INTEGRATED YIELD FORECASTING AND HARVEST SCHEDULING IN A TROPICAL PINE PLANTATION IN FIJI

A Thesis

submitted in partial fulfilment

of the requirements for the Degree

of

Doctor of Philosophy in Forestry

in the

University of Canterbury

by

Teodoro R. Villanueva

University of Canterbury

1992

ACKNOWLEDGEMENTS

I wish to thank my research supervisor Graham Whyte, whose expert advice and encouraging guidance during my stay in the School of Forestry, has made the completion of this project possible. Right from the start of identifying the research project through to preparation of the final draft of the thesis, Graham has been a very supportive supervisor. My three years of collaboration with Dr Whyte has been a very worthy and fruitful experience.

I am especially thankful to Dr R. C. Woollons of the School of Forestry for his advice and help on growth and yield modelling especially for many of his SAS programs and to Dr J. George of the Department of Management for his advice on the LP Harvest Scheduling component of the research. Thanks are also due to Bill Rea of the University Computing Centre and my friend David Baird who were very helpful in making SAS and Statistics indeed useful tools in the conduct of the various analyses in this research.

I wish to thank the Overseas Development Assistance Program of the Ministry of External Relations and Trade (MERT) for awarding me the scholarship that provided the funds for my studies and stay in New Zealand.

I give many special thanks to the Staff and other Postgraduate students of the School of Forestry for making our stay in New Zealand memorable. I personally thank Prof. Sweet, head of the School of Forestry, for his concerns and moral supports. Thanks are also due to Dave Clark for his help and advice in computing.

My colleagues, Liu Xu, Euan, Joseph, Ogweno and Noli deserve special thanks. They have been very supportive yet always probing.

I would like to acknowledge the moral support provided by my colleagues in the College of Forestry, University of the Philippines (U.P.) at Los Baños. I also thank U.P. Los Baños for allowing me to go on official leave of absence to be able to study in Canterbury.

Many friends in Christchurch have been very helpful to me and my family. I thank them all for making our stay in Christchurch enjoyable.

Finally, I personally acknowledge the support and love that my wife, Lyn, and our children Von and Ava have always provided. They have been the driving force to this endeavour.

ABSTRACT

This thesis reports on enhancements of two planning method components aimed at improving management and planning of forest plantations in the tropics. The two modular planning models subjected to detailed study are growth and yield modelling and harvest scheduling. A case study relating to Caribbean pine in Fiji is used to demonstrate the refined capability.

Growth and yield modelling has been improved by applying modern statistical and computer techniques to solve non-linear equations that describe growth of stands appropriately. Further improvements have been achieved by developing diameter distribution growth and yield models solved by a combination of parameter recovery and prediction method thereby ensuring compatibility between average stand values and diameter distribution values. In conducting improvements in growth and yield models, data manipulation and data validation procedures are described and reviewed in detail to emphasize their importance, particularly for non-linear regression fitting of equations, growth, yield and diameter distribution projection modelling.

Various growth projection equations were tested before final stand average functional forms for basal area per hectare, standard deviation of diameter at breast height outside bark, maximum diameter at breast height outside bark and survival per hectare were identified and then integrated into the growth and yield model. The precision of the equations was assessed through graphs and statistics relating to residuals. The stand simultaneous growth and yield equations

solved and used in the model consist of modified forms of different growth projection functions, such as the Gompertz, exponential and Schumacher, which were then used to derive a diameter distribution based on the Reverse Weibull probability density function. The diameter distribution growth and yield model was prepared as a simulation model to predict stand average values then, in conjunction with existing stem volume and taper equations, to derive stand and stock tables that allow disaggregation of diameter classes into log types. Three simulation models were created, one in Vax Fortran, one in PC Fortran and the other in spreadsheet format to enhance the models's portability.

The harvest scheduling model developed is a spreadsheet based LP model which is able to schedule harvests from a number of stands within a medium-term planning horizon using different logging methods with the log harvest to be delivered to different ports or utilization plants. A Fiji case study provided a demonstration of the modelling capability for fifteen stands, seven years, four logging methods and two ports.

This new kind of LP harvest scheduling model was developed with a deliberate intention to facilitate the running of it with the input from the improved growth and yield model. In developing this harvest scheduling model, the nature of LP in general was first reviewed and compared to other tools of harvest scheduling like binary search and simulation. LP harvest scheduling was found in this review to be a widely used tool and solution algorithms for which abound. A major problem with most solutions was the need to cater for sophisticated report writing and matrix generation. These two concerns were specifically addressed in the model developed as part of this study. The use of a spreadsheet as input to the LP was seen to be an efficient way of overcoming some of the major criticisms levelled at LP by potential users. The methodology developed was also advantageous because of its capability to facilitate the integration of growth and yield outputs with harvest scheduling.

It was concluded that forest planning models can be readily improved with software and hardware that developing countries can easily afford. The models reported here harness the capabilities of the now commonly employed spreadsheet as a powerful tool for easier routine input, output and sensitivity analysis, to assist decision making for harvest scheduling and to simplify managerial planning and control.

CONTENTS

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

The need to evaluate alternative methodologies for improving growth and yield modelling of tropical pine plantations and for scheduling harvests from them is vital in view of developments occurring in forestry in most tropical countries. The most important of these is that vast areas of forest plantations needed to meet the demand for wood and employment in rural areas are already established and there is a consequent pressing desirability to manage these plantations properly. The need to improve growth and yield modelling is of utmost importance because it plays a central role and provides the underpinnings for management planning. Present projections are generally still crude and uncertain in most tropical countries because of an apparent lack of commitment to data gathering for growth modelling. Similarly, management prescriptions and harvest scheduling are usually done through arbitrary rules of thumbs because planning models are not accessible if indeed they ever exist.

This chapter provides, therefore, the background to the study, a statement of the problems it intends to address, including its scope and the resultant objectives it aims to achieve.

1.1 Background of the Study

As for any planning model, modelling methodology needs to be specifically adapted. The methodology for growth and yield modelling and harvest scheduling for this study is for Caribbean pine plantations in Fiji, but it could easily serve as a prototype for other species and tropical plantations in countries that plan to invest in tree plantations.

Caribbean pine (Pinus caribaea var hondurensis), a subtropical pine tree species with wood suitable for pulp, sawlog and pole production, is an important plantation species throughout the tropics. Endemically growing in the Caribbean islands and planted in plantations in many other tropical countries, this pine species has been shown capable of growing at a wide range of elevation. A versatile plantation tree species, it grows even on poor and infertile soil (Lamb,1973). Studies are still lacking on detailed aspects that will strengthen the favourable indications and confirm the favourable growth potential for using Caribbean pine as a commercial plantation species.

In the early 1970's and in anticipation of an increasing world demand for timber products which might not be satisfied by plantation industries in temperate countries in the Asia/Pacific region, Fiji embarked on a plantation programme on a commercial scale to take advantage of early favourable indications of the growth potential of P. caribaea. This plantation resource has expanded rapidly to about 35 000 ha; its continued good prospects have been heavily dependent upon sound management. Managers of the resource have been guided along policy lines of maximum income from current stands and minimum capital expenditure for industrial development.

Through the years of administration by the Fiji Pine Commission (FPC), now Fiji Pine Limited (FPL), and in cooperation with various forestry agencies, research has been conducted to provide managers with information and tools that they can use to achieve stated goals. Two areas of research that have been continually studied are growth and yield modelling (Geiser, 1977; Manley, 1977; Broad, 1978; Wybourne,1982; and Reid, 1986) and large scale forest harvest scheduling (de Kluyver et al., 1980; Eng, 1982; and Whyte, 1987).

Many developing countries embarking on intensively managed commercial plantations can draw several valuable lessons from the early experiences of Fiji in the area of forest plantation management. Nevertheless, what has been started early in Fiji still needs refinement, as new technology and research methodology develop further. As plantations planted in this species grow rapidly with the passage of time, effective management of this species becomes urgent. Proper management of the resource relies heavily on planning models.

Before this study was undertaken, the planning models which had been developed for this resource included growth and yield models developed by various authors and a harvest scheduling model as already cited above. All of these models lack ideal properties for planning purposes, namely: the growth models have theoretical and practical deficiencies, while the harvest scheduling models were not very easy to use or understand. None of the growth models, for example, has employed the accepted form of sigmoid functions that possesses desirable mathematically and biologically sound properties.

Most of the previous models used logarithmic transformation which can be a substantial but unavoidable source of bias in growth estimates. With advances in computing, non-linear models can be solved without the necessity for log transformation. While the previous model (see, for example Whyte, 1978) is of the diameter distribution form, the approach is still deficient because the parameter prediction method was used to derive the distribution. This approach is practically deficient because it gives incompatible stand level and diameter distribution level estimates. Another practical difficulty of the previous models is that they are not implemented to be portable nor can they be easily linked to other planning models. The harvest planning models using linear programming (LP), dynamic programming (DP) and simulation are also not readily portable and too complex for the manager to understand and implement without outside help.

This study is aimed at remedying these deficiencies through providing systems specifically suited for tropical forest plantations in developing countries and through taking into account the resources available for planning that these countries normally have.

Substantial changes are occurring in the forestry sectors of many developing countries and there is a clear indication of a dire need to establish vast areas of plantation forests in many of them. Most developing countries now realize that they should no longer cut their remaining natural forests, either because they have already been exhausted or because of fear of obliteration of the resource. In the Philippines, for example, the remaining virgin forests are expected to have been logged over in the next five years. Here and in other countries, on the other hand, the growing demand for forest products continues to rise.

The second growth forest could be used to buffer supply areas, yet they cannot be relied upon to any great extent because of their poor stocking. Sources like log imports are remote possibilities but unlikely to eventuate because timbered countries would like to keep their own forest and curtail exports so as to maintain a sustainable level of it, and also because of economic considerations. Recognition of environmental concerns means the remaining virgin forests cannot be intensively logged to augment log supply.

Thus, it is likely that in the next quarter of century, the forestry sector of many developing countries will be engaged in the task of establishing and maintaining forest plantations. More and more forest lands will be devoted to plantations and even poorly stocked second growth forest areas may give way to plantations.

The preceding log supply scenario suggests an imminent dependence of most developing countries upon wood from plantation forests. The assurance of the perpetuity of the supply of logs from this type of forest still much depends, however, on how well they will have been planned and managed.

Given the nature of plantation forests that have already been established, planning for plantation management will not be an easy task let alone an inexpensive exercise. Planning needs tools to be efficient and to be able to help management attain ideals such as increasing log productivity, producing satisfactory levels of stocking and distribution of age classes ensuring continuous site and soil amelioration and providing service to society.

Planning and management system models offer a means of assisting this required process. Models, being representations of the real world, provide a means whereby alternatives can be tested without the risks and costs of experimenting on the real forest. Scenarios can also be analyzed quickly and accurately. They can be used to evaluate the effects of the present decisions on the future.

The cost and huge amount of investment involved in plantations and the fact that investments in them have to compete for capital against other investments further justifies the need for easily understood and relevant tools to aid management planning.

1.2 Statement of the Problem

The present study is based on the premise that managers need: a) more detailed stand information, specifically on diameter distribution to allow more refined planning of the use of the resource in general and felling and thinning of individual stands in particular; b) a harvest scheduling model using computer packages with which managers are familiar; and c) closer integration of these two forms of planning models.

The need for more detailed stand information is most urgent. Continuous changes in harvesting techniques, utilization standards, management approaches and silvicultural practices all contribute to this need. More detailed stand information is also needed to match wood supplies with timber characteristics and market demands. The expected product mix (pulpwood, sawlogs, veneer logs) that is harvestable from stands is defined largely by diameter limits but varies also with changes in management practices that are actually implemented.

Harvest scheduling models are needed to ensure that resources are harvested at the right time, at the right place and with the right methods. The problem that besets managers is the determination of where, when and how much to cut in order to achieve a reasonably stable and continuous supply of forest products. Systems analysis and linear programming can be used to derive optimal cutting schedules for a particular requirement and forest conditions, and they offer assistance with arriving at sound solutions to the manager's problems.

Forest management planning is now relying more and more on operations research tools. Operations research is a methodical approach to the formulation and interpretation of decision problems using mathematical analysis. Its appeal is in the way of looking at problems and in the answers it provides in the form of information about the decision process itself and which are implementable in the environment where problems exist.

Forestry planning models are still being developed which address single stand and single objective management, whereas managers should be concerned with multiple objectives at the forest level. With advances in computers, most of the growth and harvesting models should now be designed as parts or sub-models within larger models rather than independently. The challenge that this study picks up is the development of a modern growth and yield model output from which is easily integrated with a harvest optimization model. This integrated approach uses growth and yield models as the direct data generator for a harvest schedule optimization algorithm. Improvements in each of the single modular components is also a problem addressed here.

1.3 Scope of the Study

The primary focus of this study, therefore, is plantation forest regulation, specifically with respect to growth and yield modelling, harvest schedule modelling and their integration in the context of tropical forest plantations. The tropical forest plantation data analyzed in this project pertain to specific plantations in Fiji located in Lololo, Seaqaqa and Drasa, but these are simply typical of a large number of similar situations. The models and the methodology developed, therefore, can serve as prototypes for other tropical plantations. The research explores the application of recent statistical tools for growth and yield modelling, advances in operations research for harvest schedule modelling and computer technology for implementing integrated planning models.

1.4 Objectives of the Study

The overall aim of this study is to develop a methodology to improve medium term planning in pine plantations in the tropics towards the development of a decision support system that has easily interfaced modular components, one for growth modelling and the other for harvest scheduling. Specifically the objectives are:

1. to develop and evaluate growth and yield models for predicting diameter distribution of Caribbean pine plantations;
2. to develop means of implementing growth and yield models in a way that they can produce output that can be easily interfaced with other planning models;
3. to develop a spreadsheet based harvest scheduling model which can easily use the yield forecasts derived in 1 and 2 and which managers can easily understand and apply routinely;
4. to develop an integrated planning model for effectively characterizing stands of tropical pine plantations and analyzing planning options for harvesting these stands.

CHAPTER 2

REVIEW OF LITERATURE

This chapter examines recent developments in growth and yield modelling brought about by advances in statistics and computer technology. A historical review of growth and yield modelling is done first. Growth models are then reviewed as either stand average or diameter distribution models, emphasizing the importance of compatibility between whole stand and components. The general diameter distribution growth model and methods of solving probability density functions representing diameter distribution are also reviewed. The second part of the chapter examines the different approaches in harvest schedule modelling particularly operations research tools like simulation and linear programming. Advances in LP in general and its application in harvest scheduling in particular are evaluated. Such a review was done to recognize advances in computing and the current needs of managers. Efforts in interfacing planning models are also reviewed. Finally, the chapter reviews the uses of the spreadsheet environment and its potential in facilitating the needed enhancement and ease in implementing and interfacing planning models.

2.1 GROWTH AND YIELD MODELLING

Effective forest management decision-making relies greatly on accurate forecasts of realisable growth and yield. Growth modelling and yield forecasting are indispensable for updating inventory information, management planning, evaluating silvicultural options and scheduling harvests. Growth information is used as a measure of stand performance, to define how much timber has accrued in a particular stand over a specified time and the changes that have resulted from past cutting. It is also useful for answering such questions as the desirable level, structure, and composition of growing stock, the number and intensity of intermediate cuts that may be applied, the effect of initial tree spacing and the most economic rotation length (Davis, 1964). Various types of growth and yield model with strengths and weaknesses for different purposes, are discussed in the sections that follow.

2.1.1 Average Stand Models

Munro (1974) categorized growth and yield models roughly into whole stand models and individual tree models. Whole stand or stand average models consist of equations which predict the yield per unit area of the whole stand or some specified portion as a function of age, stand density, and site index. Whole stand models can be further subdivided into per unit area values only and size class distribution information, as exemplified by diameter distribution models. Individual tree models can be classified as either distance independent or distance dependent. Later classification of growth models was done by Bruce and Wensel (1987) which emphasized that the forest condition being modelled and the purpose of the modelling dictate the choice of model. The work classified models as either "process" or "empirical" models. Process models consider the biological processes that convert CO_2, nutrients and moisture into biomass through photosynthesis, and thus may also consider precipitation, hours of sunlight and other environmental processes. Empirical models are based on periodic tree measurements, with no attempt to measure every factor that may affect tree growth. Overall, the choice of model depends upon data availability, modelling objectives (including stand detail for a particular decision to be made), and upon the background and interest of the researcher.

The construction of yield tables has been oriented toward the prediction of future stand conditions which are important in the estimation of future stand values. For the sake of simplicity, early yield tables used the concept of normal stocking which is the mean stocking level of a large number of undisturbed stands or stands growing according to a specified density regime. The first yield predictions were made using normal yield tables for even-aged stands of a given species in Central Europe. Temporary plots in stands of normal stocking were used to construct tables through graphical techniques. But, the growth of stands with abnormal stocking could not be predicted from these tables without adjustment. With the advent of the computer, multiple and simple regression models were used to predict growth and yield for many combinations of age and site. Multiple regression was first used by Mackinney and Chaiken (1939) to construct variable density yield equations. Buckman (1962) introduced a very limited polynomial model where yield was obtained through mathematical integration of the growth equation over time. Before then, growth and yield had been independently developed, often resulting in illogical and inconsistent results. Since then, researchers have mostly taken into account the logical relationship which should exist between growth and yield equations: that is, the yield function represents the sum of continuous growth increments, while the growth function is the first derivative of the yield equation with respect to time. In other words, the algebraic form of yield can be derived from mathematical integration of the growth function (Clutter,1963). Clutter (1963) further elaborated on this concept and developed a compatible growth and yield model in natural loblolly pine stands. Sullivan and Clutter (1972) improved this concept more fully by developing analytically as well as numerically consistent growth and yield predictions using difference equations derived from the projection equation form of the Schumacher equation. The form of the equations is based on the consideration that a derivative-integral relationship exists between the growth function and the yield function for quantities such as stand volume and basal area (Sullivan and Clutter, 1972).

This important development brought forward a new step in yield modelling - the construction of projection equations that generated simultaneous and compatible estimates of growth and yield. With the advent of modern computing in the last 20 years or so, growth and yield models have taken the form of even more complex models consisting of sets of equations for many stand variables. Aside from the very important characteristic of compatibility , today's growth and yield models possess other desirable characteristics - namely, that of consistency, an asymptotic value and path invariance. These are a consequence of using sigmoid projection functions. Consistent models predict logically the projected stand value equal to the initial stand value when the initial age approaches the projection age. Asymptotic models possess an upper limit or asymptote on future value of a stand variable as the projection age approaches infinity. Path invariant models are able to predict stand values at a certain age irrespective of the number of steps involved in the projection.

2.1.2 Diameter Distribution Models

Diameter distribution models provide estimates of the number of trees and yield per unit area by diameter at breast height (dbh) classes. The estimate of stand volume per unit area per dbh class can be obtained by multiplying the estimated number of trees per unit area in the dbh class by an estimated volume of a tree with dbh equal to the dbh class midpoint and also sometimes predicted average height i.e. from a two dimensional tree volume equation. Summing all volumes over all dbh class provides an implicit estimate of total stand volume per unit area.

A diameter distribution is a very useful concept for describing the properties and structure of a stand of trees. From diameter data, volume can be derived, then conversion cost and product specifications can be determined (Bailey and Dell, 1973). Models that supply information about diameter distributions allow managers to plan ahead on the basis of expected diameter and volume distributions. This information is useful in scheduling appropriate equipment for harvest planning schedules. It also helps to determine raw material values, harvesting costs, product mixes, and forest management plans. Thus, accurate stand and stock table projections are vital for making sound forest management decisions (Hyink and Moser, 1983).

Diameter distribution modelling has been a viable means of predicting yields and stand structure as shown by the works of Clutter and Bennett (1965) for even aged forest stands; McGee and Della-Bianca (1967) for natural stand populations; Lohrey and Bailey (1976) and Bennett and Clutter (1968) for unthinned stands of slash pine plantation; and Baldwin and Feduccia (1988) for both thinned and unthinned stands of Loblolly pine.

Diameter distribution has also been used to model growth of thinned stands. Thinning is an important silvicultural tool applied to concentrate growth on the best and largest trees manifested in both tree size and quality and, consequently, on yield. Diameter distribution models have been used to evaluate the effects and results of different thinning regimes better. The Weibull distribution is now the most commonly used form employed for modelling diameter distributions in thinned plantations (Clutter and Jones, 1980; Bailey et al., 1981; Strub et al., 1981; Cao et al., 1982; Matney and Sullivan, 1982; Burkhart and Sprinz, 1984). In modelling thinned stands, Cao and Burkhart (1984) proposed joining different segments of cumulative distribution function (cdf) together to form a single cdf which is flexible enough to model irregularities in diameter distributions typical of many thinned stands. McTague and Bailey (1987) proposed the use of diameter distribution percentiles to describe past stand history in the absence of records on the age of thinning and the exact amount of basal area removed.

Different researchers have adopted different approaches to modelling diameter distribution, the most commonly used being the probability density function (pdf) approach. The basic assumption in diameter distribution modelling through use of probabilities is not only that the underlying diameter distribution can be adequately characterized by a pdf but that this distribution has a skewed, but normal-like shape which is ideal for depicting the diameter distribution: but any other kind of appropriate function could be used. Interest and research in describing diameter distribution in forest stands using pdf started as early as 1898 when de Liocourt described the structure of balanced uneven aged stands using a specific mathematical model for geometric series projection. Building on de Liocourt's idea Meyer (1952) suggested the use of the reverse J-shaped exponential probability density function for modelling stands of this type. Since then, pdf's which have been used to model diameter distributions include the Beta (Clutter and Bennett, 1965), Gamma (Nelson, 1964), log normal (Bliss and Reinker, 1964), Johnson's S_b (Hafley and Schreuder, 1977), and Weibull (Bailey, 1972). The choice of an appropriate pdf is usually guided by consistency and simplicity: it should be, moreover, for a single function capable of depicting a full range of unimodal continuous shape that usually characterize diameter distributions. Historically, the Weibull is preferred only because it has a closed cumulative distribution function (cdf) and it can cover the reverse J-shape with varying degrees of either positive or negative skewness. Secondly, its parameters can be easily related to shape and location features that vary in a consistent manner with stand characteristics. Thirdly, because one of the major applications of pdf is integration to obtain proportions of the stand less than a stated diameter, it must have a well defined closed form of the cumulative distribution function (Bailey and Dell,1973). This last property is less important now with the calculating power of modern computers.

Hafley and Schreuder (1977) compared six pdf's in terms of their flexibility on the skewness squared and kurtosis plane. Johnson's S_b was found most superior followed by the Weibull, but no other authors have reported similarly. Generally, the Weibull has been selected most frequently in recent researches.

2.1.2.1 Parameter Prediction Method

Given a data set and assuming that a family of distributions has been chosen, diameter distribution modelling involves estimating the parameters of a chosen pdf. The estimation of the pdf parameters for each set of data could employ procedures such as maximum likelihood, percentile or method of moments. The parameter prediction method utilizes regression techniques with the parameter values as dependent variables and the stand characteristics such as age, density, and site quality as prediction variables . This approach was employed by Smalley and Bailey (1974) for short leaf pine plantations, Dell _et al._, (1979), Feduccia _et al._, (1979), Baldwin (1982), and Manley (1977). Most researchers have assumed a linear relationship between stand variables and pdf parameters. Kuru (1989) however suggested that such an assumption has little biological foundation and may be very weak, as other variables like stand diameter variables have greater influence on stand structures. Frazier (1981) also observed that parameter prediction equations used to estimate the parameters of the pdf would typically have a coefficient of determination in the range of 0.1 to 0.2 which is too low and is indicative of inadequate understanding of the true relationship of the distribution parameters to the selected stand variable. Relationships among parameters make it very difficult to develop prediction equations that would explain a high percentage of variation in parameters. Bailey et al., (1981) circumvented this problem by predicting the 24th, 63rd, and 93rd percentiles of the Weibull pdf from stand variables then used these statistics to estimate the three parameters of the pdf. The regression equations were better than the parameter prediction equations.

2.1.2.2 Parameter Recovery Method

The parameter recovery method is a response to the need for forest modellers to have compatible estimates of whole stand and diameter distribution models. Frequently, yield estimates of these models for a given set of stand conditions could not be guaranteed to be the same when they are constructed independently. Even when constructed from the same set of data these two models did not necessarily produce the same estimate of stand yield for a given set of stand conditions (Daniels et al., 1979).

Hyink (1980a) proposed an approach, termed the parameter recovery method, the advantage of which is a mathematical compatibility of the whole stand and the diameter distribution yield models. The procedure involves the prediction of whole stand attributes (usually basal area and stand diameter variables), and use of these estimates as a basis to predict the parameters of the underlying distribution. The parameters are "recovered" from estimates of stand attributes which are expressed as functions of the expected value and the variance of the dbh distributions. The first two non-central moments of dbh distributions are examples that have straightforward interpretation corresponding to stand mean diameter and basal area. If these variables are predicted reliably they can be used as a sound basis for prediction. Hyink (1980a) discussed the theoretical and statistical framework of such an approach which was later adopted by Frazier (1981), Matney and Sullivan (1982), Cao et al., (1982), Bailey et al., (1981), Bailey et al., (1982), Cao and Burkhart (1984).

2.1.2.3 The General Diameter Distribution Yield Function

In general, at any time **T**, the yield table constructed by the diameter distribution (Strub and Burkhart, 1975 ; Frazier, 1981) is

$$Y_i = N_t \int_{D_l}^{D_u} g_i(x)\, f(x:\theta)\, dx \qquad (2.1)$$

where

$\mathbf{x}$ = tree dbh

$\mathbf{N_t}$ = number of trees per unit area surviving at $\mathbf{T}$

$\mathbf{D_l}$, $\mathbf{D_u}$ = lower and upper limits of integration respectively for that particular $g_i(x)$

$\mathbf{f(x;\theta)}$ = the pdf

θ = the parameter vector

$\mathbf{g_i(x)}$ = the i^{th} function of the tree dbh

$\mathbf{Y_i}$ = the per unit area value of the i^{th} stand attribute defined by $\mathbf{g_i(x)}$.

Hyink (1980a, 1980b) showed that any number (**k**) of pdf parameters can be solved as long as a set of **k** functions $g_1(x)$, $g_2(x)$, ..., $g_k(x)$ and the values of the corresponding stand attributes Y_1, Y_2,...,Y_k also exist. The stand attributes may be basal area, stand diameter or a statistic relevant to the distribution being considered. Nevertheless, such an approach depends finally on which parameters can be easily evaluated. Provided that a set of equations is consistent, a solution exists for each of the **k** parameters by solving **k** equations for **k** unknowns. It can be deduced that many different sets of equations can be constructed for any number of pdf's. Frazier (1981) compared two basic sets of equations to solve the parameters of the Weibull and Beta, the most commonly used pdf's in modelling diameter distribution. The first set consists of one or more volume equations in combination with non-central moment equations and the other set consists of non-central moments of the random variable **x**, which can be designated as $E(x^i)$. The latter is called the moment based parameter recovery system and will be discussed further in the sections that follow.

2.1.2.4 The Weibull pdf and the Moment Based Parameter Recovery Models

The most widely used pdf in stand growth modelling is the Weibull pdf (Pinder et al., 1978; Schreuder et al.,1979; Somers et al., 1980). Since its first use as a diameter distribution model (Bailey, 1972), the Weibull pdf has been most extensively used to model distributions of tree diameters in even aged stands. The Weibull pdf has been found to have a flexible shape, as its parameters can be easily related to stand characteristics and its cumulative distribution function can be recovered in closed form (Bailey and Dell, 1973; Schreuder and Swank, 1974, Schreuder et al., 1979).

The Weibull pdf, as it is used to represent distribution of diameters is,

$$f(X) = \left(\frac{c}{b}\right)\left(\frac{X-a}{b}\right)^{c-1} \exp\left[-\left(\frac{X-a}{b}\right)^{c}\right] \quad (2.2)$$

for $a \le X \le \infty$

$= 0$, otherwise

where

X = dbh

a = location parameter

b = scale parameter

c = shape parameter.

The cdf is

$$F(X) = 1 - \exp\left[-\left(\frac{X - a}{b}\right)^c\right] \qquad (2.3)$$

$$\text{for } a \le X \le \infty$$

$$= 0, \text{ otherwise}$$

For applications to the distribution of dbh, $b > 0$ and $c > 0$ and $a \ge 0$ are specified further.

The distribution has been found to fit data adequately for many different types of forest stands and has been widely used to model stand structure of many plantation species (Feduccia et al., 1979; Bailey et al., 1981; Strub et al.,1981; Matney and Sullivan, 1982; Schreuder et al., 1979). Bailey and Dell (1973) fitted this pdf to published diameter distribution and showed its flexibility of the pdf to model various shapes of the distribution including mound shaped for even-aged stands and reverse J-shape for severely understocked stands affected by fire and heavy cutting. Feduccia et al.(1979) improved the forecast of plantation on cutover sites of loblolly pine where no intensive site preparation was employed, by using pdf and stem taper function but this approach of estimating the parameters directly through a regression function on stand characteristics is deficient because its compatibility with stand level estimates was not ensured.

The estimation of the parameters of the pdf is based on Equation 2.1. By integrating this equation over the range of diameters, **X**, for any $\mathbf{g_i(X)}$, the total value per unit area of the stand attribute defined by $g_i(X)$ is derived. The moment based parameter recovery system is defined by letting $g_i(X)$ equal $E(X^i)$, the i^{th} non - central moment of **X**.

$$E(X^i) = \int X^i f(X:\theta)\,dx \tag{2.4}$$

Frazier (1981) used this general formula for estimating moments and showed how, for example, the first and second non-central moments are estimated by Equation 2.4 resulting in Equation 2.5. Again, since **X** represents diameter, the first and second non-central moments below are the familiar equations for the average diameter of the stand and the quadratic mean diameter squared, which is related to the mean basal area per tree.

$$E(X) = \frac{\sum_{j=1}^{N} X_j}{N} = \overline{X} \tag{2.5}$$

$$E(X^2) = \frac{\sum_{j=1}^{N} X_j^2}{N} = \overline{X^2} \tag{2.6}$$

$$= \frac{G}{0.00007854 * N} \tag{2.6a}$$

where **G** is basal area per unit area and others are as previously defined.

Frazier (1981) then used the two non-central moments to solve the two parameters of the Weibull pdf. The system of equations estimated by Equation 2.3 for the 2-parameter Weibull consists of the first two non-central moments which are,

$$\bar{X} = \int_0^\infty X \ f(X:b,c)\,dx = b\,\Gamma(1 + 1/c) \qquad (2$$

$$\overline{X^2} = \int_0^\infty X^2\ f(X;b,c)\,dx = b^2\,\Gamma(1+2/c) \qquad (2.8)$$

where Γ( ) signifies a Gamma function:

$$\Gamma(z) = t^{z-1}\,e^{-t}\,dt \qquad (2.9)$$

The estimated variance and the coefficient of variation of the distribution are then solved, respectively

$$S^2 = \overline{X^2} - (\bar{X})^2 = b^2(\Gamma(1+2/c) - \Gamma^2(1 + 1/c)) \qquad (2.10$$

$$C.V. = \frac{S}{\bar{X}} = \frac{\sqrt{(\Gamma(1 + 2/c) - \Gamma^2(1 + 1/c)}}{\Gamma(1 + 1/c)} \qquad (2.11$$

As Equation 2.11 is a function of **c** only and with estimates of $\bar{x}$ and $\overline{x^2}$ it is then possible to solve for **c**. Once **c** is obtained, the **b** parameter is then solved using Equation 2.10.

