.

Krajicek, J.E., Brinkman, K.A. and Gingrich, S.F., 1961. Crown competition – a measure of density. *For. Sci.* 7:35–42.

Kramer, H., 1967. Zur Bonitierung in der Forsteinrichtung. Mitt. Forstl. Bundesversuchsanst. Wien 77:193–208.

Krug, A.G., Nordheim, E.V. and Giese, R.L., 1984. Determining initial values for parameters of a Weibull model: a case study. *For. Sci.* 30:573–581.

Landsberg, J.J., 1986. *Physiological Ecology of Forest Production.* Academic Press, London.

Lanly, J., 1981. Manual of forest inventory. FAO Forestry Paper 27. 200 p.

Larocque, G. and Marshall, P.L., 1988. Improving single-tree distance-dependent growth models. In A.R. Ek, S.R. Shifley and T.E. Burk (eds) Forest Growth Modelling and Prediction. Proc. IUFRO Conf., 23–27 Aug. 1987, Minneapolis, MN. USDA For. Serv., Gen. Tech. Rep. NC-120, pp. 94–101.

Leak, W.B., 1968. Birch regeneration: a stochastic model. USDA For. Serv., Res. Note NE-85. 7 p.

Leak, W.B., 1969. Stocking of northern hardwood regeneration based on exponential dropout rate. *For. Chron.* 45:344–347.

Leak, W.B. and Graber, R.E., 1976. Seedling input, death and growth in uneven-aged northern hardwoods. *Can. J. For. Res.* 6:368–374.

Leary, R.A., 1979. Design. In: A generalized forest growth projection system applied to the lake states region. USDA For. Serv., Gen. Tech. Rep. NC-49, pp. 5–15.

Leary, R.A., 1980. A design for survivor growth models. In K.M. Brown and F.R. Clarke (eds) Forecasting Forest Stand Dynamics. Proc. Workshop, Thunder Bay, Ont., 24–25 June 1980. Sch. For., Lakehead Univ., pp. 62–81.

Leary, R.A., 1985. *Interaction theory in forest ecology and management.* Martinus Nijhoff/Junk/Kluwer Academic, Dordrecht. 219 p.

Leary, R.A., 1988. Some factors that will affect the next generation of forest growth models. In A.R. Ek, S.R. Shifley and T.E. Burk (eds) Forest Growth Modeling and Prediction. Proceedings of IUFRO Conference, 24–28 Aug

1987, Minneapolis, MN. USDA For. Serv., Gen. Tech. Rep. NC-120, pp. 22–32.

Leary, R.A., 1991. Near-normal, empirical, and identity yield tables for estimating stand growth. *Can. J. For. Res.* 21:353–362.

Leary, R.A., and Holdaway, M.R., 1979. Modifier Function. In: A generalized forest growth projection system applied to the lake states region. USDA For. Serv., Gen. Tech. Rep. NC-49, pp. 31–38.

Leary, R.A., Hahn, J.T. and Buchman, R.G., 1979*a*. Tests. In: A generalized forest growth projection system applied to the lake states region. USDA For. Serv., Gen. Tech. Rep. NC-49, pp. 79–89.

Leary, R.A., Holdway, M.R. and Hahn, J.T., 1979*b*. Diameter growth allocation rule. In: A generalized forest growth projection system applied to the lake states region. USDA For. Serv., Gen. Tech. Rep. NC-49, pp. 39–46.

Leech, J.W., 1993. Modelling growth and yield for forest planning. East China Forest Inventory Design and Planning Institute, Jinhua, Zhejiang, China. FAO CPR/91/151 Field Document 4, 107 p.

Leech, J.W., Correll, R.L. and Aung Kyaw Myint, 1991. Use of Hotelling's T^2 and principal coordinate analysis to assist in aggregating species for volume table construction. *For. Ecol. Manage.* 40:279–288.

Lefkovitch, L.P., 1965. A study of population growth in organisms grouped by stages. *Biometrics* 21:1–18.

Leigh, E.G., Rand, A.S. and Windsor, D.M. (eds), 1982. *The Ecology of a Tropical Forest: Seasonal rhythms and long-term changes.* Smithsonian Institution Press, Washington DC.

Lemmon, P.E. and Schumacher, F.X., 1962. Stocking density around ponderosa pine trees. *For. Sci.* 8:397–402.

Lenhart, J.D., 1988. Evaluation of explicit and implicit yield prediction in loblolly and slash pine plantations in east Texas. In A.R. Ek, S.R. Shifley and T.E. Burk (eds) Forest Growth Modelling and Prediction. Proc. IUFRO Conf., 23–27 Aug. 1987, Minneapolis, MN. USDA For. Serv., Gen. Tech. Rep. NC-120, pp. 747–753.

Leslie, P.H., 1945. On the use of matrices in certain population mathematics. *Biometrika* 33:183–212.

Leslie, P.H., 1948. Some further notes on the use of matrices in population mathematics. *Biometrika* 35:213–245.

Leuschner, W.A., 1990. *Forest Regulation, Harvest Scheduling and Planning Techniques.* Wiley, N.Y. 281 p.

Levins, R., 1966. The strategy of model building in population biology. *Am. Sci.* 54:421–431.

Lewis, N.B., Keeves, A. and Leech, J.W., 1976. Yield Regulation in South Australia *Pinus radiata* Plantations. Woods and Forests Department, South Australia. Bulletin No. 23. 174 p.

Lieberman, D., 1982. Seasonality and phenology in a dry tropical forest in Ghana. *J. Ecol.* 70:791–806.

Lieth, H. and Box, E.O., 1972. Evapotranspiration and primary productivity. *Publications in Climatology* 25:37–46.

Liu, C.-M., Leuschner, W.A. and Burkhart, H.E., 1989. A production function analysis of loblolly pine yield equations. *For. Sci.* 35:775–788.

Loetsch, F., Zöhrer, F. and Haller, K.E., 1973. Forest Inventory, Volume 2. BLV München, 469 p.

Lonsdale, W.M., 1990. The self-thinning rule: dead or alive? *Ecology* 71:1373–1388.

Lorimer, C.G., 1983. Tests of age-independent competition indices for individual trees in natural hardwood stands. *For. Ecol. Manage.* 6:343-360.

Lowell, K.E. and Mitchell, R.J., 1987. Stand growth projection: Simultaneous estimation of growth and mortality using a single probabilistic function. *Can. J. For. Res.* 17:1466–1470.

Lund, H.G., 1986. A primer on integrating resource inventories. USDA For. Serv., Gen. Tech. Rep. WO-49. 64 p.

Lynch, T.B. and Moser, J.W., 1986. A growth model for mixed species stands. *For. Sci.* 32:697–706.

MacKinney, A.L. and Chaiken, L.E., 1939. Volume, yield and growth of loblolly pine in the mid-atlantic region. U.S. For. Serv., Appalachian For. Exp. Sta. Tech. Note 33. 30 p.

MacKinney, A.L., Schumacher, F.X. and Chaiken, L.E., 1937. Construction of yields tables for non-normal loblolly pine stands. *J. Agric. Res.* 54:531–545.

MacLean, C.D. and Bolsinger, C.L., 1973. Estimating productivity on sites with a low stocking capacity. USDA For. Serv., Res. Pap. PNW-152. 18 p.

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

Mäkelä, A., 1992. Process-oriented growth and yield models: Recent advances and future prospects. In T. Preuhsler (ed.) Research on Growth and Yield with emphasis on Mixed Stands. Proc. S4.01 "Mensuration, Growth and Yield" sessions at IUFRO Centennial Meeting, Berlin, 31 Aug.–4 Sept. 1992. Bayerische Forstliche Versuchs- und Forschungsanstalt, Freising, Germany, pp. 85–96.

Mallows, C.L., 1973. Some comments on C_p. *Technometrics* 15:661–675.

Mariaux, A., 1981. Past efforts in measuring age and annual growth in tropical trees. In F.H. Bormann and G. Berlyn (eds) Age and Growth Rate of Tropical Trees: New Directions for Research. Proc. of Workshop on Age and Growth rate Determination for Tropical Trees, held at Harvard Forest, Petersham, MA, 1–3 Apr. 1980. Yale Univ. Sch. For. and Env. Studies. Bulletin 94, pp. 20–30.

Marquardt, D.W., 1970. Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics* 12:591–612.

Martin, G.L. and Ek, A.R., 1984. A comparison of competition measures and growth models for predicting plantation red pine diameter and height growth. *For. Sci.* 30:731–743.

Martin, G.L., Ek, A.R. and Monserud, R.A., 1977. Control of plot edge bias in forest stand growth simulation models. *Can. J. For. Res.* 7:100–105.

Mawson, J.C., 1982. Diameter growth estimates on small forests. *J. For.* 80:217–219.

Mayer, D.G. and Butler, D.G., 1993. Statistical validation. *Ecol. Modelling* 68:21–32.

McDill, M.E. and Amateis, R.L., 1993. Fitting discrete-time dynamic models having any time interval. *For. Sci.* 39:499–519.

McGuiness, K.A., 1984. Equations and explanations in the study of species-area curves. *Biol. Rev.* 59:423–440.

McKenney, D.W., 1990. Multiple-use planning: An application of FORPLAN to an Australian forest. *Aust. For.* 52:113–123.

McMurtrie, R.E. and Wolf, L., 1983. Above- and below-ground growth of forest stands: a carbon budget model. *Annals of Botany* 52:437-448.

McMurtrie, R.E, Rook, D.A. and Kelliher, F.M., 1990. Modelling the yield of *Pinus radiata* on a site limited by water and nitrogen. *For. Ecol. Manage.* 30:381–413.

Meldahl, R.S., 1979. Yield projection methodology and analysis of hybrid poplars based on multispaced plots. Ph.D. Thesis, Univ. of Wisconsin, Madison.

Meldahl, R.S., Eriksson, M. and Thomas, C.E., 1985. A method for grouping species-forest type combinations for the development of growth models for mixed species stands. In E. Shoulders (ed.) Proc. of the 3rd biennial southern silvicultural research conference, 7–8 Nov 1984, Atlanta, GA. USDA For. Serv., Gen. Tech. Rep. SO-54, pp. 422–428.

Mendoza, G.A. and Gumpal, E.C., 1987. Growth projection of a selectively cut-over forest based on residual inventory. *For. Ecol. Manage.* 20:253–263.

Mendoza, G.A. and Setyarso, A., 1986. A transition matrix forest growth model for evaluating alternative harvesting schemes in Indonesia. *For. Ecol. Manage.* 15:219–228.

Meng, C.H., Tang, S.Z. and Burk, T.E., 1990. A stochastic restrictions regression model approach to volume equation estimation. *For. Sci.* 36:54–65.

Mervart, J., 1972. Growth and mortality rates in the natural high forest of western Nigeria. Nigeria For. Info. Bull. (New Series) No. 22. 28 p.

Meyer, H.A., 1940. A mathematical expression for height curves. *J. For.* 38:415–520.

Meyer, H.A., 1952. Structure, growth and drain in balanced uneven-aged forest. *J. For.* 50:85–92.

Meyer, W.H., 1930. Diameter distribution series in even-aged forest stands. Yale Univ., School of Forestry, Bulletin 28. 105 p.

Michie, B.R., 1985. Uneven-aged stand management and the value of forest land. *For. Sci.* 31:116–121.

Michie, B.R. and Buongiorno, J., 1984. Estimation of a matrix model of forest growth from re-measured permanent plots. *For. Ecol. Manage.* 8:127–135.

Miner, C.L., Walters, N.R. and Belli, M.L., 1988. A guide to the TWIGS program for the north central United States. USDA For. Serv., Gen. Tech. Rep. NC-125. 105 p.

Mitchell, K.J., 1969. Simulation of the growth of even-aged stands of white spruce. Yale Univ. Sch of For. Bull No. 75. 48 p.

Mitchell, K.J., 1975. Dynamics and simulated yield of Douglas-fir. *For. Sci. Monog.* 17. 39 p.

Mitchell, K.J., 1980. Distance dependent individual tree stand models: concepts and applications. In K.M. Brown and F.R. Clarke (eds) Forecasting Forest Stand Dynamics. Proc. Workshop, Thunder Bay, Ont., 24–25 June 1980. Sch. For., Lakehead Univ., pp. 100–137.

Mitchell, K.J., 1988. SYLVER: Modelling the impact of silviculture on yield, lumber value, and economic return. *For. Chron.* 64:127–131.

Mitscherlich, E.A., 1919. [Problems of plant growth]. *Landwirtschaftliche jahrbucher* 53:167–182. [In German].

Moeur, M. and Ek, A.R., 1981. Plot, stand and cover-type aggregation effects on projections with an individual tree based stand growth model. *Can. J. For. Res.* 11:309–315.

Monserud, R.A., 1976. Simulation of forest tree mortality. *For. Sci.* 22:438–444.

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., 1989. Optimizing single-tree simulators. In H.E. Burkhart, M. Rauscher and K. Johann (eds) Artificial intelligence and growth models for forest management decisions. Proc. IUFRO Meeting, Vienna, 18–22 Sept 1989. VPI&SU, Blacksburg VA, Publ. FWS-1-89, pp. 308–321.

Monserud, R.A. and Ek, A.R., 1974. Plot edge bias in forest growth simulation models. *Can. J. For. Res.* 4:419–423.

Monserud, R.A. and Ek, A.R., 1977. Prediction of understorey tree height growth in northern hardwood stands. *For. Sci.* 23:391–400.

Moore, D.M., Lees, B.G. and Davey, S.M., 1991. A new method for predicting vegetation distributions using decision tree analysis in a geographic information system. *Env. Manage.* 15:59–71.

Moore, J.A., Budelsky, C.A. and Schlesinger, R.C., 1973. A new index representing individual tree competitive status. *Can. J. For. Res.* 3:495–500.

Moser, J.W., 1972. Dynamics of an even-aged forest stand. *For. Sci.* 18:184–191.

Moser, J.W. and Hall, O.F., 1969. Deriving growth and yield functions for uneven-aged forest stands. *For. Sci.* 15:183–188.

Mosteller, F., Siegel, A.F., Trapido, E. and Youtz, C., 1981. Eye fitting straight lines. *Amer. Statist.* 35:150–152.

Mowrer, H.T., 1989. The effect of forest simulation model complexity on estimate precision. In H.E. Burkhart, M. Rauscher and K. Johann (eds) Artificial intelligence and growth models for forest management decisions. Proc. IUFRO Meeting, Vienna, 18–22 Sept 1989. VPI&SU, Blacksburg VA, Publ. FWS-1-89, pp. 100–109.

Mowrer, H.T., 1991. Estimating components of propagated variance in growth simulation model projections. *Can. J. For. Res.* 21:379–386.

Mowrer, H.T. and Frayer, W.E., 1986. Variance propagation in growth and yield projections. *Can. J. For. Res.* 16:1196–1200.

Murphy, P.A. and Farrar, R.M., 1982. Interim models for basal area and volume projection of uneven-aged loblolly – shortleaf pine stands. *South. J. Appl. For.* 6:115–119.

Myers, N. (ed.), 1991. *Tropical Forests and Climate.* Kluwer, Dordrecht, 265 p.

Nance, W.L., Grissom, J.E. and Smith, W.R., 1988. A new competition index based on weighted and constrained area potentially available. In A.R. Ek, S.R. Shifley and T.E. Burk (eds) Forest Growth Modelling and Prediction. Proc. IUFRO Conf., 23–27 Aug. 1987, Minneapolis, MN. USDA For. Serv., Gen. Tech. Rep. NC-120, pp. 134–142.

Nelson, T.C., 1963. Basal area growth of loblolly pine stands. USDA For. Serv., Res. Note SE-10. 4 p.

Newnham, R.M., 1964. The development of a stand model for Douglas fir. Ph.D. Thesis, Univ. of B.C., Vancouver. 201 p.

Newnham, R.M., 1966. Stand structure and diameter growth of individual trees in a young red pine stand. *Bi-monthly Research Notes* 22(4):4–5.

Newnham, R.M. and Smith, J.H.G., 1964. Development and testing of stand models for Douglas fir and lodgepole pine. *For. Chron.* 40:494–502.

Nielsen, P. (ed.), 1989. Research Report 1986 and 1987, Division of Technical Services (Research and Utilization). Qld Dep. For., Res. Rep. No. 6. 135 p.

Nix, H.A. and Gillison, A.N., 1985. Towards an operational framework for habitat and wildlife management. In J. Kikkawa (ed.) Wildlife Management in the Forests and Forestry-controlled Lands in the Tropics and the Southern Hemisphere. Proc. Workshop, Univ. of Qld, 16–18 July 1984, p. 39–55.

Norberg, R.A., 1988. Theory of growth geometry of plants and self-thinning of plant populations: geometric similarity, elastic similarity, and different growth modes of plant parts. *Am. Nat.* 131:220–250.

Oavika, F., 1990. Approach to growth models in tropical rainforests. In J.K. Vanclay, Permanent sample plot data for growth modelling. FAO:DP/PNG/86/009 Working Document No 7, 17 p.

Oderwald, R.G. Simplicity in forest models. In R.A. Leary (ed.) Quantity and Quality in Forest Research. Invited papers from 19th IUFRO World Congress, Montreal, 5–11 Aug. 1990. USDA For. Serv., Gen. Tech. Rep. NC-148, pp. 11–17.

Oderwald, R.G. and Hans, R.P., 1993. Corroborating models with model properties. *For. Ecol. Manage.* 62:271–283.

Omule, S.A.Y. and Macdonald, R.N., 1991. Simultaneous curve fitting for repeated height–diameter measurements. *Can. J. For. Res.* 21:1418–1422.

O'Neill, R.V., DeAngelis, D.L., Waide, J.B. and Allen, T.F.H., 1986. A Hierarchical Concept of Ecosystems. Monograph in Population Biology, Princeton Univ. Press, N.J., 253 p.

Opie, J.E., 1968. Predictability of individual tree growth using various definitions of competing basal area. *For. Sci.* 14:314–323.

Opie, J.E., 1972. STANDSIM – A general model for simulating the growth of even-aged stands. 3rd Conf. IUFRO Advisory Group of Forest Statisticans, Jouy-en-Josas, 7–11 Sept, 1970. Institut National de la Recherche Agronomique Publ. 72–3, pp. 217–240.

Oreskes, N., Shrader-Frechette, K. and Belitz, K., 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science* 263:641–645.

Osho, J.S.A., 1991. Matrix model for tree population projection in a tropical rain forest of south-western Nigeria. *Ecol. Modelling* 59:247–255.

Osmaston, H.A., 1956. Determination of age/girth and similar relationships in tropical forestry. *Empire For. Rev.* 35:193–197.

Osunkjoya, O.O., Ash, J.E., Hopkins, M.S. and Graham, A.W., 1992. Factors affecting survival of tree seedlings in north Queensland rainforests. *Oecologia* 91:569–578.

Öttelt, K.C., 1765. Praktischer Beweis, daß die Mathesis bei dem Forstwesen unentbehrliche Dienste tue. Eisenach.

Pannell, C.M., 1989. The role of animals in natural regeneration and the management of equatorial rain forests for conservation and timber production. *Commonw. For. Rev.* 68:309–313.

Parresol, B.R., 1993. Modeling multiplicative error variance: an example predicting tree diameter from stump dimensions in baldcypress. *For. Sci.* 39:670–679.

Pastor, J. and Post, W.M., 1986. Influence of climate, soil moisture and succession on forest carbon and nitrogen cycles. *Biogeochemistry* 2:3–27.

Payandeh, B., 1974. Some applications of non-linear regression models in forestry research. *For. Chron.* 59:244–248.

Peden, L.M., Williams, J.S. and Frayer, W.E., 1973. A Markov model for stand projection. *For. Sci.* 19:303–314.

Pelz, D.R., 1978. Estimating individual tree growth with tree polygons. In J. Fries, H.E. Burkhart and T.A. Max (eds) Growth Models for Long Term Forecasting of Timber Yields. VPI&SU, Blacksburg VA, FWS-1-78, p. 172–178.

Philip, M.S., 1994. *Measuring Trees and Forests*, 2nd ed. CAB International, Wallingford, U.K., 310 p.

Phillips, O. and Gentry, A., 1994. Increasing turnover through time in tropical forests. *Science* 263:954–958.

Phillis, K.J., 1971. Management inventory of irregular broad-leaved forests of Australia. Ph.D. Thesis, Australian National Univ., Canberra. 319 p.

Pienaar, L.V. and Turnbull, K.J., 1973. The Chapman-Richards generalization of von Bertalanffy's growth model for basal area growth and yield in even-aged stands. *For. Sci.* 19:2–22.

Pinto da Costa, M.E. and Preuhsler, T., 1994. Mixed stands: research plots, measurements and results, models. Proc. IUFRO Symp. 25–29 April 1994, Lousã, Portugal. Instituto Superior de Agronomia, Lisboa.

Popper, K.R., 1958. *The Logic of Scientific Discovery.* Hutchinson, London.

Prance, G.T., 1984. Completing the inventory. In V.H. Heywood and D.M. Moore (eds) *Current concepts in plant taxonomy.* Academic Press, London, pp. 365–397.

Prance, G.T., 1992. The dilemma of the Amazon rain forests: biological reserve or exploitable resource. In B. Cartledge (ed.) *Monitoring the Environment.* Oxford Univ. Press, p. 157–192.

Prentice, I.C. and Helmisaari, H., 1991. Silvics of north European trees: compilation, comparisons and implications for forest succession modelling. *For. Ecol. Manage.* 42:79–93.

Preston, R.A. and Vanclay, J.K., 1988. Calculation of Timber Yields from North Queensland Rainforests. Qld Dep. For. Tech. Pap. No. 47. 16 p.

Pretzsch, H., 1992*a*. Wuchsmodelle für Mischbestände als Herausforderung für die Waldwachstumforschung. *Forstw. Cbl.* 111:87–105.

Pretzsch, H., 1992*b*. Zunehmende Unstimmigkeit zwischen erwartetem und wirklichem Wachstum unserer Waldbestände. *Forstw. Cbl.* 111:366–382.

Prodan, M., 1953. Verteilung des Vorrates gleichaltriger Hochwaldbestände auf Durchmesserstufen. *Allg. Forst.- u. Jagdztg.* 129:15–33.

Pütter, A., 1920. [Investigation of physiological similarities of growth processes]. *Pflüg. Arch. ges. Physiol.* 180:298–340. [In German].

Rambler, M.B., Margulis, L. and Fester, R. (eds), 1989. *Global Ecology: Towards a science of the biosphere.* Academic, Boston.

Randall, B.L., Ek, A.R., Hahn, J.T. and Buchman, R.G., 1988. STEMS model projection capability with incomplete tree list input data. *North. J. Appl. For.* 5:190–194.

Ranneby, B. and Rovainen, E., 1994. On the determination of time intervals between remeasurements of permanent plots. *For. Ecol. Manage.* 71(3), in press.

Ratkowsky, D.A., 1983. *Nonlinear Regression Modeling.* Marcel Dekker, N.Y. & Basel. 276 p.

Ratkowsky, D.A., 1990. *Handbook of Nonlinear Regression Models.* Marcel Dekker, N.Y. & Basel. 241 p.

Rayner, M.E., 1992. Evaluation of six site classifications for modelling timber yield of regrowth karri (*Eucalyptus diversicolor* F. Muell.). *For. Ecol. Manage.* 54:315–336.

Rayner, M.E. and Turner, B.J., 1990. Growth and yield modelling of Australian eucalypt forests. I. Historical development. *Aust. For.* 53:224–237.

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

Reed, D.D., Liechty, H.O. and Burton, A.J., 1989. A simple procedure for mapping tree locations in forest stands. *For. Sci.* 35:657–662.

Reed, K.L., 1980. An ecological approach to modelling the growth of forest trees. *For. Sci.* 26:33–50.

Reineke, L.H., 1927. A modification of Bruce's method of preparing timber yield tables. *J. Agric. Res.* 35:843–856.

Reineke, L.H., 1933. Perfecting a stand density index for even-aged stands. *J. Agric. Res.* 46:627–638.

Reinhardt, E.D., 1982. The influence of site quality on the height–diameter relationship of western larch. Unpublished M.Sc. Thesis, Univ. of Montana. 42 p.

Reinhardt, E.D., 1983. Using height/diameter curves to estimate site index in old-growth western larch stands. Montana Forest and Conservation Expt Stn. School of For., Univ. of Montana. Res. Note 20. 3 p.

Rennolls, K., Geary, D.N. and Rollinson, T.J.D., 1985. Characterizing diameter distributions by the use of the Weibull distribution. *Forestry* 58:57–66.

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

Reynolds, M.R. and Chung, J., 1986. Regression methodology for estimating model prediction error. *Can. J. For. Res.* 16:931–938.

Reynolds, M.R., Burk, T.E. and Huang, W., 1988. Goodness-of-fit tests and model selection procedures for diameter distribution models. *For. Sci.* 34:373–399.

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

Roberts, M.R. and Hruska, A.J., 1986. Predicting diameter distributions: a test of the stationary Markov model. *Can. J. For. Res.* 16:130–135.

Robichaud, E. and Methven, I.R., 1993. The effect of site quality on the timing of stand breakup, tree longevity, and the maximum attainable height of black spruce. *Can. J. For. Res.* 23:1514–1519.

Romesburg, H.C. and Mohai, P., 1990. Improving the precision of tree height estimates. *Can. J. For. Res.* 20:1246–1250.

Rorres, C., 1978. A linear programming approach to the optimal sustainable harvesting of a forest. *J. Environ. Manage.* 6:245–254.

Rudolf, P.O., 1930. A comparison of several of the growth percent methods of predicting growth. *J. For.* 28:28–33.

Schmoldt, D.L., Martin, G.L. and Bockheim, J.G., 1985. Yield based measures of northern hardwood site quality and their correlation with soil-site factors. *For. Sci.* 31:209–219.

Schneider, 1853. Jahrbuch zum Forst- und Jagdkalender für Preußen.

Schönau, A.P.G., 1987. Problems in using vegetation or soil classification in determining site quality. *S. Afr. For. J.* 141:13–18.

Schönau, A.P.G. and Aldworth, W.J.K., 1991. Site evaluation in black wattle with special reference to soil factors. *S. Afr. For. J.* 156:35–43.

Schreuder, H.T. and Swank, W.T., 1974. Coniferous stands characterized with the Weibull distribution. *Can. J. For. Res.* 4:518–523.

Schreuder, H.T., Boos, A.D. and Hafley, W.L., 1979. Application of robust estimation techniques in forestry. In W.E. Frayer (ed.) Forest Resource Inventories. Proc. SAF/IUFRO Workshop, Colorado State Univ., Fort Collins, 23–26 July 1979. Dep. For. and Wood Sciences, Colorado State Univ., pp. 106–119.

Schreuder, H.T., Gregoire, T.G. and Wood, G.B., 1993. *Sampling Methods for Multiresource Forest Inventory.* Wiley, N.Y., 446 p.

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

Schumacher, F.X. and Chapman, R.A., 1954. Sampling Methods in Forestry and Range Management. Duke Univ. Sch. of For. Bull. 7 (revised). 222 p.

Sedjo, R.A., 1987. FORPLAN: An evaluation of a forest planning tool – a summary. In T.W. Hoekstra, A.A. Dyer and D.C. Le Master (eds), FORPLAN: An Evaluation of a Forest Planning Tool, Proc. Symp. Nov 4–6 1986, Denver CO. USDA For. Serv., Gen. Tech. Rep. RM-140, pp. 161–162.

Shifley, S.R., 1987. A generalized system of models forecasting Central States growth. USDA For. Serv., Res. Pap. NC-279. 10 p.

Shifley, S.R., Moser, J.W. and Brown, K.M., 1982. Growth and yield models for the elm-ash-cottonwood type in Indiana. USDA For. Serv., Res. Pap. NC-218. 16 p.

Shifley, S.R., Ek, A.R. and Burk, T.E., 1993. A generalized methodology for estimating forest ingrowth at multiple threshold diameters. *For. Sci.* 39:776–798.

Shoulders, E. and Tiarks, A.E., 1980. Predicting height and relative performance of major southern pines from rainfall slope and available soil moisture. *For. Sci.* 26:437–447.

Shugart, H.H., 1984. *A theory of forest dynamics.* Springer Verlag, N.Y.

Shugart, H.H. and West, D.C., 1977. Development of an Appalachian deciduous forest succession model and its application to assessment of impact of the chestnut blight. *J. Environ. Manage.* 5:161–180.

Shugart, H.H., Hopkins, M.S., Burgess, I.P. and Mortlock, A.T., 1980. The development of a succession model for subtropical rainforest and its application to assess the effects of timber harvest at Wiangaree State Forest, New South Wales. *J. Environ. Manage.* 11:243–265.

Sievänen, R. and Burk, T.E., 1993. Adjusting a process-based growth model for varying site conditions through parameter estimation. *Can. J. For. Res.* 23:1837–1851.

Sievänen, R., Burk, T.E. and Ek, A.R., 1988. Construction of a stand growth model utilizing photosynthesis and respiration relationships in individual trees. *Can. J. For. Res.* 18:1027–1035.

Silva, J.N.M., Carvalho, J.O.P.de, Lopes, J.do C.A., Almeida, B.F. de, Costa, D.H.M., Vanclay, J.K. and Skovsgaard, J.P., 1994. Growth and yield of a tropical rain forest of the Brazilian Amazon 13 years after logging. *For. Ecol. Manage.* 71(3), in press.

Simon, G., 1976. Computer simulation swindles, with applications to estimates of location and dispersion. *Appl. Statist.* 25:266–274.

Skovsgaard, J.P., 1994. [Structure and development of unthinned stands of sitka spruce (*Picea sitchensis* (Bong.) Carr.) in Denmark]. Thesis, Royal Vet. and Agric. Univ., Copenhagen. [In Danish with English abstract].

Smith, J.L. and Burkhart, H.E., 1984. A simulation study assessing the effect of sampling for predictor variable values on estimates of yield. *Can. J. For. Res.* 14:326–330.

Smith, N.J. and Hann, D.W., 1984. A new analytical model based on the −3/2 power rule of self-thinning. *Can. J. For. Res.* 14:605–609.

Smith, P.L., 1979. Splines as a useful and convenient statistical tool. *Amer. Statist.* 33:57–62.

Smith, W.B., 1983. Adjusting the STEMS regional forest growth model to improve local predictions. USDA For. Serv., Res. Note NC-297. 5 p.

Snedecor, G.W. and Cochran, W.G., 1980. *Statistical Methods*, 7th ed. Iowa Univ. Press. 507 p.

Snee, R.D., 1977. Validation of regression models: Methods and examples. *Technometrics* 19:415–428.

Soares, P., Tomé, M., Skovsgaard, J.P. and Vanclay, J.K., 1994. Evaluating a growth model for forest management using continuous forest inventory data. *For. Ecol. Manage.* 71(3), in press.

Soest, J.van, Ayral, P., Schober, R. and Hummel, F.C., 1965. The standardization of symbols in forest mensuration. Univ. of Maine, Maine Agric. Expt Stn, Tech. Bull. 15, 32 p.

Solomon, D.S., Hosmer, R.A. and Hayslett, H.T., 1986. A two stage matrix model for predicting growth of forest stands in the northeast. *Can. J. For. Res.* 16:521–528.

Spurr, S.H., 1952. *Forest Inventory.* Ronald, N.Y. 476 p.

Spurr, S.H., 1962*a*. A measure of point density. *For. Sci.* 8:85–96.

Spurr, S.H., 1962*b*. Growth and mortality of a 1925 planting of *Pinus radiata* on pumice. *N.Z. J. For.* 8:560–569.

Stacey, E.W. and Mihram, G.A., 1965. Parameter estimation for a generalized gamma distribution. *Technometrics* 7:349–358.