The system of equations needed to solve the parameters of the Weibull with three parameters is more complex as the same number of equations and the same number of attributes are needed. This system of equations consisting of three non-central moments of the distribution of **X** as a function of **a**, b, and **c** has proved difficult to solve because of convergence problems (Frazier, 1981). An alternative proposal was to reduce the problem to the 2-parameter Weibull. This was done by considering **a** to correspond to the smallest possible value of dbh in the stand and set this parameter as a function of the minimum value. The problem is thus reduced to a 2-parameter Weibull.

The Weibull pdf used to represent distribution of diameters is not without inadequacies. Cao and Burkhart (1984), in addressing the inadequacy of the Weibull pdf to represent multimodal or irregular diameter distributions, developed a methodology that put different equations and cdf's together to form a single smooth cdf flexible enough to model irregular distributions. Their approach, which was based on a modified Weibull with five parameters that required five percentiles was found superior especially for diameter distributions of thinned stands.

Another form of the Weibull pdf, the Reverse Weibull pdf has also been tried to model diameter distribution because of some of the Weibull pdf's inadequacies. The graph of this pdf starts at a finite maximum "anchoring point" specified by the location parameter and moves towards the origin as X becomes infinitely small. The Reverse Weibull pdf, as it is used to represent distribution of diameters is,

$$f(X) = \left(\frac{c}{b}\right)\left(\frac{a - X}{b}\right)^{c-1} \exp\left[-\left(\frac{a - X}{b}\right)^{c}\right] \qquad (2.12)$$

for $a \geq X \geq -\infty$

$= 0$, otherwise

where

X = dbh

a = location parameter

b = scale parameter

c = shape parameter.

The cdf is

$$F(X) = \exp\left[-\left(\frac{a - X}{b}\right)^{c}\right] \qquad (2.13)$$

for $a \geq X \geq -\infty$

$= 1$,

for $X \geq a$

with terms as defined above. The mean and variance of this distribution are,

$$\bar{X} = a - b\,\Gamma(1 + 1/c) \qquad (2.14)$$

$$S^2 = b^2\,(\Gamma(1+2/c) - \Gamma^2\,(1 + 1/c)) \qquad (2.15)$$

Kuru (1989), finding that the maximum diameter can be more readily modelled and be closely associated with changes in stocking saw the potential utility of the reverse Weibull pdf. By setting the location parameter **a** as some function of the distribution of maximum diameter the study worked on the reverse Weibull distribution. The study found out that estimates of diameter distribution can be made more precise and accurate through the adoption of this pdf.

Equating the location parameter to the maximum diameter is not without problems, however, because clearly there is doubt about the estimated **Dmax** being equal to the true population Dmax. The resultant bias is further exacerbated when one projects the distribution with transition functions, since these variables are obviously affected by genetics, mortality, silviculture and microsites. The precision of the diameter distribution projection, therefore, may be extremely coarse.

Improvement in modelling was also brought about by the adoption of the parameter recovery method to solve the parameters of the pdf. The main advantage of the parameter recovery method is the ability to predict compatible whole stand and diameter distribution estimates of the stand attributes defined by the moments. In this system, consequently, the parameters of the pdf will be sensitive even to small changes in stand attributes.

It can thus be deduced that a crucial step in diameter distribution modelling, or in any model-building process for that matter, is sound estimation of the parameters. Bailey and Dell (1973), Ek et al., (1975), Strub et al., (1981), Frazier (1981) and Abernethy (1981) addressed different techniques, each with their accompanying advantages, disadvantages and problems in estimating the Weibull parameters. Burk and Newberry (1984) investigated further the possibility of recovering all three Weibull parameters considering the first three non-central moments and Zarnoch and Dell (1985) evaluated two methods of estimating the three parameters by computer simulation and field data comparison using maximum likelihood and percentile estimators. Other more elaborate studies on solving the location parameter **a** are illustrated in Kuru (1989) and Xu (1990). The amount and importance of efforts placed along this line of mathematical statistics should be evaluated carefully in proper perspective. Proper data acquisition and specification of mathematical theory for building growth models are also important steps that need to be given careful attention. Throughout this study, these three aspects of growth and yield modelling are given appropriate prominence.

2.2 HARVEST SCHEDULE MODELLING AND LINEAR PROGRAMMING

2.2.1 Description of harvest scheduling problems

Planning the future sequence of harvests of timber on a forest is one of the more difficult tasks for a forest manager to accomplish successfully, yet it is also one of the most relevant, because the achievement of the temporal and spatial scheduling of harvest operations means that the manager has a control of quantities such as growing stock volumes, growth rates, cash flows, present worths and returns on investment.

This thesis places emphasis on harvest scheduling, assuming that one of the major purposes of forestry is still to supply wood. Timber harvest scheduling dominates other planning efforts because of the historical and economic importance of timber as a commodity resource. Traditionally these schedules were looked at as a means to ensure the even flow of raw material products. It started in early European forestry where the concern was continuous production and self-sufficiency in timber products. The fear of a timber famine was the reason to organize forest regulation so that an even flow of timber could be supplied forever (Davis and Johnson, 1987).

Timber harvest schedules have found new uses. Today they are increasingly important because they provide a relevant means to describe and value a forest. Timber harvest scheduling models have also provided an ecologically sound concept for multi-resource analysis (Alston and Iverson, 1987).

The timber harvest scheduling problem consists of deciding when, where, and how much raw material to cut in order to attain all management objectives to acceptable degrees; all these decisions have strong irreversible economic impacts on investments, profits, benefits, and industrial activities.

Different factors govern the cutting schedule that best satisfies the objective to maximize yield or value from a forest. These factors include area of the forest, present volume and growth of the resource, rotation age or cutting cycle, number of cutting periods included in the schedule, and whether or not it is desirable to have the yields increase, decrease or remain constant in succeeding periods. Any change in these factors will affect the maximum total yield that can be scheduled to be harvested in the forest. All these components need to be quantified, a knowledge that has led to improved information systems and/or research to supply basic data (Kidd et al., 1966).

The harvest scheduling problem is not a trivial one. It is difficult, because it involves the long term nature of the timber production process in a way that introduces much uncertainty. As such, an appropriate level of uncertainty concerning future and biological conditions needs to be recognized and considered. Furthermore as timber is grown on large areas, scheduling problems expand to levels that involve almost unlimited numbers of possible cutting strategies.

The end result of the decision on where, when and how much to cut controls the efficient allocation of the factors of production like labour, capital and natural resources. By having a schedule, the manager can vary these factors and can determine how sensitive the schedule and total yield are to such changes. Efforts can then be directed appropriately as a consequence.

Harvest scheduling requires data that are relevant and accurate, as resulting schedules can be no better than the information and data used to construct them. These data include growth and yield, prices, costs, machine capabilities and existing management policies.

Growth and yield data are of crucial importance. The importance of the construction of yield tables so as to reflect present and expected net harvestable volumes per unit area cannot be overemphasized and is a critical step in harvest scheduling (Leak, 1964). Growth and yield data provide input to drive harvest schedule models. Growth and yield are usually inputted to planning models as discrete data. What planners have to do is to predict in a separate step the yield which is realisable in a particular stand. The difficulty with this method is the inefficiency with which models are used and which may also result in errors due to data handling, because any time there is update in the growth model due to changes in the stand initial condition, new growth data have to be generated and incorporated in the harvest planning model which will be reformulated and re-run.

2.2.2 Approaches to modelling the harvest scheduling problem

Traditional harvest scheduling models fall into one of two categories: 1) area control and 2) volume control. In area control, the area that will be harvested and regenerated is the same in each year or period as that which would be harvested in a fully regulated forest. If this is done the resultant volume harvested is defined by the timber on the area scheduled for cutting each year. In volume control, the essential decision is how much volume to cut each year depending on the total resource volume or its increment or both. The areas to cut are then chosen to satisfy this volume.

These traditional techniques cannot be used very successfully because forests are less uniform than the theoretical normal forest. Therefore, techniques to solve and analyze harvest scheduling problems and to produce efficient and workable solutions have flourished since the 1960's. The techniques include binary search, simulation and linear programming.

2.2.2.1 Binary Search

Binary search belongs to a class of simulation techniques. It uses forest inventory data and appropriate growth models to find the maximum even flow of volume or discounted net value that can be sustained over a finite planning interval. The constraints include certain harvest flow and ending inventory restraints. Two properties lead to the name binary search: (1) there is only one decision variable per period, the level of harvest, and (2) there are only two choices in the problem, either increase or decrease that harvest. As there are many aspects decided outside the model, binary search is considered a heuristic, that provides a shortcut through reducing the scheduling problem to a few decision variables and then exploiting the sequential nature of timber stand development to find the harvest levels that meet certain constraints. Its advantages are low cost per run and the ability to recognize the inventory in greater detail. The disadvantages are its inability to consider alternative management intensities, consider constraints beyond harvest flow and inability to find the optimal harvest schedule.

Some examples of binary search models developed thus far are:

SIMAC (Simulating Intensively **M**anaged **A**llowable **C**ut) searches for the maximum even flow harvest over 10-40 periods subject to meeting restrictions on the inventory remaining at the end of the planning horizon (Sassaman _et al_., 1972).

ECHO (EConomic Harvest Optimization) finds the maximum discounted net revenue for a forest under a situation in which stumpage price per unit received in a period is a function of the quantity sold. It is implemented by equating discounted marginal net return between periods i.e. similar to equating volume between periods in the usual iterative approach (Schmidt and Tedder, 1981). Its feature is that one period's prices or demand function is dependent on the previous period's.

TREES (Timber Resource Economic Estimation System) with the objective of finding the maximum harvest volume that can be sustained over some periods subject to timber flow constraints.

Johnson and Tedder (1983) outlined the binary search approach implemented in **TREES** (Timber Resource Economic Estimation System). These procedures are outlined in Figure 2.1.

With the objective of finding the maximum harvest volume that can be sustained over periods subject to timber flow constraints, **TREES** :

1. provides, an initial estimate of the harvest level along with an amount to increase (or decrease) the harvest if more (or less) can be harvested than the initial estimate;
2. determines the source of the first period harvest;

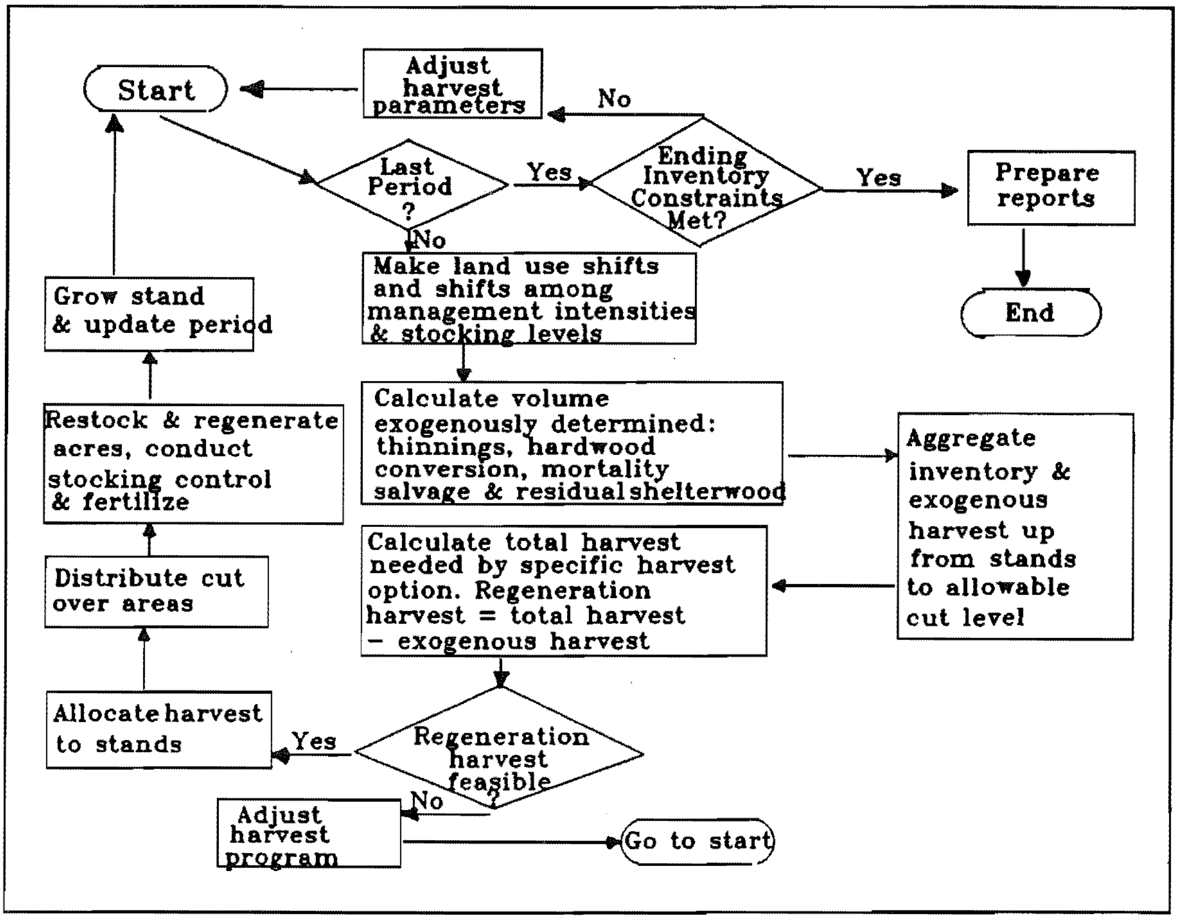


Figure 2.1 Binary search approach implemented in TREES.

3. deducts any intermediate harvest or predetermined harvest from the estimated total harvest;
4. allocates remainder of the harvest from stands according to stand priority rules that have been provided but if regeneration harvests exhaust the inventory, the initial harvest level is lowered by the decrement and the process begins again, whereas if the regeneration harvest can be met in the first period without exhausting the inventory, the harvest is taken, the inventory is updated to the second period and the process begins again.

TREES and **ECHO** answer decisions on two aspects in each time period - the amount of harvest and the order of the stands to be harvested. Both models presume that the harvest priority can be pre-specified to simplify the search for an optimal schedule. In **TREES**, for example the harvest priorities may be the oldest trees, the slowest growing trees or the highest value trees. Once the harvest priority is determined the amount to be harvested in each period is found through a binary iterative search started by specifying an initial guess for the total quantity to be harvested in the first period.

Binary search can be of two types: 1) ordinary binary search has one decision variable, i.e. the amount of harvest that can be sustained over the planning horizon; 2) sequential binary search has as many variables as periods in the planning horizon: i.e. the amounts that can be sustained starting at each period and going for selected periods into the future.

The major disadvantage of binary search is that it can consider only one criterion at a time. This major drawback of binary search was overcome by Hoganson and Rose (1984). By using heuristic simulation, alternative intensities of and optimal stand priority for harvest were found. Given an objective function, the price of stumpage in each period was varied until a set of prices was found for the timber harvest such that the best time to harvest each stand to maximize its present net worth on an individual basis is also the best time to harvest the stands in aggregate to meet the overall harvest constraints.

2.2.2.2 Simulation

One of the most popular simulation harvest scheduling model is **FORMAN** (**FOR**est **MAN**agement) being used by some provincial governments in Canada (Jamnick, 1990). The **FORMAN** model is a simulation model without any complex statistical models or mathematical relationships. As a bookkeeping device, it permits users to describe a resource in quantitative dynamic terms , to specify harvesting/silvicultural activities and to track the changes in the resource over time in response to these activities (Jamnick, 1990). **FORMAN** does not have explicit harvest flows, it uses operable limits to determine stand type eligibility for harvest, and it reports a solution for a given management scenario. Nor does it have an explicit objective function although it may be implicit in the harvest rules which are necessary inputs to the model (Jamnick, 1990).

Simulation-based techniques are basically descriptive. In them, the scenarios are specified and models are run to form details of the activities for specified scenarios. They are computationally easy but less detailed than mathematical programming models. Better scenarios may remain untested as tests are not exhaustive. However, simulation models, such as **FORMAN** when compared with LP models, are more appropriate to use where the harvest scheduling problem is relatively simple and limited to finding sustainable harvest levels for the silvicultural activities included in the model.

One other example of a simulation model is **IFS** (Interactive Forest Simulator). A modified form of this model to incorporate costs and revenue for implementing the prescribed strategies was used as the forest estate model in a previous application in Fiji. It was used to evaluate the long term consequences of continuing to implement a short term bucking model and a medium term LP model. It was used for coordination and for generation of more detailed information about the whole resource (Whyte, 1989).

2.2.2.3 Linear Programming

Of the mathematical programming techniques, linear programming is by far the most widely used in timber harvest scheduling (Curtis, 1962; Leak, 1964; Loucks, 1964; Kidd et al., 1966; Ware and Clutter, 1971; Nautiyal and Pearse, 1967; Navon, 1971 and; Clutter 1968). Applied to forestry and in general terms, linear programming is concerned with the problem of planning the complex of interdependent plantation activities for best possible use. It is a technique of specifying how to use limited resources available to managers, how best to utilize machine capacities and to meet demand requirements while at the same time obtaining a particular objective such as least cost, highest profit, or least time when these resources have alternative uses. It is a technique that systematizes for certain conditions the process of selecting the most desirable courses of action thereby giving management information for making a more effective decision about the resources under control.

Relying basically on mathematics, it is a method of optimizing a linear function (**x1, x2, ..., xn**) when the variables **x1, x2,..., xn** are subject to a set of linear constraints. It is a mathematical technique that provides a maximum or a minimum solution to a linear equation when the variables in the equation are restricted within certain limits.

A mathematical programme exists when the objective and restrictions in a decision problem can be algebraically formulated as (Daellenbach, et al., 1983)

$$\text{Maximize(Minimize)} \ Z = c^{\tau} x$$

subject to restrictions

$$A\underline{x} = b; \qquad b > 0;$$

and

$$x \geq 0.$$

where

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \text{ is an } n \times 1 \text{ vector;}$$

$$c^{\tau} = (c_1, c_2, \ldots, c_n) \text{ is a } 1 \times n \text{ vector;}$$

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{pmatrix}$$ is an m x 1 vector

$A = (a_{ij})$ is an m x n matrix.

This is a general formulation with no assumption about the form of the function of the constraints. If they are linear then the problem is one of linear programming.

In a specific formulation, volume or value can be the item to be maximized. The constraints may include hectares available in each inventory category, the volume harvested in each period or the amount that the harvest can fluctuate between periods, the minimum inventory that must be left at the end of the planning horizon and the maximum or the minimum hectares or volume per period that can be harvested from particular age classes or groups of age classes and financial constraints like logging cost or transport cost.

Johnson and Scheurman (1977) described two mathematical structures that can represent forest harvest scheduling linear programs. In Model I formulation, the area of existing timber regeneration harvested each period in an inventory category forms a management unit the integrity of which is retained throughout the planning horizon. Each activity in the linear program represents a possible management regime for a particular management unit with its associated inputs and outputs, over the entire planning horizon. In Model II the existing timber in each inventory category forms a management unit until it is regeneration harvested. Thus for the Model I formulation, a management regime represents a sequence of intermediate and regeneration harvests and an associated cultural treatment regime that is throughout the planning horizon whereas for Model II such will only be throughout the life of the stand i.e the identity of the stand is lost once it is cut (Johnson and Tedder, 1983). In intensive plantations however, neither Model I nor Model II is entirely relevant and neither is adequately sensitive. Both these models assume an additional component of the utility of the forest, that is, the value of the inventory left at the end of the planning horizon aside from the discounted net revenue from timber harvests over the planning horizon. Such assumption is not relevant in intensive forest plantations.

Five elements can be defined to clearly specify a management problem, and which make it amenable to LP modelling:

1. an objective to be pursued;
2. restrictions on its pursuit;
3. alternatives which are open to management or levels at which resources are to be used;
4. the contribution of each alternative or level to the objective and the technological coefficients and;
5. the relationships between the alternatives and the restrictions.

These elements are sufficient to allow the best solution to be recognized. Harvest scheduling involves many data, interactions among which are too complex to be solved by inspection or simple computation.

The use of LP forces an increased understanding of the problem, improved profits and proof and reassurance that current practices are in fact correct (Wardle, 1965).

Aside from these, LP has a number of advantages because: 1) it has proved itself in many industrial corporations in the U.S.A. (Ware and Clutter,1971); 2) it has been the basis of many planning models like Timber RAM (Navon, 1971), MASH (Gibson et al., 1974) and FORPLAN (Kent et al., 1991); 3) it uses standard computer programs for optimization; and 4) it is able to use and incorporate economic factors like value and costs and discount rates.

Another advantage of LP is that it can serve as a means of learning more about the problem. This is accomplished by comparing optimal solutions of various LP formulations to examine the impacts of the changes in assumptions which are both required and questionable (Hoganson and Rose, 1984). Therefore, in addition to the optimal solution to the problem, the use of linear programming gives information about variation in the optimum and changes in the restrictions which provide critical guidance on the direction which management should take. The value of this sensitivity analysis is in providing a means of reducing the lack of certainty in LP models. Thus, while certainty may not be tenable in forest resource planning problems, this drawback is solved by parametrically changing some of the data values to evaluate coefficients of the objective function or RHS values.

The use of LP to develop a model to develop a framework within which a forest planning problem can be conceptualized and modelled is very much emphasized in many applications of this mathematical programming tool (Kent et al., 1991).

Evaluation of Existing LP harvest scheduling models

International. An early application of LP to the industrial plantation forest regulation problem was reported by Curtis (1962). It was apparently first used by Buckeye Cellulose Co. of Perry Florida to schedule optimally the harvest of annual cutting blocks within a fixed rotation, area control regulatory system. The objective was to maximize net present value (NPV) of future cash flows subject to restrictions by periods, regeneration areas and volumes harvested (Curtis, 1962). The shortcoming of this application was that it could not attempt to see what the optimum schedules would produce, once policy statements and assumptions are changed. In this and in other succeeding applications, the drawback was the difficulty of reprogramming the model as conditions on prices, costs and technology changed. Indeed it was noticeable that the early models reported optimal values as if to imply that the solutions are themselves the end of modelling. Techniques had hindered early modellers from appreciating the value of gaining insights resulting from the various experiments on them, usually performed once the model is running. This benefit is what this study wants to facilitate. In one succeeding application, Leak (1964) reported its application in industrial forest management to provide estimate of (1) maximum yields under specific conditions (b) areas to be cut or thinned by age classes, operating cycles and other categories so as to achieve maximum yield and (c) the effects of different restrictions or cutting policies upon estimated allowable cut.

Other early LP harvest scheduling model capabilities in U.S.A. were Timber RAM and MaxMillion. Kidd et al.,(1965) preceded them in applying LP to the regulation of timber harvests. In this last application, LP allowed forest managers to assess the impact of a change in managerial constraints prior to actually making the change. In scheduling reforestation investment, LP provided a superior solution than capital budgeting and certain rules of thumb (Teeguarden and von Sperber, 1968).

Forest harvest scheduling problem can become extremely large. From a practical point of view, this is the most troublesome characteristic of LP harvest scheduling models (Nautiyal and Pearse, 1967) and they are the more difficult to reprogram when new data arrive.

New Zealand. An optimising forest estate modelling system called Forestry Oriented Linear Programming Interpreter (FOLPI) (Garcia, 1984) was developed in New Zealand which finds the management strategy that optimises a user-defined objective function subject to structural and user defined constraints. The system has been developed to be complementary to a simulation model called Interactive Forest Simulator (IFS) (Garcia, 1981a) and thus can use the same input data. The IFS/FOLPI system is used as one component of the Conversion Planning Project Team Model System where it is linked to a stand prediction model that provides the data on yield needed by the models. A distinct shortcoming of FOLPI, like many other previous LP models reviewed, was its inability to have quick and less cumbersome re-runs when model assumptions change. This was experienced by Manley and Threadgill (1991) in the use of the model in developing forest valuation methodology for the sale of plantation forests in 14 corporation districts in New Zealand.

Broad (1985) used a mixed integer linear programming technique to model resource flows in a system comprising industrial plantation forests and subsequent wood processing and marketing activities.

A regional harvest planning and resource allocation model (REGRAM) is being developed to determine the thinning and clearfelling programme for a number of forests in a region (McGuigan, 1992). Using a combination of simulation and linear programming, the model consisted of: a) a database to enable the user to define crops, locations, resources and processes, b) a simulator for individual forest, c) an optimiser to determine optimal harvest and resource allocation strategy and d) a reporter to generate reports for a forest. The model has many significant features, for example, close integration of simulation and optimisation, flexibility of modelling anything from individual stands to entire regions that may consist of several forests and processing locations and the special way of treating time by having single year periods at the start and multi-year time periods at the end of the model. Two major models, one for Nelson/Marlborough and the other for Central North Island have been built with REGRAM.

Tropical Plantations. There have been very few applications of LP in scheduling timber harvest in tropical plantations. One reported by De Kluyver and Whyte (1980), demonstrated the applicability of LP in formulating and solving a large-scale forest harvest scheduling in Fiji. Compared with two other models, one heuristic and the other a compact decomposition LP, a large LP model was found "most useful in identifying common features among good solutions" to a harvest scheduling problem in a pine plantation in Fiji. They then suggested that solutions from LP can be altered and fine tuned through the other methods because LP harvest scheduling problems tend to be large and computer dependent. Further, the study emphasized the value of extensive sensitivity analysis and the use of the model as a framework for decision making, a feature also given importance in LP models developed to improved long-term management plans for a forest plantation in Tanzania (Kowero and Dykstra, 1988). Such emphasis on the role of LP model was lacking, for example, in the log allocation and transport model of Araño and Bonita (1977).

2.3 The New Class of LP Models - The Spreadsheet LP Model

Various phases of LP may include (Turner et al., 1977):

1. a stand generation phase which generates simulated alternative management strategies for each forest stand or other crop aggregation;
2. a matrix generation phase in which the output from the first phase together with

information on constraints, such as supply levels to be attained, areas of each stand and budgets are put in a form suitable for input to a standard LP computer package;

3. an optimization phase in which the strategies which satisfy all constraints and optimize the objective are selected; and
4. a report writing phase in which the optimal solution is tabulated and reported in a form suitable for managers to assimilate.

An optimiser, matrix descriptor and generator, report writer and data manager are the major parts of LP packages (Welch, 1987). There has been a lot of work on the development of the LP algorithm. The simplex method of solving general LP problems has been translated into many computer languages and implemented in many codes.

Matrix descriptor and generator programs to translate the model from the modeller's algebraic form into an algorithmic form MPS format are required. Jones and Carmona (1987) studied some of the drawbacks of matrix generators which included non-generality of the conversion tool and the difficulty of verifying, documenting and modifying the models. Gordon (1987) presented the disadvantages of the use of matrix generators, viz.,

a. the development of matrix generators is a time consuming and error prone process., even with the help of special purposes languages;

b. the matrix generated is difficult to validate because the output is extensive and intended for machine processing, not for human comprehension;

c. the relationship between the model and the matrix generator is often unclear and abstract and thus it is difficult to determine whether the matrix generated conforms to the modeller's intentions;

d. the model must be documented as must also the matrix generator and the relationship between model and generator;

e. whenever a change in the data is made, the matrix generator needs also to be modified (while changing the model may require hours, the modification and revalidation of the matrix generator may require days, a particularly unsatisfactory outcome when the model is undergoing constant revision, as in planning applications); and

f. the casual notation for the model and the hard translation from model to matrix generator makes it impossible to provide help for many critical steps.

Aside from these, matrix generators are mostly limited to mainframe computers and domain specific. While they are standard in the main frame, they cannot be assumed in the microcomputer because of the limited random access memory (Sharda, 1986).

Computer packages use three different ways of representing a problem: a) natural; b) compact; and c) spreadsheet (Wasil et al.,1989). The natural way represents the model so that it closely resembles the traditional "paper and pencil" formulation which lists the objective function and constraints. The compact form represents the model in a way that data are stored as a matrix of coefficients and parameters. The spreadsheet form presents the model so that the data and coefficients are placed in cells from which tables of relationships are created. Cells are referenced to create constraints and objective function.

The new breed of models make use of computer packages which can solve LP problems in linear algebraic form, a form with which users are very familiar. In the same manner they can easily be interfaced with each other. Popular spreadsheet models can help in better preparation and delivery of LP harvest scheduling models and can provide a good mechanism for problem specification and presentation to improve the overall quality of the modelling project.

Yield data are a major ingredient in directing outputs from timber harvest scheduling models. A programming modelling approach should use dynamic growth functions as input instead of using yields to represent how inventory changes over time in response to various silvicultural treatments can be accommodated (Alston and Iverson, 1987).

The early harvest models had difficulties with this approach. For example, as changes in the data occurred, appropriate changes had to be made in the cutting schedule which required solving the problem again using new data and developing an entirely new schedule. Kidd et al.(1966) noted that it is not wise to adhere to a schedule for 50 years that maximizes a property's net worth, as that schedule would be a correct interpretation only in the unlikely case that all assumptions remained valid for the entire 50 years. The problem should be formulated and solved with the best and most up to date available information. If the best present information is used, the resulting solution should be the best obtainable at the time. As better information becomes available the problem can be reformulated and a new solution obtained. The problem can also be reformulated with different or varied restrictions. This subscribes to the philosophy that planning should be a perpetual process with continually response to new economic and biological information (Whyte, 1990, pers. com.)

Ware and Clutter as early as 1971 recommended that timber harvest scheduling and rescheduling could be made inexpensive by having data in an input file and then entering them into the mathematical programming system. This proposition is predicated on the fact that successful harvest scheduling normally requires repeated solutions at short time intervals because the basic input parameters, like prices and costs, are subject to frequent changes (Ware and Clutter, 1971).

Modelling can now move more easily towards such model interfacing. For example, timber and transportation models can be interfaced, the solution to one model serving as an input for the other and an iterative procedure followed (Weintraub and Navon, 1986). This can be costly and may lead to suboptimal solution, however, if both transportation and timber resources management activities, for example, are represented explicitly in the same model. Weintraub and Navon (1986) showed the use of LP for managing timber integrated with a mixed integer program for planning the development and use of a transportation network.

There is another similar trend towards interactive modelling, where one can change the model by adding new variables and constraints as the situation dictates, without compromising the solution method. Another example of progress is the need to consider the use of a Geographic Information System (GIS) in tactical planning because of the close association between decision making models and GIS.

Decision Support Systems (DSS) have evolved from all these developments. DSS are flexible integrated software for accessing, retrieving and generating reports on data base information plus simulation and decision models for conducting further analyses and automated goal seeking. They are further characterized by (1) output displays in tabular, graphic and map forms; (2) having an interactive mode of operation, ultimately dependent on human judgment and expertise for final decisions; and (3) providing rapid feedback on the consequences of management alternatives offered to the decision maker (Covington et al., 1988).

Planning models should be built within computer packages with which users are familiar. For this reason, spreadsheets, now the most popular computer package, are seen as a very appropriate environment in which to build planning models. One advantage of having a spreadsheet based planning model in microcomputers is the greater possibility of actual application and use of the models. The spreadsheets are widely recognized as a flexible, robust tool for managerial decision making. The use and acceptance of the spreadsheet is legendary. It is interactive and screen oriented and converts the memory of the computer into a large matrix . Numbers and formulae can be stored in the cells of this matrix. Once relationships are established between variables and cells, what if and what's the best strategy analysis can be conducted. So this is close to optimization already. Also data entry and editing features are convenient.

While other analytical and programming solutions may be elegant, they may not be able to offer the realism that table driven spreadsheets offer to a problem as sophisticated as harvest scheduling. Table driven formulation of an LP harvest scheduling problem allows modelling of a complex real world problem (Winter, 1989).