Stage, A.R., 1963. A mathematical approach to polymorphic site index curves for grand fir. *For. Sci.* 9:167–180.

Stage, A.R., 1973. Prognosis model for stand development. USDA For. Serv., Res. Pap. INT-137. 32 p.

Stage, A.R., 1976. An expression for the effect of aspect, slope and habitat type on tree growth. *For. Sci.* 22:457–460.

Stage, A.R. and Renner, D.L., 1988. Comparison of yield-forecasting techniques using long-term stand histories. In A.R. Ek, S.R. Shifley and T.E. Burk (eds) Forest Growth Modelling and Prediction. Proc. IUFRO Conf., 23–27 Aug. 1987, Minneapolis, MN. USDA For. Serv., Gen. Tech. Rep. NC-120, pp. 810–817.

Stage, A.R. and Wykoff, W.R., 1993. Calibrating a model of stochastic effects on diameter increment for individual-tree simulations of stand dynamics. *For. Sci.* 39:692–705.

Stage, A.R., Crookston, N.L. and Monserud, R.A., 1993. An aggregation algorithm for increasing the efficiency of population models. *Ecol. Modelling* 68:257–271.

Steffen, W., 1994. *GCTE News* No. 6, March 1994.

Sterba, H. and Monserud, R.A., 1993. The maximum density concept applied to uneven-aged mixed-species stands. *For. Sci.* 39:432–452.

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

Strub, M.R., Green, E.J., Burkhart, H.E. and Pirie, W.R., 1986. Merchantability of loblolly pine – An application of nonlinear regression with a discrete dependent variable. *For. Sci.* 32:254–261.

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

Swain, D.E. and Turner, B.J., 1988. Modelling the growth of jarrah (*Eucalyptus marginata*). In J.W. Leech, R.E. McMurtrie, P.W. West, R.D. Spencer and B.M. Spencer (eds) Modelling Trees, Stands and Forests. Proc. Workshop, Aug. 1985, Univ. of Melbourne. School of Forestry, Univ. of Melbourne, Bulletin No. 5, pp. 298–309.

Swaine, M.D. and Whitmore, T.C., 1988. On the definition of ecological species groups in tropical rain forests. *Vegetatio* 75:81–86.

Swaine, M.D., Hall, J.B. and Alexander, I.J., 1987. Tree population dynamics at Kade, Ghana (1968–1982). *J. Tropical Ecology* 3:331–345.

Terborgh, J., 1986. Keystone plant resources in the tropical forest. In Soule, M.E. (ed.) *Conservation Biology: the science of scarcity and diversity.* Sinauer, Sunderland, pp. 330–344.

Tham, Å., 1989. Prediction of individual tree growth in managed stands of mixed *Picea abies* (L.) Karst. and *Betula pendula* Roth & *Betula pubescens* Ehrh. *Scand. J. For. Res.* 4:491–512.

Tomboc, C.C., 1977. Growth, yield and economic rotation of Bagras (*Eucalyptus deglupta*) pulptimber in the PICOP plantations (Mindanao). Part 2: yield prediction model. *Sylvatrop Philipp. For. Res. J.* 2:117–126.

Tufte, E., 1983. *The Visual Display of Quantitative Information.* Graphics Press, Chesire CT. 197 p.

Tufte, E., 1990. *Envisioning Information.* Graphics Press, Cheshire CT. 126 p.

Turner, J., Thompson, C.H., Turvey, N.D., Hopmans, P. and Ryan, P.J., 1990. A soil technical classification for *Pinus radiata* (D. Don) plantations. I. Development. *Aust. J. Soil Res.* 28:797–811.

Turvey, N.D., Booth, T.H. and Ryan, P.J., 1990. A soil technical classification for *Pinus radiata* (D. Don) plantations. II. A basis for estimation of crop yield. *Aust. J. Soil Res.* 28:813–824.

Ure, J., 1950. The natural vegetation of the Kiangaroa as an indicator of site quality for exotic conifers. *N.Z. J. For.* 6:112–123.

Usher, M.B., 1966. A matrix approach to the management of renewable resources, with special reference to selection forests. *J. Appl. Ecol.* 3:355–367.

Usher, M.B., 1971. Developments in the Leslie matrix model. In J.N.R. Jeffers (ed.) *Mathematical Models in Ecology.* Blackwell Scientific, Oxford, pp. 29–60.

Usher, M.B., 1976. Extensions to models, used in renewable resource management, which incorporate an arbitrary structure. *J. Environ. Manage.* 4:123–140.

Valentine, H.T. and Campbell, R.W., 1975. A simulation model of gypsy moth - forest interaction. *For. Sci.* 21:233–238.

Valsta, L.T., 1988. Optimizing species composition in mixed even-aged stands. In A.R. Ek, S.R. Shifley and T.E. Burk (eds) Forest Growth Modelling and Prediction. Proc. IUFRO Conf., 23–27 Aug. 1987, Minneapolis, MN. USDA For. Serv., Gen. Tech. Rep. NC-120, pp. 913–920.

Vanclay, J.K., 1988*a*. A stand growth model for cypress pine. In J.W. Leech, R.E. McMurtrie, P.W. West, R.D. Spencer and B.M. Spencer (eds) Modelling Trees, Stands and Forests. Proc. Workshop, Aug. 1985, Univ. of Melbourne. School of Forestry, Univ. of Melbourne, Bulletin No. 5, pp. 310–332.

Vanclay, J.K., 1988*b*. PLATIPUS Physiology: Design of a plantation growth and conversion simulator incorporating silviculture and wood quality. In A.R. Ek, S.R. Shifley and T.E. Burk (eds) Forest Growth Modelling and Prediction. Proc. IUFRO Conf., 23–27 Aug. 1987, Minneapolis, MN. USDA For. Serv., Gen. Tech. Rep. NC-120, pp. 998–1005.

Vanclay, J.K., 1989*a*. A growth model for north Queensland rainforests. *For. Ecol. Manage.* 27:245–271.

Vanclay, J.K., 1989*b*. Modelling selection harvesting in tropical rain forests. *J. Tropical Forest Science* 1:280–294.

Vanclay, J.K., 1989*c*. Site productivity assessment in rainforests: an objective approach using indicator species. In Wan Razali Mohd., H.T. Chan and S. Appanah (eds) Proceedings of the Seminar on Growth and Yield in Tropical Mixed/Moist Forests, 20–24 June 1988, Kuala Lumpur. Forest Research Institute Malaysia, Kepong, pp. 225–241.

Vanclay, J.K., 1991*a*. Modelling changes in the merchantability of individual trees in tropical rainforest. *Commonw. For. Rev.* 70:105–111.

Vanclay, J.K., 1991*b*. Aggregating tree species to develop diameter increment equations for tropical rainforests. *For. Ecol. Manage.* 42:143–168.

Vanclay, J.K., 1991*c*. Mortality functions for north Queensland rainforests. *J. Tropical Forest Science* 4:15–36.

Vanclay, J.K., 1991*d*. Compatible deterministic and stochastic predictions by probabilistic modelling of individual trees. *For. Sci.* 37:1656–1663.

Vanclay, J.K., 1992. Modelling regeneration and recruitment in a tropical rainforest. *Can. J. For. Res.* 22:1235–1248.

Vanclay, J.K., 1993. Lessons from the Queensland rainforest: A century striving for sustainability. In P.J. Wood, J.K. Vanclay and Wan Razali Wan Mohd (eds) The Quest for Sustainability: 100 years of silviculture and management in the tropics. Proceedings of Tropical Silviculture Workshop, IUFRO Centennial conference, Berlin, 1–3 September, 1992. Forest Research Institute Malaysia, Kepong, pp. 69–89.

Vanclay, J.K., 1994. Sustainable timber harvesting: Simulation studies in the tropical rainforests of north Queensland. *For. Ecol. Manage.* 69(1–3), in press.

Vanclay, J.K. and Henry, N.B., 1988. Assessing site productivity of indigenous cypress pine forest in southern Queensland. *Commonw. For. Rev.* 67:53–64.

Vanclay, J.K. and Preston, R.A., 1989. Sustainable timber harvesting in the rainforests of northern Queensland. In Forest Planning for People, Proc. 13th biennial conf. Institute of Foresters of Australia, Leura, NSW, 18–22 Sept. 1989. IFA Inc, Sydney, pp. 181–191.

Vanclay, J.K. and Preston, R.A., 1990. Utility of Landsat Thematic Mapper data for mapping site productivity in tropical moist forests. *Photogrammetric Engineering and Remote Sensing* 56:1383–1388.

Vandermeer, J., 1989. *The Ecology of Intercropping.* Cambridge Univ. Press, U.K., 237 p.

Vries, P.G.de, 1986. *Sampling Theory for Forest Inventory: A teach-yourself course.* Springer, Berlin, 399 p.

Vuokila, Y., 1965. Functions for variable density yield tables of pine based on temporary sample plots. *Commun. Inst. For. Fenn.* 60(4):1–86.

Wadsworth, F.H., Parresol, B.R. and Colon, J.C.F., 1989. Tree increment indicators in a subtropical wet forest. In Wan Razali Mohd., H.T. Chan and S. Appanah (eds) Proceedings of the Seminar on Growth and Yield in Tropical Mixed/Moist Forests, 20–24 June 1988, Kuala Lumpur. Forest Research Institute Malaysia, Kepong, pp. 205–213.

Wahlenberg, W.G., 1941. Methods for forecasting timber growth in irregular stands. USDA Tech. Bull. 796. 55 p.

Walters, J.W., Hinds, T.E., Johnson, D.W. and Beatty, J., 1982. Effects of partial clearing on diseases, mortality and regeneration of Rocky Mountain aspen stands. USDA For. Serv., Res. Pap. RM-240. 12 p.

Wan Razali Mohd., 1989. Modelling tree mortality in mixed tropical forests of peninsular Malaysia. In Wan Razali Mohd., H.T. Chan and S. Appanah (eds) Proceedings of the Seminar on Growth and Yield in Tropical Mixed/Moist Forests, 20–24 June 1988, Kuala Lumpur. Forest Research Institute Malaysia, Kepong, pp. 96–105.

Warren, W.G., 1981. Basic statistical methods in forestry research: use, misuse and prognosis. In Proc. XVII IUFRO Congress, Japan. Div 6, pp. 108–120.

Watson, R.M. and Nimmo, J.M., 1992. The use of permanent plots for monitoring vegetation change in Somalia: Problems, opportunities, recommendations. In H.G. Lund, R. Päivinen and S. Thammincha (eds) Remote Sensing and Permanent Plot Techniques for World Forest Monitoring. Proc. IUFRO Workshop, 13–17 Jan. 1992, Pattaya, Thailand. Faculty of Forestry, Kasetsart Univ., Bangkok, pp. 149–156.

Webb, L.J. and Tracey, J.G., 1967. An ecological guide to new planting areas and site potential for hoop pine. *Aust. For.* 31:224–240.

Webb, L.J., Tracey, J.G., Williams, W.T. and Lance, G.N., 1967. Studies in numerical analysis of complex rainforest communities. I. A comparison of methods applicable to site/species data. *J. Ecol.* 55:171–191.

Webb, L.J., Tracey, J.G., Williams, W.T. and Lance, G.N., 1970. Studies in numerical analysis of complex rainforest communities. V. A comparison of the properties of floristic and physiognomic-structural data. *J. Ecol.* 58:203–232.

Weber, L.A., Ek, A.R. and Droessler, T.D., 1986. Comparison of stochastic and deterministic mortality estimation in an individual tree based stand growth model. *Can. J. For. Res.* 16:1139–1141.

Weck, J., 1951. Wachstumsdiagnosen. *Forstarchiv* 22:135–138.

Weck, J., 1955. *Forstliche Zuwachs- und Ertragskunde.* Neumann, Radebeul, Germany, 160 p.

Weibull, W., 1951. A statistical distribution function of wide applicability. *J. Appl. Mech.* 18:293–296.

Weisberg, S., 1985. *Applied Linear Regression*, 2nd. ed. Wiley, N.Y., 324 p.

Wenk, G., Antanaitis, V. and Ŝmelko, Ŝ., 1990. *Waldertragslehre*. Deutscher Landwirtschaftsverlag, Berlin, 448 p.

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

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., 1987. A model for biomass growth of individual trees in forest monoculture. *Annals of Botany* 60:571–577.

West, P.W., 1993. Model of above-ground assimilate partitioning and growth of individual trees in even-aged forest monoculture. *J. Theor. Biol.* 161:369-394.

West, P.W., 1994. Application of regression analysis to inventory data with measurements on successive occasions. *For. Ecol. Manage.* 71(3), in press.

West, P.W., Ratkowsky, D.A. and Davis, A.W., 1984. Problems of hypothesis testing of regressions with multiple measurements from individual sampling units. *For. Ecol. Manage.* 7:207–224.

West, P.W., Stocker, G.C. and Unwin, G.L., 1988. Environmental relationships and floristic and structural change in some unlogged tropical rainforest plots of north Queensland. *Proc. Ecol. Soc. Aust.* 15:49–60.

Westoby, M., 1984. The self-thinning rule. *Advances in Ecological Research* 14:167–225.

White, J., 1981. The allometric interpretation of the self-thinning rule. *J. Theoretical Biology* 89:475–500.

Whitmore, T.C., 1989. Guidelines to avoid remeasurement problems in permanent sample plots in tropical rain forests. *Biotropica* 21:282–283.

Whitmore, T.C., 1990. *An Introduction to Tropical Rainforests*. Clarendon, Oxford. 226 p.

Wigley, T.M.L., 1993. Climate change and forestry. *Commonw. For. Rev.* 72:256–264.

Wilson, F.G., 1951. Control of stocking in even-aged stands of conifers. *J. For.* 49:692–695.

Wood, G.B., 1990. Choosing a portable data recorder. *Aust. For.* 53:173–181.

Wright, R.G. and van Dyne, G.M., 1971. Comparative analytical studies of site factor equations. In G.P. Patil, E.C. Pielou and W.E. Walters (eds) *Statistical Ecology*, Volume 3. Penn. State Univ. Press, pp. 59–105.

Wyatt-Smith, J. and Vincent, A.J., 1962. The swing from qualitative to quantitative assessment of individual tree crown parameters in the Malayan Forest Service. *Malay. Forester* 25:276–291.

Wykoff, W.R., 1986. Supplement to the user's guide for the stand prognosis model – Version 5.0. USDA For. Serv., Gen. Tech. Rep. INT-208. 36 p.

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., Crookston, N.L. and Stage, A.R., 1982. User's guide to the stand prognosis model. USDA For. Serv., Gen. Tech. Rep. INT-133. 112 p.

Yoda, K., Kira, T., Ogawa, H. and Hozami, K., 1963. Self thinning in overcrowded pure stands under cultivated and natural conditions. *J. Biol. Osaka City Univ.* 14:107–129.

Zeide, B., 1980. Plot size optimization. *For. Sci.* 26:251–257.

Zeide, B., 1987. Analysis of the 3/2 power law of self-thinning. *For. Sci.* 33:149–166.

Zeide, B., 1989. Accuracy of equations describing diameter growth. *Can. J. For. Res.* 19:1283–1286.

Zeide, B., 1993. Analysis of growth equations. *For. Sci.* 39:594–616.

Glossary

Symbols

These symbols are based on the IUFRO standard (van Soest *et al.* 1965). Note that capital letters refer to totals per unit area, while lower case symbols generally refer to individual tree values. Greek letters are used to denote unknown population values and estimated coefficients in equations.