Prior to bringing the power of mathematical programming to the spreadsheet environment, there had been various applications which could have contributed to the current integrative capability of MP and spreadsheet. Spreadsheet-like DSS can be a significant aid in production planning by providing better decisions in less time and effort. In a laboratory experiment to determine the effectiveness of a production planning DSS built from spreadsheets, Sumichrast (1990) found how possible solutions to a problem can be studied in much less time. Parekh (1990) illustrated the use of a spreadsheet for capacity/inventory planning which was very simple, very basic and a useful simulation tool that can be updated and expanded. Cornwell and Modianos (1990) described some aspects of using spreadsheets for simulation modelling in two application problems, one a fixed-time simulation to compare a rental plan against a purchase plan of a new forklift, and the second a variable time model to compare two replacement policies of drill bits for a drill press. Ogweno (1988) implemented an optimal equipment replacement model in a microcomputer spreadsheet that could serve as a financial planning tool in timber harvesting projects. Fisher (1986) showed how the spreadsheet can provide the capability for creating and analyzing deterministic simulation models.

There have been, then, several applications of the spreadsheet to several forms of quantitative analysis for decision support. Since linear programming is a basic quantitative tool widely used in OR approaches,it is logical to have it implemented in a spreadsheet (Ho, 1987)

The spreadsheet has been widely recognized as a flexible, robust tool for management decision making. In order to fully exploit its popularity it seems reasonable to make efforts bringing the power of basic mathematical programming (MP) tools into the spreadsheet environment. Bringing the power of MP to the spreadsheet has been accomplished in three ways: (1) augmenting existing spreadsheet capabilities with mathematical programming based optimization software; (2) directly modifying the spreadsheet itself; (3) designing solution methodology that can be implemented using existing spreadsheet capabilities, i.e. to use spreadsheet macros to implement MP techniques. Its possible drawback, however, is that more computation time is needed than when programmed in more traditional computer languages outside the spreadsheet environment.

There are three classes of optimization that can be used with spreadsheets (Sharda, 1988).

(1) Programs can simply accept a problem formulation from a spreadsheet file. This allows the LP user only to take advantage of the spreadsheet features relevant to problem input e.g. MICROLP, MPS-PC, RAMLP.

(2) Programs can read LP problems from a spreadsheet file and also store the optimal solution in such a file. JANUS, for example is a utility program with LPS-867 which converts a spreadsheet file into a format accepted by LPS-867 and then transforms the problem solution from LPS-867 into a spreadsheet file.

(3) Programs can reside in the memory within the spreadsheet program. Here, the user creates the spreadsheet, activates the optimization algorithm, returns to the spreadsheet and makes it appear that the spreadsheet has optimization capabilities. This is useful if accomplished in real time. One receives data, converts data in LP form, solves and prepares results all in real time, thus taking advantage of timely information.

There are other advantages of optimization in the spreadsheet (Sharda, 1986). Firstly, optimization in a spreadsheet is seen as interactive, screen oriented software which converts the memory of a computer to a large matrix containing rows and columns. The computer can store the numbers as well as the formulae in the cells in the matrix, and if a cell is changed, all other cells affected by the change are automatically recalculated. Secondly, models developed using spreadsheets are close to optimization models anyway. One just needs to specify which cells are decision variables, which cells/rows are constraints and which relationship denotes the objective function. An optimization algorithm can then perform the necessary computations (Fisher, 1986). Thirdly, there is ease of problem specification. Fourthly, data and information management in these models can be used to explore relationships using the graphic utilities, automatic and manual recalculations, and inbuilt functions that can be used to move, copy and insert rows and columns to improve the spreadsheet layout (Jones and Carmona, 1987). Spreadsheets now have features not envisaged in the early days - graphics, word processing, data base management and macro command language. Model documentation can use mnemonic labels, short comments,and more extensive textual explanation can be incorporated into the same support (the electronic spreadsheet) as the model and in the same way as the model is written on to it i.e. through use of the same simple input and editing facilities. The methodology does not separate model generation and report writing.

In a spreadsheet-based harvest scheduling model, Leefers (1991) combined optimization with simulation to illustrate the use of readily available spreadsheet packages to develop LP-based models that include timber yield variability using Monte Carlo simulation. In the process, the study confirmed some strengths of harvest schedule modelling in the spreadsheet environment e.g., easy to use and facilitate communication. Its weakness is that they are rigid with detailed equations and formats that may be difficult to adjust, for example, if management strategies have to be changed. However, as the study pointed out, forest management strategies tend to be well defined, making them easy to capture in a structured template. This feature then becomes a strength, because model structure remains consistent.

From the foregoing review, it was seen that one harvest scheduling model may be preferred over another, the choice depending on the characteristics of the harvest scheduling problem, available resources and objectives of the analysis (Johnson and Tedder, 1983). The properties of harvest scheduling problem in tropical plantations being driven by prices, costs and yields and which consists of many decision variables and constraints related to future forest structures make LP a more appropriate tool than the other tools reviewed here for harvest scheduling.

Simulation models cannot model in a single run a harvest schedule that simultaneously generates a non-declining primary harvest and guarantees that the secondary harvest will be at least a certain percent of the other harvest. The secondary harvest cannot also be directly constrained to a desired level. LP on the other hand can be formulated to direct whatever set of activities and outputs the user desires through the inclusion of constraints (Jamnick, 1990).

It has been emphasized in this review that LP modelling does not end with obtaining optimal solutions after having the model run. The greatest value of modelling is its potential for sensitivity analysis and as a decision framework whereby the effects of changes in assumptions and data can be examined. The results of these kinds of analysis may be the most valuable information resulting from LP modelling. It provides answers to the questions on the values of change in the constraints or introducing new activities. It is this information which provides critical guidance on the direction managers should take, particularly those that do not involve clear cut choices among simple alternatives but rather the reconciliation of alternatives which conflict one with another and are variously affected by restrictions on management. The course of action in these circumstances is not immediately apparent. It is with this background that the spreadsheet shows great potential as an appropriate environment for LP-based harvest scheduling models.

CHAPTER 3

METHODS

This chapter sets out in detail the materials used and the methods employed for the two major modular components of the study: i) modelling growth and yield and ii) harvest scheduling. The discussion in this chapter focuses on various aspects of modelling, primarily on data validation, description of the models used and the derivation, evaluation and selection of the stand level and diameter distribution growth and yield equations. It also discusses the nature of the data that were required in the development of the harvest scheduling model including the linear programming (LP) mathematical formulation. The chapter ends with a case study to explain the general nature of the steps involved in the analysis and construction of the harvest scheduling model.

3.1 GROWTH AND YIELD MODELLING

The algebraic differential equation (ADE) used to describe changes over time in stand and diameter statistics necessary for this study has the general form of a state space function, (see Clutter et al., 1983). That is,

$$Y_2 = f\ (\ Y_1,\ T_1,\ T_2,\ \theta,\ MR\ ) \qquad (3.1)$$

where,

Y_2 = value of a continuous variable defined for a tree or stand at age T_2;

Y_1 = value of the same variable at age T_1;

T_1 = tree or stand age at initial measurement;

T_2 = tree or stand age at next remeasurement;

θ = set of parameters of the equation and;

MR = management regime.

In this state space approach, at a given time, the future state of the variable and the transition functions or changes in the state variable are a function of the initial state of the variable, time elapsed, management inputs and prevailing environment. In using projection equations of this functional form, it was appropriate that real growth series data available from a system of permanent sample plots (PSP) for tropical plantation studied be used to obtain sample estimates of the parameters of equations that best described the growth and yield of the selected stand variables. These estimates were then used to derive diameter distributions.

3.1.1 Data set for modelling growth and yield

Data set. The data set used in this study comes from permanent sample plot records applicable to measurements from years 1968 to 1985 in the forests of Lololo, Drasa and Seaqaqa in Fiji. These plantations consist of both thinned and unthinned stands.

There were three sources of data: (i) Manley (1977) with 18 plots established in Lololo and Drasa and measured between

1964-1967; (ii) Wybourne (1982) with 54 plots from a thinning and spacing trial established in 1971 in Seaqaqa and; (iii) FPC (1978) matched inventory plot records collected from 1968 to 1977 in the above-named plantation forests. In total, the 320 plots available were able to yield useful data on stand diameter, stocking and basal area statistics. All plots were of sufficient size to hold at least 20 trees, with as many as 200 in a few. Projection data in periods that included the occurrence of a cyclone were excluded if the mortality was more than 200 trees per hectare. For the plots in thinned stands, data in intervals that included a thinning were excluded. In using the data from these sources the effect of thinning was modelled only from measurements in intervals that did not include the year of thinning. Data from Wybourne (1982) did not have maximum diameter at breast height outside bark (**Dmax**) so a modelling estimation procedure was used to derive Dmax for these plots.

Original measurements on each plot contain data on diameter at breast height ($\mathbf{Dbh}_{ob}$) measured for each tree. Such raw data were not available to this study. Instead, derived diameter statistics like maximum, minimum and variance of $\mathbf{Dbh}_{ob}$'s for single plots were used to estimate the parameters of a probability density function for modelling diameter distributions. Not many areas of tropical plantations would have the same relevant plot measurements and would likely provide only limited data that have limited potential for analysis and applications, but such information is vital and every encouragement should be given to its acquisition.

The three forest sources contained data, however, that were eminently suitable, after transformation, to form a data base for the purpose of constructing a growth and yield model as envisioned for this study. Ideally, however, such a data base should be created from the original tree measurements and not from derived diameter statistics solved. Such raw data were not universally available for this study, but were of sufficient coverage to validate the reliability of the diameter distribution estimates.

An ideal data base should also have height measurements for development and validation of height equations. In addition, an acceptable number of sectional measurement of trees should be taken for constructing and validating compatible tree taper and volume equations. These aspects, however, have been thoroughly investigated elsewhere by, for example Geiser (1977) and Broad (1979) and did not warrant repeating here.

From the three sources, the following data were collected for individual plot:

1) forest locality e.g. Lololo or Drasa or Seaqaqa;
2) year planted;
3) plot number;
4) age at measurement or remeasurement;
5) mean diameter at breast height outside bark of living trees inside the plot;
6) maximum diameter at breast height outside bark of living trees inside the plot;

7) variance of diameters at breast height outside bark of living trees inside the plot;

8) minimum diameter at breast height outside bark of living trees inside the plot;

9) net basal area per hectare of trees inside the plot;

10) living stems per hectare; and

11) management regime conducted e.g. thinning intensity.

The data collected formed the set described in Table 3.1. Slightly different data structures were formed when each stand or diameter variable was modelled because of the validating procedures subsequently undertaken. Nevertheless the data set characterized by Table 3.1 has always been the starting data set for all modelling work reported here. In essence, there were as many data set structures created as the number of variables modelled. Because of the nature of the differences in the data set structures for the different models, any later attempt at treating the models (even net basal area per hectare and mortality relationships) as systems of equations for simultaneously estimating their parameters, could not be accomplished even with PROC SYSNLIN available in SAS. The simultaneous solution of the two equations would have been a useful procedure, especially, if one dependent variable predicted by one model i.e. stocking, were to be used as an independent or explanatory variable in the other models; for example, in the net basal area per hectare equation or in any of the diameter variables. For example, N_2, was not used as an independent variable because in using the projection equation, it has to be specified, which is not possible if it still has to be predicted. Thus in this study, a starting stocking, which is a constant was validly substituted as an independent variable instead of the predicted stocking.

The original data were transformed into yield projection data format. A SAS program shown in Figure 3.1 was used to accomplish the creation of projection data from the yield data. All possible growth intervals (AI) were created with this program. From this structure, two other data structures were made, one with a no overlapping interval (NI) and one with only the longest interval of measurement (LI).

LI and NI data structures were also used to derive the models but the models derived when using them were poorer than the ones which used an all interval data structure. Therefore, only the all-interval data structure is reported here in detail.

Table 3.1 Summary statistics for the base data set

VARIABLE	N	MINIMUM	MAXIMUM	MEAN	C.V.
T	426	3	17	10	2.0508
Dmean	426	5	34	21	0.9725
Dmax	426	9	55	33	0.6409
Dmin	426	1	28	12	1.7873
Dvar	426	3	75	19	1.1539
G/ha	375	2.5	46.8	22.35	0.3777
N/ha	426	222	2152	748	0.4286

where,

C.V. = coefficient of variation;

T = age of the stand at time of measurement or remeasurement, years;

Dmean= mean plot diameter at breast height outside bark, cm;

Dmax = maximum plot diameter at breast height outside bark, cm;

Dmin = minimum plot diameter at breast height outside bark, cm;

Dstd = standard deviation of plot diameter at breast height outside bark, cm;

G/ha = net basal area per hectare, m^2/ha;

N/ha = stocking, stems per hectare.

```
DATA  YIELD;
INPUT  AGE YIELD;
GE2=LAG(AGE); YIELD2=LAG(YIELD);
CARDS;

    4.0   15.0;
    5.0   26.5;
    .     .
    .     .
    30.0  89.0;
    DATA HOLD; SET YIELD;

AAGE=AGE2-AGE;
IF AAGE GT 0 THEN DELETE;
ELSE DO;
  PUT AGE2 1-4. 1 AGE 6-9. 1 YIELD2 11-14. 1 YIELD 16-19. 1;
END;
PROC PRINT DATA=HOLD;
(THEN EDIT  .LOG AND RENAME)
```

Figure 3.1 SAS program to produce interval projection data format from yield data format

Table 3.2 Variables in the projection data set

VARIABLE	DESCRIPTION
FOR	Name of forest containing the plot, either Lololo, Seaqaqa or Drasa.
YEAR	Year the stand was planted
PNO	Plot number
T_1	Age at time of measurement
T_2	Age at time of remeasurement
$Dmean_1$	Arithmetic mean plot diameter at breast height outside bark of all trees in the plot at the time of measurement, cm
$Dmean_2$	Arithmetic mean plot diameter at breast height outside bark of all trees in the plot at the time of remeasurement, cm
$Dstd_1$	Standard deviation of plot diameter at breast height outside bark at time of measurement, cm
$Dstd_2$	Standard deviation of plot diameter at breast height outside bark at time of remeasurement, cm
$Dmin_1$	Minimum diameter at breast height outside bark at time of measurement, cm
$Dmin_2$	Minimum diameter at breast height outside bark at time of remeasurement, cm
$Dmax_1$	Maximum diameter at breast height outside bark at time of measurement, cm
$Dmax_2$	Maximum diameter at breast height outside bark at time of remeasurement, cm
N_1	Number of stems per hectare at time of measurement, N/ha
N_2	Number of stems per hectare at time of remeasurement, N/ha
G_1	Net basal area per hectare at time of measurement, m^2/ha
G_2	Net basal area per hectare at time of remeasurement, m^2/ha

The resulting rows of stand and diameter projection data are listed in the file Table 3.2. Some data lacked some variables (column) because they were not available from the original data sources. With some missing data it was found appropriate to describe the data values in column format rather than in list format in the INPUT statement of the subsequent SAS program that used the data set. Initial runs using the latter format caused problems in reading data sets with missing data.

The plot data used as the example in Table 3.3 are typical of the other permanent sample plots used in this study, having been measured more than twice. As mentioned earlier, a

Table 3.3 Sample plot measurement example and transformation to yield projection data format.

T	G/ha	N/ha	Thinning
5	18.5	1087	
6	23.6	1087	
7	29.3	1087	
7	22.6	815	yes
8	25.8	815	
8	23.1	667	yes
9	25.5	667	
10	28.1	667	
11	29.9	667	
12	32.4	667	
13	33.8	667	

Corresponding projection data format

T_1	T_2	G_1	G_2	N_1	N_2
5	6	18.5	23.6	1087	1087
5	7	18.5	29.3	1087	1087
6	7	23.6	29.3	1087	1087
7	8	22.6	25.8	815	815
8	9	23.1	25.5	667	667
8	10	23.1	28.1	667	667
8	11	23.1	29.9	667	667
8	12	23.1	32.4	667	667
8	13	23.1	33.8	667	667

projection data structure consisting of all possible growth intervals was used. This resulted in very substantial autocorrelation among observations. If a permanent sample plot has been measured **n** times the possible number of T_1 and T_2 combinations is however less than $\binom{n}{2}$ as can be deduced from perusal of Table 3.3, because of the exclusion of combinations of measurements that included thinning within the projection interval.

Initial Validation of Data Set. Prior to any model estimation the data were verified and screened to ensure mensurationally sound data. Examples of data validation include ensuring that N_2's are not greater than N_1's, G_2's are greater than G_1's and the T_2's are greater than T_1's. Observations were also deleted if the decrease in stocking was more than 200 trees per hectare from successive measurements. Residuals were also used to detect outliers. Outliers were observations that had residuals greater than 3.5 standard deviations from zero.

Because of the different scales of measurements for the various stand and diameter variables to be modelled, and the need to have a uniform value at which residuals were to be declared, there was a need to standardize screening of residuals. Standardizing residuals is an attempt to utilize the concept of standard normal deviates (Anscombe and Tukey, 1963). If ϵ_i is a normal random variable with mean zero and variance σ^2, then $\epsilon_{i/\sigma 2}$ is a standard normal random variable. Hence a standardized residual S_i is defined as

$$S_i = \frac{\text{Residual}}{\text{MSE}} \qquad (3.2)$$

where,

Residual = (Observed value) - (Fitted Value) and

MSE = Residual mean square.

As a rule, in all the model fitting routines, observations which have value of S_i greater than 3.5 were considered to be outliers. Such outliers indicated the need to examine all the corresponding items of data. Observations corresponding to S_i greater than 3.5 were not deleted automatically. If after review, they were indicative of obviously erroneous data that could not be corrected objectively, only then were they deleted.

For growth and yield data there are various sources of errors which may produce outliers: e.g. incorrect reading of measuring instruments, wrong recording and wrong calculation of derived values are probably the biggest source of error in the data set used here. Other less likely ones may result from measuring the wrong part of a tree or measuring a wrong tree. Measurements may be properly conducted but errors may still occur if conditions for measurement are not properly met. In this study, there was not always the possibility of direct checking, but the data had all been screened routinely in Forest Inventory System (**FIS**), a system used by the Fiji Pine Commission analysts (Patel, 1985). Errors did still appear in the data, however, and much effort was made to put these right.

Data validation is a vital pre-cursor to fitting models. Failure to do so will result either in development of an inappropriate model or derivation of an unnecessarily imprecise model. As yield models are almost always the driving component for most other planning models, their accuracy and precision are of the utmost importance. The central role that growth and yield models play in many other planning models is thus properly recognized here.

Validation procedures also included manually checking the data pertaining to different variables one by one. PROC UNIVARIATE, a SAS procedure to summarize data, was also used to analyze general trends and extreme values. Graphs of the data, too, could be utilized to verify that outliers caused by inaccurate data recording or inaccurate measurements were recognized.

Two basic data base files were created and are appended as Appendix D. BAREA.DAT contains the data on basal area and DIAMETER.DAT contains the data on diameter variables.

Sample graphs of diameter variables through time in Figure 3.2 to Figure 3.4 indicate the general growth trends in the variables. Such graphs were used to confirm the outliers declared by the use of the standardized residual criteria. The summarized results of the data validation procedures that have been conducted are shown in Table 3.4. Observations were paired in the projection data sets.

Table 3.4 Summary of the results of the data validation conducted for the projection data sets.

Variable	Total No. of Pairs	Outliers Identified	Outliers Corrected	Outliers Rejected
Basal Area	1053	8	-	8
$Dmax_{ob}$	1146	6	-	6
$Dstd_{ob}$	1082	12	-	12
Survival	90	2	-	2

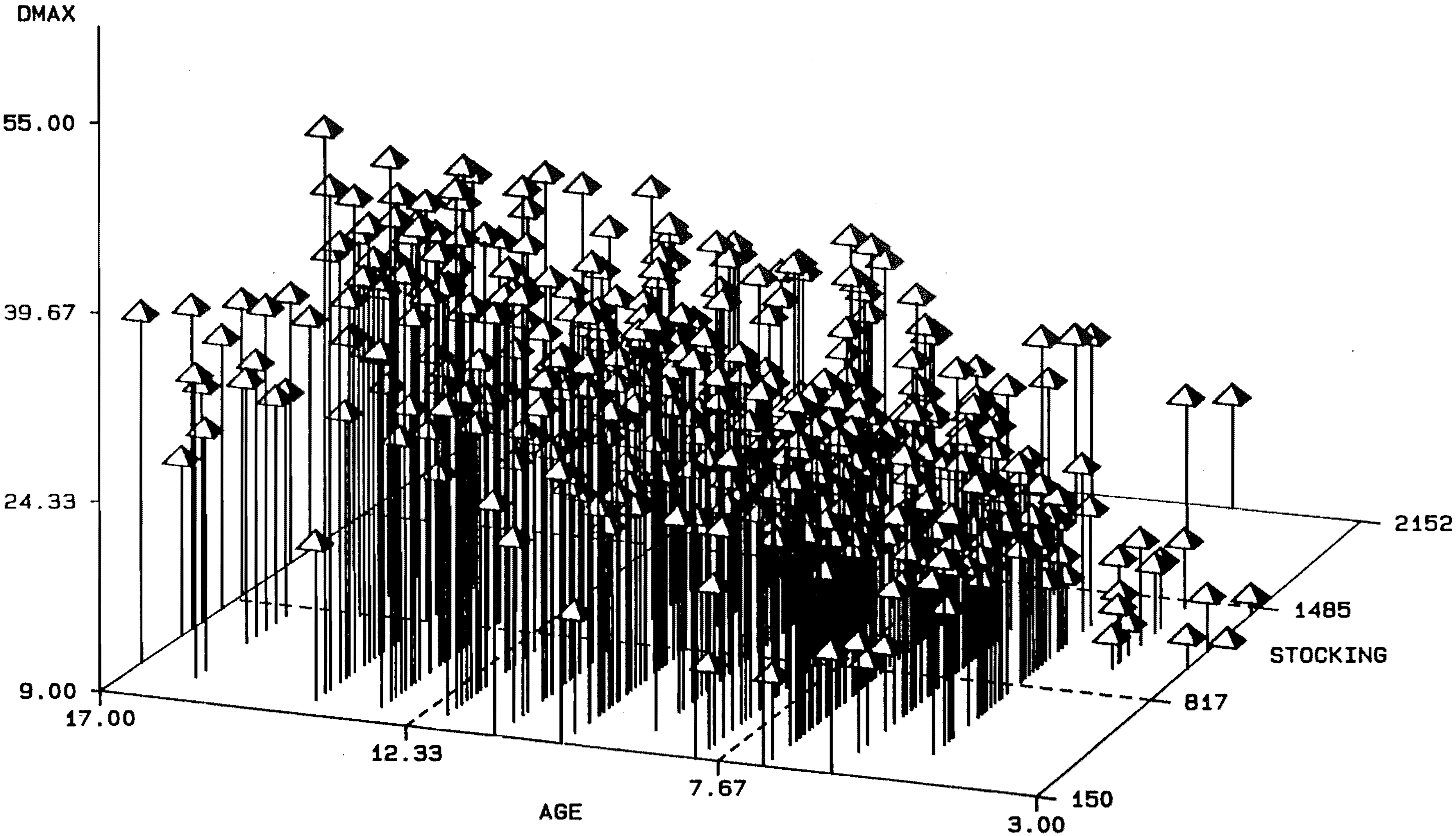


Figure 3.2 Graph of maximum diameter against stand age at varying initial stand stocking

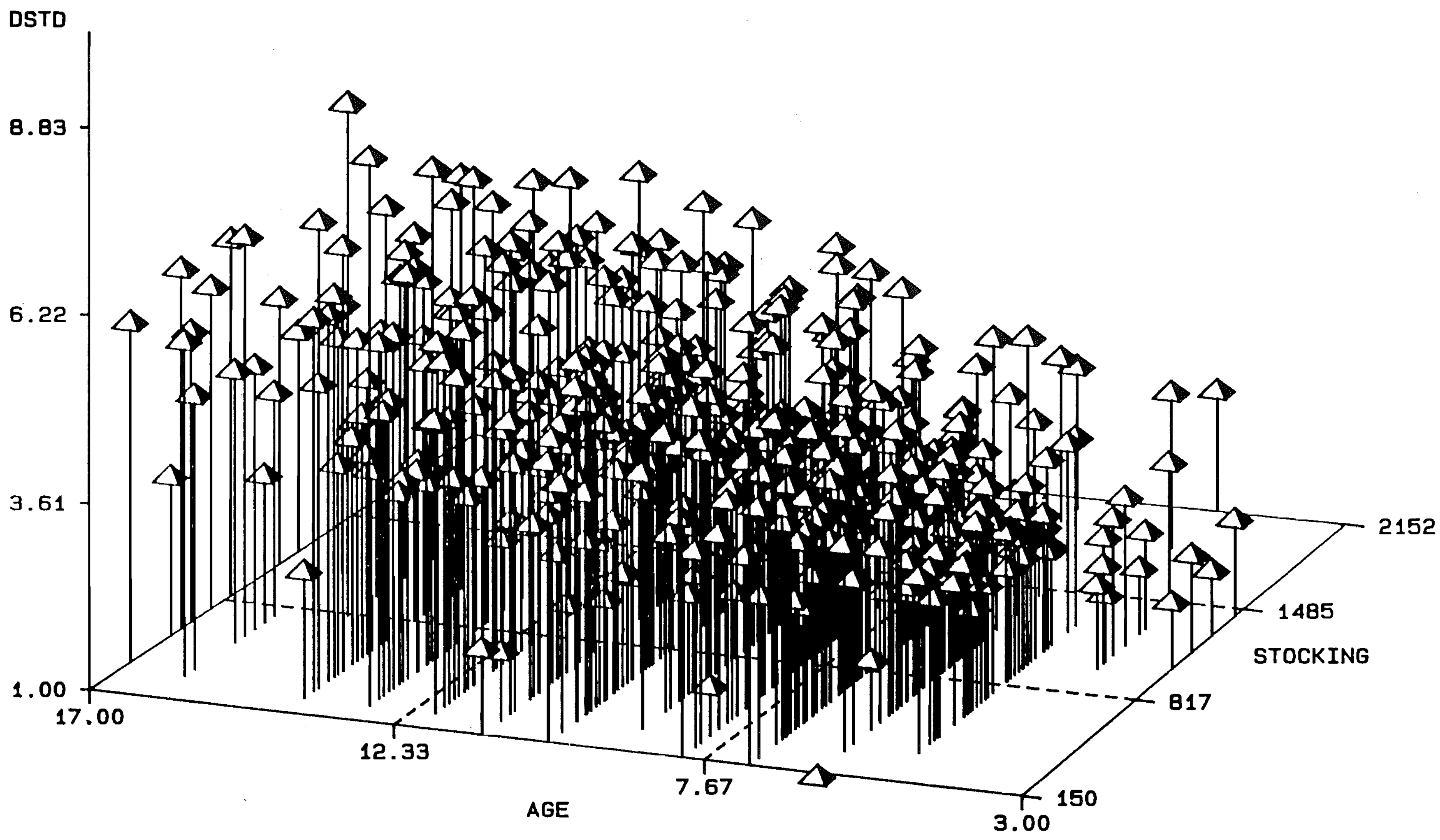


Figure 3.3 Graph of diameter standard deviation against stand age at varying initial stand stocking

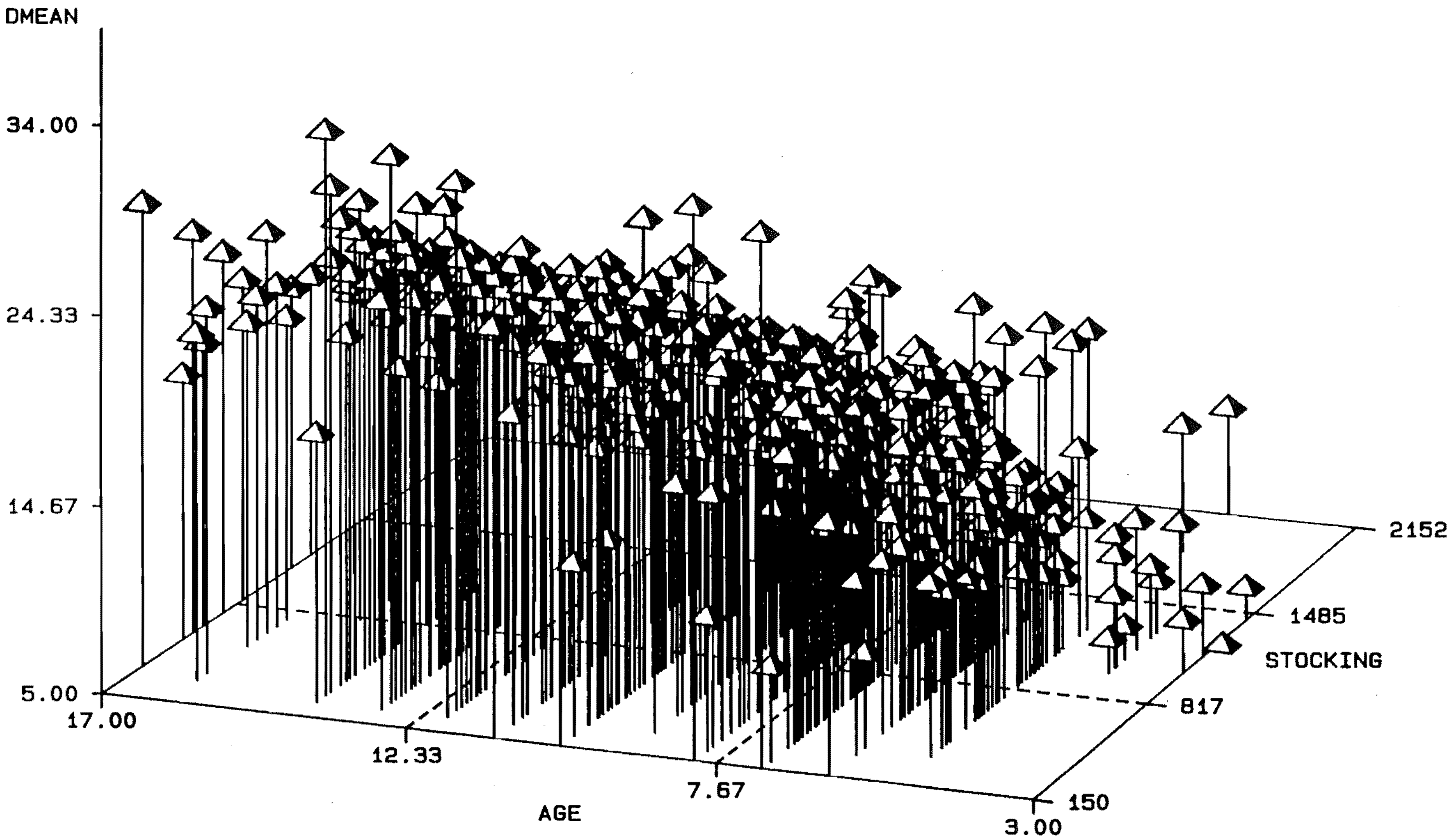


Figure 3.4 Graph of mean diameter against stand age at varying initial stand stocking

3.1.2 Growth and yield modelling procedures

Stand level growth and yield modelling. This phase involved the development of stand average models to estimate projected stand attributes based on the validated data sets. The number of stand and diameter attributes that needed to be estimated was equivalent to at least the number of parameters of the probability density function selected to model diameter distribution.