α, β, γ	Parameters to be estimated.
a_{leaf}	Leaf area of a tree (m^2).
b	Biomass of an individual tree.
c_a	Horizontal projection of crown area (m^2).
C_{cf}	Crown competition factor, $C_{cf} = 100\Sigma(\pi c_{wo}^2 /4)/10\ 000$, the potential crown area per hectare, expressed as a percentage. Canopy closure occurs at about $C_{cf}=100$, and high values indicate more competition.
c_l	Live crown length, the distance from the top of the tree to the lowest live branch.
c_r	Crown ratio, the ratio of full live crown length to total tree height.
c_w	Crown width (m).
c_{wo}	Crown width of an open-grown tree with the same diameter as the tree under consideration. Usually estimated as $c_{wo}=\beta_0+\beta_1 d^{\alpha}$.
Δd	Change in d (or any other variable) in a specified period. The notations ∂d and $\mathrm{d}d/\mathrm{d}t$ are also used where the function can be integrated to give a yield equation.
Δd_{max}	Expected maximum diameter increment (cm y^{-1}) for a given species–stand–site combination (i.e. the maximum on the Δd vs d curve).
D	Sum of diameters in a stand, $D=\Sigma d$.

d	Diameter at breast height (1.3 m) or above buttress, over bark, of an individual tree (cm).
$\overline{d}$	Arithmetic mean diameter, $\overline{d} = \Sigma d/N$
d_c	Binary variable which takes the value 1 if the tree's diameter d exceeds the specified girth limit for harvesting (cutting), and 0 otherwise.
d_g	Quadratic mean diameter, $d_g = \sqrt{(\Sigma d^2/N)}$
d_{max}	Maximum diameter that a given species is expected to attain.
$d_{0.2}$	Diameter at one fifth of tree height.
e	Natural number, $e \approx 2.71828$
ϵ	Error in estimate, $\epsilon = Y - \hat{Y}$.
E	Environmental factors, including aspect (E_a), elevation (E_e) and slope (E_s).
f	Form factor, $f = v/gh$
G	Stand basal area $G = \Sigma g$ (m^{2}ha^{-1}).
g	Basal area of an individual tree $g = \pi d^2/40\,000$ (m^2).
$G_{>d}$	Basal area of trees larger than the subject tree (m^{2}ha^{-1}).
h	Height of an individual tree (m).
$\overline{h}$	Mean height (m) of trees within a defined area.
h_b	Breast height, 1.3 m or 4.5 ft above ground level.
h_{max}	Maximum stand height (m), estimated from the biggest trees in a stand, or as the asymptote of a height–diameter curve.
h_T	Top height, the average height of a defined number of thickest trees per unit area (m).
k	The constant $\pi/40\,000$ used to convert diameter to basal area, $g = kd^2$.
l	Index of competition based on the estimated light available to an individual tree.
ln	Natural logarithm to base e.
$\log_{10}$	Logarithm to base 10.
$\boldsymbol{M}$	A transition matrix. Bold symbols refer to matrices and vectors.
N	Number of trees within a defined area (ha^{-1}).
N_r	Number of recruits at a specified threshold size, r, (ha^{-1}).
n	Number of items in a sample.
p	Probability that an individual tree will grow to the next size class, die, be harvested, or experience some other event.
s	An estimate of the standard error σ.
S	An estimate of site productivity.
$S_{\Delta d}$	An index of site productivity indicated by the average growth rate (adjusted for competition) of individual trees of several selected species, called growth index (Vanclay 1989c).
S_h	An index of site productivity based on the mean height of trees in the stand (m).

$S_{h,d}$	An index of site productivity given by the expected tree height (m) at some index diameter, sometimes called site form (Vanclay and Henry 1988).
$S_{h,t}$	Site index, an index of site productivity given by the expected top (or predominant) height (m) at some index age.
S_{soil}	Binary variable which takes the value 1 when the soil type *soil* occurs on the plot, and 0 otherwise.
t	Time or age (years).
t_l	Time since last logging (years).
t_s	Time since last silvicultural treatment (years).
V	Volume over bark of all the trees within a defined area (m^{3}ha^{-1}).
v	Volume over bark of a single tree (m^3).
v_u	Volume under bark of a single tree (m^3).
X	Explanatory variable (sometimes called an independent variable).
Y	Response variable (sometimes called a dependent variable).
$\hat{Y}$	Estimated value of a response variable, $Y=\hat{Y}+e$.
Z_{spp}	Binary variable which takes the value 1 if the species *spp* occurs on the plot, and 0 otherwise.

Common and Botanical Species Names

alder, blush	*Sloanea australis*
alder, buff	*Apodytes brachystylis*
ash	*Fraxinus* spp.
ash, alpine	*Eucalyptus delegatensis*
ash, mountain	*Eucalyptus regnans*
ash, red	*Alphitonia whitei*
ash, silver	*Flindersia bourjotiana*
aspen	*Populus tremuloides*
aspen, lemon	*Acronychia acidula*
bagras	*Eucalyptus deglupta*
basswood, ivory	*Polyscias australiana*
basswood, white	*Polyscias murrayi*
beech	*Fagus* spp.
beech, southern	*Nothofagus* spp.
birch	*Betula* spp.
bleedingheart, native	*Omalanthus populifolius*
bollywood	*Litsea leefeana*
boxwood, Macintyre's	*Xanthophyllum octandrum*
butternut, rose	*Blepharocarya involucrigera*
cedar, eastern red	*Juniperus virginiana*
cherry	*Prunus* spp.
cottonwood	*Populus deltoides*

cudgerie, brown	*Canarium baileyanum*
dipterocarp	Dipterocarpaceae
elm	*Ulmus* spp.
fig	*Ficus* spp.
fir, Douglas	*Pseudotsuga menziesii*
fir, white	*Abies concolor*
gardenia, brown	*Randia fitzalanii*
gum, spotted	*Eucalyptus maculata*
gum, varnished	*Eucalyptus vernicosa*
heather	*Calluna* spp.
jarrah	*Eucalyptus marginata*
kamarere	*Eucalyptus deglupta*
larch, western	*Larix occidentalis*
laurel, cinnamon	*Cryptocarya cinnamomifolia*
laurel, ivory	*Cryptocarya angulata*
laurel, rusty	*Cryptocarya mackinnoniana*
mahogany, American	*Swietenia* spp.
maple	*Acer* spp.
maple, Queensland	*Flindersia brayleyana*
maple, red	*Acer rubrum*
maple silkwood	*Flindersia pimenteliana*
maple, sugar	*Acer saccharum*
meranti	*Shorea* spp.
milkwood, hard	*Alstonia muellerana*
oak, blush silky	*Bleasdalea bleasdalei, Opisthiolepis heterophylla*
oak, brown silky	*Darlingia darlingiana*
oak, brown tulip	*Argyrodendron trifoliolatum*
oak, northern silky	*Cardwellia sublimis*
pine, cypress	*Callitris* spp.
pine, hoop	*Araucaria cunninghamii*
pine, jack	*Pinus banksiana*
pine, klinkii	*Araucaria huntsteinii*
pine, loblolly	*Pinus taeda*
pine, lodgepole	*Pinus contorta*
pine, maritime	*Pinus pinaster*
pine, Mexican	*Pinus patula*
pine, ponderosa	*Pinus ponderosa*
pine, radiata	*Pinus radiata*
pine, red	*Pinus resinosa*
pine, Scots	*Pinus sylvestris*
pine, shortleaf	*Pinus echinata*
pine, slash	*Pinus elliottii*
pine, white	*Pinus strobus*
quandong, tropical	*Elaeocarpus largiflorens*

rapanea	*Rapanea achradifolia*
redwood	*Sequoia sempervirens*
rubber tree	*Hevea braziliensis*
sal	*Shorea robusta*
salwood, brown	*Acacia aulacocarpa, A. mangium*
salmon bean	*Archidendron vaillantii*
satinash, Kuranda	*Syzygium kuranda*
satinash, white Eungella	*Syzygium wesa*
spruce, white	*Picea glauca*
sterculia, tulip	*Sterculia laurifolia*
stringybark, messmate	*Eucalyptus obliqua*
tamarind, brown	*Castanospora alphandii*
tamarind, pink	*Toechima erythrocarpum*
vitex	*Vitex acuminata*
walnut, blush	*Beilschmiedia obtusifolia*
walnut, yellow	*Beilschmiedia bancroftii*
wattle	*Acacia* spp.

Terminology

These definitions are based largely on the IUFRO-SAF standard (Ford-Robertson 1977), except where otherwise indicated.

Accretion: survivor growth plus the increment on trees that died between the first and second measures. Also called gross growth of initial volume.

Accurate: how close an estimate is to the true value, thus implying precision and freedom from bias.

Allometric: a relationship which maintains constant proportions so that $Y = \alpha X^{\beta}$ or $\ln Y = \ln \alpha + \beta \ln X$ (e.g. the volume of a cube is the cube of its side, irrespective of its size).

Anamorphic curves: a series curves or equations scaled so that each is a simple constant times the base curve. For example, the high curve might be 1.2 times the average, and the low curve might 0.8 times the average (cf. polymorphic).

Basal area of a tree (g): the cross-sectional area of a tree stem (including the bark) at breast height. The basal area of a stand (G) is the sum of the cross-sectional areas at breast height of all trees on a defined area.

Bias: the difference between the expectation of a sample estimator and the true population value, systematically distorting results. May arise in sampling, measurement or estimation due to poor calibration of an instrument (e.g. a stretched tape), or by favouring (perhaps unintentionally) one outcome over others.

Binary variable: a variable that takes the value zero or one.

Breast height: standardized point for measuring trees, usually 1.3 m or 4.5 ft above ground level, depending on the country.

Cohort: a group of individuals which are similar in some respect. For modelling, a cohort of trees would normally be the same species, and have similar size.

Continuous forest inventory (CFI): a method of dynamic sampling using permanent plots which are often systematically located and invisibly marked so that they are treated the same as the rest of the forest.

Crown competition factor: an index of competition based on the potential open-grown crown area of a tree, expressed as a percentage, $C_{cf} = 100\Sigma(\pi c_{wo}^2/4)/10000$. Canopy closure occurs at about $C_{cf} = 100$, and high values indicate more competition (Krajicek *et al.* 1961).

Crown ratio: the ratio of full live crown length to total tree height.

Cutting cycle: the planned, recurring lapse of time between successive harvests in a forest stand.

Dbh (*d*): diameter at breast height (1.3 m or 4.5 ft, depending on country) over bark. Trees with buttresses are usually above the buttress instead of a breast height. Dbh and other diameters may be measured with a calliper or a diameter tape.

Deliquescent: tree growth habit usually found in "broad-leaved" trees, where the stem is not continuous from the ground to the tip of the tree, but merges into the crown either gradually or abruptly. Cf. excurrent = coniferous.

Diameter tape: A tape measure graduated in π units (e.g. cm) so that the diameter may be read from the tape when it is placed around the circumference of the stem.

Dominant: one of 4 crown classes (dominant, co-dominant, intermediate, suppressed) based on relative status of forest trees. Dominant trees have their crowns in the upper part of the canopy and are largely free-growing.

Dynamic inventory: inventory on successive occasions to detect change, usually by sampling with permanent plots.

Even-aged: forest stand composed of trees of approximately the same age (i.e. generally the maximum range in age should be less than 10 years).

Expansion factor: the number of trees (per hectare or per plot) represented by a single record during simulation by a tree list model.

Experiment: A planned inquiry to obtain new information or to confirm or refute an hypothesis or the outcome of a previous experiment, using a formal procedure (the experiment design) to control factors which may influence the outcome.

Explanatory variable: a variable which is used as a basis to predict other details which may be more difficult to measure. E.g. diameter may be an explanatory variable in a volume equation to predict tree volume. Also known as an independent or regressor variable.

Extrapolation: prediction beyond the range of the calibration data (cf. interpolation).

Fecundity: the number of individuals produced for each mature tree in a matrix model (cf. recruitment).

Forest: a plant community predominantly of trees and other woody vegetation, growing more or less closely together.

Girth: the measurement around the stem of a tree or log. When you measure a tree with a diameter tape, you measure the girth to get an estimate of the diameter.

Growth: the change in size of an individual or stand of trees (cf. yield).

Harvesting: refers to both *logging* (felling and extraction of timber) and *thinning* (felling in an immature crop primarily to stimulate growth of the residual trees). I refer to thinning in immature pure, even-aged stands, and to logging in uneven-aged mixed stands.

Heterogeneous: implies that members of a sample or population differ to a greater or lesser extent, in respect of some or all parameters of interest (cf. homogeneous).

Heteroscedastic: heterogeneous variance.

Homogeneous: implies that all the members of a sample or population are similar in respect of some or all of their parameters. E.g. homogenized milk is treated so that it retains the same composition throughout (i.e. the cream doesn't float to the top). Cf. heterogeneous.

Ingrowth: individuals entering a size class in a size class model (cf. upgrowth, recruitment).

Interpolation: prediction within the range of calibration data (cf. extrapolation).

Inventory: measuring and recording the number, size, condition, etc. of one or more species in a forest, generally above a specified size limit, either by total enumeration, or by sampling using plots.

Light-demanding: tree species which are intolerant of shade and require sunlight to grow satisfactorily.

Logging: felling and extraction of timber from a forest.

Markov assumption: that the transition probabilities in a Markov chain depend only on the state of the system and not on any external factors or past events.

Markov chain: a representation of a system as a finite number of discrete states.

Mixed forest: forest comprising two or more species, generally with the major species comprising <80% of a stand.

Model: a simplified representation of some aspect of reality (not to be confused with the normative meaning of the word, something worthy of being imitated). Specifically in a statistical context, a model is a formalized expression of a theory. Generally, a growth model may include a series of mathematical equations, the numerical values embedded in those equations, the logic necessary to link these equations in a meaningful way, and the computer code necessary to implement the model on a computer.

Movement ratio: the proportion of individuals in a class advancing to the next size class, calculated as expected increment divided by class width of a size class model (cf. upgrowth).

Natural basal area: the maximum or limiting basal area which an undisturbed stand will tend towards. This limit may be site-specific, so it serves as a measure of site productivity.

Natural forest: forest established by natural regeneration (cf. plantation). In this book, refers to forests managed with selection harvesting systems which maintain an uneven-aged structure.

Parameter: a quantitative characteristic of an individual or population. E.g. the mean, variance, and constants describing an equation fitted to data.

Pioneer: a plant species capable of invading bare sites and persisting there until displaced by succession. Thus early successional stands may be dominated by these species. Typically, these species have seeds that are light and easily dispersed by wind or animals, and individuals are often fast-growing and short-lived.

Plantation: a forest established artificially by sowing or planting (cf. natural forest). This book assumes that plantations are even-aged and single species, even though this is not always true in practice.

Point sample: a form of sampling with probability proportional to size (PPS) using a horizontal variable radius plot defined by a central point and an angle gauge.

Polymorphic curves: series of curves of different shapes (cf. anamorphic). Typically used for modelling height growth, when these curves increase monotonically and do not cross.

Precise: the degree to which estimates are clustered about their own average, or the repeatability of a measurement or estimate.

Predominant height: the average total height of a specified number of the tallest trees per unit area. Often determined as the mean of the tallest tree on each of several non-overlapping sampling units. Cf. top height.

Pure: stand composed principally of one species (i.e. over 80% by stem number, basal area or volume).

Qualitative variable: a variable that may take $n+1$ possible values (e.g. 1, 2, 3; A, B, C; or dominant, intermediate, suppressed; etc.), and which is represented in regression analyses by a set of n binary variables z, with $z_i = 1$ when the qualitative variable takes the value i, and zero otherwise.

Rainforest: evergreen forest at least 30 m high, rich in epiphytes and woody lianas, occurring in areas where temperature, humidity and rainfall are high throughout the year (e.g. over 18°C and over 100 mm rain each month, see e.g. Whitmore 1990).

Recruitment: trees that have entered a particular size category, usually the smallest in the model, during a specified period. Also known as ingrowth, especially in size class models.

Regeneration (natural): renewal of a forest stand by self-sown seed or vegetative means including coppice, suckers and lignotubers. Artificial regeneration included sowing and planting.

Response variable: a variable predicted from one or more explanatory variables. Also known as a dependent variable.

Sample: a subset of a population used to estimate characteristics of that population.

Serial correlation: correlation between successive observations on the same sample unit. E.g. if a tree grows faster than average during one period, it is likely to continue to be faster during subsequent periods, so errors from the expected growth trends (e=y-ŷ) will generally be positive.

Shade tolerant: tree species able to survive and grow under shade.

Silviculture: the science and art of cultivating forest crops (cf. agriculture). More specifically, the theory and practice of controlling the establishment, composition, structure and growth of forests.

Site class: an objective classification of site productivity into several classes (cf. site quality).

Site index: the top or predominant height of an even-aged forest stand at a specified index age, often estimated for stands of other ages using a height–age curve.

Site productivity: the potential of a specified species and site to produce wood. Note that it is the inherent capability of the site, which may not be achieved by all silvicultural regimes. Site index, site quality and site class are approximate measures of the true site productivity.

Site quality: a subjective appraisal of site productivity, often by a visual assessment into relative (i.e. good-poor) classes (cf. site class).

Stand: a group of trees having sufficient uniformity in composition and spatial arrangement to constitute a silvicultural entity or sampling unit.

Stand table: a table showing the number of trees by species and diameter classes, generally per unit area of a stand.

Static inventory: inventory on a single occasion to record the present state of the forest.

Stationary assumption: assumption that the transition probabilities in a Markov chain do not change over time.

Stratum: a subdivision of the population which is more homogeneous with respect to the variable of interest than the population as a whole. Plural is strata.

Survivor growth: growth on trees that are alive (and above a specified measurement limit) at both the first and second measures on a permanent plot. Does not include recruitment. Differs from accretion in that it does not include increment on trees that died between the two measures.

Thinning: the felling of trees in an immature stand, primarily to stimulate growth of the residual trees. May include both non-commercial (the thinned trees may be too small or defective to have any commercial value) and commercial thinning.

Top height: the average total height of a specified number of the thickest trees per unit area. Cf. predominant height.

Total height (h): the vertical distance from the ground to the highest point on the tree.

Uneven-aged: stand composed of intermingling trees that differ markedly in age.

Upgrowth: individuals moving from one size class to another in a size class model (cf. movement ratio, ingrowth).

Usher assumption: defining the states of a matrix model or Markov chain so that individuals can progress at most, one size class during a single time-step.

Variable: a characteristic that may vary from one individual in a population to another, and which relates to some property of the individual. E.g. height and diameter of trees.

Variance: a measure of the variability of a sample or population.

Voucher specimen: a plant sample taken to assist identification, for lodgement in an herbarium. Specimens should, if possible, include several leaves joined to a branchlet, any fertile material (buds, flowers and fruits), and any other distinctive parts (e.g. bark). These should be pressed, dried and mounted on paper, or preserved in alcohol. All parts should be clearly labelled with tree and plot numbers, location, date and collector.

Yield: the final size of an individual or group of individuals at the end of a given period (cf. growth).

Yield class: a measure of site productivity based on the maximum mean annual volume increment.

Solutions to Selected Exercises

Chapter 2

2.1. It is unlikely that whole stand models will prove useful in forests where there are many tree species and sizes, unless users are happy to have the species lumped into a few groups. The main limitation is that the stand has to be adequately described with a few parameters, both for input to the model, and as model output (estimates of these parameters at some future date). If suitable however, the data requirements are not too demanding, and models can be built from remeasured plot data in which the individual trees are not identified. Of course, it is preferable that individual trees are identified, are this will help to reduce errors.

2.2. The formulae for volume growth are:

$$\mathit{Survivor\ growth} = \sum^{s} (v_2 - v_1)$$

$$\begin{aligned} \mathit{Accretion} &= \sum^{s} (v_2 - v_1) + \sum^{d} (v_2 - v_1) + \sum^{r} v_2 \\ &= \left(\sum^{s} v_2 + \sum^{r} v_2 \right) - \left(\sum^{s} v_1 + \sum^{d} v_1 \right) + \sum^{d} v_2 \\ &= V_2 - V_1 + V_m \end{aligned}$$

2.3. Your graphs for total height (h) and stand basal area (G) should look like the ones in Fig. S.1. You should also have drawn the increment curves.

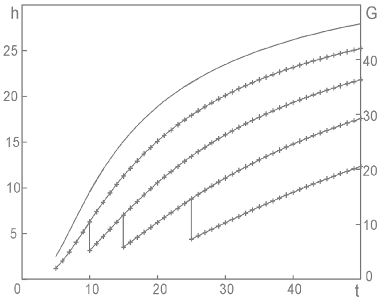


Fig. S.1.

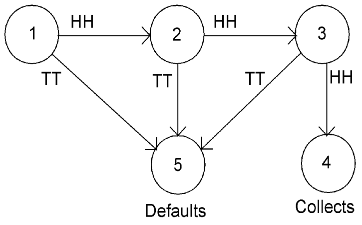


Fig. S.2.

2.4. If you concluded that the chance of the spinner collecting his payout was $(0.25)^3$, then you're wrong and you presumably didn't construct the Markov chain. Your chain should look like Fig. S.2, and your matrix and steady state should be:

$$\begin{bmatrix} 0.5 & 0 & 0 & 0 & 0 \\ 0.25 & 0.5 & 0 & 0 & 0 \\ 0 & 0.25 & 0.5 & 0 & 0 \\ 0 & 0 & 0.25 & 1 & 0 \\ 0.25 & 0.25 & 0.25 & 0 & 1 \end{bmatrix}^n = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.125 & 0.25 & 0.5 & 1 & 0 \\ 0.875 & 0.75 & 0.5 & 0 & 1 \end{bmatrix}$$

Notice that the columns in both matrices sum to one. There are two absorbing states, the spinner collecting (state 4) or defaulting (5) on his payout. The spinner has a 0.125 chance (column 1, row 4) of collecting when he begins, and once he has thrown some pairs of heads, his odds improve.

2.5. Your transition matrix should look like this:

$$\boldsymbol{M} = \begin{bmatrix} 0.76 & 0 & 0 & 0 \\ 0.08 & 0.95 & 0.03 & 0 \\ 0.13 & 0.03 & 0.93 & 0.04 \\ 0.03 & 0.02 & 0.04 & 0.96 \end{bmatrix}$$

Notice that the columns sum to one, and that there are no absorbing states (i.e. no entry is exactly 1.0). The plantation area in 10 years is expected to be 6274 ha, estimated by doing:

$$\begin{bmatrix} 0.76 & 0 & 0 & 0 \\ 0.08 & 0.95 & 0.03 & 0 \\ 0.13 & 0.03 & 0.93 & 0.04 \\ 0.03 & 0.02 & 0.04 & 0.96 \end{bmatrix}^2 \begin{bmatrix} 14400 \\ 3350 \\ 21600 \\ 2400 \end{bmatrix} = \begin{bmatrix} 8317 \\ 6274 \\ 22324 \\ 4835 \end{bmatrix}$$

You can find the eventual areas by repeatedly squaring the matrix to find the steady state. You should get 0, 21, 34 and 45% in forest, plantation, agriculture and other land uses respectively. This equilibrium emerges between 2^7 and 2^8 cycles, so it may take 640 years to appear.

2.6. The vector of probabilities is [0.2177, 0.2539, 0.3822, 0.1462].

Chapter 3

3.1. Six alternatives are summarized in Table S.1.

(a) The simplest alternative is to assume that each tree continues to grow for the next 25 years at the same rate observed during the last 5. This is likely to be an overestimate, as we might anticipate increasing competition.

(b) The simplest form of stand table projection uses a single movement ratio estimated as the mean increment divided by the class width (0.423). It too is likely to overestimate, and it proliferates fractions of trees. Remember to initialize your stand table with the d_5 data. Notice that you lose some precision when you form the stand table. Assuming that all trees are at the class midpoint gives $d_5 = 28.8$ and $G_5 = 1.58$, while the true values are 29.0 and 1.56 respectively.

(c) A better result is obtained by computing a movement ratio for each class. I used the following ratios [0.387, 0.500, 0.457, 0.380, 0.326, 0.266, 0.206, 0.146]. The first 4 values are based on the mean class increment, the others from extrapolating the linear trend [0.500, 0.457, 0.380] and estimating the ratios as $0.655 - 0.006d_i$. A more conservative alternative is to assume "no data, no growth" (c′). Note that you should use increment of the coming period, not of the past period, so there are data only for four classes (and 10 observations in class 1). These predictions can be made with stand table projection or a matrix model.

(d) Some authors recommend the use of actual movements rather than ratios estimated from increments. This gives [0.500, 0.400, 0.333, 0.333], and the 0.333 looks like an asymptote, so I used it for the other classes too (but give results for the "no data, no movement" assumption in d′).

(e) If you plot the increments against tree size, you will see a pattern (Fig. S.3). This is no surprise, as I synthesized the data using $\Delta d = d/3 - d^2/180$, and added some random noise. If you fit this equation to the data, you get $\Delta d = 0.336d - 0.00568d^2$, close to the original function. We can use this to

Table S.1. Stand table projection with different assumptions.

Method of projection	Stems in size class (midpoint in cm dbh)								$\bar{d}$ (cm)	G (m^2)
	15	25	35	45	55	65	75	85		
Initial stand table in year 5	5	8	4	3	1				29.0	1.56
a) Same growth rate			6	5	5	5			50.1	4.36
b) Single movement ratio	0.3	1.7	3.8	5.1	4.6	3.1	1.7	0.6	49.9	4.51
c) Movement ratio for each class	0.4	1.2	3.3	5.9	5.6	3.2	1.1	0.2	49.4	4.34
c′) No data, no movement	0.4	1.2	3.3	5.9	10.2				46.5	3.74
d) Actual movement from class	0.2	1.8	5.2	5.8	4.3	2.4	1.0	0.4	47.1	3.99
d′) No data, no movement	0.2	1.8	5.2	5.8	8.1				44.5	3.43
e) Increments from regression				11	10				48.8	3.98
f) Regression with serial correlation			3	9	9				48.9	3.94

predict increments for each individual tree. Notice that the size distribution tends to "bunch up", because the trees grow more slowly as they get bigger.

(f) Trees do not grow at the rate predicted in (e), but show some variation. This prediction assumes that the relative difference between observed and expected increments during the first period is maintained during the next 25 years.

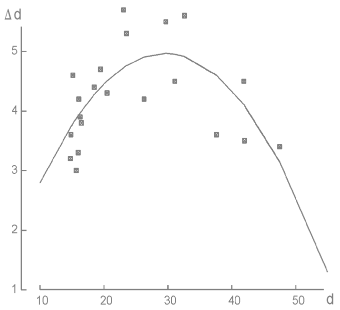


Fig. S.3.

These are all estimates of the future stand structure, and we do not know which is closest to the truth, which remains unknown. Estimate (f) may be one of the better estimates, but it is likely to overestimate since it is based on growth at a relatively low stand basal area and extrapolated to a stand with twice the density. The trends observed in Table S.1 may be specific to this particular case and to the pattern of diameter increments assumed.

However, the tendency for size class models to propagate fractions is general (cf. the single-tree approaches a, e and f).

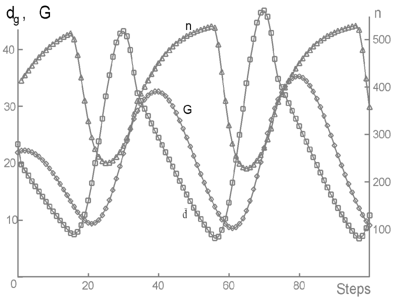


Fig. S.4.

3.2. If you attempted this exercise, you will have had to decide what to do with negative numbers in your stand table. I have re-set any negative numbers in the state vector at every time step, but have allowed "negative recruitment". If the number in the smallest class is non-zero, and the basal area is high, a single projection will reduce the number more than can be attributed to mortality and upgrowth, because a negative amount of recruitment is being added to it. This is difficult to avoid unless the matrix is split into its two components (Eqn 3.1) and that detracts from the mathematical elegance of the matrix method. So I have allowed negative recruitment, and regard it a limitation of the model. Figure S.4 illustrates the undisturbed development of a stand. Three lines trace the total number of trees (*n*), the stand basal area (*G*) and the mean diameter (d_g). All three curves increase to an asymptote, crash rather abruptly, and begin to increase again. This example was made using 40 trees in each size class, but similar patterns appear from many starting states.

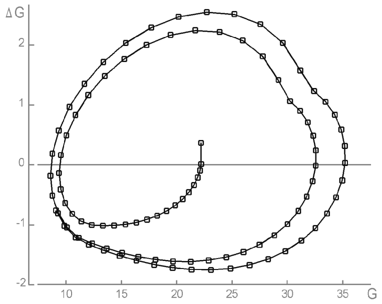


Fig. S.5.

There is not much evidence of convergence to a steady state. The authors suggest that undisturbed development tends toward an all-aged stand with an equal number of trees in each size class. Figure S.5 illustrates how stand basal area develops, starting from a state with 40 trees in each size class. The trajectory starts in the centre, spirals outwards, and soon converges into a closed loop which would appear as a wave in Fig. S.4. The same sort of loop appears for many starting conditions, both inside and outside of the loop.

The model is convenient, but rather simplistic. It may provide reasonable short-term predictions for stand conditions similar to those in the database on which it was calibrated, but extrapolations in space and time need to be interpreted with caution. It is rarely possible to assert that any model is "good", but we can say that a model is adequate for certain purposes.

3.3. The top row of the matrix (fecundity) is positive, so recruitment will be positively correlated with stand density (i.e. open stands will have little regeneration, while dense stands will have lots of regeneration); not a very realistic proposition. If you make projections from any starting condition, you will find a steady decline in tree numbers and stand basal area; the steady state is a zero vector and the eigenvalue is $\lambda = 0.97$. You cannot simulate a sustainable harvest with this model, as for all stands, recruitment is insufficient to replace mortality.

Chapter 4

4.4. Your answer should contain the following concepts:

Definitions

Each *tree_record* has a *species*, *diameter*, *expansion_factor*.
The *tree_list* may contain many *tree_records*.
A *lookup_table* indicates the *species_group* used in each function for each *species*.
Another *lookup_table* gives the equation coefficients for each *species_group*.

Functional Code

```
For each tree in the stand:
    Look up the species_group for the diameter increment equation.
    Increment the diameter by adding a prediction from the diameter
        increment equation for this species_group.
    Look up the species_group for the mortality function.
    Reduce the expansion factor by multiplying by the survival rate
        predicted by the mortality function for this species_group.
Next tree.
Add some tree_records for recruitment, setting diameter to the
    threshold, and predicting species and expansion factors.
If harvesting is simulated this year:
    Reduce expansion factors of big trees to account for harvesting.
    Reduce expansion factors of small trees to account for damage.
End harvesting.
Return to main program.
```

This is a very simple outline, but contains the important concepts of a tree list model. It represents a model with a one-year time step, invoked from another program which handles all input and output. This model would be invoked repeatedly to get multi-year predictions. Harvesting could be handled in a separate module. The model could be improved by allowing record splitting to account for variability in diameter increment. The definition of a tree_record could also be expanded to include other tree characteristics (e.g. height, merchantability, etc.).

Chapter 5

5.4. You cannot solve this question by simulation studies (e.g. with Equation 2.1, p. 19), because your results will depend entirely on the assumptions you make, and you cannot predict what variation will occur in the real data. One plot should remain at 17 $m^2\,ha^{-1}$, and another could be near the optimum (11 $m^2\,ha^{-1}$). There is little point thinning below about 3 $m^2\,ha^{-1}$, where trees are effectively open-grown. If your main interest was to fit an equation like 2.1, one reasonable solution might be plots at 5, 7, 11, 15 and 17 $m^2\,ha^{-1}$ (the equation is constrained to pass through the origin, and is linear for small values of *G*, so there is not a big need to sample there). Remember that treatment should be consistent across plots (i.e. do not thin some from above and others from below, but thin all the same way), or results may be confounded.

Chapter 6

6.1. Current annual increment:

$$\frac{dY}{dt} = \frac{95e^{-0.1t}}{1+19e^{-0.1t}}$$

Mean annual increment:

$$\frac{\Delta Y}{\Delta t} = \frac{50\left(1+19e^{-0.1t}\right)^{-1}-2.5}{t}$$

It is quite difficult to determine the intersection of the CAI and MAI curves analytically, but if you plot them (Fig. S.6) you can see that they intersect near $t = 42$.