While a three parameter reverse Weibull probability density function was proposed to model the distribution of the diameter, only two variables i.e. the shape and the scale parameters, needed to be solved by the parameter recovery method, because the location parameter could be derived as a function of the maximum diameter. Since only two parameters were to be estimated through the method of moments approach, the first and second moments of the distribution are therefore related to two stand attributes, namely the mean stand diameter and stand basal area. These variables formed the basis for solving the parameters. Thus, maximum diameter, mean stand diameter and net basal area per hectare are the main variables that need to be modelled to solve the parameters of the distribution, together with a survival or mortality equation in order to project stand tables. Mean stand diameter proved to be a very difficult variable to model successfully. Previous researchers have had similar experiences (Kuru, 1989; Xu, 1990). All the general forms of the yield projection equations listed in Table 3.5 were tried and the failure to fit any acceptable functions to model this diameter variable resulted in a different approach to solve the problem. Instead of modelling mean diameter directly, it was estimated from the following relationship (Clutter <u>et</u> <u>al</u>., 1983, p. 72).

$$Dmean = \sqrt{(Dq^2 - (\frac{n-1}{n}) Dstd^2)} \qquad (3.3$$

where,

Dmean = stand arithmetic mean diameter at breast height outside bark in cm,

Dq = stand quadratic mean diameter in cm,

= $[10000*(\mathbf{g}/0.7854)]^{1/2}$

where **g** is mean tree basal area in m^2,

Dstd = diameter standard deviation in cm.

The use of Equation 3.3 required additional projection equations for diameter standard deviation and mean tree basal area per hectare in order to estimate the mean diameter. The mean tree basal area was estimated from estimates of net basal area per hectare (**G/ha**) and stocking per hectare (**N/ha**). **G/ha** and **N/ha** are variables that measure stand density. They are variables which are basic to all growth and yield prediction. Attempts to model mean tree basal area directly were not successful, and so the implicit derivation, G/N was employed. Again, all forms of possible equation were tried but none was found to be acceptable based on criteria that had been set for this study. Thus, the stand attributes that needed to be modelled were:

a) stand net basal area per hectare;

b) stand diameter at breast height outside bark variables;

   1) diameter standard deviation;

   2) maximum diameter and;

c) tree survival/ha.

Mathematically compatible growth and yield projection equations for the above variables were derived for the corresponding data set using PROC NLIN, the non-linear least squares procedure in the Statistical Analysis System (SAS). When the derivatives of the functions could be easily derived, the estimation of the parameter was done by the Gauss-Newton or Marquardt option. Otherwise DUD (Does not Use Derivatives) was used with sacrifice on speed of solution.

Non-linear regression estimation techniques are relatively new tools for growth modelling. The approach here is set out because modelling techniques with this tool are still rapidly evolving. Previous modellers have recognized the non-linear forms of functions that could represent growth of trees and stands satisfactorily, but were limited to the then available computational algorithms (often having to resort to transformation to linearize nonlinear functions). This approach produced biased models.

Non-linear solution routines without resorting to transformation are therefore powerful tools for growth modellers. Given a non-linear equation, the sample data were fitted to the chosen form by estimating the values of the parameters that minimize the sum of squared residuals. Non-linear estimation is an iterative process; success in obtaining convergence in the solution can be ensured by providing good initial guesses of the parameters. The efficiency of convergence depends therefore upon the adequacy of the initial estimates, usually available from previous experience (Woollons, 1989). Graphs, the study of which can help interpretation of the function in terms of the parameters, may well assist, therefore, in the choice of an initial set of parameters. When values of the parameters are unknown, the SAS statement options called PARMS and BOUNDS were invoked. These options limit the range and the steps within which the program would iterate to solve the values of the parameters that minimise the sums of squared residuals.

The non-linear algorithm is quite straightforward. Given initial parameter estimates, the sum of squared residuals (RSS_1) is solved. The values of the parameters are then changed according to the bounds and steps specified, and then a new sum of squared residuals (RSS_2) is derived. The new sum of squared residuals is compared to the old sum of squared residuals. The procedure is continued iteratively until no further reduction in the sum of squared residuals can be found. In SAS, this point called convergence occurs when the change in the sum of squares on successive iterations is smaller than some previously specified value. This procedure could be shortened up by using efficient algorithms like Gauss-Newton, Marquardt or steepest descent, all of which are fully described in SAS Manuals (SAS Institute, 1985, pp.1135-1193) and in Bates and Watts (1988, pp.78-83). These procedures usually require the partial derivatives of the model with respect to the parameters to have already been solved and provided in the program. Otherwise a derivative-free method, DUD, which does not use numerical approximation to derivatives, can be employed.

Various other procedures were used along the way before and after PROC NLIN, including PROC SORT, for sorting the data; PROC PLOT, for plotting various graphs used in data validation and in deriving regression coefficients and; PROC UNIVARIATE, for analyzing the normality of the data and the residuals.

Several forms of different model functions were tested that describe each of the different attributes of the stand which needed to be described. The fits of the models were mainly assessed through study of the values and characteristics of the residual sums of squares (RSS) and residual mean squares (RMS), preference being for the smaller and more normally and randomly distributed ones. Because of the nature of the data set for growth modelling, the errors were correlated and therefore tests of independence of residuals were not included. The usual t-test and analysis of variance outputs are also inappropriate analytical tools on their own. The plots of residuals, therefore, served as the main diagnostic tool to assess the fit of the model and the randomness of the residuals. Probability plots were also created to assess the normality of the residuals.

All available general forms of yield projection equation were tested as set out in Table 3.5. These general forms are all compatible with their corresponding growth functions. They are, therefore, the integral of the corresponding growth functions (Woollons, 1989) which are shown in Table 3.6.

Table 3.5 General form of projection equations. (Source: Wollons, R.C. 1989. Advanced Growth and Yield Modelling. Lecture Notes. Univ. of Canterbury.)

I. Yield functions

A. Schumacher

$$Y_2 = \exp(\ln(Y_1)(T_1/T_2) + \alpha(1-(T_1/T_2)))$$

$$Y_2 = \exp(\ln(Y_1)(T_1/T_2)^{\beta} + \alpha(1-(T_1/T_2)^{\beta}))$$

B. Gompertz

$$Y_2 = \exp(\ln(Y_1)\exp(-\beta(T_2-T_1)) + \alpha(1-\exp(-\beta(T_2-T_1))))$$

$$Y_2 = \exp(\ln(Y_1)\exp(-\beta(T_2-T_1)+\gamma(T_2^2-T_1^2)) + \alpha(1-\exp(-\beta(T_2-T_1)+\gamma(T_2^2-T_1^2))))$$

C. Weibull

$$Y_2 = Y_1\exp(-\beta(T_2^{\gamma}-T_1^{\gamma})) + \alpha(1-\exp(-\beta(T_2^{\gamma}-T_1^{\gamma}))$$

D. Morgan-Mercer-Flodin

$$Y_2 = Y_1\left(\frac{(\gamma+T_1^{\delta})}{(\gamma+T_2^{\delta})}\right) + \alpha\left(1-\frac{(\gamma+T_1^{\delta})}{(\gamma+T_2^{\delta})}\right)$$

E. Chapman-Richards

$$Y_2 = (\alpha/\gamma)^{[1/(1-\beta)]}(1-(1-(\gamma/\alpha)Y_1^{(1-\beta)})\exp(-\gamma(1-\beta)(T_2-T_1))^{[1/(1-\beta}$$

$$Y_2 = (\alpha/\gamma)^{[1/(1-\beta)]}(1-(1-(\gamma/\alpha)Y_1^{(1-\beta)})(T_2-T_1)^{\gamma(1-\beta)})^{[1/(1-\beta)]}$$

F. Umemura

$$Y_2 = e^{-\beta(T_2-T_1)}(Y_1(1+\beta(T_2-T_1)+Y_1(T_2-T_1)) + \gamma/\beta^2(1-e^{-\beta(T_2-T_1)}(1+\beta(T_2-T_1)))$$

G. Hossfeld

$$Y_2 = \frac{1}{((1/Y_1)(T_1/T_2)^{\gamma}+(1/\alpha)(1-(T_1/T_2)^{\gamma}))}$$

Table 3.6 Growth functions

A. Scumacher

$$dY/dT = Y/T(\alpha - lnY)$$

$$dY/dT = Y/T(\alpha - \beta lnY)$$

B. Gompertz

$$dY/dT = \beta Y(\ln(\alpha) - \ln(Y))$$

$$dY/dT = (\beta + \gamma T)Y(\ln(\alpha) - \ln(Y))$$

C. Weibull

$$dY/dT = \delta\gamma T^{\delta-1}(\alpha - Y)$$

D. Mercer-Morgan-Flodin

$$dY/dT = \delta T^{\delta-1}(\alpha - Y)/(\delta + T^{\delta})$$

E. Chapman-Richards

$$dY/dT = \alpha Y^{\beta} - \gamma Y$$

$$dY/dT = (\alpha Y^{\beta} - \gamma Y)/T$$

F. Umemura

$$dY^2/dT^2 = \beta dY/dT + \gamma$$

G. Hossfeld

$$dY/dT = (\alpha\beta\gamma Y)/(T(\alpha\beta + T^{\gamma}))$$

As a general example, the yield equation

$$\ln(Y) = \alpha + \beta/T \tag{3.4}$$

where,

Y = a response variable;

T = stand age and;

α, β = parameters

was differentiated with respect to **T** to derive the growth equation

$$dY/dT = Y/T(\alpha - \ln Y) \tag{3.5}$$

or

$$dY/(Y(\alpha - \ln y)) = dT/T \tag{3.6}$$

By separating **Y** from **T** and integrating both sides,

$$\int_{G_1}^{G_2} \frac{dG}{G(\alpha - \ln G)} = \int_{T_1}^{T_2} \frac{dT}{T} \tag{3.7}$$

$$[\ln(\alpha - \ln Y]_{Y_1}^{Y_2} = [\ln T]_{T_1}^{T_2} \tag{3.8}$$

or,

$$\frac{\alpha - \ln Y_2}{\alpha - \ln Y_1} = \frac{T_1}{T_2} \tag{3.9}$$

producing the corresponding difference equation

$$\ln(Y_2) = \ln(Y_1)(T_1/T_2) + \alpha(1-(T_1/T_2)) \qquad (3.10)$$

Clutter et al. (1983) listed the advantages of this form of equation, namely that:

1. the equations are compatible in that the integration of growth over any period will equate exactly to the corresponding yield estimate;
2. the equation is consistent in the sense that when T_2 equals T_1, then Y_2 equals Y_1.
3. there is an upper limit which means that Y approaches α as T approaches ∞.
4. the projection is invariant in that the projected value is not affected by the number of steps over the period of projection.

The equation that best described the behaviour of a variable was then selected based on the goodness of fit as exhibited by the characteristics of the residuals.

The precision of the general equation form selected was further improved by modifying the coefficients through the addition of other variables. These additional variables were used either to modify the exponent term of the general equation selected or to modify the upper asymptote or to modify both the exponent term and the asymptote. Modified candidate equations were again evaluated in terms of residual sums of squares and patterns of residuals. Among the variables tested to improve precision, initial stocking and dummy variables for locality contributed most to the improvement of equations especially in maximum diameter and net basal area per hectare.

The modelling process done in this study had been guided by the practical rules for modelling growth and yield projection equations set out in Appendix A and summarized here in Figure 3.5.

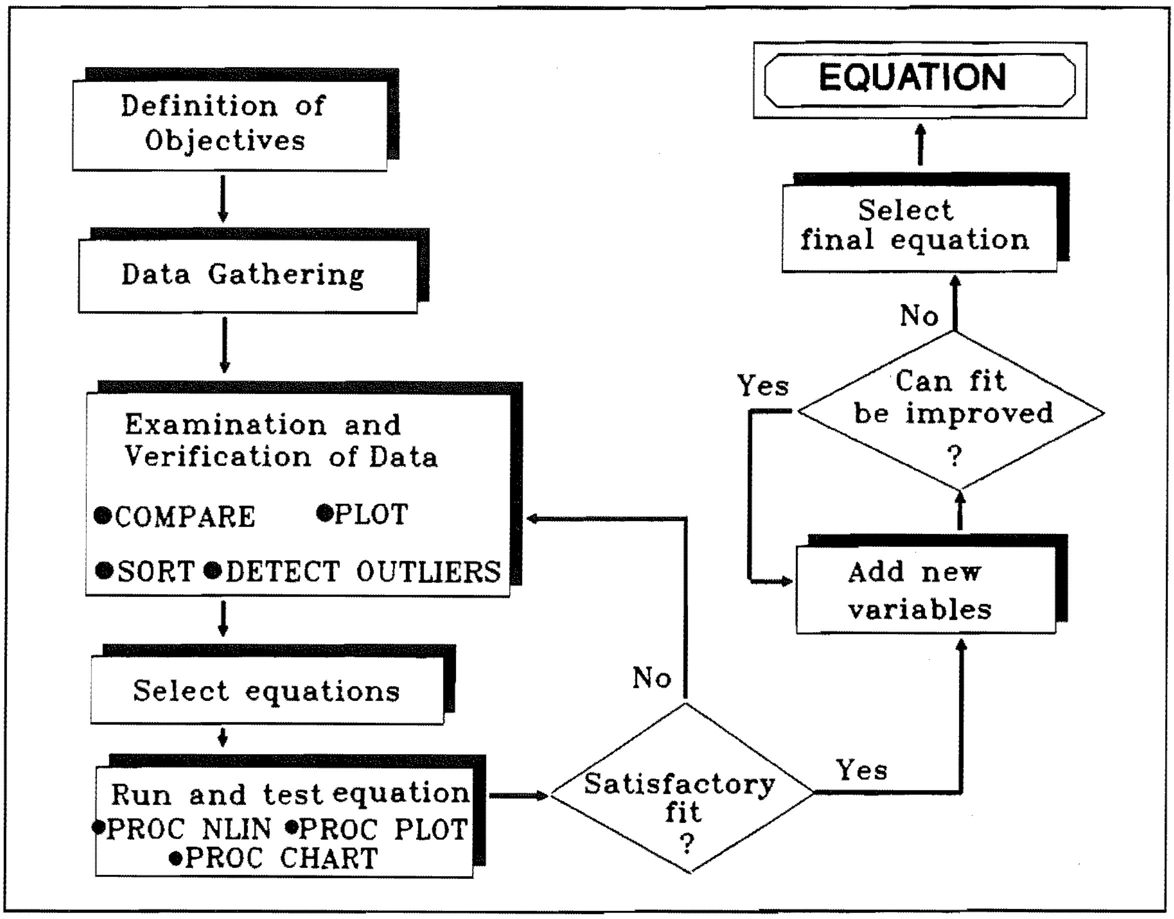


Figure 3.5 Flow diagram for fitting non-linear equations to data

Diameter distribution growth and yield modelling. As for stand level growth modelling where the most suitable functions had first to be selected, modelling the growth of a diameter distribution starts with specifying an appropriate form of probability distribution function (pdf). The one most widely used for this purpose is the Weibull pdf, as explained in Chapter 2.1, where its advantages are fully discussed.

The further advantage of reversing the distribution is that it then provides more information on the larger trees, which represent the more important output of yield forecasts, especially if they are to be used for harvest scheduling. This form of pdf was used, therefore, to test its possible application for this species. Specifically, the form of the Reverse Weibull Distribution function is

$$F(X) = \exp\left[-\left(\frac{a - X}{b}\right)^{c}\right] \quad \text{for } a \geq X \geq -\infty \tag{3.11}$$

$$F(X) = 1, \quad \text{for } X \geq a \tag{3.12}$$

with a probability density function

$$f(X) = \left(\frac{c}{b}\right)\left(\frac{a - X}{b}\right)^{c-1} \exp\left[-\left(\frac{a - X}{b}\right)^{c}\right] \quad \text{for } a \geq X \geq -\infty \tag{3.13}$$

$$= 0, \quad \text{otherwise}$$

where,

X = dbh, the continuous variable being modelled;

a = location parameter;

b = scale parameter;

c = shape parameter.

The parameters of the above Weibull distribution were solved using the Method of Moments technique which estimates parameters of probability distribution functions where some properties of the distribution function are equated to its moments; for example, the mean and standard deviation of the function equate to the first and second non-central moments respectively. The method of moments procedure was adopted largely because of the availability of a tested algorithm to estimate the parameters of the distribution (Strub and Burkhart, 1975; Frazier, 1981; Newby, 1980; Garcia, 1981b; Burk and Newberry, 1984).

Derivation of Stand and Stock Tables. The projection of stand and stock tables required the derivation of $\mathbf{dbh_{ob}}$ class frequencies based on the solved cumulative distribution function (**cdf**). The $\mathbf{dbb_{ob}}$ class frequencies were solved from,

$$m_{ij}=N_{ij}*[\exp[-(\frac{a-U}{b})^{c}]-\exp[-(\frac{a-L}{b})^{c}]] \qquad (3.14)$$

where

L and **U** are lower and upper limit of the diameter class.

$\mathbf{m_{ij}}$ = dbh_{ob} class frequency;

$\mathbf{N_{ij}}$ = estimated surviving trees, **N/ha** and;

a, **b**, **c** are parameters of the **cdf**.

If an appropriate height equation is used, the volume per class can be reliably determined by solving the volume function with height, diameter and age as independent variables. An existing appropriate taper equation was then used to solve the volume and number of log grades by diameter classes (Broad, 1978).

Stand and stock table disaggregation of volumes and numbers of stems per hectare into log classes can be projected at any future age. These projections have properties that make them compatible with stand values projected using stand average projection equations.

Growth and yield simulation system. Three computer programs were written to implement the diameter distribution growth and yield projection model derived in this study. One was written in Vax version FORTRAN, a second in PC version FORTRAN and the third on a spreadsheet template program.

In all three implementations, the growth and yield simulation system starts by allowing the user to enumerate initial stand conditions at an initial age. It then asks the user the age to which projection is wanted. Prior to simulation the user has the option to view all the initial inputs and to edit them, if necessary. Otherwise, the values are confirmed and the simulation starts. When simulation proceeds, projected stand conditions are listed based on the stand projection equation solved for net basal area per hectare, maximum diameter at breast height outside bark, standard deviation of diameter at breast height outside bark and stem survival with specific forms shown in Chapter 4.1.

An estimate of the distribution together with its associated estimated stocking allows the system to generate a stand table with a diameter class that the user specifies.

An existing precise height model of form

$$h=\exp(b_0-b_1/d+b_2(T)) \quad (3.15)$$

where,

h = height of tree in metres

d = diameter at breast height over bark in cm,

T = age of stand in years

b_0, b_1, b_2 are least-squares regression coefficients,

was used to estimate the height at the diameter class midpoint.

The height and diameter class midpoint are then used to determine the tree class volume using the equation

$$v=a_0+a_1d^2h \quad (3.16)$$

where,

v = volume inside bark in m^3

a_0, a_1 are least squares regression coefficients.

The other component of the system is the breakdown of volume into log assortment classes defined by small end diameter and length. This needs the determination of diameter at any point along the length of the log through use of a compatible taper equation. The compatible taper equation used to determine diameter at any point in the log is that derived by Broad (1978) of form,

$$d_{ib}(l') = \sqrt{(4X10^4/v/h(b_1(l'/h)+b_2(l'/h)^2+b_3(l'/h)^3+b_4(l'/h)^4+b_5(l'/h)^5))} \quad (3.17)$$

where,

$d_{ib}(l')$ = diameter inside bark, in cm, l' metres from the tip of the stem;

b_1, b_2, ..., b_5 are existing least-squares regression coefficients (see Broad, 1978).

The diameter distribution growth and yield model was solved and implemented as a simulation model. Its implementation as a spreadsheet simulation model to provide input to a harvest scheduling model is discussed in Chapter 4.2. The growth and yield projection system, YIELD, that was developed was designed so that it produces output that can be easily interfaced with the harvest scheduling model development as discussed in the sections that follow.

3.2 HARVEST SCHEDULE MODELLING

The harvest scheduling model was developed in this study as a multi-period single resource model. This medium term forest level planning is aimed at identifying the sequence of harvests over the planning horizon while satisfying constraints and meeting other management objectives specified. The latter prerequisite has been met by considering other objectives as constraints. The resultant is a harvesting schedule which specifies the hectares of a stand to be harvested in a particular locality employing a logging method, and route for logs to port and most importantly the timing of these actions.

The constraints include realistic capabilities and resource levels expressed as total amounts, increases and decreases over time of the resources used in the different operations, even-flow of harvest, sustainability of wood supplies for each resource, meeting demands for logs from the forests, feasible capital investment and amounts that can be spent in the different operations and desired levels of application of labour intensive logging methods.

The other constraints include some conditions which can be implemented or enforced i.e. ending forest structure, restriction on ages of clearfelling, upper and lower bounds of the resources, required age class distribution of the forest or of the cut at any time. An ideal ending forest structure, traditionally a target normal forest or a fully regulated forest (Johnson and Davis, 1986) has became an appropriate target ending forest structure.

3.2.1 Data Set for Harvest Schedule Modelling

The data used for the harvest scheduling model are derived from a study by de Kluyver et al.(1980) and were generated as a result of the School of Forestry's involvement in several training and research projects for the Fiji Pine Commission. These authors used the data base to formulate and solve a large scale forest harvest scheduling problem. Their study used a traditional LP formulation and solution, the disadvantages of which have been discussed in Chapter 2.2 and alternatives to which are being addressed in this study. The data described here were used in the spreadsheet based harvest scheduling model called **HARVEST**. Its composition can be gauged from Tables B-1 to B-6, or alternatively, some of them like yields and prices, can actually be prepared from functional relationships. The entries in the tables in yield and prices are actually derived from functions, an approach being emphasized in this study so that changes in the data base can be facilitated. The data used by **HARVEST** and shown in Appendix B consist of the following data bases.

The Yield Data Base. The yield data base consists of data that the growth and yield model generated for the different initial crop conditions specified. Each of the yield values in each period in the yield data base is generated through running **YIELD** for each of the initial conditions of the stands. These values were then related to other variables of the model like stand areas and age. The advantage of using formulae relating output of a growth and yield system with other variables and not pre-formed data as input to the LP model is considerable (Villanueva and Whyte, 1992). Moreover, the data base can easily be updated if new and better crop inventory measurements are available. Consequently any change in the yield input data results in a new model formulation which could yield a new harvesting schedule. Table 3.7 sets out the initial stand conditions of the fifteen stands which were scheduled for harvest in the case study as a demonstration example.

Table 3.7 Initial stand conditions.

Stand	Age (yrs)	G (m2/ha)	N (N/ha)	Maximum Dbh_{ob} (cm)	St. Dev. Dbh_{ob} (cm)
1	17	44.0	1181	41	6.1
2	18	46.0	1181	41	6.3
3	19	46.6	1180	43	6.5
4	22	49.0	1178	44	7.0
5	20	48.0	1179	43	6.7
6	15	42.0	1183	40	5.6
7	15	40.0	1180	39	5.0
8	13	39.0	1180	37	5.0
9	18	45.0	1180	42	6.3
10	13	35.0	1190	35	4.5
11	11	35.0	1185	34	4.5
12	12	37.0	1180	36	5.0
13	11	34.0	1190	33	4.3
14	10	32	1190	33	4.5
15	9	30	1190	30	4.0

The Price Function. The price function derives the different prices of logs at a given harvest age within the planning horizon. The function reflects the dependency of log price upon the crop age at time of harvest. Formulating the problem using age-independent prices and costs is possible but not recommended, because the solution from such a formulation possesses undesirable features (de Kluyver et al., 1980). The age distribution of the stands of the case study area indicated a wide range, 9 years being the youngest and 22 years being the oldest, with an average of 14.8 years. This range of ages and a planning horizon of 7 years indicated the need for prices and costs to be projected up to age 28 because harvesting the oldest stand (age 22) at the end of the planning horizon was still an option. The function estimates the value for each combination of log age and price. An implicit price-log size function in the form

$$Price_{age} = f(T, \underline{\theta}) \qquad (3.18a)$$

where,

T is age of the stand and

$\underline{\theta}$ is the set of parameters of the estimating function which was used to derive log prices. The same function was incorporated easily in the spreadsheet LP harvest model, thus making update of the model due to change in prices much easier. Any change in prices of the logs was easily incorporated and recognized by the model. The specific form of the equation is:

$$Price = -22.50 + 4.76(T) - 0.126(T^2) \qquad (3.18b)$$

<u>The Logging Cost Function</u>. The logging costs function estimates the total costs that would be incurred in felling, extracting, and preparing a m^3 of log ready for loading on to a truck. Initially, an effort was made to develop a single equation for all the four methods through aggregated modelling and use of dummy variables. The aggregate model however was no better than any of the individual models. It was decided, therefore, that four separate functions for each of the four methods would be developed and utilized.

Thus, costs were determined for each period in the planning horizon and for the different logging methods. While the derivation of log price and logging method functions are not main concerns of this study, the ability to derive a reasonable function illustrates their value as inputs to other planning models. Having tried functionalizing of price and cost, this study has also initially explored possible improvements in methodology through model aggregation as set out in Whyte <u>et al</u>. (1992).

The implicit form of the equation that was used to model logging cost is,

$$LC_{ijk} = f(\text{logging method, cropage, } \underline{\theta}) \qquad (3.18c)$$

where,

$\underline{\boldsymbol{\theta}}$ is the set of parameters for the cost function

A related data base is that for logging methods, including: a) proportion of each area that is suitable for clearing by individual logging methods and b) number of hectares of a stand that can be cleared by a full year's application of a logging method. For each stand, data of this kind were prepared to reflect factors affected by its condition and topographic class.

The Transport Cost Data Base. The transport costs reflect transporting logs from stands to the different port destinations. It was decided to retain a discrete data base, because it is not foreseen that port and utilization plant locations, and therefore transport distances, which dictate transport cost, will change. Similarly, stand locations are fixed, so that the average transport cost per unit volume remains unaffected at least by port and site distances unless road re-routing is done. The inclusion of other variables like maintenance and insurance costs that may affect transport cost, can result in an even more comprehensive model, but it was considered that those aspects were beyond the scope of this study.

The Port or Utilization Plant Requirements Data Base. Port or utilization plant requirements are represented as volumes of logs that they can accommodate in any one year, thus reflecting limits on what can be transported to them.

3.2.2 Methodologies for Harvest Schedule Modelling

Influence Diagram. **HARVEST**, like any system, consists of variables and their interactions; understanding them enhances the decision-making, planning and control abilities of managers with particular responsibilities for harvest planning. Consisting mainly of variables that can be influenced either directly or indirectly by the decisions managers make, **HARVEST** was designed as a decision support system for the regulation of plantation harvest flows.

The model can be easily understood by examining its components which included its **objectives**, the **decisions** to be made and the **systems environment**. These are discussed individually below.

The **objectives** that the model intended to achieve can be classified as primary and secondary. Its primary objective was the maximization of total net discounted financial returns. In the demonstration example, logging was from 15 areas for a planning period of 7 years using 4 possible harvesting methods and potential routing of log shipments through 2 ports. Its secondary objectives were: (1) to make full and sensitive use of the yields forecasted by the growth and yield forecasting model **YIELD**, the separate modular component of the forest plantation regulatory system described in Chapter 3.1; (2) to maximize use of labour intensive cutting methods and (3) to maximize port and plant utilization while minimizing cost of transporting logs.

The **decisions** that were to be made in **HARVEST**, therefore, answer the following questions: (1) which areas and what volumes will be cut in which stands in which year? i.e. a harvesting schedule; (2) what mix of harvesting methods to use? and (3) to which port should the harvest be routed?. This can be understood from perusal of Figure 3.6 where the simplified interactions of the variables are shown.

The **systems environment** consisted of the variables over which the manager had minimal control, yet they were the variables that largely affect the decisions managers take. These variables included market demand, prices and interest rates, labour costs, machine fixed costs and supplies, topography and port capacities. Most of these exogenous variables were random in character which implies that their values were subject to considerable uncertainty.

A display of the decision variables (enclosed in boxes), the intermediate variables (enclosed in circles), the exogenous variables (neither directly nor indirectly preceded by a decision variable) and outcome attribute (discounted net present value) pertaining to the harvest scheduling problem, along with the dependent relationships among them (represented by arrows or influence lines), resulted in the influence diagram in Figure 3.7. The influence diagram represents these components in proper juxtaposition and is a record of how the system works. The direction of the arrows and the signs in their ends represent the sign of the general form of the dependent relationships between the variable at the tail of the arrow and the variable at its tail. The + sign indicates that the variables changes in similar direction whilst a - sign

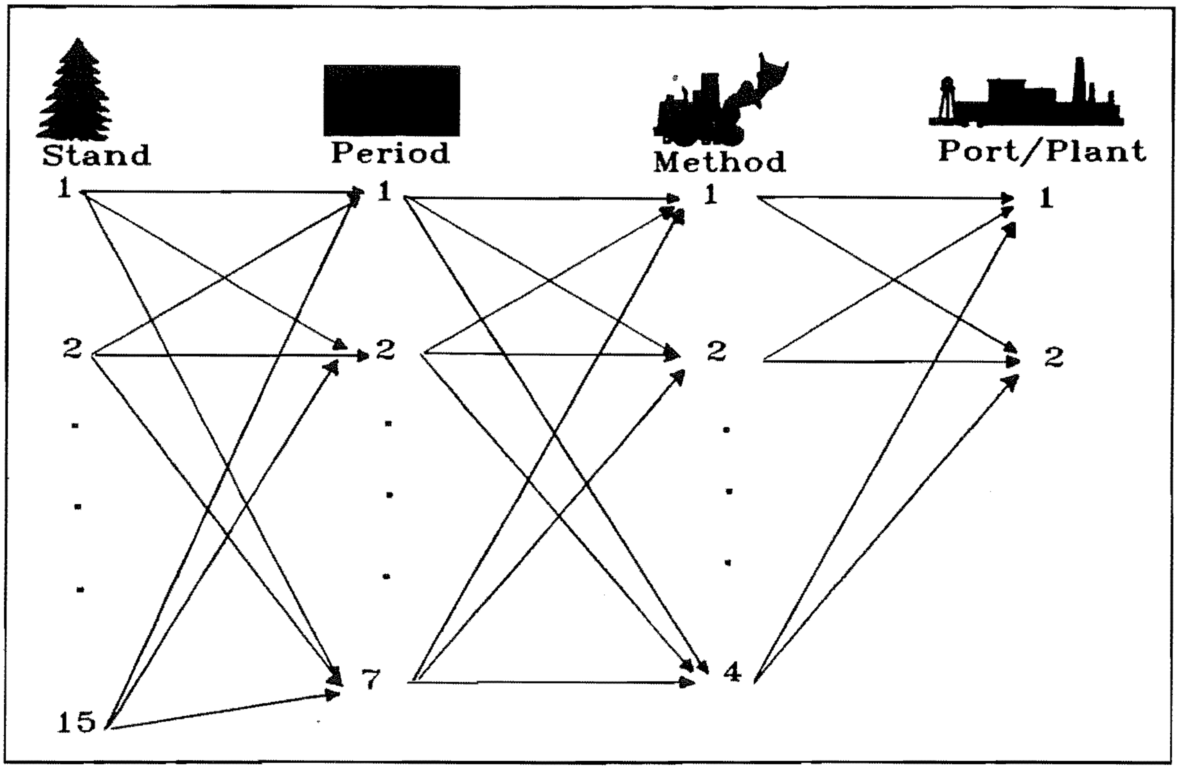


Figure 3.6 Simplified interactions in harvest scheduling problems.

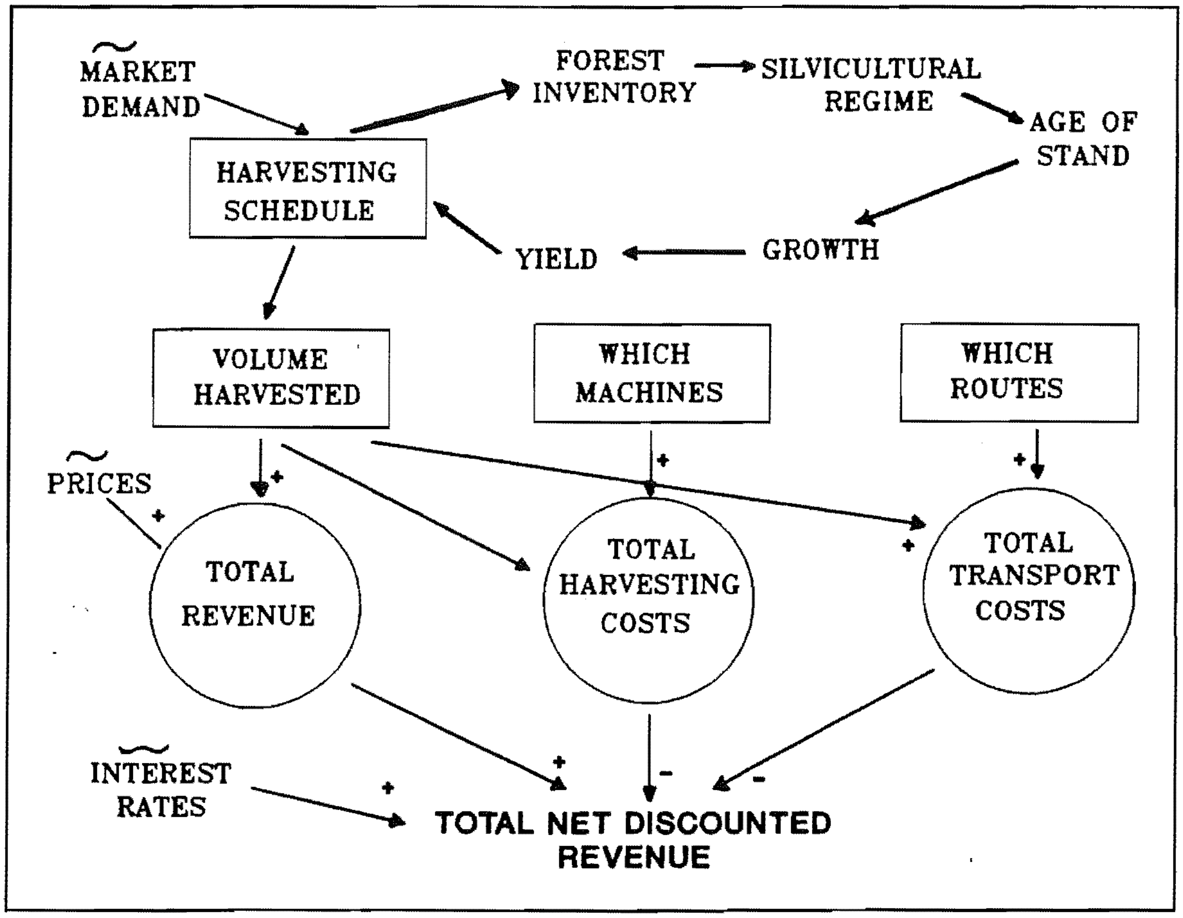


Figure 3.7 Influence diagram of HARVEST.

indicates opposite direction of change; thus for example an increase in revenue, assuming other things equal, results from an increase in harvest and thus produces an increase in profit. Wiggly arrows indicate that the dependency of one variable on the level of another variable is uncertain in magnitude. The usefulness of the influence diagram prior to quantitatively formulating model has been well cited (Coyle, 1977; pp. 63-93); it is especially useful for explaining **HARVEST**, which seeks profit optimization over time, a characteristic peculiar to dynamic optimization systems. The other aspect that demonstrates the dynamic form of this is the presence of the exogenous time series variable which also drives the variable in the loop connecting growth, yield, volume and schedule of harvest. This loop is indicated by darker arrows in the diagram.

For harvest schedule modelling it is useful because: (1) it is used to display the harvest scheduling problem and to frame the concept of the model; (2) it explains a great deal about the information and its structure that must be available for decisions to be made and (3) it serves as a framework for expressing more specifically the exact nature and direction of the influence and relationships within the system.

The influence diagram was used to define a system boundary sufficient for the purpose for which the model was proposed to be built. Thus, through this diagram, the variables to include and exclude from this specific harvesting model could be selected.

The Mathematical Model. The harvest scheduling problem was formulated as a linear programme (LP). Such formulation also served as a basis for the construction of the spreadsheet model. As a management problem it has the elements which make it amenable to LP modelling: thus there is an objective to be pursued, there are restrictions in its pursuit, alternatives which are open to management and levels at which resources are to be used. The contribution of each alternative or level to the objective and the technological coefficients and the relationships between the alternatives and the restriction can then be evaluated.

I. Variables

The indices and variables with their corresponding symbols used in the mathematical formulation of the model are set out below.

A. Indices

i = area index, $i = 1, 2, \ldots I$;

j = year index, $j = 1, 2, \ldots J$;

k = method index, $k = 1, 2, \ldots K$;

m = port index, $m = 1, 2$;

B. Variables

r_i = number of hectares available for cutting in area i;

g_{ij} = total yield of area i in m^3, if cut in year j;

X_{ijkm} = hectare of area i cut in year j using method k with logs brought to port m;

Y_{ijm} = m^3 of logs from area i to port m in year j;

D_j = combined annual requirements of both ports in m^3;

d_{j1} = annual port requirements in year j of port 1 in m^3;

dj_2 = annual port requirements in year j of port 2 in m^3;

N_{ijk} = equivalent number of annual applications of machine intensive cutting method k in year j, area i (k=2, 4);

M_{ijk} = equivalent number of annual applications of labour intensive cutting method k in year j, area i (k=1, 3);

f_{ik} = number of hectares that can be cleared in one full year's application of machine intensive cutting method k in area i (k = 2, 4);

e_{ik} = number of hectares that can be cleared in one full year's application of labour intensive cutting method k in area i (k=1, 3);

F_{ik} = maximum or minimum number of hectares to be cleared using machine intensive method k in area i (k=2,4);

E_{ik} = maximum or minimum number of hectares to be cleared using labour intensive method k in area i (k=1,3);

P_j = price per m^3 realised in year j;

C_{ijm} = transport cost per m^3 from area i to port m in year j;

n_{ijk} = cost of applying cutting method k for one full year in area i in year j (k = 2, 4) and;

m_{ijk} = costs of applying cutting method k for one full year in area i in year j (k = 1, 3);

II. Mathematical formulation

The harvest scheduling model defined in a mathematical formulation using the indices, variables and notation listed above consists of functions relating the variables in forms that translate their relationships into mathematical inequalities. The formulation meets the assumptions of linear programming viz., linearity, divisibility, non-negativity and deterministic variables.

The subsequent mathematical formulation consists of the following objective function and constraints.

A. Objective Function

$$Max Z = R - LC_1 - LC_2 - TC \qquad (3.19)$$

where,

$$R = \sum_{j=1}^{J} P_j \left(\sum_{i=1}^{I} \sum_{k=1}^{K} \sum_{m=1}^{M} g_{ij} X_{ijkm} \right) \qquad (3.20)$$

represents the total gross revenue. This gross revenue represents the sum of all revenues from cutting each of the planning units.

$$LC_1 = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} n_{ijk} N_{ijk} \qquad (3.21)$$

$$LC_2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} m_{ijk} M_{ijk} \qquad (3.22)$$

$$TC = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{m=1}^{M} c_{ijm} \left(\sum_{k=1}^{K} g_{ij} X_{ijkm} \right) \qquad (3.23)$$

Equation 3.21 represents the total logging costs for using labour intensive logging methods while 3.22 represents the total logging costs for using machine intensive logging methods. Equation 3.23 represents the total transport costs.

The costs and revenues are assumed to be incurred at mid-year annual periods over the planning horizon of seven years. Profits and costs incurred in the future were valued in comparison to the present by discounting future revenues and costs at an appropriate rate of interest, r_t which is

$$r_t = (1+i)^{-j} \qquad (3.24)$$

where,

i = discount factor in % * 10^{-2}

j = year when the crop is harvested

B. Constraints

The constraints restrict the values that the decision variables can assume. They represent the factors that have significant limiting effect on the selection of any harvesting schedule.

1) Area constraints

This constraint ensured that each stand can be clearcut only once during the planning horizon.

$$\sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{m=1}^{M} \left(\frac{X_{ijkm}}{r_i}\right) \leq 1 \qquad \text{for } i=1,2,..I \qquad (3.25)$$

Since there were 15 stands in the case study example, there were 15 constraints of this type.

2) Cutting method constraints

Four types of logging method are included in the model. The use of each method was based on two factors. The first factor is the product type to be harvested and the second factor is topography. Their application can be summarized as set out below.

Table 3.8 Harvesting method applications based on topography and product to be harvested.

Harvesting Method	Product	Topography (% slope)
Skidder	Sawlog	≤ 45
Cable Yarder	Sawlog and/or Pulplog	>45
Plastic chute	Pulplog	>20
Manual	Pulplog	≤ 20

From the above table either of the machine intensive logging methods can be used to harvest sawlogs. The use of the skidder, however, is restricted to flatter areas. Logging by cable yarder, the other machine logging method needs to be used in steeper terrain for both pulpwood and sawlogs. Plastic chutes were used in steep areas to harvest short length pulpwood, posts or poles. On flatter areas a manual method was used to stack bundles of short wood which are picked up by

Based on these considerations, two constraints on logging methods were formulated.

a) Cutting method constraints within stands:

$$r_i \sum_{j=1}^{J} \sum_{m=1}^{M} X_{ij2m} \leq F_{i2}; \qquad \text{for } i=1,2,..I \qquad (3.26)$$

$$r_i \sum_{j=1}^{J} \sum_{m=1}^{M} X_{ij4m} \geq F_{i4}; \qquad \text{for } i=1,2,..I \qquad (3.27)$$

b) Overall maximum use of labour intensive methods:

$$\sum_{i=1}^{I} N_{ij2m} \leq 1; \qquad \text{for } j=1,2,..J \qquad (3.28)$$

3) Port requirements constraints

Port 1 constraints require lower limit (minimum) of volumes to be transported to it in the form of ≥ constraints (see Equation 3.30). Port 2 constraints (Equation 3.31) also require an upper limit to volume that can be transported to it, in the form of ≤ constraints. Overall there are combined annual port requirements shown in Equation 3.32. These annual port requirement constraints are typical of harvest flow constraints as explained in Jonhson and Scheurman (1977). The harvest flow constraints also prescribe a constant increase (or decrease) of harvest from one period to the next. While the harvest flow constraints are able to constrain an even flow of harvest, the port constraints in the model are only able to ensure volumes are harvested, albeit unevenly, in each year in the planning horizon. From another viewpoint it can also be compared to the demand constraints.

$$\sum_{k=1}^{M} (g_{ij}X_{ijkm}) - \sum_{m=1}^{M} Y_{ijm} = 0 \qquad \text{for } j=1,2,..J \text{ and } i=1,2,..,I. \qquad (3.29)$$

$$\sum_{i=1}^{I} Y_{ij1} \geq d_{j1}; \qquad \text{for } j=1,2,..J \qquad (3.30)$$

$$\sum_{i=1}^{I} Y_{ij2} \leq d_{j2}; \qquad \text{for } j=1,2,..J \qquad (3.31)$$

$$\sum_{i=1}^{I} \sum_{m=1}^{M} Y_{ijm} \geq D_j \qquad \text{for } j=1,2,...J \qquad (3.32)$$

4) Periodic harvest regulation

$$\sum_{i=1}^{I} (g_{ij+1} \sum X_{ij+1km}) \geq \sum_{i=1}^{I} (g_{ij} \sum X_{ijkm}); \qquad j=1,2,...J. \qquad (3.33)$$

This constraint is stated as a non-declining yield constraint but is also another way of defining relationships among harvests in the different periods, i.e. the harvest in a period cannot vary more than a certain percentage from the harvest in the preceding period. For example if a harvest cannot be increased or decreased by more than 20% of the preceding period then the following constraints are written,

$$H_{j+1} \geq 0.8*H_j \qquad (3.34)$$

which restrain harvest in period j+1 to be not less than 80% of the harvest in period j: that is harvest can drop only 20%.

$$H_{j+1} \leq 1.2*H_j \qquad (3.35)$$

The above inequality restrains harvest in period j+1 to be not more than 120% of the harvest in period j; that is harvest can increase only 20%. These constraints in general form are written as

$$H_{j+1} \geq (1-\alpha)*H_j \qquad \text{for all } j \qquad (3.36)$$

and

$$H_j+1 \leq (1+\beta)*H_j \qquad \text{for all } j \qquad (3.37)$$

where α and β are the permitted proportional increases or decreases. For a non declining yield constraint, $\alpha=0$ and β is unspecified. Therefore the constraints will be written as

$$H_{j+1} \geq H_j \qquad \text{for all } j, \qquad (3.38)$$

These constraints fall under the category of volume control constraints. Another constraint falling under this category is a constraint of equal periodic cut typical of a fully regulated forest and written as

$$H_{j+1} = H_j \qquad \text{for all } j. \qquad (3.39)$$

A more direct way of injecting this volume control constraint is to set an upper and lower limit on the absolute amount of harvest in the periods and written as

$$H_j \geq L \qquad \text{for all } j \qquad (3.40)$$

and

$$H_j \leq U \qquad \text{for all } j, \qquad (3.41)$$

and where L and U are the minimum and the maximum that can be harvested in each period respectively.

7) Non-negativity constraints

This constraint requires that the decision variables can only take positive values.

III. Model Implementation and solution

Linear programming solution algorithms abound. The advances in computing have contributed much to this development. What was started by Dantzig fifty years ago has been improved considerably by people who have been working on the simplex algorithm.

There have been many choices on which LP solution package should be used to solve the model in this study. What has been chosen is the one which can most easily interface with the growth functions. The next chapter describes in detail how planning data were efficiently used by the harvest schedule model. Results of the integrated growth and yield and harvest schedule modelling are then discussed.

CHAPTER 4

RESULTS AND DISCUSSION

This chapter presents the results of the various modelling efforts done in growth and yield modelling, harvest scheduling and their interfacing. First, it discusses the results of modelling the growth of the various stand and diameter variables and then it proceeds to show how the variables were used to solve the diameter distribution model. The results of validation of the model with the use of an independent data set and evaluation through sensitivity analyses are also shown and discussed. This chapter continues with a discussion on the results of a case study used as an application of the microcomputer spreadsheet-based harvest schedule model that have relied heavily upon the yield model implemented in the same programming environment. Finally, the overall performance of the interfaced models is assessed.

4.1 Stand Projection Equations Solved

4.1.1 Stand Net Basal Area Per Hectare

In modelling the growth of net stand basal area per hectare, the basic data set (BAREA.DAT) was validated by checking basal area outliers after a first fit of a selected functional form, as explained in Chapter 3. In determining outliers, Equation 3.2 was used to standardize the residuals. This transformation was also used in detecting outliers in all equation-fitting analyses.

The result of the validation procedure created the final projection data set that was reduced in size and which is described in Table 4.1. The table shows the dependent and the independent variables that were used in the model.

Table 4.1 Description of the final projection data set used to model net basal area per hectare.

VARIABLE	MEAN	STD. DEV.	MAXIMUM	MINIMUM
G_1	22.7	8.3	50.4	1.8
G_2	29.0	8.7	60.1	2.3
N_1	777	288	2152	100
T_1	8.3	2.3	13	3
T_2	11.1	2.5	17	4

All the general forms of projection equation listed in Table 3.4 were tested systematically using first the forms that are more commonly used (for example Schumacher and Gompertz), and then the less common ones (for example logistic and others). Initial runs of the most acceptable function, namely the Schumacher form, showed that behaviour of the growth of net basal area per ha was also affected by initial stocking index ($I = N_1/1000$) and forest locality (S). An ideal equation was therefore one with both of these predictor variables. Therefore all possible forms of equations with these variables modifying the parameters of their general form were tested.

The modelling process revealed that the growth of net basal area per ha was best described by the Schumacher equation with its exponent modified by a function of initial stocking and differences in locality. The explicit form of the equation is

$$G_2 = \exp(\ln(G_1)(T_1/T_2)^{\beta+\gamma S_2+\delta I} + \alpha(1-(T_1/T_2)^{\beta+\gamma S_2+\delta I})) \qquad (4.1)$$

where,

S_2 = dummy variable for locality Seaqaqa, else Lololo or Drasa;

I = index for stocking which is $N_1/1000$ where N_1 is initial stocking;

α, β, γ, δ are coefficients estimated by non-linear least squares;

T_1, T_2, G_1, and G_2 are as defined in Table 3.2.

The above equation shows that modifications incorporated in the general form of the equation account for the differences in responses exhibited by the different forest localities namely Lololo, Drasa and Seaqaqa. Initial runs had each locality represented by a dummy variable, but the results showed that the coefficient for one dummy variable was insignificant and that the growth of net stand basal area per hectare in the two localities, Lololo and Drasa, did not actually exhibit enough difference, which meant that they could be combined as one locality. However, growth of net stand basal area per hectare in Seaqaqa differed significantly, so that any model developed needed to be sensitive to the different growth paths in only two localities. The use of the dummy variables also allowed the adoption of a single growth projection equation for the forests, one that was still able to account for their differences, rather than using separate equations for each. The additional predictor variable in the model, I, is a stocking index. The inclusion of this variable accounted for much variation induced by having widely different stand densities at any age. This circumvented the problem of developing separate models for thinned and unthinned stands. Thus, Equation 4.1 is an attempt to represent, in one functional model form, the variations in the behaviour of stands that could have been described less efficiently by six models, one model for each of the thinned and unthinned stands of each of the three forest localities. This finding indicated possible improvement in the growth and yield modelling studies in tropical pine plantations. Much modelling research which could be done in these countries should strive for efficiency in utilizing available data, and at the same time consider the accuracy warranted.

Table 4.2 shows the results of the analysis of variance (ANOVA) and the estimates of the parameters of the function chosen to fit the data best. While the corresponding standard errors of the parameters shown are very small with respect to the estimate of the parameters and indicate significance of the parameters, such statistics were used only in a relative way, since no complete reliance can be placed on them as they come from a highly correlated data set. The real indication of precision can be gleaned from the graphs in Figures 4.1 and 4.2 which show that the model meets the assumption of randomness and normality of the residuals which fall mainly within $\pm$ 4 m^2/ha.

The PROC UNIVARIATE statistics in Table 4.3 provides proof that the equation provides an unbiased precise estimate of net stand basal area per hectare.

Table 4.2 Parameter summary and ANOVA for net basal area per hectare projection equation.

PARAMETER	ESTIMATE	ST. ERROR
α	4.2490	0.02467
β	0.5473	0.02679
δ	0.0981	0.01518
γ	0.2953	0.03302

SOURCE	DF	SS	MS
Model	4	945836.72	236459.18
Error	1041	3759.81	3.61
Total	1045	949596.54	

Table 4.3 Summary of characteristics and distribution of residual values for the net basal area per hectare projection model.

Mean......................	-0.0380		
Standard Deviation........	1.8973		
Skewness..................	0.0108		
Kurtosis..................	2.4175		
T:MEAN = 0...............	-0.6486	Prob >/T/...	0.51
SIGN RANK................	-8042.5	Prob > /S/..	0.40
D:NORMAL..................	0.0759	Prob > /D/..	< .01

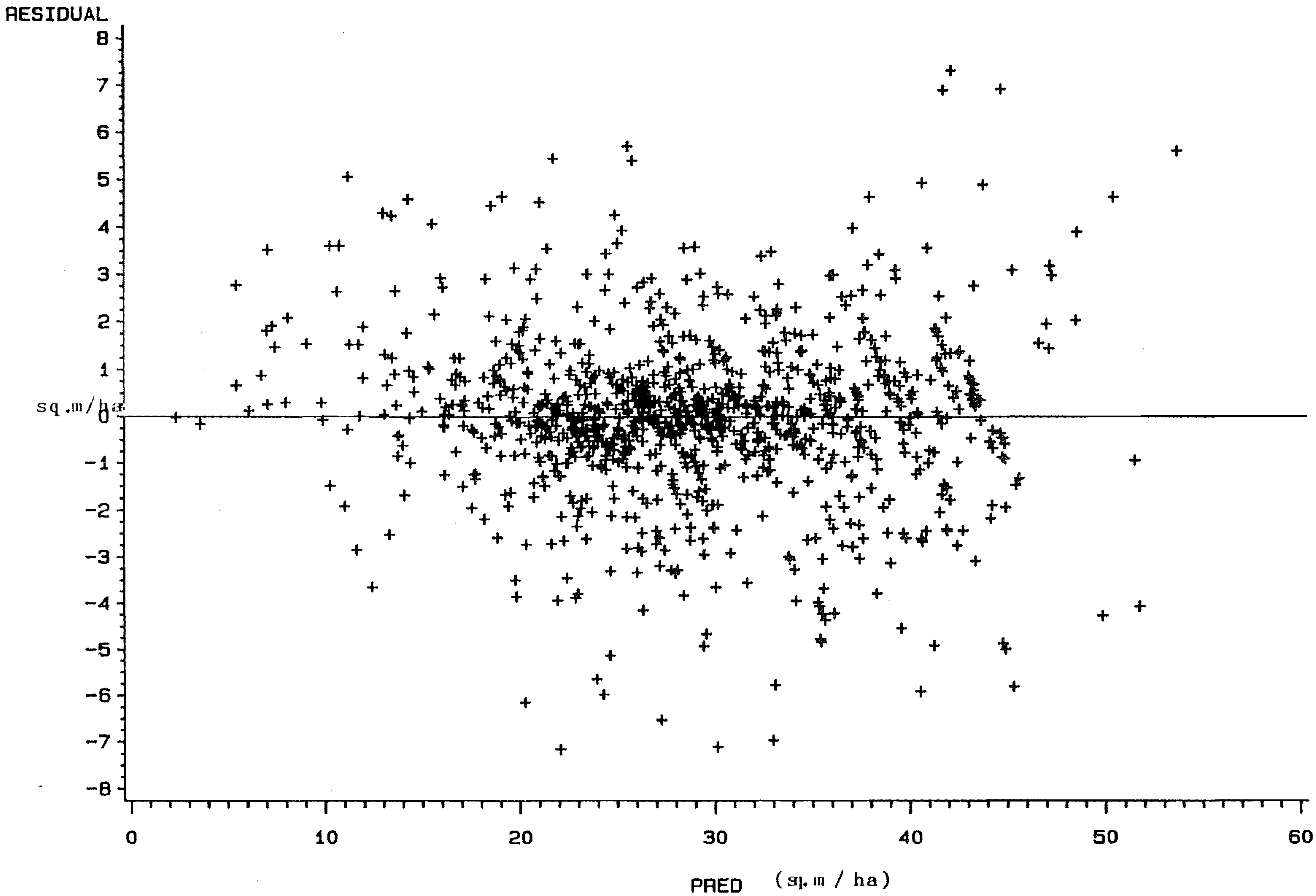


Figure 4.1 Plot of residuals for net basal area per ha projection equation.

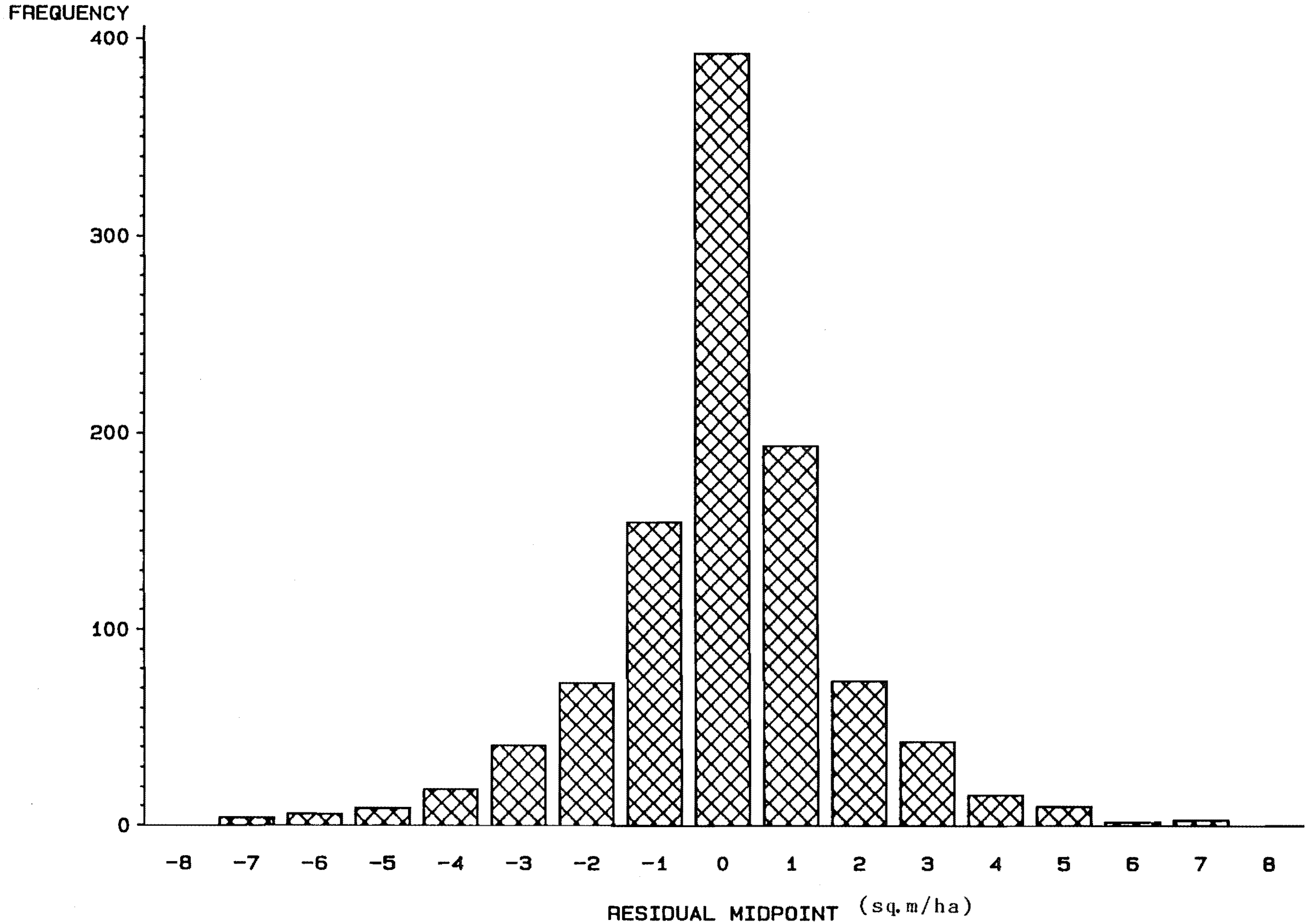


Figure 4.2 Frequency distribution of residuals for net basal area per ha projection equation.

4.1.2 Mean Tree Basal Area

Mean net basal area per tree was not able to be successfully modelled. It was derived, therefore, from projected values of the net basal area per ha and stocking. For any projection period the corresponding mean net basal area per tree was solved by the following relationship.

$$\overline{g_t} = \frac{G_t}{N_t} \tag{4.2}$$

where,

$\overline{g}$ = projected mean net basal area per tree in period t;

N_t = projected stocking at period t; and

G_t = projected net basal area per hectare in period t.

The growth of mean tree basal area as projected by the use of this approach is shown in Figure 4.3. Stocking which was used to estimate the values in this figure is projected by the mortality model in Section 4.3 while net basal area per hectare is projected by the model discussed in the previous section.

4.2 Stand diameter variables projection equation solved

4.2.1 Mean diameter

Mean diameter was found to be a variable that was quite difficult to model. All the equation types listed in Table 3.4 were tried in an attempt to describe the behaviour of the growth of this variable, but no functional form passed the minimum criteria deemed necessary for accepting a model. This variable was, therefore, estimated by the relationship

$$Dmean = \sqrt{(Dq^2 - (\frac{n-1}{n})Dstd^2)} \qquad (4.3)$$

where,

Dmean = stand mean diameter in cm,

Dq = stand quadratic mean diameter in cm,

= $[10000*(g/0.7854)]^{1/2}$

where **g** is mean tree basal area in m^2,

Dstd = diameter standard deviation in cm,

as cited in Equation 3.4. The projected growth of **Dmean** with the use of this approach is shown in Figure 4.4.

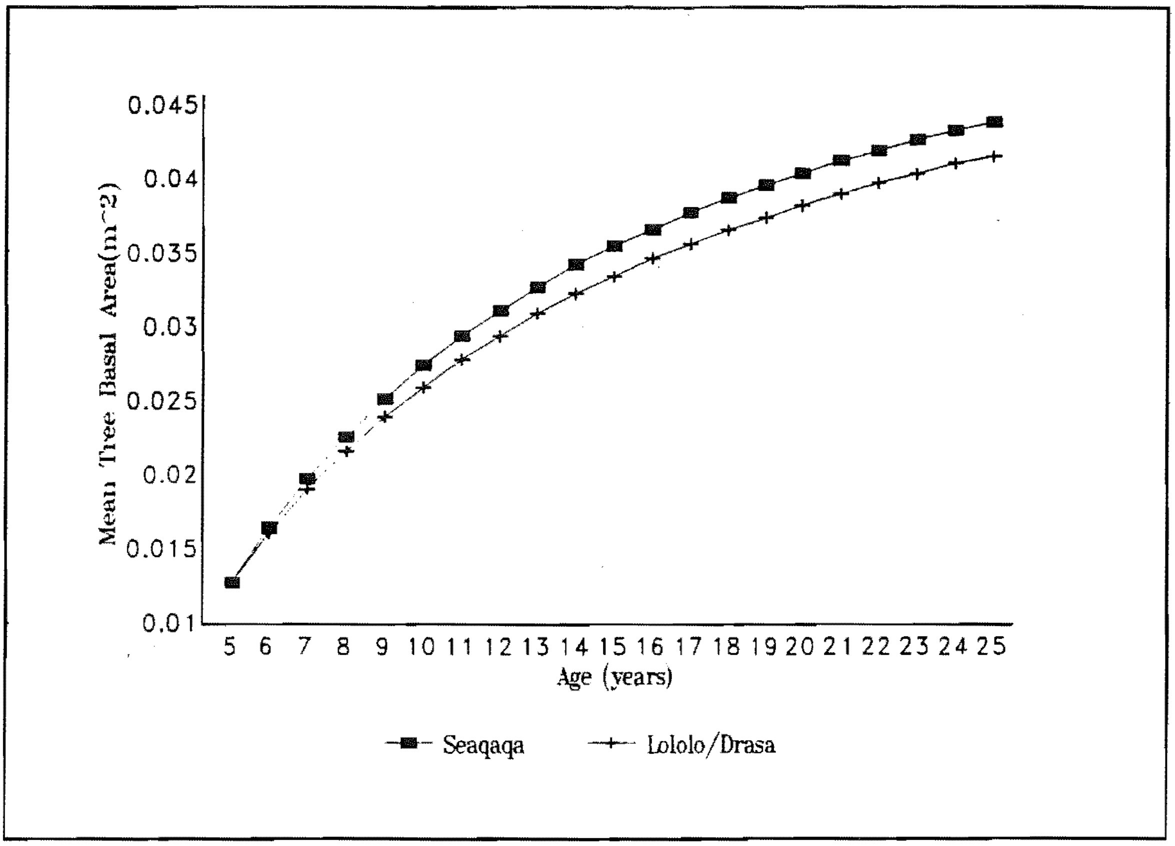


Figure 4.3 Predicted mean tree basal area.