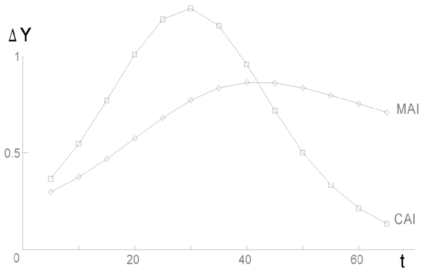


Fig. S.6.

Periodic annual increment (10-year period):

$$\frac{\partial Y}{\partial t} = \frac{5}{1+19e^{-0.1t}} - \frac{5}{1+19e^{-0.1(t-10)}}$$

6.2. You should find $R^2 = 0.95$ with the linear model $Y = -22 + 11X$. The R^2 does not tell you everything about a model !

6.3. Using ten random values ($0<X<1$) for each of X_1, X_2 and X_3, and 20 trials, I found R^2 ranged from 0.5 to 0.995 with a median of 0.9 (i.e. half my R^2s were ≥ 0.9). This should be a good warning that R^2 is not a good indicator of the quality of a model !

6.4. The data were generated artificially, but are based loosely on equations by Vanclay (1991b). The potential explanatory variables d, $G_{>d}$ and G were generated on a grid with random perturbations. Expected diameter increments were generated from the equation

$$\ln\hat{\Delta}d = -2.1 - 0.1G_{>d} + 1.0\ln d - 0.0333d \tag{S.1}$$

and the response variable Δd was generated by adding a random element so that the Δds varied by about 40% above and below the expected values. It is clear from Equation S.1 that the increments were calculated only from diameter d and basal area in larger trees $G_{>d}$, so the stand basal area G should not be required in the fitted model. I fitted the model

$$\ln\Delta d = \beta_0 + \beta_1 G_{>d} + \beta_2 \ln d + \beta_3 d$$

and obtained a good fit with the following estimates:

Parameter	β_0	β_1	β_2	β_3
Estimate	−2.07	−0.117	1.11	−0.0390
Standard error	0.82	0.011	0.30	0.0073

These estimates are significant ($P \leq 0.01$) and lie within one standard error of the original values used in Equation S.1.

Your model should be rather similar. There is clearly some correlation between basal area G and increment Δd, but it contributes little to the model after the inclusion of the terms above. Try it, and you will find it is not significant, and has the wrong sign (+ve, suggesting that tree growth should increase as competition increases !). You can see from Fig. 6.6 that the suggested model fits the data well, and can be extrapolated safely.

6.5. The correlation coefficients (r) are

Variable	d	$G_{>d}$	G
$G_{>d}$	−0.70		
G	−0.21	+0.75	
Δd	+0.10	−0.66	−0.72

It is noteworthy that these simple correlation coefficients do not reflect the utility for modelling, as we have already seen in Exercise 6.2. The correlation for d is small because it is a curvilinear relationship (see Fig. 6.6), rather than a simple straight line. This problem may be avoided with partial correlations, but it's easier and better to plot the data, model and residuals.

Chapter 7

7.2. If you simply calculated mean basal area increments for each plot (Table S.2), you may have had difficulty assigning plots to three productivity classes. You also failed to notice the big range in initial basal area, or didn't think about the effect that this might have on increments. Have a look at Fig. 2.2 again.

Table S.2.

Plot	ΔG	r'	$S_{h,d}$
2	0.12	−0.06	14
6	0.17	−0.03	14
1	0.21	+0.01	17
4	0.16	+0.01	17
3	0.24	+0.04	19
5	0.25	+0.05	19

I fitted an equation to the individual remeasurements:

$$\ln\Delta G = -4.60 + 2.02\ln G - 0.165G$$

and computed the mean residual for each plot ($r'=\sum\{\Delta G-\Delta\hat{G}\}$). These give a better indication of site productivity, and adjust for stand basal area.

The data were synthesised from Equation 2.1, and the values of site form used to generate the data were 14, 17 and 19. The data and original equations are illustrated in Fig. S.7. The data do not fit these lines exactly because some random variation was simulated. The equation given above generates a line similar to that for $S_{h,d} = 17$, but is a little more curved.

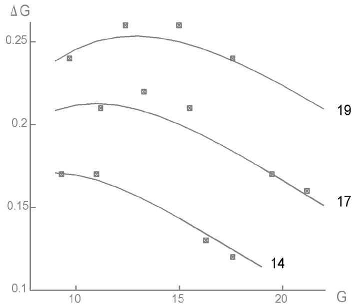


Fig. S.7.

7.3. Visual inspection quickly reveals that species *B* and *C* are correlated with site productivity. Discard species *A* and *G* because they are very rare and common respectively. Note that species *D* and *E* contain the same information ($D=1-E$), so ignore one of these. Four species are left, and it is easy to see that species *B* occurs only on the poor sites, and *C* only on good sites. Since the site form of these six plots is known (see above), the following prediction equation can be formulated

$$\hat{S}_{h,d} = 17 - 3B + 2C$$

Visual inspection is easy here because there are few species and few plots. If there were more plots and species, you could calculate, for each species, the mean site form of plots where the species occurs, and the mean site form of plots where it is absent, and look for species that maximized the difference between these two means.

If you tried to use regression analysis to find the solution, you may have had problems with multicollinearity (e.g. $A+F=G$). You can reduce these problems by some preliminary screening (e.g. remove rare and common species).

Chapter 8

8.1. The parameter estimates are:

Model	β_0	β_1	R^2	Furnival index	PRESS
Δd	0.508	0.0258	0.8742	0.0342	0.0153
d_1	0.508	1.0258	0.9999	0.0342	0.0153
Δg	0.829	0.0386	0.9865	0.0347	0.0155

Notice that although the Δd and d_1 models are effectively the same ($d_1 = d+\Delta d = d+(\beta_0+\beta_1 d) = \beta_0+(1+\beta_1)d$), their R^2s are very different. The equivalent calculation for basal area increment gives conflicting results (e.g. $\Delta g = 2kd\Delta d = 2kd(\beta_0+\beta_1 \mathrm{d})$, with $k = 7.853\times10^{-5}$, suggests $\beta_0 = 0.798$ and $\beta_1=0.0406$, which differ from the regression estimates), because the models imply different error structures ($e_{\Delta g} = 2kde_{\Delta d}$). The Δd and d_1 models are equally good, and the Δg model is slightly inferior. You can see this from the Furnival index or the plot of residuals, but not from R^2. The Furnival index and the PRESS have been adjusted for units (i.e. divided by $2k\overline{d}$), and since they both represent the sum of errors (standard error and prediction error sum of squares respectively), smaller values are better.

8.2. Inspection of residuals (and the Box-Cox transformation) indicate that a logarithmic transformation is required to stabilize the variance. The following is one possible solution:

$$\ln\Delta d = -3.957 + 1.650\ln d - 0.0005480 d^2 - 0.5558\ln G$$

Compare it with your own solution.

8.3. Residuals indicate a logarithmic transformation. One possible model is:

$$\begin{aligned}\ln\Delta d = -3.555 &+ 0.7866 \ln d - 0.0002413\, d^2 \\ &+ 0.01850\, G_{r\%} - 1.180\times10^{-7}\, G_{r\%}^2 d^2\end{aligned}$$

This differs substantially from Howard and Valerio's (1992) model (they used α=50), but seems to fit the data better (Fig. S.8, dotted lines join data with same $G_{r\%}$, dashed lines show Howard & Valerio's model, solid lines are the above equation). Heavy logging is unlikely to stimulate the growth of the largest trees, as they are likely to suffer damage from logging or exposure, so the response implied in the above equation is biologically reasonable.

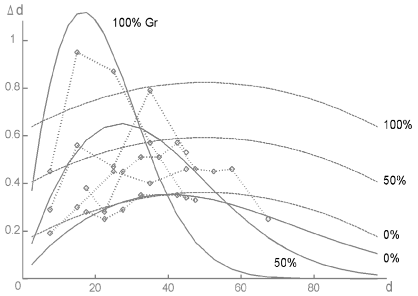


Fig. S.8.

Chapter 9

9.1. Insufficient information is given to correct for the varying measurement intervals. One good description of these data is

$$P = (1 + \exp[-1.45 + 0.0132d - 0.664 \ln d])^{-1}$$

Notice that this implies an average mortality slightly less than 1% per year, and that mortality is lowest for trees of about 50 cm dbh, and increases for smaller and larger trees.

9.2. We cannot calibrate a model to predict the proportion of cut stems that are merchantable, because of insufficient data and too much noise in the data. The average (87% merchantable) is as good as any.

The proportion of trees harvested can be predicted as

$$P = (1 - \exp[11.0 - 0.135d])^{-1}$$

and the incidence of damage can be predicted as

$$P = (1 - \exp[1.73 + 0.041d])^{-1}$$

Notice that the former (harvest) refers to the proportion of trees in the initial stand, while the latter (damage) relates to the proportion of trees in the residual (initial−harvest) stand.

Chapter 10

10.2. The trick is to recognize a class for the newly-dead trees, so that for example, the stand table might contain small trees, big trees and newly-dead trees. Assuming that both movement and mortality are 10%, and that every death will give rise to one recruit, the Usher matrix will be

$$\begin{bmatrix} 0.8 & 0 & 1 \\ 0.1 & 0.9 & 0 \\ 0.1 & 0.1 & 0 \end{bmatrix}$$

and a simulation commencing with 10 trees in the smallest class will lead to:

Cycle	0	1	2	3	4	5	6	∞
Small	10	8	7.4	6.82	6.366	6.002	5.711	4.545
Big	0	1	1.7	2.27	2.725	3.089	3.380	4.545
Dead	0	1	0.9	0.91	0.909	0.909	0.909	0.909

Notice that there are always 10 trees in the stand (counting the newly dead), and that it takes an additional cycle for the recruits to appear in the stand table. The newly-dead is not an absorbing state, so it does not indicate the total deaths, but merely the deaths in the current cycle. Since it is a Markov matrix (columns sum to 1), a steady state exists, and the system quickly converges.

10.3. The data were generated using Equations 10.1 and 10.2, and display the characteristic dichotomy (none, some) that suggests a two-stage approach should be used:

$$P = (1 + \exp[-(0.187 - 0.0414G)])^{-1}$$

$$\ln N_r = 5.18 - 0.589 \ln G$$

There are insufficient data to build more complex relationships, but both these equations provide good estimates of the effect of stand basal area (G and $\ln G$ respectively; there are not enough data with small G to fit $\ln G$ in the probability equation). These equations indicate the 5-year relationships; the annual relationships will have different estimates for β_0.

Chapter 11

11.2. The model can be implemented with different time-steps, and West (1981) observed that time steps longer than 5 years lead to biased estimates. Simulations up to age 80 allowed some trees to grow to excessive sizes (the likely maximum attainable diameter seemed to be 130 cm), and West (1981) remedied this by imposing a maximum growth rate of 2 cm/y.

Oderwald and Hans (1993) examined this model, and questioned why site index did not enter the prediction equations, and why growth increases with stand density (see Chapter 6, p. 106, for more discussion of this issue).

Mortality is not density-dependent, and mortality estimates become negative for ages over 88 years. Negative increments were possible for some trees at low stockings.

11.3. West (1981) reported several tests and concluded that his model produced accurate estimates of future conditions of stands for simulations up to 21 years in length. Reynolds (1984) also examined these data and found no evidence of bias in the model. He computed the prediction interval for future basal area as -0.19 ± 8.66 $m^2 ha^{-1}$, which implies that users may be 95% confident that in a single future prediction, the error will lie in the interval -8.85 to $+8.47$ $m^2 ha^{-1}$. Figures 11.2 and 11.3 are based on West's data.

Chapter 12

12.1 See Vanclay (1994) for a discussion of these issues.

12.4 There are many ways to examine optimum cutting limits. One way is to do it analytically (e.g. Exercise 6.1), but it isn't easy to do this. The time-of-passage approach is easier, and should give a reasonable result it the class width is small. The result will depend on the assumptions made. Fig. S.9 uses 1 cm size classes and assumes that d=1 at t=0 and that $v=\beta d^2 h$. The numbers 1–5 refer to Equations 8.1–8.5 respectively. Four of the equations suggest very similar optima, while Eqn 8.5 suggests a higher cutting limit because of the way this growth curve approaches the asymptote Δd more slowly. Recall that Eqns 8.4–8.5 were calibrated to Eqn 8.1 in the range $5 \leq d \leq 90$, and note that the MAI curves agree closely in this range. You should not attempt to judge which equation is "better" since three of the equations were not fitted to data but were simply fitted to approximate equations 8.2 and 8.3. But you should note the similarities and differences between the equations.

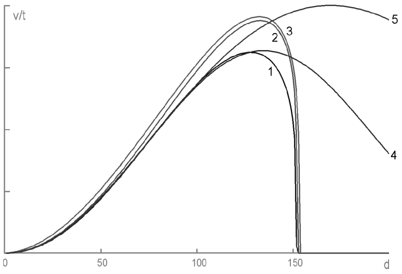


Fig. S.9.

Mortality may be taken into account by adding an additional column to your calculation, indicating the cumulative probability that a tree will survive to that size. This probability can be used to reduce the expected volume in each size class of the time-of-passage calculation. Mortality, and longer cutting cycles will tend to reduce the optimal cutting limit (e.g. Table 12.2).

Author Index

Adams 232, 252
Adlard 81, 252
Aikman 172, 252
Aitkin 96, 121, 123, 124, 126, 179, 190, 252
Alder 35, 50, 51, 70, 71, 79, 89-91, 94, 150, 157, 161, 194, 220, 221, 252
Aldworth 147, 268
Allison 51, 58, 160, 255
Amateis 158, 263
Anscombe 126, 125, 252
Arney 94, 159, 176, 252, 253
Arvanitis 86, 257
Assmann 42, 142, 209, 253
Avery 94, 226, 253
Bailey 23, 112, 253, 256
Barclay 61, 253
Bare 145, 230, 232, 253
Baur 134, 253
Beck 142, 253, 254
Beetson 87, 86, 253
Belcher 11, 77, 194, 219, 253
Bella 157, 253
Bertalanffy 19, 21, 27, 108-112, 163-167, 170, 253, 267
Bevege 177, 253
Biging 60-62, 253, 254
Binkley 31, 45, 253
Bolsinger 154, 263
Bonnor 151, 253
Borders 22, 120, 253
Bosch 46, 48, 194, 253
Bossel 41, 42, 156, 160, 246, 250, 251, 253, 254
Botkin 6, 7, 70, 71, 137, 161, 164, 165, 174, 176, 196, 241, 245, 247, 254, 256
Bowling 25, 254
Box 7, 88, 146, 254, 262, 294
Bradley 142, 254, 261
Bragg 65, 254
Brand 72, 253, 254
Brewer 148, 254
Brickell 198, 259
Bruner 45, 254
Buchman 10, 254, 262, 267
Buckman 19, 254
Buell 19, 254
Bunnell 6, 254
Buongiorno 48, 49, 194, 245, 254, 264
Burk 54, 141, 160, 218, 250, 254, 257, 258, 264, 268, 269
Burkhart 20, 74, 75, 94, 151, 226, 227, 235, 253, 254, 256, 263, 269, 270

Butler 207, 211, 216, 263
Byrne 99, 254
Cajander 148, 254
Campbell 38, 69, 90, 94, 153, 160, 176, 182, 254, 255, 271
Candy 151, 261
Cannell 66, 161, 255
Canonizado 138, 255
Carleton 150, 255
Carlson 198, 257
Carmean 146, 147, 255
Carron 36, 49, 255
Cassell 45, 255
Chai 40, 255
Chaiken 17, 263
Chapman 21, 35, 109, 226, 255, 269
Christie 15, 254, 257
Chung 206, 214, 268
Clutter 7, 13, 17, 19, 20, 28, 51, 58, 76, 160, 253, 255, 256, 270
Cochran 88, 240, 270
Corns 150, 256
Crawley 179, 256
Curtis 91, 141, 154, 256
Czarnowski 146, 256
Daniels 60, 74-76, 256
Daubenmire 148, 149, 256
Davis 236, 238, 256, 274
Dawkins 35, 81, 89, 94, 227, 256
Dell 23, 253
Diggle 60, 161, 257
Dobbertin 60-62, 253
Doyle 71, 196, 256
Draper 124, 211, 256
Drew 76, 148, 169, 172, 256
Duerr 15, 136, 256
Dzialowy 151, 258
Edwards 15, 178, 179, 257, 259
Efron 218, 257
Ek 10, 59, 61, 67, 76, 112, 157, 159, 162, 163, 176, 181, 201, 202, 232, 235, 239, 252, 257, 264, 265, 267, 269, 273
Elston 207, 257
Enright 47, 257
Esser 146, 257
Evans 146, 257
Farr 146, 257
Farrar 21, 109, 265
Ferguson 176, 198, 199, 255, 257
Field 81, 89, 256
Fins 96, 257
Fiske 153, 255
Flewelling 172, 256
Ford 58, 60, 161, 255, 257
Ford-Robertson 257, 279
Foster 153, 257
Fowler 86, 257
Fox 152, 258
Fralish 147, 258
Frayer 71, 210, 225, 239, 265, 266
Freese 153, 258
Furnival 21, 117, 126, 127, 157, 258, 294
Gadow 23, 258
García 4, 26-28, 27, 28, 109, 175, 258
Gentry 247, 267
Gertner 8, 71, 86, 151, 207, 210, 239, 258
Gevorkiantz 15, 136, 256
Gillison 3, 244, 258, 266
Goelz 141, 258
Gong 218, 257
Goodwin 77, 152, 219, 258
Goulding 204, 258
Gower 149, 258
Graber 40, 178, 261
Gregoire 153, 206, 258, 269
Grey 145, 147, 258
Grimes 40, 258
Grosenbaugh 117, 236, 237, 258, 259
Gumpal 12, 18, 138, 264
Hafley 23, 259, 268
Hägglund 235, 259

Hahn 52, 164, 165, 240, 259, 262, 267
Haight 230, 232-234, 238, 245, 259
Hall 21, 109, 265, 271
Hamilton 15, 178, 179, 181, 187, 198, 259
Hann 172, 187, 188, 215, 228, 235, 259, 269
Hans 206, 207, 266, 296
Hara 161, 259
Harris 146, 257
Hart 175, 259
Hartshorn 47, 175, 177, 260
Havel 137, 148, 260
Helmisaari 167, 267
Henry 139, 140, 272, 277
Herrick 35, 260
Hett 201, 260
Hettmansperger 240, 260
Higgins 49, 260
Hilt 77, 163, 169, 260
Hof 230-234, 238, 245, 260
Holdaway 52, 219, 253, 260, 262
Hool 31, 260
Howard 40, 171, 191, 260, 295
Hruska 45, 268
Hulst 45, 260
Husch 17, 36, 260
Hutchinson 79, 91, 93, 260, 267
Hyink 23, 24, 260
Jackson 146, 260
Jardine 149, 260
Jeffers 31, 33, 260, 271
Johnson 23, 25, 157, 236, 238, 256, 260, 261, 273
Johnston 145, 254, 261
Jones 153, 255, 268
Joyce 245, 260
Keenan 151, 261
Kennard 87, 261
Kimmins 6, 7, 246, 261
Korsgaard 39, 40, 261
Krajicek 59, 261, 280
Kramer 138, 261
Krieger 41, 42, 156, 160, 254
Krug 25, 261
Landsberg 6, 246, 261
Lanly 90, 261
Larocque 63, 261
Layton 61, 253
Leak 40, 176-178, 192, 201, 261
Leary 11, 15, 51, 52, 71, 72, 74, 135, 159, 160, 164, 165, 207, 239, 259, 261, 262
Leech 32, 128, 129, 176, 257, 262
Lefkovitch 46, 262
Leigh 96, 262
Lemmon 60, 262
Lenhart 28, 76, 256, 262
Leslie 46, 194, 201, 262, 271
Leuschner 80, 234, 236, 238, 262, 263
Levins 250, 262
Lewis 137, 262
Lieberman 96, 262
Lieth 7, 146, 262
Liu 206, 207, 210, 263
Loetsch 2, 226, 228, 239, 263
Lonsdale 172, 263
Lorimer 61, 159, 263
Loucks 147, 201, 258, 260
Lowell 49, 168, 263
Lund 83, 263, 273
Lynch 24, 263
Macdonald 142, 266
MacKinney 15, 17, 263
MacLean 154, 263
Mader 147, 263
Mäkelä 6, 263
Mallows 125, 263
Mariaux 80, 263
Marquardt 124, 263
Marshall 63, 261
Martin 61, 112, 159, 162, 163, 263, 268
Mawson 163, 263
Mayer 207, 211, 216, 263
McDill 158, 263

McGuiness 90, 263
McKenney 239, 264
McMurtrie 65, 264, 271
Meldahl 70, 128, 161, 264
Mendoza 12, 18, 48, 138, 264
Meng 220, 264
Mervart 35, 264
Methven 138, 268
Meyer 23, 138, 232, 264
Michie 45, 48, 49, 194, 232, 254, 259, 264
Mihram 23, 270
Miner 77, 264
Mitchell 6, 7, 49, 63, 64, 67, 157, 168, 175, 176, 263, 264
Mitscherlich 109, 140, 264
Moeur 235, 264
Mohai 94, 268
Monserud 59, 61, 67, 76, 137, 142, 149, 157, 159, 172, 176, 181, 201, 202, 209, 230, 233, 239, 257, 259, 263, 265, 270
Moore 59, 75, 146, 152, 265, 267
Morrier 151, 253
Moser 21, 23, 24, 45, 46, 69, 109, 178, 254, 255, 260, 263, 265, 269
Mosteller 115, 265
Mowrer 71, 76, 207, 210, 225, 239, 265
Murphy 21, 109, 265
Myers 248, 265
Nance 60, 265
Nelson 19, 240, 259, 265
Newnham 54, 62, 63, 176, 178, 265
Nielsen 65, 265
Nimmo 81, 273
Nix 3, 266
Norberg 172, 266
O'Neill 248, 266
Oavika 171, 266
Oderwald 103, 206, 207, 266, 296
Ogden 47, 257
Omule 142, 266
Opalach 230, 232, 253
Opie 61, 69, 159, 160, 176, 255, 266
Oreskes 204, 266
Osho 56, 266
Osmaston 35, 266
Osunkjoya 175, 266
Öttelt 138, 266
Pannell 247, 266
Parresol 121, 266, 272
Pastor 246, 248, 266
Payandeh 113, 266
Peden 45, 266
Pegg 40, 258
Pelz 60, 267
Philip 39, 94, 267
Phillips 247, 267
Phillis 237, 267
Pienaar 137, 256, 267
Pinto da Costa 10, 267
Pluth 150, 256
Popper 206, 267
Prance 94, 247, 255, 267
Prentice 167, 267
Preston 11, 12, 52, 152, 231, 237, 267, 272
Pretzsch 59, 67, 68, 159, 181, 247, 267
Preuhsler 10, 263, 267
Prodan 23, 267
Pütter 108, 267
Randall 196, 229, 267
Ranneby 96, 267
Ratkowsky 111, 117, 206, 267, 274
Rayner 69, 154, 267
Reed 51, 92, 153, 176, 182, 268
Reineke 15, 22, 69, 137, 172, 174-176, 268
Reinhardt 139, 140, 268
Renner 172, 270
Rennolls 25, 268
Reynolds 23, 25, 153, 206, 207, 214, 258, 268, 297

Richards 21, 109, 267, 268
Roberts 45, 268
Robichaud 138, 268
Romesburg 94, 268
Rorres 45, 46, 268
Rovainen 96, 267
Rudolf 16, 268
Rust 96, 257
Schmoldt 142, 145, 268
Schneider 16, 268
Schönau 147, 150, 268
Schreuder 3, 23, 80, 118, 226, 228, 239, 259, 268, 269
Schumacher 17, 20, 60, 163, 226, 262, 263, 269
Sedjo 238, 269
Setyarso 48, 264
Shifley 10, 69, 128, 157, 159, 163, 178, 196, 197, 216, 219, 229, 254, 257, 269
Shoulders 153, 264, 269
Shugart 6, 71, 196, 256, 269
Shumway 138, 270
Sia 40, 255
Sibson 149, 260
Sievänen 7, 54, 160, 250, 269
Silva 94, 269
Simon 72, 269
Skovsgaard 172, 206, 269, 270
Smith 38, 62, 124, 151, 172, 211, 219, 227, 235, 252, 256, 265, 269, 270
Snedecor 88, 240, 270
Snee 204, 216, 217, 270
Soares 206, 211, 215, 270
Soest 270, 275
Solomon 49, 270
Sorrensen 58, 257
Sprinz 20, 254
Spurr 15, 60, 177, 270
Stacey 23, 270
Stage 7, 11, 69-74, 77, 136, 146, 169, 170, 172, 182, 220, 221, 257, 270, 274
Steffen 250, 270
Sterba 137, 172, 176, 270
Stone 87, 261
Stout 138, 270
Strub 190, 256, 270
Sullivan 20, 270
Swain 77, 219, 271
Swaine 128, 150, 169, 177, 271
Swank 23, 268
Sweet 99, 254
Synnott 79, 89-91, 94, 252
Teck 77, 260
Tedder 238, 261
Terborgh 245, 271
Tham 59, 61, 271
Tiarks 153, 269
Tomboc 5, 271
Tracey 147, 150, 273
Trousdell 142, 253
Tufte 250, 271
Turnbull 137, 267
Turner 69, 77, 147, 148, 219, 267, 271
Turvey 148, 152, 271
Ure 149, 271
Usher 43, 45-48, 56, 203, 271
Valentine 182, 271
Valerio 40, 171, 191, 260, 295
Valsta 232, 271
van Dyne 147, 274
Vanclay 6, 11, 12, 19, 49, 52, 58, 64, 70, 74, 75, 94, 128-131, 137-143, 150, 152, 160, 161, 166-168, 180-184, 187-190, 195, 199, 201, 231, 230, 237, 253, 266-272
Vandermeer 246, 272
Vincent 94, 274
Vries 2, 226, 239, 272
Vuokila 15, 272
Wadsworth 161, 272

Wahlenberg 16, 37, 36, 37, 39, 273
Walters 187, 258, 264, 273, 274
Wan Razali 181, 273
Warren 132, 273
Watkinson 172, 252
Watson 81, 273
Webb 147, 149, 150, 273
Weber 182, 273
Weck 50, 49, 109, 273
Weibull 23, 273
Weisberg 115, 121-126, 206, 211, 218, 273
West 65-67, 89, 119, 120, 157, 178, 196, 216, 221, 222, 228, 241, 269, 273, 274
Westoby 172, 274
White 172, 274
Whitmore 92, 128, 271, 274, 282
Wigley 247, 274
Wilson 21, 117, 175, 176, 258, 274
Wolf 65, 264
Wood 96, 269, 272, 274
Wright 147, 255, 274
Wyatt-Smith 94, 274
Wykoff 11, 74, 77, 161, 162, 167, 166, 167, 170, 220, 270, 274
Yoda 172, 174, 175, 274
Zeide 21, 111, 166, 172, 175, 228, 274

Subject Index

absorbing state 44
Acacia 147, 279
  aulacocarpa 169, 279
  mangium 189, 279
accretion 17
Acer
  rubrum 153, 257
  saccharum 165, 241
Acronychia acidula 143, 277, 169, 180
Africa
  east 51, 252
  Ghana 150, 161, 177, 252, 262, 271
  Nigeria 56, 264, 266
  South Africa 147
  Swaziland 146, 257
  Uganda 256
age 1, 4, 14, 46, 106, 111, 120, 172, 177, 212
  maximum 70, 174
agroforestry 243, 246
allocation rules 160
Alphitonia whitei 143, 277
Alstonia muellerana 169, 278
America
  Amazon 90, 255, 267, 269
  Brazil 255, 269
  Canada 63, 77, 150, 152, 256, 261
  central 47, 71, 196
  Costa Rica 40, 171
  north 11, 15, 142, 149, 241
  USA 31, 54, 62, 71, 74, 77, 99, 138, 139, 146, 147, 149, 153, 168, 172, 187, 196, 238, 245
anabolism 108, 109
Anisoptera thurifera 171
Araucaria 67, 138
  cunninghamii 47, 147, 150, 178, 273, 278
  huntsteinii 47, 278
Archidendron vaillantii 150, 279
Argyrodendron trifoliolatum 143, 278
Asia
  Indonesia 48, 264
  Malaysia 39, 41, 42, 156, 187, 255, 261, 272, 273
  Papua New Guinea 47, 171, 257
  Philippines 5, 18, 138
  Sarawak 40, 255, 260
aspen 147, 235, 258, 273, 277
assumption
  Markov 44, 45
  stationary 45
  Usher 46
Australia 11, 69, 71, 196, 219

  - Queensland 11-12, 19-20, 64, 87, 101, 127, 137, 140-141, 147-152, 160, 181, 184, 186, 195, 199, 228, 231, 237
  - South 137
  - Tasmania 151, 152, 221, 273
  - Victoria 67, 69
  - Western 38, 137
- basal area (G) 276, 279
  - in larger trees ($G_{>d}$) 87, 113, 123, 128, 132, 133, 158-162, 166-169, 179, 180, 276, 292
  - natual (G_{max}) 88, 89, 137, 138, 141, 282
- *Beilschmiedia*
  - *bancroftii* 12, 279
  - *obtusifolia* 169
- benchmarking 204
- bias 105, 279
- binary variable 104, 280
- birch 40, 196, 261, 277
- *Bleasdalea bleasdalei* 127, 150, 278
- *Blepharocarya involucrigera* 150, 180, 199, 277
- bootstrapping 218
- Brazil 255, 269
- Britain 15, 46
- buffer 89
- buttress 93
- *Callitris* 19, 87, 102, 140, 141, 160, 201, 271, 272, 278
- *Calluna* 33, 278
- Canada 63, 77, 150, 152, 256, 261
- *Canarium baileyanum* 52, 278
- *Cardwellia sublimis* 12, 52, 144, 143, 189, 278, 199
- *Castanospora alphandii* 143, 279
- cedar, eastern red 163, 277
- censored data 229
- checklist
  - model evaluation 206
  - model suitability 10
  - plot measurements 95
  - scatterplots 113
- cherry 196, 277
- circular plots 90, 257
- climax vegetation 148, 150
- coefficient
  - anomalous 105
  - of determination 125
  - partitioning 67
- cohort 34
- collinearity 123, 124
- competition 58
  - area potentially available 59
  - cone-of-light 68
  - crown area 62
  - crown competition factor 166, 197, 275, 280
  - influence zone 59
  - limitations 61
  - mortality 172
  - one-sided 66, 161, 240
  - self-thinning 174
  - size-distance 60
  - sky-view 58-61, 64, 68, 77
  - two-sided 161
- confidence limits 239, 240
- Costa Rica 40, 171
- cottonwood 197, 269, 278
- cross-validation 115, 127, 218, 257
- crown
  - competition factor 59, 66, 123, 166, 176, 197, 198, 275, 280
  - length 54, 64, 77, 164, 176, 275, 280
  - ratio 61, 62, 67, 68, 164, 166, 221, 275, 280
  - width 63, 68, 176, 197, 198, 202, 275
- *Cryptocarya*
  - *angulata* 143, 278
  - *cinnamomifolia* 150, 278
  - *mackinnoniana* 143, 278
- cutting cycle 35, 229
  - analysis 236, 237