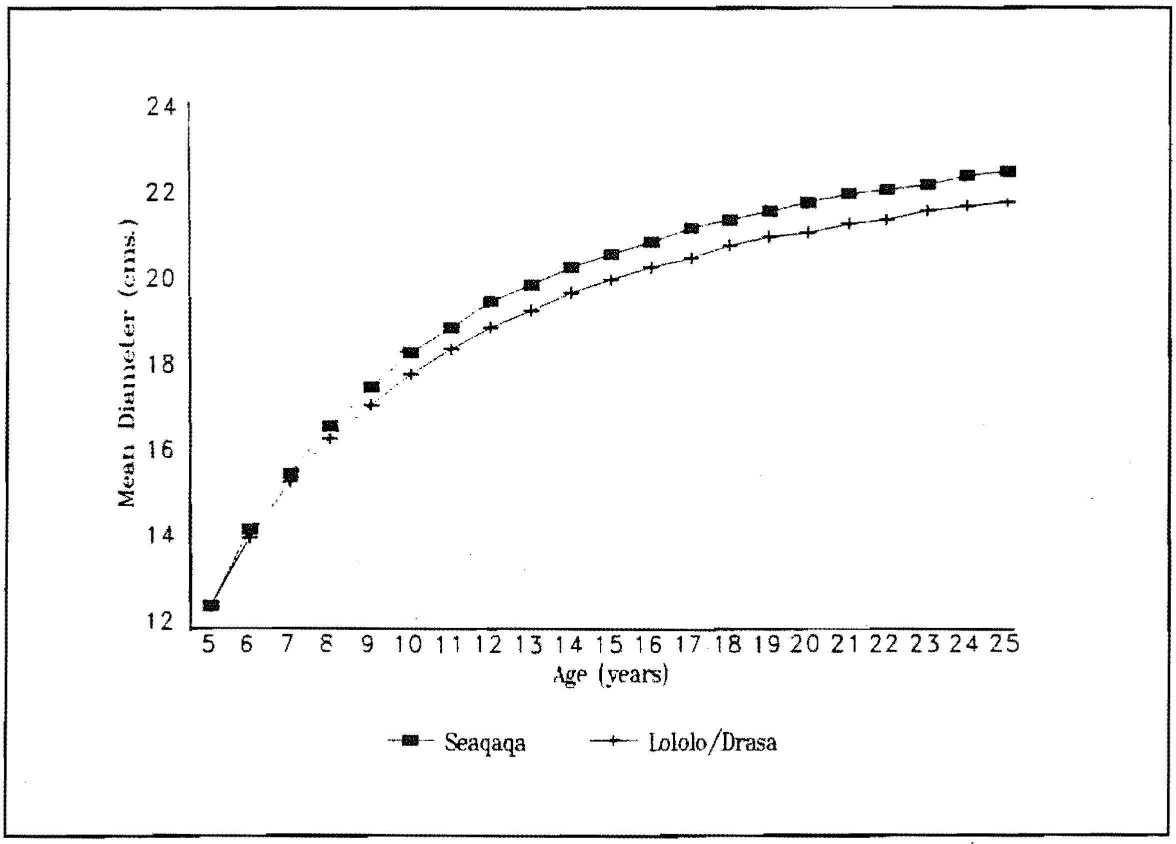


Figure 4.4 Predicted growth of mean diameter.

4.2.2 Diameter standard deviation

The validation procedure prior to finally selecting the function that best describes the change in growth of diameter standard deviation over time resulted finally in the creation of the data set outlined in Table 4.4 to model this variable.

Table 4.4 Description of the final data set used to model diameter standard deviation.

VARIABLE	MEAN	STD. DEV.	MAXIMUM	MINIMUM
$Dstd_1$	3.89	1.07	6.85	1.00
$Dstd_2$	4.57	1.11	7.42	1.41
T_1	8.24	2.33	13.0	3.0
T_2	10.80	2.44	14.0	4.0

Again, all forms of the various equations were tested. The modelling routines revealed that this variable could be best modelled by a modified form of the Gompertz growth function, namely:

$$Dstd_2 = \exp(\ln(Dstd_1)\exp(A) + \alpha(1-\exp(A))) \qquad (4.4)$$

where,

$$A = -\beta(T_2 - T_1)$$

α and β are coefficients estimated with non-linear least-squares and

$Dstd_1$, **$Dstd_2$**, **T_1** and **T_2** are as defined in Table 3.2.

Table 4.5 shows the results of the ANOVA, the estimates of the parameters and their corresponding standard errors. The values indicated significance of the parameters. The graphs in Figure 4.5 and Figure 4.6 show that the model meets the assumption of randomness and normality of the error term. The statistics in Table 4.6 provide evidence that the equation provides an unbiased precise estimates of diameter standard deviation. Almost all residuals lie within ±1.00 cm.

Table 4.5 Summary of characteristics and distribution values for the residuals of the diameter standard deviation projection model.

Mean......................	0.01315		
Standard Deviation........	0.36368		
Skewness..................	0.50519		
Kurtosis..................	0.85401		
T:MEAN = 0................	1.18398	Prob > /T/...	0.2366
SIGN RANK................	2267	Prob > /S/...	0.3053

Table 4.6 Parameter summary and ANOVA for diameter standard deviation projection equation.

PARAMETER	**ESTIMATE**	**ST. ERROR**
α	2.3906	0.0466
β	0.0653	0.0034

SOURCE	**DF**	**SS**	**MS**
Model	2	23430.28	11715.00
Error	1069	141.71	0.1325
Total	1071	23572.00	

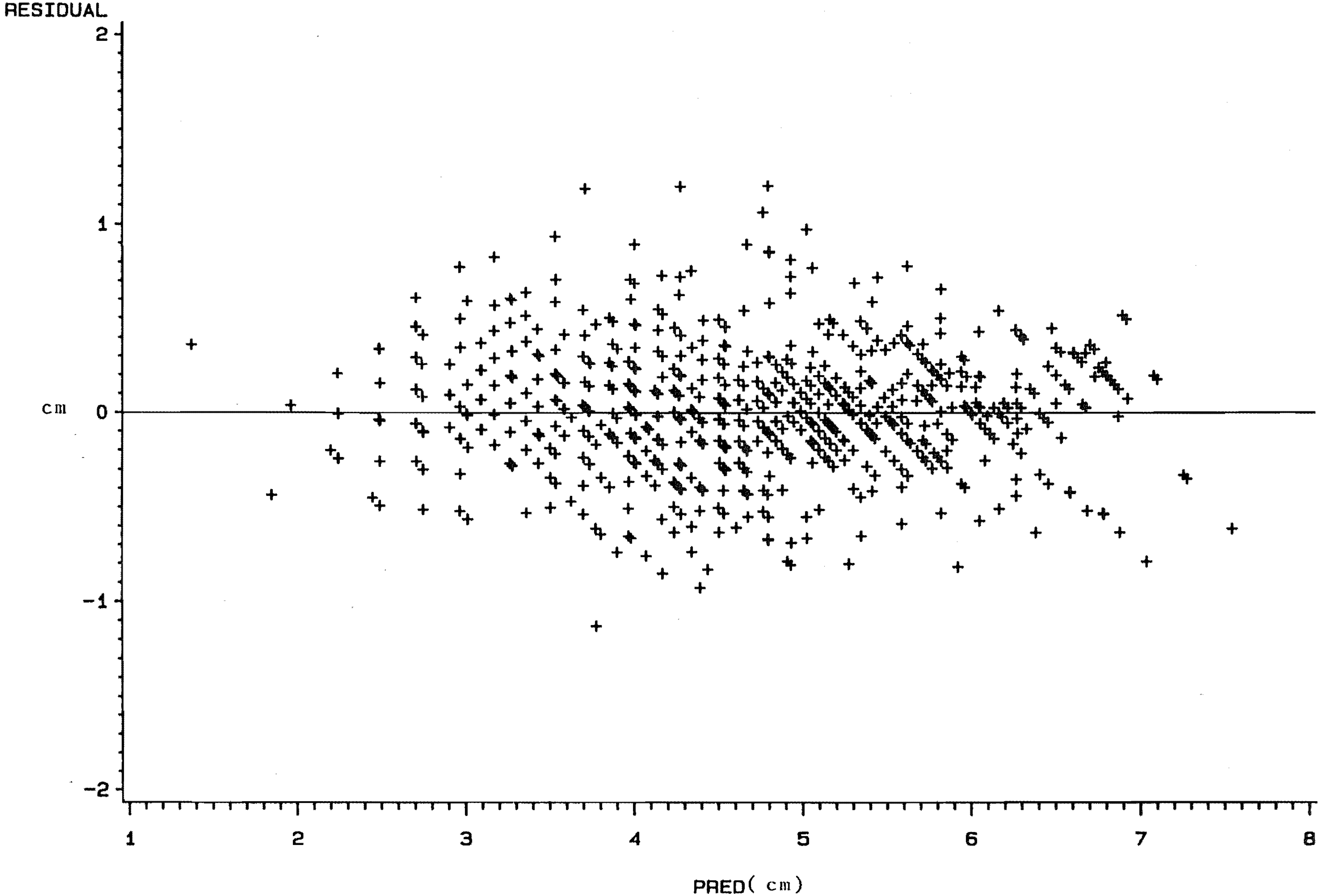


Figure 4.5 Plot of residuals for diameter standard deviation projection equation.

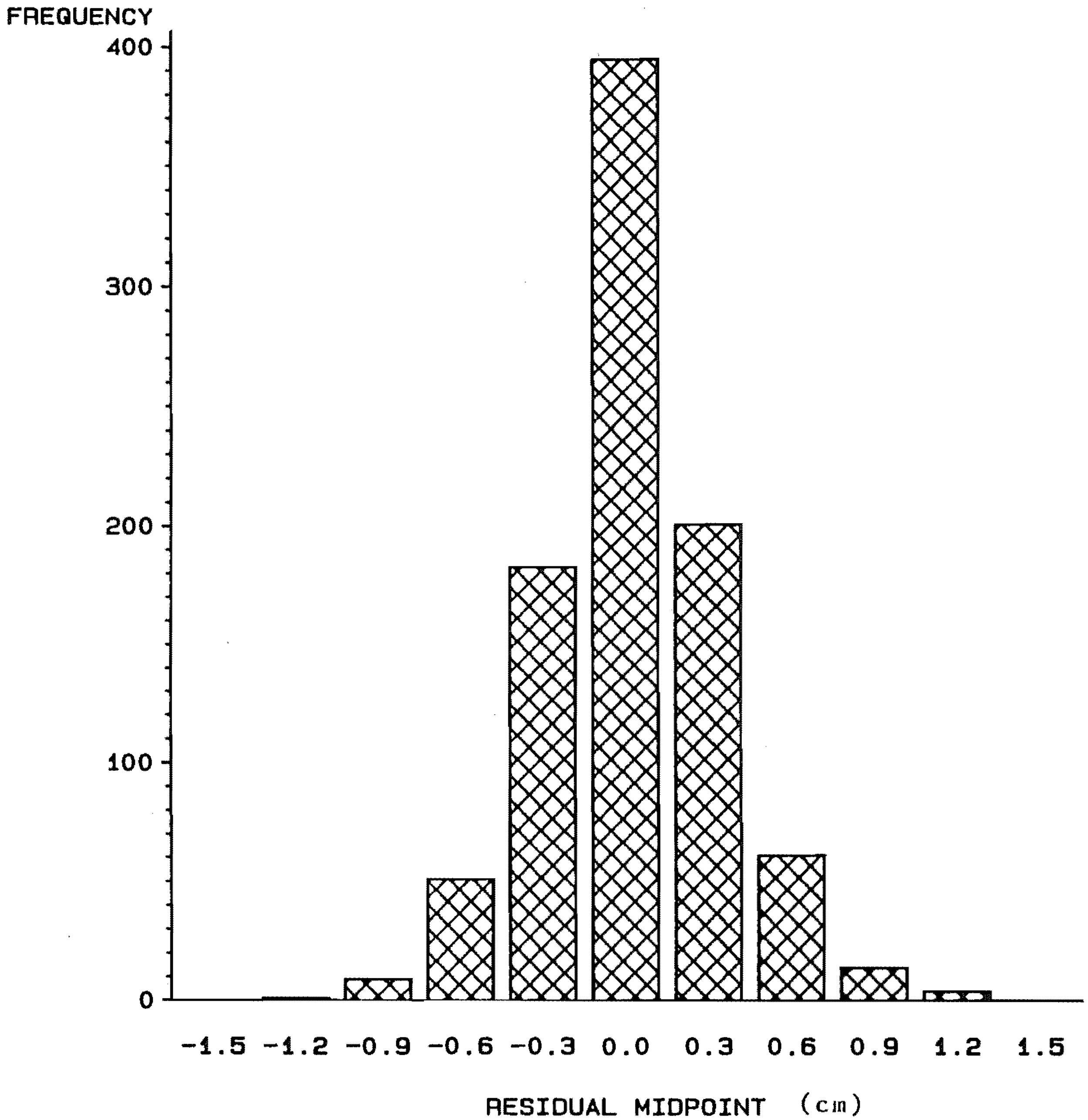


Figure 4.6 Frequency distribution of residuals for diameter standard deviation projection equation.

4.2.3 Maximum diameter

The validation process conducted to assess the representativeness of the data set used to model maximum diameter resulted in the data base summarized in Table 4.7.

Table 4.7 Description of the final data set used to model maximum diameter.

VARIABLE	MEAN	STD.DEV.	MAXIMUM	MINIMUM
$Dmax_1$	30.69	8.40	53.0	9.0
$Dmax_2$	35.36	7.72	55.0	13.0
T_1	8.3	2.34	13.0	3.0
T_2	10.8	2.41	14.0	4.0
N_1	768.3	298.9	2152.0	150.0

This variable was best modelled by the modified form of the Schumacher equation, the specific form of which is shown below.

$$Dmax_2 = \exp(\ln(Dmax_1)(T_1/T_2)^{\beta+\gamma S_2} + (\alpha+\gamma S_2+\delta I)(1-(T_1/T_2)^{\beta+\gamma S_2})) \qquad (4.5)$$

where,

S_2 = dummy variable for locality;

I = index for stocking which is **N**/1000 where **N** is initial stocking/ha;

α, β, γ, δ are coefficients estimated by non-linear least-squares;

$Dmax_1$, $Dmax_2$, T_1, and T_2 are as defined in Table 3.4.

Table 4.8 shows the results of the ANOVA and the estimates of the parameters and their corresponding standard errors. The graphs in Figure 4.7 and Figure 4.8 show that the model meets the assumption of randomness and normality of the residuals which fall mainly within $\pm$ 4.0 cm. The statistics in Table 4.9 provide evidence that the equation provides unbiased, precise estimates of maximum diameter.

Table 4.8 **Parameter summary and ANOVA for the maximum diameter projection equation.**

PARAMETER	ESTIMATE	ST. ERROR
α	4.1287	0.0383
β	0.7396	0.0277
δ	0.1797	0.0096
γ	-0.1878	0.0232

SOURCE	DF	SS	MS
Model	4	1411375.11	352843.77
Error	1136	2993.88	2.63
Total	1140	1414369.00	

Table 4.9 **Summary of characteristics and distribution values of the residuals for the maximum diameter projection model.**

Mean......................	0.2971		
Standard Deviation........	1.6210		
Skewness..................	0.5331		
Kurtosis..................	2.1497		
T:MEAN = 0................	0.6188	Prob > /T/...	0.5361
SIGN RANK................	-8845	Prob > /S/...	0.4263
D:NORMAL..................	0.0581	Prob > /D/...	< .01

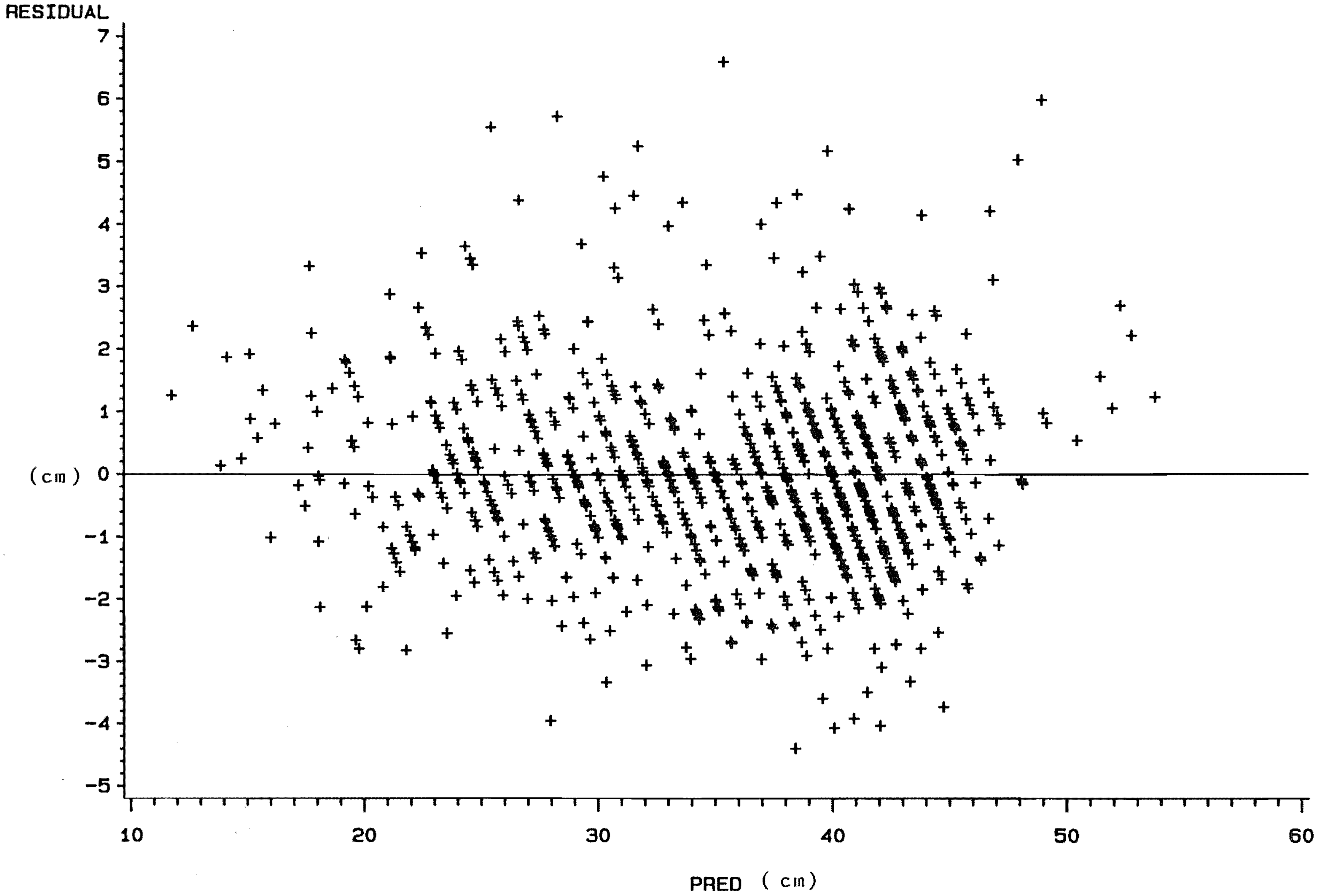


Figure 4.7 Plot of residuals for maximum diameter projection equation

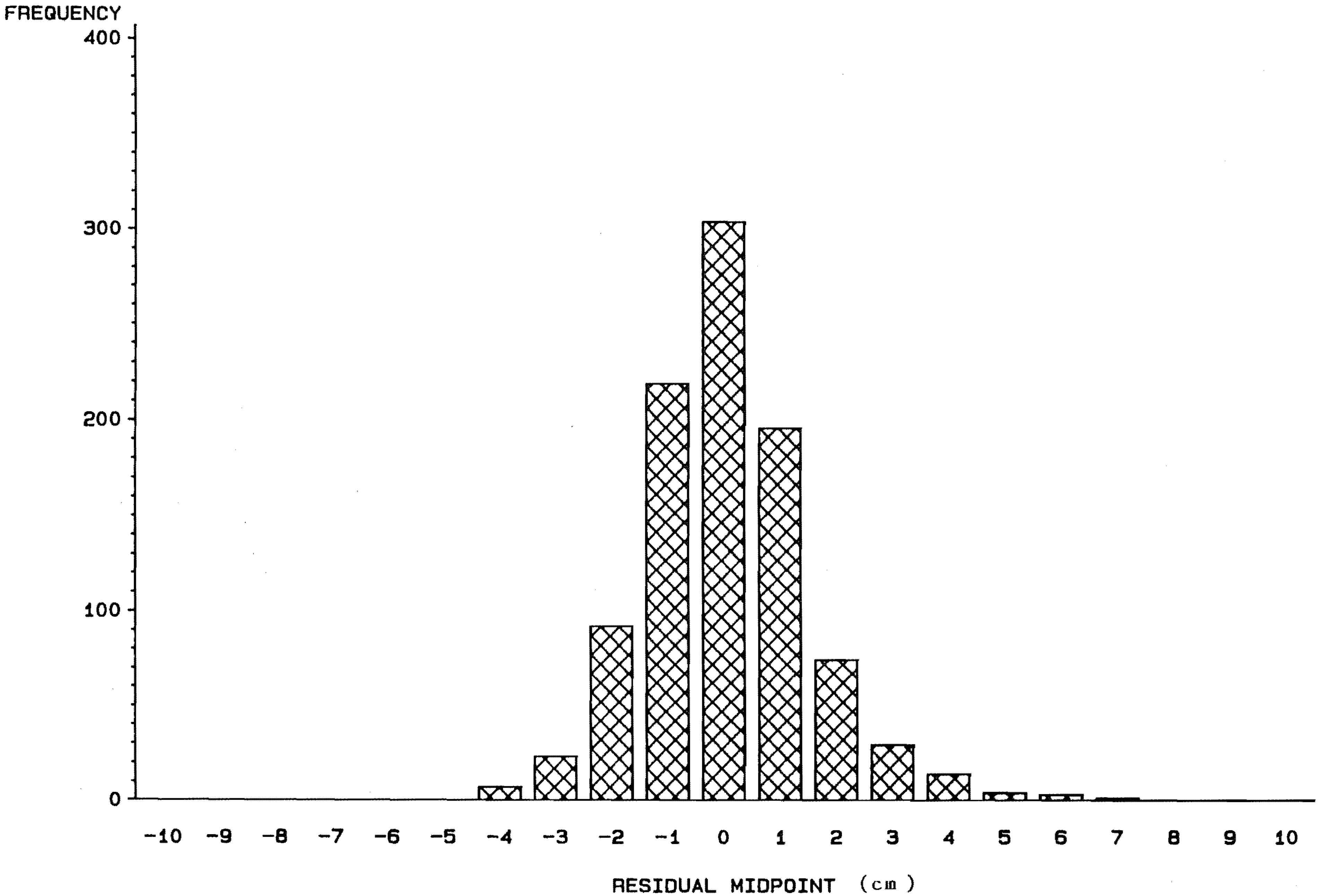


Figure 4.8 Frequency distribution of residuals for maximum diameter projection equation.

4.3 Modelling Mortality

Mortality has almost always been found a difficult variable to model successfully and this data set was no exception, as no single model for all plots could be found. One reason for the difficulty was that only very few data are available on mortality trends since many of the existing stands have either been thinned or badly affected by cyclone or hurricane damage, so that very few measurements could be taken to reflect natural mortality reliably. The data used to derive the mortality model for this study, therefore, used only the data on thinned plots after having ascertained through a comprehensive study of routine continuous forest inventory information that competition per se is not a problem. The difficulty in fitting the models arises from the complicated behaviour of the total data base as a result of the effects of hurricanes and cyclones. Consequently, a new data set structure was created to model mortality. This new data set structure consisted of a modified longest interval (LI) which includes measurements only between intervals where there is a change in stocking, and the last interval. For example, the plot data on stocking on the left (Table 4.10), was converted to the projection set on the right.

This new data structure had more observations than the longest-interval data structure. With a longest-interval data structure no model was appropriate to explain the behaviour of the variable under study, so it was decided that only the observations in the thinned stands would be used with the data structure as in Table 4.10 below. This also assumed that

Table 4.10 Sample measurements of stocking and transformation into projection data format.

Yield format

T	N/ha
5	2000
6	2000
7	2000
8	2000
9	1900
10	1900

Projection format

T_1	T_2	N_1	N_2
5	8	2000	2000
5	9	2000	1900
5	10	2000	1900
9	10	1900	1900

natural competition _per se_ is not a problem. Table 4.11 below describes the final data set used to derive the mortality function.

Table 4.11 Description of the final data set used to model mortality.

VARIABLE	MEAN	STD. DEV.	MAXIMUM	MINIMUM
N_1	679	224	1211	297
N_2	667	222	1161	297
T_1	7.75	2.2	13	5
T_2	10.3	2.9	14	6

The model that was found to describe stand mortality best, a form of the inverse exponential equation, is shown below.

$$N_2 = N_1 \exp(-\gamma (T_2 - T_1)) \qquad (4.6)$$

where,

γ = is a regression parameter obtained from non-linear least squares and;

N_1, N_2, T_1, and T_2 are as defined in Table 3.4.

Table 4.12 shows the results of the ANOVA and the estimates of the parameters and their corresponding standard errors. The graphs of the residuals which lie mainly within ± 35 trees per ha are shown in Figure 4.9 and Figure 4.10. These graphs together with Table 4.13 provide evidence that the equation provides a relatively unbiased precise estimates of net stocking per hectare.

Table 4.12 Parameter summary and ANOVA for mortality equation.

PARAMETER	ESTIMATE	ST. ERROR	
γ	0.00775	0.00139	
SOURCE	**DF**	**SS**	**MS**
Model	1	44150146.4	44150146.4
Error	89	52006.6	584.3
Total	90	44202153.0	

Table 4.13 Summary of characteristics and distribution values for the mortality model.

Mean......................	-2.56		
Standard Deviation........	18.38		
Skewness..................	-0.86		
Kurtosis..................	-0.1726		
T:MEAN = 0................	-1.30	Prob > /T/...	0.1 41
SIGN RANK.................	-121	Prob > /S/...	0.6
D:NORMAL..................	0.25	Prob > /D/...	.0

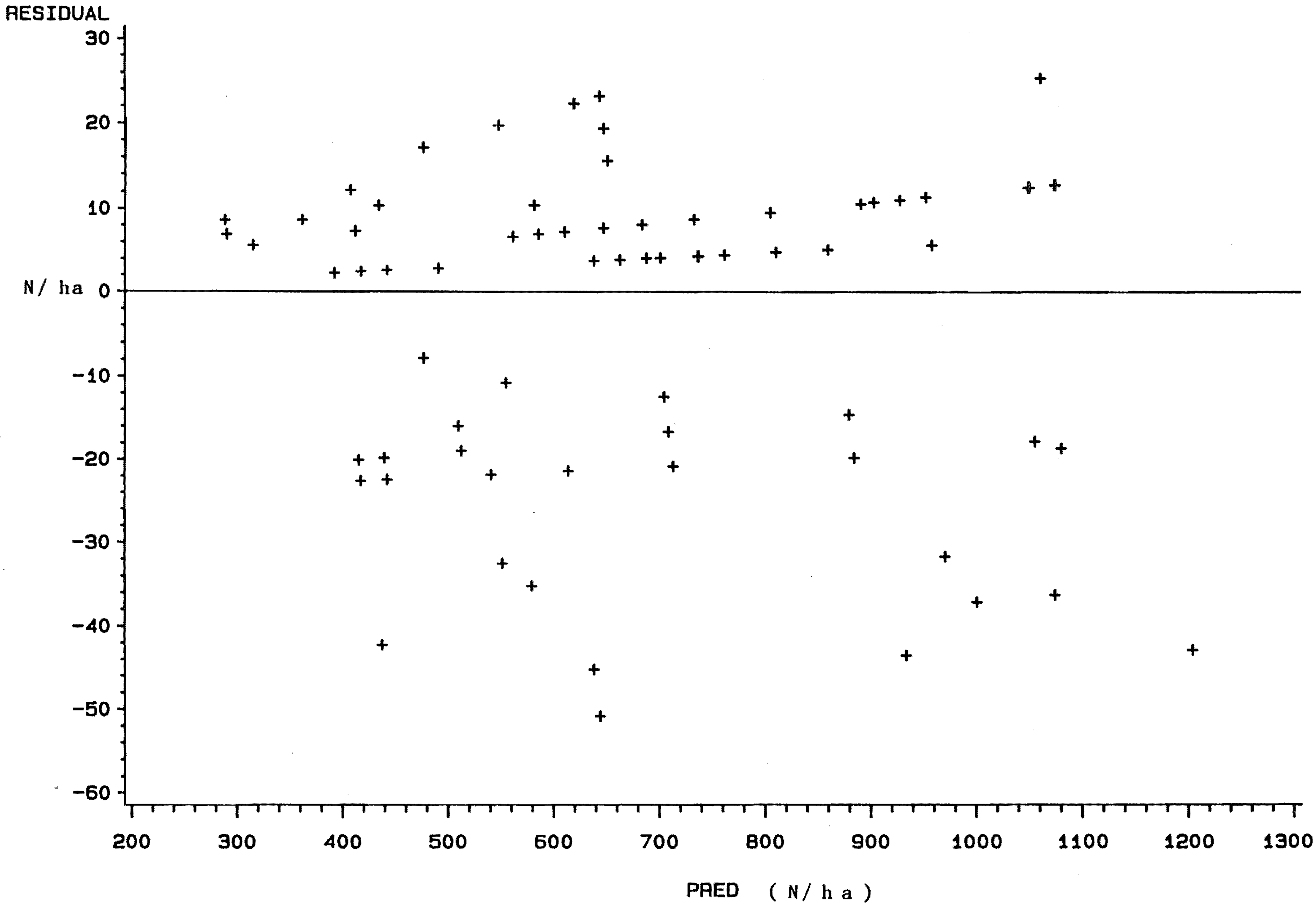


Figure 4.9 Plot of residuals for mortality projection equation.