cutting limit 47, 48, 183, 231, 230, 297
damage, harvesting 185
Darlingia darlingiana 143, 278
data 79
  censorship 196, 229
  checklist 95
  collection 79
  electronic recorder 96
  entry 97
  increment 96, 122
  mortality 89
  passive monitoring 88
  requirements 42, 54, 79
  treatment response 88
deciles 50, 51
decrements 95, 122
Denmark 90
density-dependent mortality 85, 88, 89, 172, 174-180
deterioration 187
diameter increment 156
  allocation rules 160
  competition 159
  future diameter 158
  modifier function 159
digital elevation model 152, 153
dipterocarp 18, 39-41, 138, 156, 187, 260, 278
  meranti 40, 278
diseases 173
distribution
  Johnson's S_B 23
  normal 75
  Weibull 23
eco-physiology 246
eigenvalue 47
eigenvector 47
Elaeocarpus largiflorens 143, 279
elevation 85, 166
equation
  autocatalytic 110
  Bertalanffy 21, 108-112, 163-167
  Chapman-Richards 21, 109
  empirical vs theoretical 107
  Gompertz 110
  growth 18
  monomolecular 110
  Schumacher-Clutter 17, 20
  yield 17
error 218
  budget 215
  characterizing 211
  critical 214
  distributions 118, 157, 170, 179, 190
  linear regression 108, 119
  multiplicative 122
  structures 157, 158, 215
Eucalyptus 128, 141, 152
  deglupta 5, 271, 277, 278
  delegatensis 151, 261, 277
  maculata 93, 278
  marginata 219, 271, 278
  obliqua 221, 241, 279
  regnans 67, 69, 128, 160, 277
  vernicosa 128, 278
Europe 15, 148, 149, 247, 267
  Britain 15, 46
  Denmark 90
  Germany 15, 67, 263, 273
  Nordic 96
expansion factor 50, 52, 57, 58, 72-75, 280, 290
experiments 88
explanatory variable 104, 108
  basal area in larger trees ($G_{>d}$) 161
  competition index 61
  diameter increment 160
  harvesting 186
  independence 123
  linear regression 116
  mortality 179
  recruitment 194
explicit prediction 76
extrapolation 56, 82-87, 106, 112, 117, 138, 244, 281
F-ratio 124

fecundity 46
Ficus 80, 278
fir
  Douglas 62, 63, 177, 252, 253, 264, 265, 278
  white 62, 278
fire 64, 94, 95, 101
Flindersia 12
  bourjotiana 12, 143, 277
  brayleyana 12, 143, 184, 278
  pimenteliana 12, 93, 143, 168, 169, 180-184, 188, 189, 199, 200, 222, 242, 278
FORCYTE model 6, 7, 246, 261
forest
  beech-birch-maple 40
  conifer 51, 64, 76, 94, 140, 157, 172, 245, 252, 253, 259, 268, 280
  deciduous 80, 177, 269
  dipterocarp 18, 41, 156, 260
  elm-ash-cottonwood 197, 269
  mixed eucalypt 219
  mixed hardwood 48, 142
  mixed swamp 40, 255
  rainforest 11, 12, 71, 132, 150, 186, 187, 190, 196, 252, 269, 272-274, 282
  tropical moist 1
  tropical montane 196
FORMIX model 41, 42, 56, 102, 156
FORPLAN 238, 239, 264, 269
Furnival index 126
future diameter 158
$G_{>d}$ 87, 113, 123, 128, 132, 133, 159-162, 166-169, 179, 180, 276, 292
gap dynamics 70
GCM model 248
geocentric view 135
geographic information system 265
Germany 15, 67, 263, 273
Ghafosim model 161, 252
Ghana 150, 161, 177, 252, 262, 271
goodness-of-fit 116, 117, 157, 181, 207, 268
GROPE model 71, 220, 221, 252
ground vegetation 148
growth 4, 17, 106
  changing rank 17, 50, 130, 179, 180
  percent 16
  survivor 17
  tables 16
growth index 143, 144, 143-145, 150-152, 154, 155, 168, 179, 276
habitat 104, 148, 149, 166, 175, 198, 245, 265, 266, 270
half-splitting 218
harvesting 183
  damage 185
heteroscedastic data 121
heuristic simulation 236-238, 236
Hevea braziliensis 80, 279
homogeneous variance 121, 130
implicit prediction 76
increment 8, 53, 62, 74, 111, 156
  data 122
  diameter 8, 35, 50, 70
  height 68
  in mortality model 181
  mean annual 142
  periodic annual 132, 141, 154, 291
  relative to measurement error 96
  stand basal area 19
  stand volume 19, 109, 110
  stem volume 64
  threshold 176

indicator plants 93, 148-150, 154
Indonesia 48, 264
ingrowth 8, 17, 48, 92-94, 196, 269, 281, 283, 284
interpolation 54, 82, 83, 162, 227, 226, 281
inventory 80
  continuous forest 80
  dynamic 3
  static 2
J-distribution 39
JABOWA model 7, 57, 70, 71, 161, 164, 174, 196, 245, 247
jackknifing 218
Juniperus virginiana 163, 277
keystone species 245
Kiambram model 71
larch, western 139, 140, 268, 278
leading desirables 35
leaf area index 196
liberation thinning 184
linear equation 107, 111, 118, 166, 178
linear programming 67, 238, 241, 261, 268, 238
  FORPLAN 238, 239, 264, 269
LINKAGES model 246
Litsea leefeana 143, 277
logging 95, 182, 185, 233, 281
  damage 185
  plot remeasure 96
  selection 48
  time since 16, 18, 187, 188, 18, 187
logistic function 178
Malaysia 39, 41, 42, 156, 187, 255, 261, 272, 273
Mallows' C_p 125
maple, red 153, 257, 278
marking trees
  paint vs nails 92
Markov
  assumption 44, 45, 48, 281
  chain 29, 30, 29-32, 44, 43, 44, 254, 260, 281, 283, 284, 286
  matrix 30, 44, 46, 296
matrix
  estimation 49
  Leslie 46
  Markov 43
  transition 43
  Usher 46
maximum likelihood estimation 118, 190
measuring
  diameter 93, 94
  height 94
  multiple stem 93
  tape 93
meranti 40, 278
merchantability 187, 189
merchantable height 138
model 4
  benchmarking 204
  choosing a 9
  cohort 50
  criticism 204
  decile 51
  deterministic 7
  for prediction 6
  for understanding 6
  FORCYTE 6, 7, 246, 261
  FORMIX 41, 42, 56, 102, 156
  GCM 248
  Ghafosim 161, 252
  GROPE 71, 220, 221, 252
  JABOWA 7, 57, 70, 71, 161, 164, 174, 196, 245, 247
  Kiambram 71
  LINKAGES 246
  non-spatial 68
  NORM 11-12, 52, 130, 168, 169, 179, 180, 183, 185, 187, 189, 198-201, 237
  percentile 51
  PLATIPUS 64, 65, 271
  process 6
  Prognosis 7, 11, 72-74, 77, 166, 169, 170, 187,

- 198, 199, 209, 220, 224, 233, 257, 270, 273, 274
  - recruitment 192
  - regeneration 192, 201
  - single-tree 6, 8, 57
  - size class 6, 34
  - spatial 58
  - stand table projection 35
  - state space 26
  - STEMS 11, 74, 77, 194, 219, 253, 258, 260, 267, 270
  - stochastic 7, 69
  - succession 6
  - whole stand 5, 14
  - whole stand distribution 23
  - whole stand transition matrix 29
- modifier function 159
- mortality 8
  - catastrophic 182
  - data 89
  - density-dependent 85, 88, 89, 172, 174, 176, 180
  - explanatory variables 179
  - growing space 176
  - logistic function 178
  - relative size 179
  - self-thinning 175
  - threshold increment 176
- movement probabilities 49
- movement ratio 36
- multi-modal size distributions 31
- nails 92
- natural basal area (Gmax) 88, 89, 137, 138, 141, 282
- New Zealand 51, 149, 255-257
- Nigeria 56, 264, 266
- non-linear regression 111, 117, 266
- NORM model 11-12, 52, 130, 168, 169, 179, 180, 183, 185, 187, 189, 198-201, 237
- *Nothofagus* 257, 277
- Occam's razor 103, 127
- *Omalanthus populifolius* 277, 180
- open-grown 59, 60, 63, 68, 88, 160, 176, 197, 198, 275, 280, 291
- *Opisthiolepis heterophylla* 127, 150, 278
- optimization
  - dynamic 231
  - static 230
- outliers 94, 114, 118, 122, 168
- paint 92, 93
- Papua New Guinea 47, 171, 257
- parameter recovery 23, 25, 38
- passive monitoring 88, 101, 106, 124
- *Pentaclethra macroloba* 47
- percentiles 51
- permanent plot 3, 80, 85
  - independent observations 119
  - procedures 92
- pests 173
- Philippines 5, 18, 138
- phytocentric view 134
- pine
  - cypress 19, 20, 87, 102, 140, 141, 160, 201
  - hoop 47, 147, 150, 178
  - jack 235, 258, 278
  - klinkii 47, 278
  - loblolly 21, 60, 153, 190, 210, 241, 278
  - lodgepole 255, 265, 278
  - Mexican 146, 147, 278
  - ponderosa 62, 149, 188
  - radiata 27, 51, 137, 146, 147, 149, 152, 177, 278
  - red 54, 163, 235, 278
  - Scots 46, 278
  - shortleaf 21, 153, 197, 278
  - slash 153, 177, 278
  - white 21, 146, 147, 153
- *Pinus*
  - *banksiana* 235, 258, 278
  - *contorta* 255, 265, 278

- *echinata* 21, 153, 197, 278
- *elliottii* 153, 177, 278
- *patula* 146, 147, 278
- *pinaster* 38, 278
- *ponderosa* 62, 149, 188
- *radiata* 27, 51, 137, 146, 147, 149, 152, 177, 278
- *resinosa* 54, 163, 235, 278
- *strobus* 21, 146, 147, 153
- *sylvestris* 46, 278
- *taeda* 21, 60, 153, 190, 210, 241, 278

plant functional attributes 244
PLATIPUS model 65, 64, 271
plots
- buffers 89
- checklist 95
- edge bias 61
- experimental 88
- homogeneous 90
- mapping trees 92
- measurements 92, 95
- minimum girth 93
- number required 89
- numbering trees 92
- orientation 90
- passive monitoring 88, 106
- permanent 80, 85
- range of conditions 86
- remeasurements 96
- shape 90
- size 90
- subjective location 86
- visibility 89

point sample 90
Polyscias 277
- *australiana* 277, 169, 180
- *murrayi* 277, 180

precision 105, 282
principal components 151
probability density function 23
- Johnson's S_B 23
- normal 75
- Weibull 23

Prognosis model 7, 11, 72-74, 77, 166, 169, 170, 187, 198, 199, 209, 220, 224, 233, 257, 270, 273, 274
Pseudotsuga menziesii 62, 63, 177, 278
quadratic spline 38
Queensland 11, 12, 19, 20, 64, 87, 101, 127, 137, 140, 141, 147-152, 160, 181, 184, 186, 195, 199, 228, 231, 237
R^2 82, 125-127, 132, 157, 213, 214, 218, 291, 292, 294
Randia fitzalanii 150, 278
Rapanea achradifolia 150, 279
re-calibration 205
recruitment 8, 192
- 2-stage model 198
- dynamic model 194
- static model 194

rectangular plots 90
redwood 46, 253, 279
regeneration 192
- model 201

regression
- linear 116
- non-linear 117
- stepwise 116

remote sensing 68, 151, 152
- Landsat 152, 258, 272

research needs 243
- competition 62, 162
- defining 215
- growth relationships 245
- mortality 181
- site productivity 147, 154

residuals 112-116, 122-126, 143, 144, 207, 211, 212, 215, 219, 293, 294
- studentized 114, 122

response variable 104, 108
robust estimators 118
sampling 80, 84, 226, 283
- error 91
- large trees 86
- strategies 123, 227

Sarawak 40, 255, 260

scatterplots 113
self-thinning 172-175, 263, 266, 269, 274
Sequoia sempervirens 46, 279
serial correlation 35, 67, 69, 74, 120, 158, 169, 170, 211, 283, 288
shade tolerant 42, 47, 62, 128, 138, 196, 202, 283
Shorea 40, 278
  robusta 80, 279
silvicultural treatment 18, 96, 100, 173, 182, 184, 185, 200, 277
single-tree model 5, 69, 72, 74, 76, 202
site assessment
  climate 145
  floristic classification 148
  geocentric 135
  geographic information 152
  geology 147
  height-age 134
  indicator species 149
  multiple estimates 153
  ordination 149
  phytocentric 134
  remote sensing 151
  soils 147
  topography 146
site form 19, 87, 102, 139-142, 277, 293, 294
site index 14-20, 27, 62, 64, 67, 68, 105, 134-142, 145-149, 152, 154, 160, 163, 164, 188, 212, 216, 227
site productivity 134
  growth index 143
  height-diameter 139
  natural basal area 137
  site form 139
  site index 134
  stand height 138
  volume production 142
site quality 18, 86, 88, 100, 137, 141, 180, 195, 201, 220, 221, 226, 283
size class model 5, 76, 281-284
size, relative 60, 162, 178, 202
Sloanea australis 52, 277
soil 147
  depth 147, 152, 153, 147
  nutrient status 147
  parent material 147, 151, 152, 168, 188, 254
  physical characteristics 95
  type 85, 86, 104, 187, 277
South Africa 147
species
  groups 127
  identification 127
  pivotal 245
  presence/absence 151
  species–area curve 90
spruce 63, 67, 68, 202, 232
stand table projection 35
  critical upgrowth 38
  J-curve 39
  largest tree 40
  quadratic splines 38
states of markov chain/matrix
  absorbing 44
  steady 45
  transient 44
stationary assumption 45
steady state 47
stem analysis 79
STEMS model 11, 74, 77, 194, 219, 253, 258, 260, 267, 270
stepwise regression 116, 123, 162
Sterculia laurifolia 143, 279
stochastic model 7, 8, 43, 69, 182, 261
stratification 226
stratum 2, 86, 100, 212, 216, 226, 235, 236, 283
sugar maple 165, 241
suppression 45, 48, 63, 71, 136, 172, 173
survival 54, 67, 68, 168, 178-181, 187, 222, 233, 247, 266, 290

sustainable
  diameter distribution 229, 232
  yield 234
Swaziland 146, 257
Swietenia 80, 278
Syzygium
  kuranda 143, 279
  wesa 189, 279
temporary plots 2, 15, 80, 86, 101, 145
thinning 88, 89, 182, 281, 284
  competition index 61
  from below 162
  liberation 184
  self-thinning 172
  stand parameters 180
threshold
  increment 62, 176
  recruitment 196
timber stand improvement 182, 184
time of passage 35
Toechima erythrocarpum 143, 279
top height 24, 26-28, 160, 175, 181, 208, 276, 282, 284
topography 85, 90, 93
transient state 44
transition function 26
transition matrix 29, 31, 43, 47, 48, 194, 222, 257, 264, 276, 286
treatment response 88
tree list models 55, 57, 71, 75
tripling 73, 72, 168, 169, 211
Uganda 256
understorey 150, 201, 255, 265
upgrowth 35-39, 42, 45, 49, 194, 281, 282, 284, 289
USA 31, 54, 62, 71, 74, 77, 99, 138, 139, 146, 147, 149, 153, 168, 172, 187, 196, 238, 245
Usher assumption 46, 284
Usher matrix 43, 45-48, 56, 203, 295
validation 204
variable 284
  binary 104, 280
  explanatory 104, 108, 280
  qualitative 104, 282
  response 104, 108, 283
variance
  heteroscedastic 121
  homogeneous 121
  non-constant 114
  propagation 210
variance-covariance matrix 22, 64
verification 204
virtual reality 249
Vitex acuminata 150, 279
voucher specimen 94, 284
Weibull 23-25, 37, 38, 51, 76, 112, 232, 253, 261, 268, 273
weighted regression 121, 157
western larch 139, 140, 268, 278
Xanthophyllum octandrum 52, 143, 277
yield 4, 106
yield prediction 234
  cutting cycle analysis 236
  explicit 76
  heuristic simulation 236
  implicit 76
  linear programming 238
  time of passage 35
  yield scheduling 236
yield table 4, 15
  normal 15
  variable density 15, 272

Subject Index