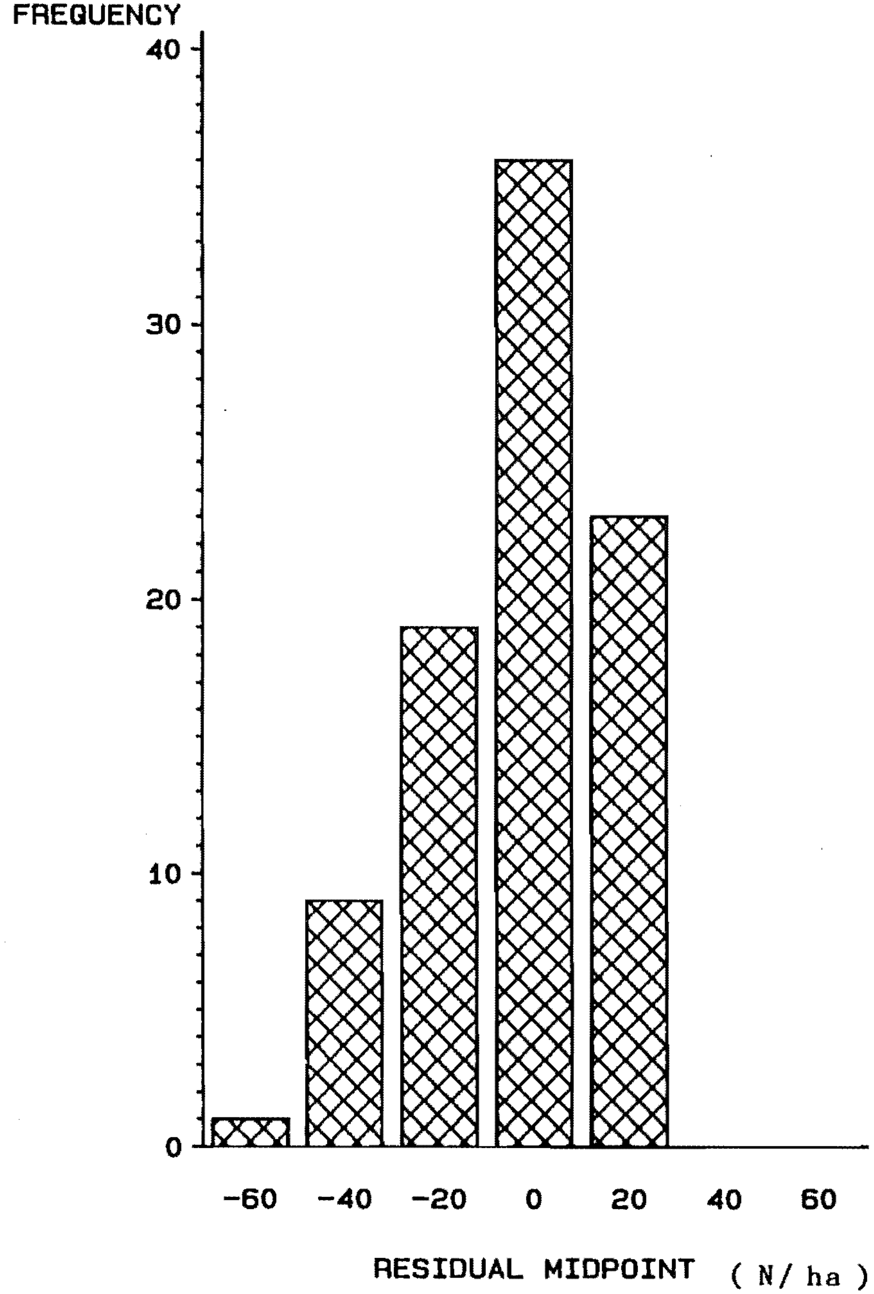


Figure 4.10 Frequency distribution of residuals for mortality projection equation.

Assessment of the Graphical Techniques. The graphical techniques of analyzing residuals from fitting equations to data are easy to invoke but need careful interpretation before deciding which equation should be selected. Graphs helped to detect general trends, outliers and non-normality or deviation from normality of the distribution of errors. Care was observed in assessing residuals for example by standardizing them. Because of the subjectivity that may come in interpreting the charts and graphs, numerical statistics from PROC UNIVARIATE NORMAL were also used. Custom-built frequency distributions of residuals which previous researchers have found very useful were used only with utmost care and in conjunction with the other data generated by PROC UNIVARIATE NORMAL.

4.4 Diameter Distribution Modelling

The estimation of the parameters of the reverse Weibull pdf through the parameter recovery method was done via a computer program in two ways: a) one is available in FORTRAN and b) the other is available in a spreadsheet environment. The stand attributes predicted were used to recover the estimate of the parameters through the method of moments techniques in which the mean diameter and the mean basal area were related to the first two non-central moments of the distribution thus providing at least two equations for the two parameter estimation system.

4.4.1 Diameter class height formula

The existing height equation to determine the average height, (**h**) of a diameter class was of the form (Equation 4.7) recommended by Geiser (1977), as shown below.

$$h = \exp[2.7376 - 8.8562/dbh_{ob} + 0.0460(T)] \qquad (4.7)$$

where,

dbh_{ob} = diameter at breast height outside bark in cm and;

T = age of crop in years.

Various studies reported in Whyte (1987) have shown that this is a robust predictor of height in all forest localities.

4.4.2 Tree volume formulae

The existing equations to derive tree volume outside and inside bark, v_{ob} and vib, for the different localities were used in the form set out below, from Broad (1979).

1) For Lololo/Drasa

$$v_{ob} = 0.00661 + 0.34081(dbh_{ob}/100)^2 h \qquad (4.8a)$$

$$v_{ib} = -0.01173 + 0.28618(dbh_{ob}/100)^2 h \qquad (4.8b)$$

2) For Seaqaqa

$$v_{ob} = 0.01663 + 0.31648(dbh_{ob}/100)^2 h \qquad (4.9a)$$

$$v_{ib} = -0.00854 + 0.26180(dbh_{ob}/100)^2 h \qquad (4.9b)$$

where,

dbh_{ob} = diameter at breast height outside bark in cm and;

h = height in metres.

4.4.3 Tree Taper Equations

In disaggregating the total stem volumes by diameter class into potential log assortments a taper equation was used to solve the volume for any section of the log. The equation derived by Broad (1979) was of the form

$$d' = \sqrt{\left(\frac{V}{0.7854 \times 10^4 h}\right)(\beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \beta_4 X^4 + \beta_5 X^5)} \qquad (4.10)$$

where,

d' = diameter at h';

h = total height of tree in m;

h' = chosen intermediate height within the tree in m;

X = (h - h') / h ;

and the coefficients vary with locality, as set out in Table 4.14 below.

Table 4.14 Coefficients for the taper equations for predicting diameter outside bark for the different localities.

LOCALITY	β_1	β_2	β_3	β_4	β_5
SEAQAQA	1.00990	1.10349	0.00000	0.00000	0.76332
LOLOLO	0.72712	1.84104	0.09104	0.00000	0.00000
DRASA	0.72712	1.84104	0.09104	0.00000	0.00000

This form of taper equation provides a means of estimating reliably the diameter at any point along the length of the stem. The lengths of the logs into which stems are cut as specified is shown in Table 4.15.

Table 4.15 Log assortment classes used.

CLASSES	Sed_{ib} (mm)	Length (m)
Pulpwood1	>355	5.5
Pulpwood2	149 - 355	5.5
Sawlog1	>355	3.3
Sawlog2	149 -355	3.3
Sawlog3	149 - 355	2.2 - 5.49
Chips	70 - 149	2.2 - 5.49
Waste	<70	-

For each log class, volume was then estimated using the following expression.

$$v_s = \frac{v}{h} \int_{l_1}^{l_2} \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \beta_4 X^4 + \beta_5 X^5 \, dl \qquad (4.11)$$

where

v_s = volume of the section of the log in m^3;

v = volume of a tree in the class in m^3 derived from Class volume/Class Frequency;

X = (h - h') / h in m,

h being the total length of log and

h' being the predetermined length of section;

$\beta_1, \ldots \beta_5$ are coefficients that vary in locality as reflected in Table 4.14.

l_1 and l_2 are the distances from the tip of the tree of the small and large ends of the log respectively.

Integration and deduction, i.e. subtracting appropriate sectional volume from total volume was done in such a way that volume by section can be determined.

No new coefficients were solved: this study accepted the existing coefficients for the compatible taper and volume equations as they had proved reliable in regular use (see Whyte, 1987) up until 1986. Thus,

a) given dbh class and estimated total height, an appropriate equation can be used to estimate diameter inside or outside bark at any point along its length;

b) volumes of any specific log lengths can be calculated by integration and;

c) the sum of the volumes of the individual sections always equals the volume predicted by the total tree volume equation.

The resulting disaggregations of class volumes into log assortments, as exemplified in the stand and stock tables in Appendix F, represent information that is very relevant to plantation project planning and managing; they are especially helpful in assisting managers in the kinds of decisions they have to take (Whyte,1989): i.e. the amounts of a steady flow of log types by size classes that should be supplied in the long term to one or more processing plants, either existing or yet to be established, and the total costs and returns involved in maintaining that level of flow.

4.4.4 Assessment and implementation of the model

Assessment of the goodness of fit. Procedures for testing how well the sample data conform to a given distribution have been elaborated by d'Agostino and Stephens (1986). There are two graphical tests. The first is the empirical cumulative distribution function (ecdf) plotting. The ecdf of a random sample $x_1, x_2, \ldots, x_n$ drawn from a distribution with cdf $F(x)$ is defined as

$$F_n(x) = \frac{\#(x_j \leq x)}{n}, \qquad -\infty < x < \infty \tag{4.12}$$

where $\#(x_j \leq x)$ is the number of x_j's less than or equal to x, often called the ecdf. The plot of the ecdf is done on an arithmetic graph plotting paper using $Fn(x)$ as the ordinate and the ith ordered value of the sample of xj as abscissa. To assess how well a particular statistical distribution fits the data, the ecdf of the sample and the cdf of the hypothesized distribution are plotted on the same graph i.e. overlaying one on the other.

The second graphical tool is called probability plotting. The major drawback of the ecdf plot above is the difficulty of judging visually the closeness of the curved ecdf to the cdf curve. The probability plot provides a means of testing the goodness of fit by judging if the set of points deviates from a straight line. The probability plot is a plot of

$$Z_i = G^{-1}(F_n(X_i)) = G^{-1}(p_i) \quad \text{on} \quad X_{(i)} \tag{4.13}$$

where $G^{-1}(.)$ is the inverse transformation of the standardized distribution of the hypothesized distribution and p_i is $(i - 0.5)/n$

Other researchers have used residual analysis and plotting for each diameter class, the residual which is,

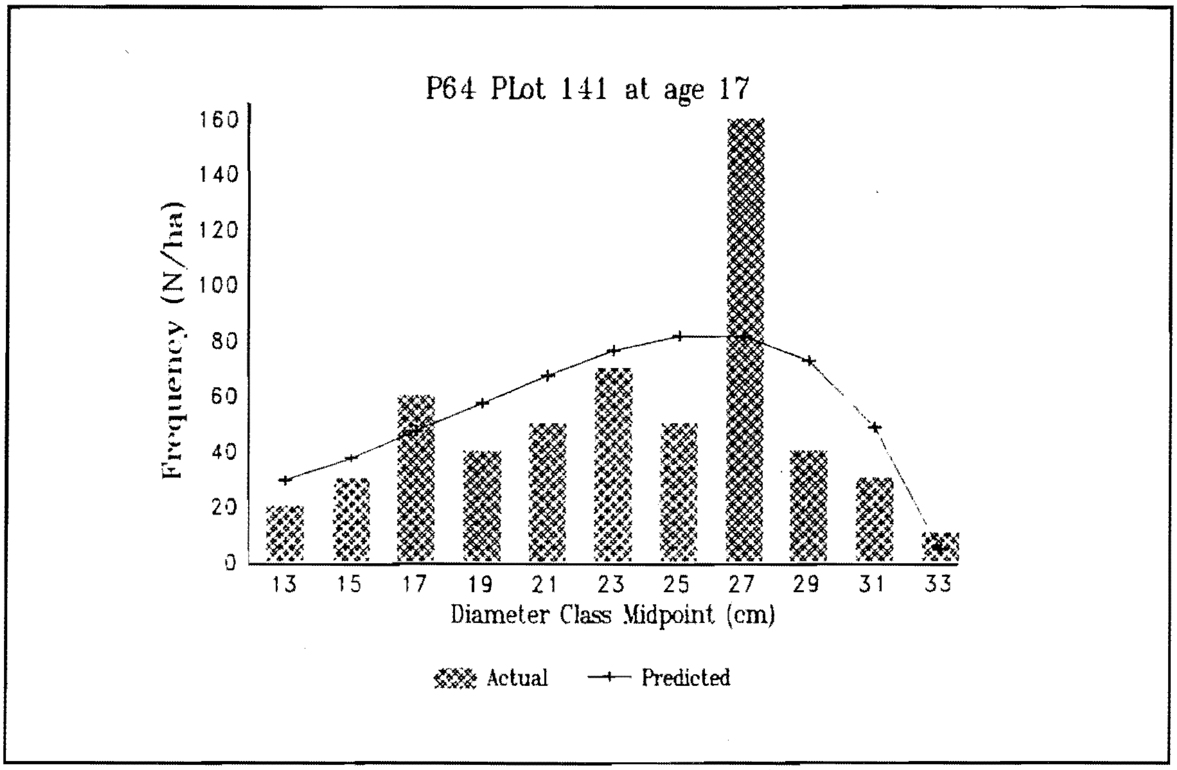
$$\text{Residual} = F(x_i) - F_0(x_i) \tag{4.14}$$

where,

$F(x_i)$ = actual cdf and;

$F_o(x_i)$ = hypothesized cdf.

Application of these procedures in assessing a fitted diameter distribution requires data on diameter of trees in the plot so that $F(x_i)$ can be solved. Such detailed data were not available for this study, and so, the above tests could not be performed on a large data set. However, in adherence to the philosophy that computer simulation experiments and consequent plans developed from them cannot be better than the models they use, the diameter distribution model developed in this study was validated by using a small set of independent data from Cromarty (1981). For this process, plots 57 (stratum P65) and 41 (stratum P64) in Drasa were used. The validation is not comprehensive but was conducted here to illustrate the principle. The results of the validation in Figure 4.11 show how well the model predict diameter distributions at various ages.



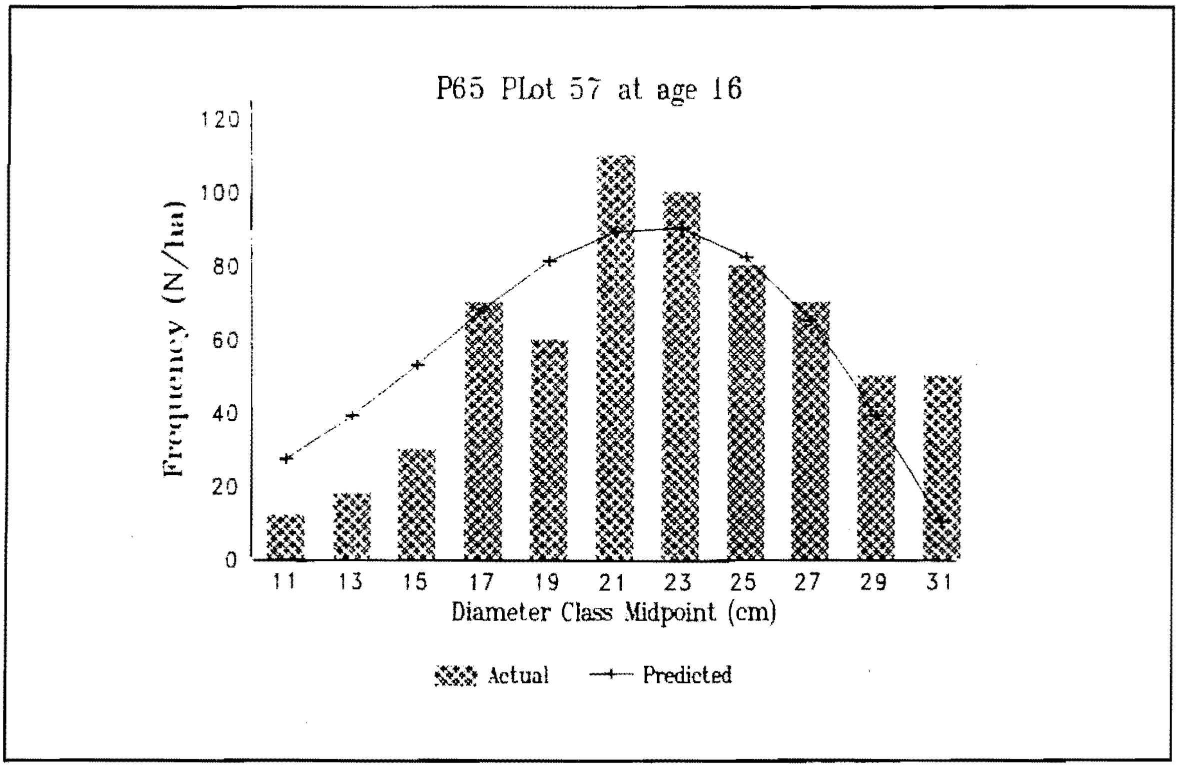


Figure 4.11 Plots of observed diameter distribution and distribution produced by YIELD.

Implementation of the growth and yield model. The model was implemented in such a way that it can:

1) provide the user with optional computational forms depending upon the kind of data available (for example, stand or tree values);
2) read data on trees in the plot and calculate the initial stand statistics from these data;
3) read in stand data provided by the user, the minimum being forest locality, initial age, net basal area per hectare, net stocking, maximum diameter, standard deviation of diameter and minimum diameter (the last simply as a check);
4) project future stand statistics, then output all stand statistics to the desired projection age;
5) calculate the Weibull parameters;
6) disaggregate tree length into log classes for each dbh_{ob} class; and
7) link the projected stand volume to the harvest scheduling model.

Three programming options, one in Vax FORTRAN, one in PC FORTRAN and another in spreadsheet environment were used to implement the diameter distribution growth and yield model. the FORTRAN versions require a relevant compiler while the spreadsheet version requires spreadsheet software. It was discovered that there are advantages in implementing the model in spreadsheet form: e.g. relative ease of use, capability to easily link its output to the harvest schedule model without sophisticated programming and a very convenient user-oriented environment using menus and graphs. Thus, the various simulations and sensitivity tests in the sections that follow were performed in the spreadsheet versions. Implementation and coding of the three versions of the growth simulation had been guided by the flowchart diagram shown in Figure 4.12.

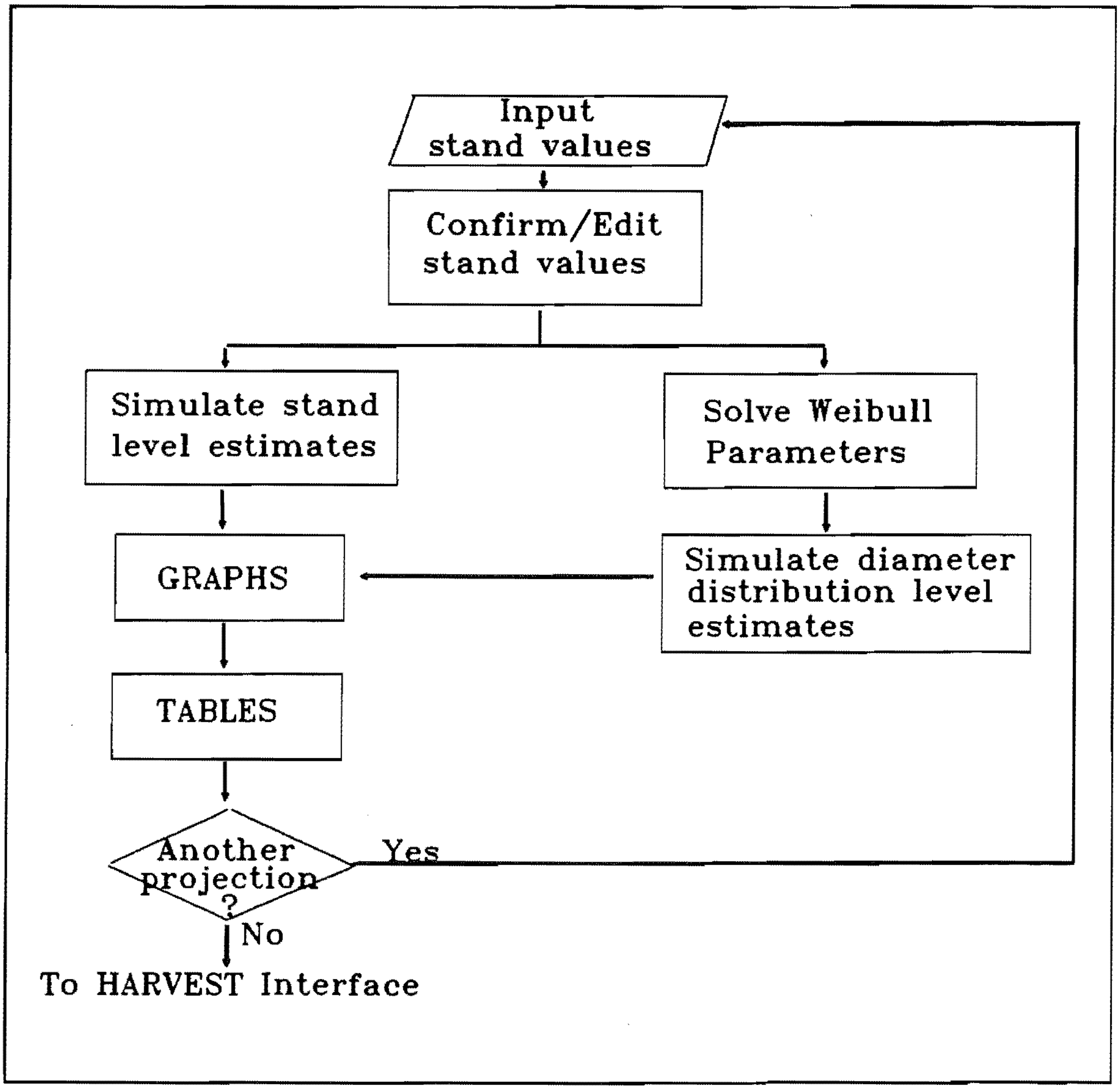


Figure 4.12 Schematic diagram summarizing the programming of YIELD.

4.4.5 Evaluation of the growth and yield model

By plotting the projected values of the stand attributes, it was seen that the projection models were forecasting mathematically and biologically realistic values. In order to test the sensitivity of the model to changes in initial stand condition, experiments with the diameter distribution growth models were conducted. First, the model was run with an assumed initial stand condition. Basal area and stocking were then varied systematically and the model run again so that the effects of these changes on growth could be evaluated. The various experiments conducted for each locality include:

1) initial conditions (base case)
2) stocking increased by 10% (first case)
3) basal area increased by 10% (second case)

Figure 4.13 and 4.14 and Tables 4.16 to Table 4.19 display the result of these experiments. It can be seen from these results that the distribution varied distinctively between the two localities. It can also be seen that the stocking projected by the average stand projection models is always more than the total of trees in the stand and stock projection table. This is not a sign of incompatibility, but simply because trees below 6.0 cm $\mathbf{dbh_{ob}}$ are not included in the stand and stock table projection. The estimate of the location parameter **a**, the maximum diameter at a given age, corresponds to the upper limit of the last diameter class in the stand and stock table as expected.

Table 4.16 **Projected values for stand level at three sensitivity tests [LOCALITY: Lololo].**

Sensitivity Test I

Age	G/ha	Dmean	Dmax	Dstd	N/ha
5*	**15.0***	12.08	**20.00***	**4.00***	**1180***
10	30.6	17.41	28.83	5.29	1177
15	39.4	19.63	33.20	6.47	1173
20	44.9	20.79	35.88	7.49	1170
25	48.6	21.49	37.71	8.32	1166

Sensitivity Test II

Age	G/ha	Dmean	Dmax	Dstd	N/ha
5	15.0	11.44	20.00	4.00	**1300***
10	31.2	16.69	28.57	5.29	1296
15	40.2	18.83	32.79	6.47	1292
20	45.8	19.92	35.37	7.49	1289
25	49.6	20.56	37.12	8.32	1285

Sensitivity Test III

Age	G/ha	Dmean	Dmax	Dstd	N/ha
5	**16.5***	12.73	20.00	4.00	1180
10	32.2	17.90	28.83	5.29	1177
15	40.8	20.02	33.20	6.47	1173
20	46.1	21.12	35.88	7.49	1170
25	49.7	21.71	37.71	8.32	1166

where,
Age = years; **G/ha** = m^2/ha; **Dmean** = cm; **Dmax** = cm; **Dstd** = cm; **N/ha** = stems/ha.

* Values in bold in the first table indicate initial inputs. The subsequent tables have the same initial inputs except N/ha and G/ha in the second and third table respectively.

Table 4.17 **Projected values for stand level at three sensitivity tests [LOCALITY: Seaqaqa].**

Sensitivity Test I

Age	G/ha	Dmean	Dmax	Dstd	N/ha
5*	**15.0***	12.08	**20.00***	**4.00***	**1180***
10	32.3	18.0	33.4	5.29	1177
15	41.8	20.4	40.0	6.47	1173
20	47.5	21.5	43.9	7.49	1170
25	51.3	22.2	46.4	8.32	1166

Sensitivity Test II

Age	G/ha	Dmean	Dmax	Dstd	N/ha
5	15.0	11.44	20.0	4.00	1300*
10	32.9	17.3	33.1	5.29	1296
15	42.6	19.5	39.5	6.47	1292
20	48.4	20.6	43.2	7.49	1289
25	52.2	21.2	45.6	8.32	1285

Sensitivity Test III

Age	G/ha	Dmean	Dmax	Dstd	N/ha
5	**16.5***	12.73	20.00	4.00	1180
10	33.9	18.5	33.4	5.29	1177
15	43.1	20.7	40.0	6.47	1173
20	48.6	21.8	43.9	7.49	1170
25	52.3	22.5	46.4	8.32	1166

where,
Age = years; **G/ha** = m^2/ha; **Dmean** = cm; **Dmax** = cm; **Dstd** = cm; **N/ha** = stems/ha.

* Values in bold in the first table indicate initial inputs. The subsequent tables have the same inputs except N/ha and G/ha in the second and third table respectively.

Table 4.18 **Comparative stand tables derived from projected stand level variables [LOCALITY: Lololo].**

AGE (years)	DIAMETER CLASS (cm)	CLASS FREQUENCIES (N/ha) TEST I	TEST II	TEST III
5	7.5	266	343	221
	12.5	520	574	503
	17.5	321	282	397
	22.5	0	0	0
	∑ N/ha	1107	1199	1121
15	7.5	74	96	68
	12.5	169	209	158
	17.5	283	328	273
	22.5	332	356	336
	27.5	238	227	255
	32.5	45	32	52
	∑ N/ha	1141	1248	1142
25	7.5	70	89	67
	12.5	129	158	125
	17.5	201	236	196
	22.5	256	288	255
	27.5	256	269	261
	32.5	172	161	181
	37.5	29	17	32
	∑ N/ha	1114	1217	1116

Table 4.19 **Comparative stand tables derived from projected stand level variables [LOCALITY: Seaqaqa].**

AGE (years)	DIAMETER CLASS (cm)	CLASS FREQUENCIES (N/ha) TEST I	TEST II	TEST III
5	7.5	265	342	220
	12.5	519	573	502
	17.5	321	282	397
	22.5	0	0	0
	∑ N/ha	1105	1197	1119
15	7.5	72	96	66
	12.5	180	222	170
	17.5	287	328	280
	22.5	302	321	306
	27.5	210	207	221
	32.5	88	78	96
	37.5	13	9	15
	∑ N/ha	1152	1261	1154
25	7.5	73	94	70
	12.5	142	173	138
	17.5	211	244	207
	22.5	243	267	243
	27.5	218	227	221
	32.5	150	147	154
	37.5	72	65	75
	42.5	17	12	18
	∑ N/ha	1126	1229	1126

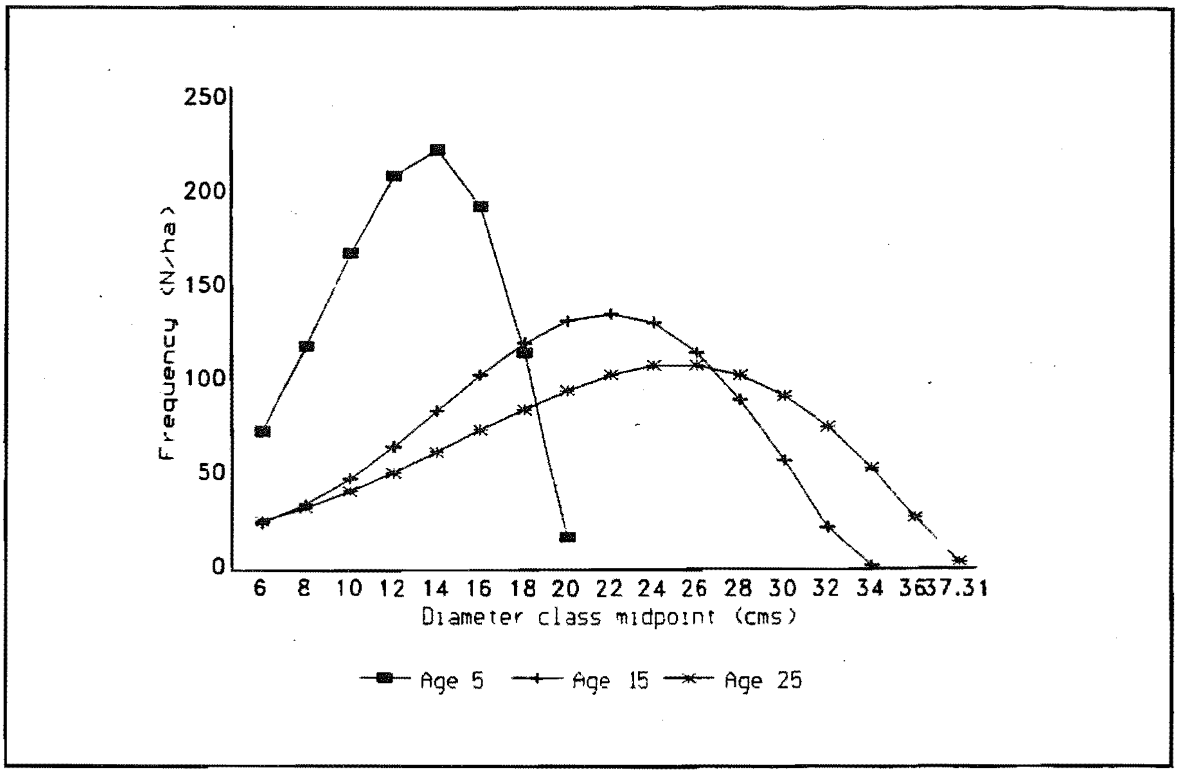


Figure 4.13 **Comparative frequency distributions at different ages.**

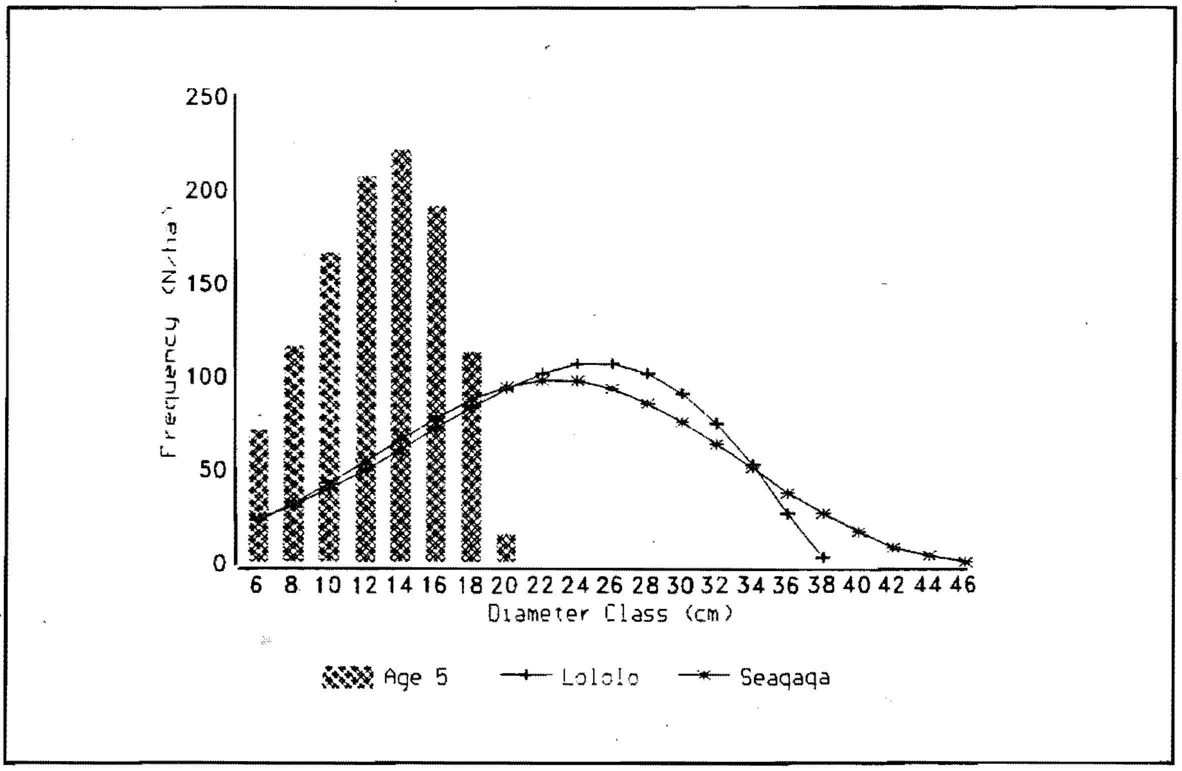


Figure 4.14 **Comparative frequency distributions at projected age in the two different areas.**

4.5 Spreadsheet LP Harvest Scheduling Model

The following discussion relates to the results of the modelling for harvest scheduling with spreadsheet based LP software.

The purpose of developing this particular methodology in this way is to serve the needs of decision makers more readily. In order to assist the user, the kind of system developed consisted of linking a problem generator, a solver, and a report writer together in ways with which decision makers are familiar. The models should also be interactive to allow users, through a visual display terminal, to enter data, manipulate them, derive solutions, retain such solutions for comparison with other formulations with different objectives and to conduct other sensitivity analysis. To be interactive, the model should allow users to view different solutions and permit easy analyses of problem options. Furthermore, the modelling should be flexible in the sense that a base model can be expanded to generate any desired complexity of model through specifying only data tables or functional relationships. There should also be a capability to update and analyze results from the solutions without resorting to complete re-runs of the model.

The model that was developed attempted to explore the use of a spreadsheet interface to facilitate development of large-scale harvesting models and at the same time attain the modelling aims above. Arden (1980), as cited by Sharda (1988), points out that the limitations of linear programming do not lie in the computational algorithm but in the amount of work required to develop the model to prepare large problems and to use the answers. Thus, LP algorithms in different source codes abound and specific problems that need to be addressed today are, in the first instance, relevant to formulating the problem and, secondly, to interpreting and reporting results of user modelling.

The Harvest Scheduling Model. The model that was developed for this study simulated a large worksheet where columns of numbers are summed for harvest planning along the lines of the general mathematical formulation in Chapter 3. In this worksheet, each row and column is identified as a cell with a unique cell address identified by a column letter and row number. In developing this spreadsheet model for harvest planning, cells may contain (1) descriptions, (2) numbers, or (3) formulae. Descriptions, also called **labels**, are the text used to annotate parts of the table like REVENUE, COSTS, VOLUMES, PERIODIC HARVESTS and DISCOUNT FACTORS. The numbers, also called **values**, are the actual data. The **formulae** direct the model to perform calculations; for example, SUM CELLS A5 TO A10. Formulae feature regularly and prominently throughout the spreadsheet model. Wherever possible, formulae were input to drive the LP model rather than raw data: e.g. inclusion of a growth and yield model, a log age-price function and logging methods cost functions in the harvest scheduling model. Since the contents of any cell can be calculated from, or copied to any other cell, a total of one column can be used as a detail item in another column. For example, the formula from a cell in the yield table (not raw discrete data, but formula computed from average yield/ha provided by the yield functions and stand areas provided in the initial stand conditions) can be carried over to the revenue column and used to calculate the revenue from that yield. If data in the initial stand conditions change, in this case stand average yield and area, its total yield changes, which is then automatically copied to the revenue cell so that the total in the revenue column changes subsequently. Thus, a different schedule may eventuate due to any stand or other change. If this had to be carried out with traditional LP modelling, any data change would require recalculating and changing the coefficients and then re-running the model for each time there was a change in the data.

The automatic "ripple" effect allowed this study to create a plan, plug in different assumptions and immediately see the impact on the bottom line. This "what if" capability has made the spreadsheet an indispensable tool for budgets, planning, forecasting, financial statements and many other equation based tasks. Managers have become more and more familiar with their capabilities, and extension into LP is logical. It is more powerful than conventional sensitivity analysis.

Every spreadsheet has the capability of creating a two-dimensional matrix of rows and columns. In order to summarize data, totals from various parts of the spreadsheet can be summed to another part of the spreadsheet. Recent improvements in the capabilities of spreadsheets now allow the model to be built as series of various pages, each page dynamically linked one to the other. Dynamic linking allows data in one spreadsheet file (or in one page) to automatically update another spreadsheet file (or another page of the same spreadsheet).

The harvest scheduling model developed is much like a template, containing as many formulae as possible and as few user data. It contains mostly descriptions and formulae with most cells set to zero. Thus, the starting template of a model contains cells containing zeros which are actually the starting current values of the formula contained by them.

4.6 The Case Study

The model was tested to schedule harvests in an area of FPL plantations. The case study area consisted of 3800 ha of mature timber resource ready for harvesting. The object of modelling was to provide managers with various optional harvest schedules close to the maximum net discounted profits over a planning horizon. This case study adopted a planning horizon of seven years with yearly planning periods.

Outputs of the new growth and yield models and data on the stands were dynamically linked with the other data of the harvest scheduling model. These can easily be modified when more up to date data are available: for example, the age of the crops and, with the availability of better inventory information, initial crop conditions and stand area too.

The LP model can determine an optimum schedule of logs to be harvested and to be brought from several forests over a period of years using various types of logging and delivery to more than one destination near the resource.

4.6.1 Spreadsheet Optimization

There are several computer software packages that can solve linear programming problems. Sharda (1984) made an extensive survey of LP packages which are available for the microcomputer. This type of software is similar to what Kent (1989) referred to as software that actually loads and solves an LP problem. A second category of computer software packages includes those that utilize user input to structure the LP problem and give out the results after the problem has been solved on a separate optimizer. The steps for the latter include data input, matrix generation, solving and report writing as distinct steps. These types are very commonly implemented on mainframe computers and are exemplified in FORPLAN, Timber RAM and MUSYC. The same structures can be assumed in LP packages designed for the microcomputer because features like matrix generators and report writers which are additional users of random access memory (RAM) and auxiliary storage space can now be comparably accommodated on the PC in the manner that until recently only the mainframe could do. As this study implements the model on a microcomputer, the matrix generator and report writer were automatically incorporated so that it was easier to understand and use by managers. Convenient data entry and editing features possessed by spreadsheets make it an attractive alternative to matrix generation programs. The block angular structure of LP harvesting models like **HARVEST** make the development of LP matrices in the spreadsheet easy because the commands of the spreadsheet allow blocks of numbers or relationships to be copied from one location to another. This capability has yet to be taken advantage of widely because of the limitation of the number of columns corresponding to the number of decision variables that may be accommodated in existing software packages. Also the matrix format is a complicated format of LP models. Thus, instead of formulating the model in a matrix format in a spreadsheet, the model was formulated as a set of tables with which managers are familiar. For example, the objective function to be maximized in this formulation is a formula cell that is derived from a succession of different tables containing revenues, logging costs, transport costs, yield and hectares corresponding to each of the decision variables that resulted from the interactions of the variables shown in Figure 3.6.

In this study the timber harvest scheduling problem was modelled and solved by linear programming algorithms integrated with spreadsheets. The microcomputer used was an IBM compatible personal computer based on INTEL 80386 microprocessor. It was equipped with a 2 Megabyte Random Access Memory (RAM) and operating under the Microsoft Disk Operating System (MS-DOS) version 3.3.

The Linear Programming package for spreadsheet optimization used in this study is Beeline[1]. Beeline is a sophisticated package which provides a linear programming interface to many spreadsheet package that are currently in widespread use, including Lotus 1-2-3[2], Quattro Pro[3] and VP

---

[1] Copyright ©Ashley Software Inc., New Zealand.

[2] Copyright ©Lotus Development Corp., U.S.A.

[3] Copyright © Borland International, U.S.A.

Planner[4]. Beeline has built-in worksheet macros that are invoked to define variables, constraints and objective functions relevant to the LP problem.

The sophistication of Beeline as a linear programming package implemented on the spreadsheet does not neglect other extra facilities which may be common to some other LP packages. In using the revised Simplex algorithm, it includes other in-built facilities, some of which can be invoked by the user and others of which are automatically performed when the algorithm solves the problem. These facilities include:

a) detection and removal of redundancies thereby reducing model size and time to solve;
b) ability to allow for any starting solution that may be reasonably close to optimality or for any starting values thereby allowing quick runs even when data have been modified or an objective function changed for a re-solve;
c) use of generalized upper bounds in place of constraints representing a bound on the sum of variables thereby making solution times shorter; and
d) availability of facilities for sensitivity analysis to determine limits (ranges) within which one of the coefficients of the objective function or the right-hand-side coefficients can be changed to predict the numerical effect on the solution.

Other than the added facilities of the algorithm, the

[4] Copyright ©Paperback Software, U.S.A.

'Solve' macro of Beeline was used to make a copy of the worksheet, check the worksheet for errors like non-linearity and solve the LP problem. Solutions and messages are loaded back into the worksheet.

All of the above features are well documented in the Beeline Manual (Ashley Software, 1989).

The spreadsheet program used in this study to formulate **HARVEST** is the latest version of VP Planner 3D. The program is similar to Lotus 1-2-3, with sophisticated windowing, database, graphics, communications, spreadsheet and programming functions. It has one option that allows users to create a spreadsheet on a single page where entries can be identified by columns and rows and another option to create a spreadsheet consisting of various pages where entries are identified by row, column and page, thereby creating a 3-dimensional spreadsheet. The model developed in this study was prepared on a single page spreadsheet but could easily have been produced as a 3-dimensional spreadsheet model. The use of 3-dimensional spreadsheets makes data linkage and organization easier because each cell in the spreadsheet has an X, Y and Z reference. For example, a spreadsheet of net revenue items by planning period uses two dimensions, but net revenue items by period by crop requires three dimensions. While the 3-dimensional spreadsheet is clearly superior for consolidating data, it lacks some of the flexibility since all pages should have essentially the same structure.

Spreadsheet command language or macros and graphics functions were used to develop **HARVEST**, and to interface it with the output of the growth and yield model.

4.6.2 Implementation of **HARVEST**

The model, as implemented in the spreadsheet, was organized into sections each section comprising one or more tables (see Appendix B and C). Spreadsheet optimization is much more convenient to use than the traditional LP, because the usual matrix is dispensed with and familiar tables used instead. Figure 4.15 shows how the model was organized in sections. The logical flow of the model is shown in Figure 4.16. This can be compared with the structure of FORPLAN which is typical of most, if not all, LP package developed for natural resource management applications, that is shown in Figure 4.17. The most distinct difference between the two structures is in the way the inputs and outputs are prepared. In traditional LP's, matrix generators and report writers are essential. In the model developed for this study, there is no distinction between the model output and the model input. Both basically use the

same tables. When the model is run, the solver goes back to the spreadsheet tables which are updated based on the results of the solution.

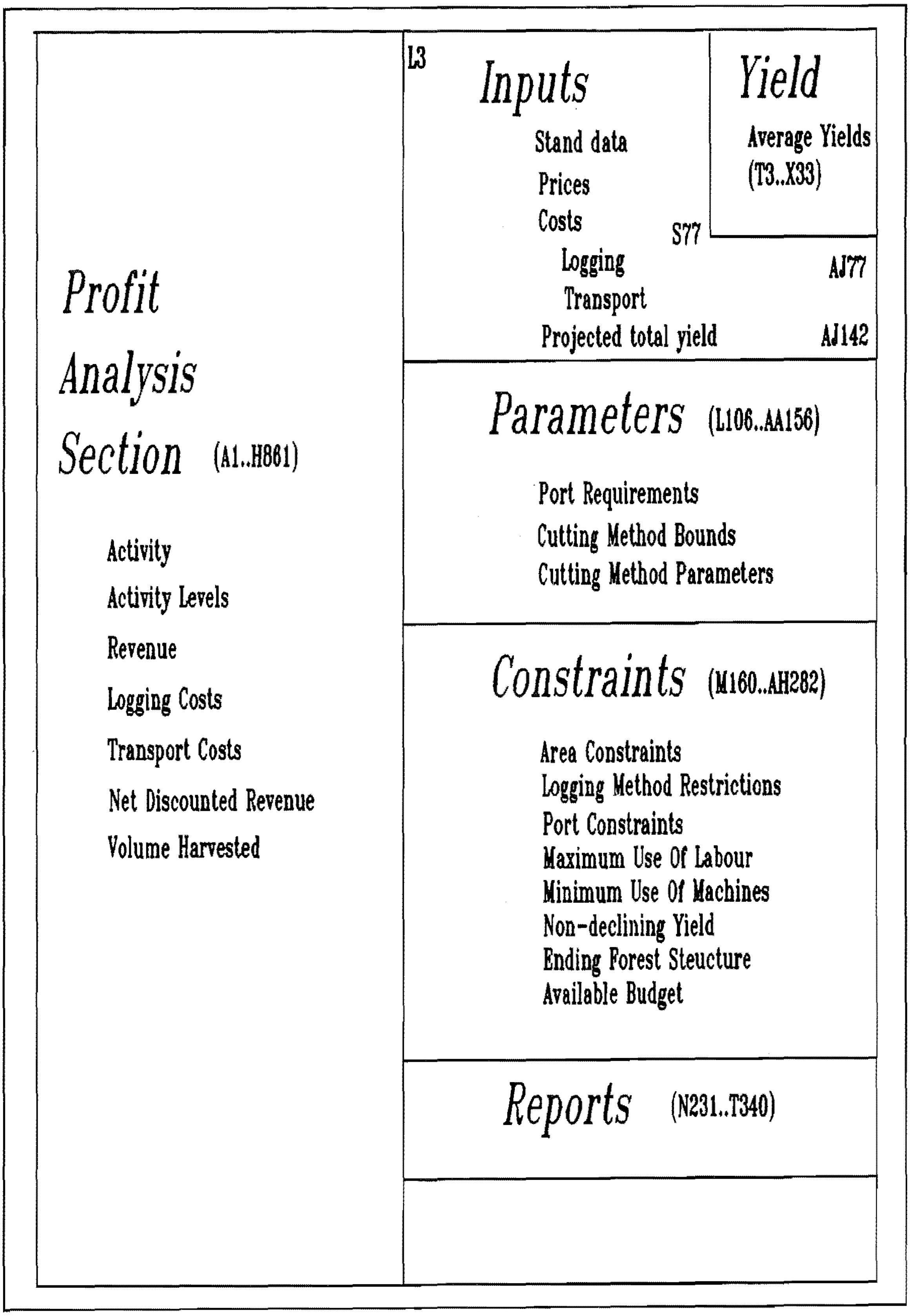


Figure 4.15 Spreadsheet structure of HARVEST.

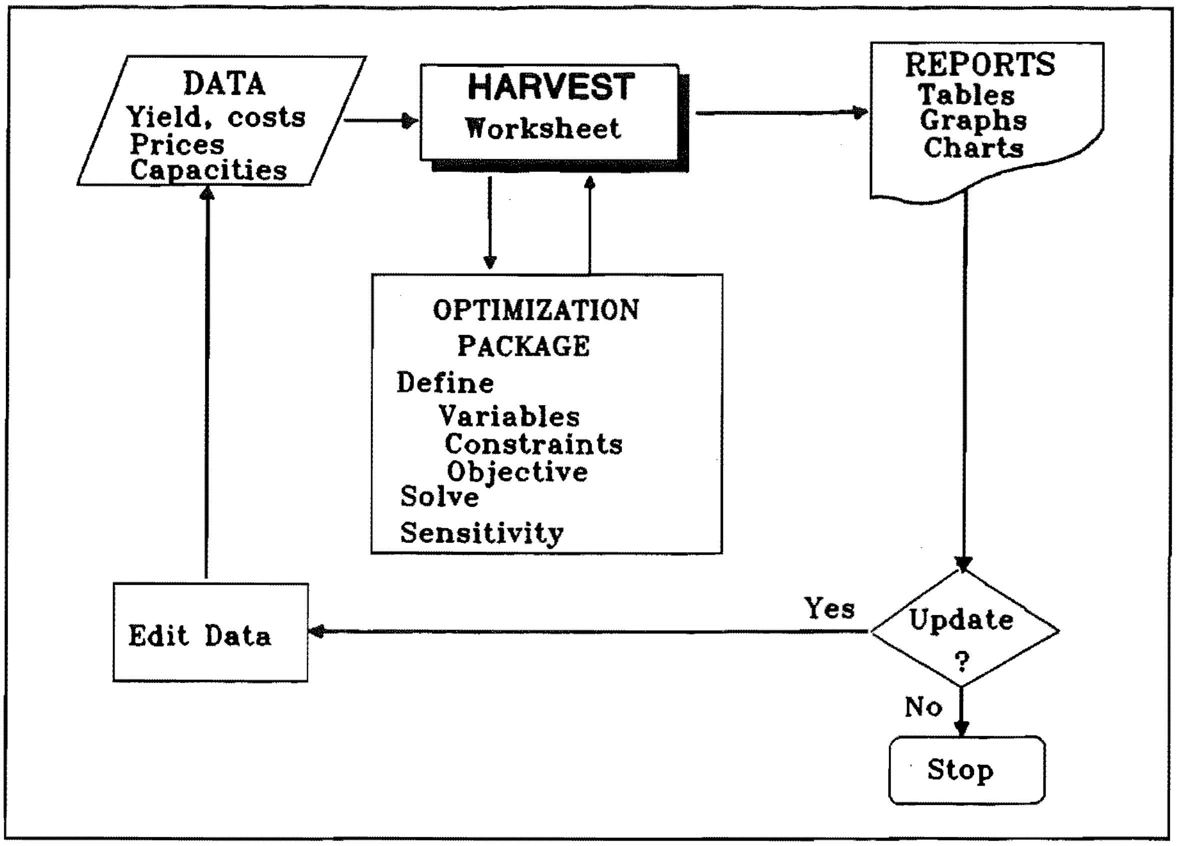


Figure 4.16 Schematic diagram of HARVEST.

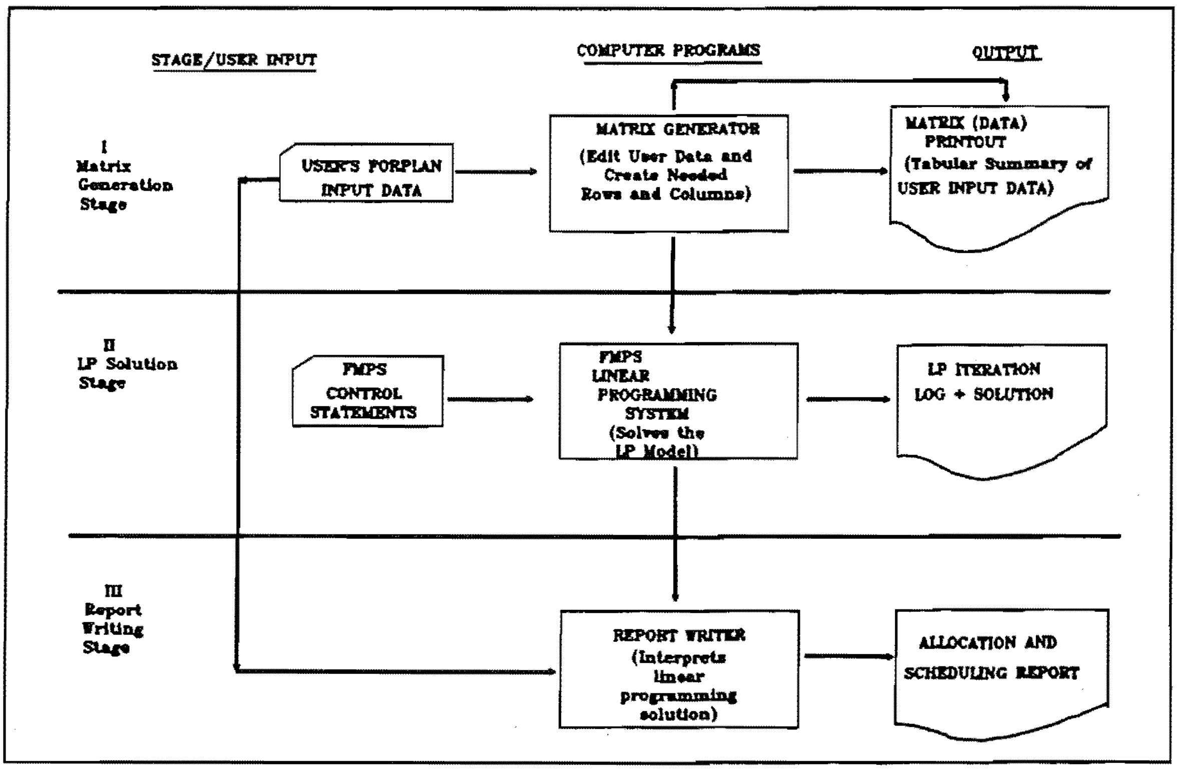


Figure 4.17 System structure of FORPLAN (from Kent, et al., 1991)

The model **HARVEST** was organized into six sections namely:

a) Input/Output section;

b) Profit Analysis Section;

c) Summary Report Section

d) Parameter Section;

e) Constraints Section;

f) Yield Section

g) Macro and Graph Section.

The Input/Output Section is where the data are entered and modified and is the only interactive part prior to the solution of the model. As cells are interrelated, once data are entered or modified in an Input/Output section, all other cells expressed as functions of it are automatically modified or updated. In this section, the user can specify initial crop conditions which the interfaced yield model will use to generate yields in the modelling. The input section is also the section where the parameters of the LP model may be input or changed. These parameters may include discount rates, prices, costs and capacities. The output section provides a summary of the different levels of output that may be generated from the model. The optimal solution provides the results to be generated in forms of tables ready for use and readable to planners.

The Profit Analysis Section (Table 4.20) consists of a table used to calculate for each decision variable the net discounted revenue which would accrue if that decision were undertaken; it is a function of (a) revenue from logs if that decision is undertaken, (b) logging cost, (c) transport cost if

Table 4.20 Condensed initial template of the profit analysis section of HARVEST.

A	B	C	D	E	F	G	H
			HARVEST SCHEDULE AND PROFIT ANALYSIS SECTION				
ACTIVITY	PROP.	AREA	REVENUE	LOGGING COSTS	TRANS. COSTS	NET DIS. REV.	TOTAL HARVEST
X_{1111}	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00
X_{1112}	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00
X_{1121}	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00
.							
.							
X_{2111}	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00
.							
.							
X_{3111}	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00
.							
.							
$X_{15,111}$	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00
.							
.							
$X_{15,742}$	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00
TOTAL		AREA (ha)	REVENUE ($)	LOGGING COSTS ($)	TRANS. COSTS ($)	NET DIS. REV. ($)	TOTAL HARVEST (m^3)

that decision is undertaken and (d) the discount factor. Since there were 840 decision variables in this case study, there were as many net discounted revenues, one for each decision. The sum of these is contained in one cell which can then be defined as the cell to maximize, the objective function. There are also cells which total the volume harvested, the logging costs and the transport costs, again one for each, any of which can also be defined as the cell to maximize or minimize for a chosen objective function. A summary report section is shown in Table 4.21.

Table 4.21 Initial template of summary report section of HARVEST

SUMMARY REPORT

ITEM	TOTAL	LIMIT	BUDGET (SURPLUS/ SHORTFALL)
TOTAL COSTS ($)			
LOGGING COSTS ($)			
TRANSPORT COSTS ($)			
TOTAL REVENUE ($)			
TOTAL NET REVENUE ($)			
TOTAL NET DISCOUNTED REVENUE ($)			
TOTAL VOLUME CUT (m^3)			
TOTAL AREA CUT (ha)			

From Table 4.21, cells that contain expressions as functions in one form or another of the decision variables are labelled below.

LOGGING COSTS appear in the cell containing the formula that sums all logging costs for harvesting crops with all logging methods for the whole planning horizon. This is actually the interpretation of Equation 3.21 and Equation 3.22. This cell picks up data from the yield, logging costs and logging method parameter data bases.

TRANSPORT COSTS are in the cell containing the formula that sums all costs for transporting logs to two ports. This is actually the interpretation of Equation 3.23. This function picks up the cells from the yield and transport cost data base.

TOTAL COSTS are in the cell containing the formula that sums the logging and transport costs.

TOTAL REVENUE is in the cell that contains the formula that sums all the revenues that are generated from cutting the stands scheduled year by year for harvesting throughout the whole planning horizon. This cell picks up data from the log price and the yield data bases.

TOTAL NET REVENUE is in the cell that contains the formula that deducts the total costs from the total revenue.

TOTAL NET DISCOUNTED REVENUE is in the cell that contains the formula that discount the total net revenue realized throughout the whole planning horizon and picks up the cell containing the discount factor.

TOTAL VOLUME CUT is in the cell that contains the formula that computes the total volume harvested from all stands harvested throughout the whole planning horizon.

Any of these cells can be marked as goals to be optimised in the model. The model is designed to proceed solving a goal using the spreadsheet solved with another goal. The algorithm can start with any starting basis, for example a previous solution.

Macros were also written to clear the template to start a new run with a different objective. These macros are invoked before implementing Beeline to solve the new, (re)formulated problem. This is very useful because the model was built to have a capability to generate alternatives by analyzing goals alternately, an essential feature if the model is to be of real use to possible users whose objectives may differ. Also there are many instances when objectives to be analyzed may be changed depending, for example, on the data base or functions which could be considered more reliable by the user. For example, a goal to minimize harvesting costs generated from available projection functions, could be considered in a new formulation. Objective functions that may relate to volume can also be chosen, but these are less desirable than objectives that relate to financial criteria and take into account the cash flows over the planning horizon. Thus, in most instances the runs made are usually maximizing net revenue while secondary objectives that relate to physical quantities like volume are more rightly placed as one of the harvest regulating constraints. But the opportunity is there to alternate them readily.

The macros created in this study increased the ability of the model to make re-runs after any new re-formulation and to incorporate new changes in the other variables like cutting method parameters and prices. This is especially important at the forest level where harvesting plans can change rapidly depending on current environments and on some uncertain elements (the exogenous variables in the **HARVEST** over which the manager has little or no control) such as weather (affecting logging method parameters) demand and sales (affecting prices).

The Constraint Section contains all the tables that are used to restrain the levels of the activities that the model can assume; for example, the use of certain logging methods, cutting of area limiting logs to be brought to ports and regulating the harvest. All these relevant tables are included and detailed in Appendix C. The cells in the tables are functions of the cells containing the value for the decision variables.

The Yield Section contains the formulae that are used to derive yield forecasts for the model. The characteristics of the models used to derive values used in the formulae contained in this section have been described more fully in sections 4.1 to 4.4.

The Summary Report Section contains forms that can be readily assimilated in reports derived from an LP optimizing model for planning the development of industrial plantations in three levels, as outlined in Dargavel (1978). Level 1 reports are a set of reports on information directly obtained from the standard LP output e.g. total net discounted revenue, total costs, total volume harvested, etc. Level 2 reports are used to analyze the schedule to bring out the harvest for each period from the different stands. These can cover areas harvested by each method, volume cut from each stand etc. Level 3 reports are detailed analysis of the distribution of the volume harvested by log assortment classes per period. Because the solution outputs to the problem are also basically the input tables to the problem (stored in spreadsheet form), report generation is easy.

The Parameters Section contains any constant data for a given problem; for example, cutting method and port capacities. There is still flexibility to change these parameters because the locations of the cells can easily be accessed.

The Macro and Graph Section is where the customized macros and the commands to generate options for graphs are contained. Totals can be calculated, results can be presented in graphical form and tables developed that can be processed by word processors.

The logical flow of the model depicted in Figure 4.16 shows the different steps that are taken in using the model. The components include input data for initial stand conditions, updated costs and prices, updated machine capabilities, updated port requirements, updated discount rates, selected objective, update of **HARVEST** worksheet, optimization of the worksheet, sensitivity analysis, generation of reports including tables, graphs.

The steps in the analysis include a description of the present crop resource conditions through **YIELD**. This is accomplished by inputting initial conditions for each stand. Updating of price forecast by a price function can also be done in the input zone. Updating costs can also be done in the input zone. This process then creates an updated worksheet which can be interfaced with the spreadsheet optimizer. Once in the spreadsheet optimizer, appropriate cells are then defined as variables, constraints or an objective. Also in this step, the optimizer can be invoked to solve the current spreadsheet. With a properly formulated spreadsheet model, the optimizer solves the problem and returns the solution to the spreadsheet. While in the spreadsheet, reports can be generated with the normal reporting capabilities of the spreadsheet.

Table 4.22 contains the spreadsheet LP solutions of formulations for three regimes. Run I is the base case with maximization of net discounted revenue as the objective function. Run II is the same formulation except that the objective is maximization of volume while Run III is similar to the base case except that there is an additional constraint of non-declining yield. The table is an aggregate of the information available in the Profit Analysis Section, where individual values for each decision variable are shown.

Table 4.22 Summary reports of three sensitivity runs.

Items(Totals)	SENSITIVITY TESTS		
	RUN I	RUN II	RUN III
Area Harvested (ha)	3802	3802	3802
Vol. Harvested (m^3)	1689000	1821000	1565000
Revenue ($'000)	36770	40439	33262
Logging Costs ($'000)	11041	12468	11000
Transport Costs ($'000)	6717	10653	5754
Net Disc. Rev. ($'000)	13000	10654	12444

Aside from the customized output tables such as Table 4.22, each of the tables prepared on the spreadsheet to run the model can themselves be used as output. This is one advantage of spreadsheet optimization in that the input and output are not distinct and that, unlike traditional LP models where inputs and outputs are distinct, sophisticated matrix generators and report writers are needed to be understood by the planners. A complete analysis excluding spreadsheet preparation should take less than 10 minutes on an IBM compatible 386 machine.

To evaluate changes in policy and data and to demonstrate the capability of the model for quick sensitivity analysis the inputs can be easily changed in the input zone and the model re-run. A macro is written that will enable the user to go directly to the input/output zone by pressing one of the options in the menu. In the case of changing the constraints options in the menu. In the case of changing the constraints the user can go to the constraint section and modify constraints or right hand sides found in that section. Constraints on periodic harvest regulation can have standard formulation, for example, as discussed in Chapter 3. Three possible constraints on periodic harvest regulation were readily interchangeable providing the user more flexibility for testing various options. Each can be made active at a time by pressing the keys that will activate the macro.

Aside from tabular outputs/inputs, the GRAPH component of the model can be used to produce graphs, for example showing periodic volume harvested, area cut and proportion of area cut. Graphs may also be used to compare periodic harvest for different cases. Graphs of this sort are more useful to decision makers than the tables and summary tables in the spreadsheet. Figure 4.18 is an example of a graphical output.

Like many other LP models, the model developed in this study possesses deficiencies which all such models should generally have: for example, a) limitation in accounting for the uncertainties that should be attached to the coefficients of the model; b) limitations with respect to model size; and c) a single criterion of optimality. The influence of changed coefficients can of course be evaluated through sensitivity analysis. The tabular approach here improves the digestibility of impacts of change through producing more compact models that represent relationships in a more succinct way. This also makes the model easier to implement as well understand. Various other criteria can be examined in the objective function because of

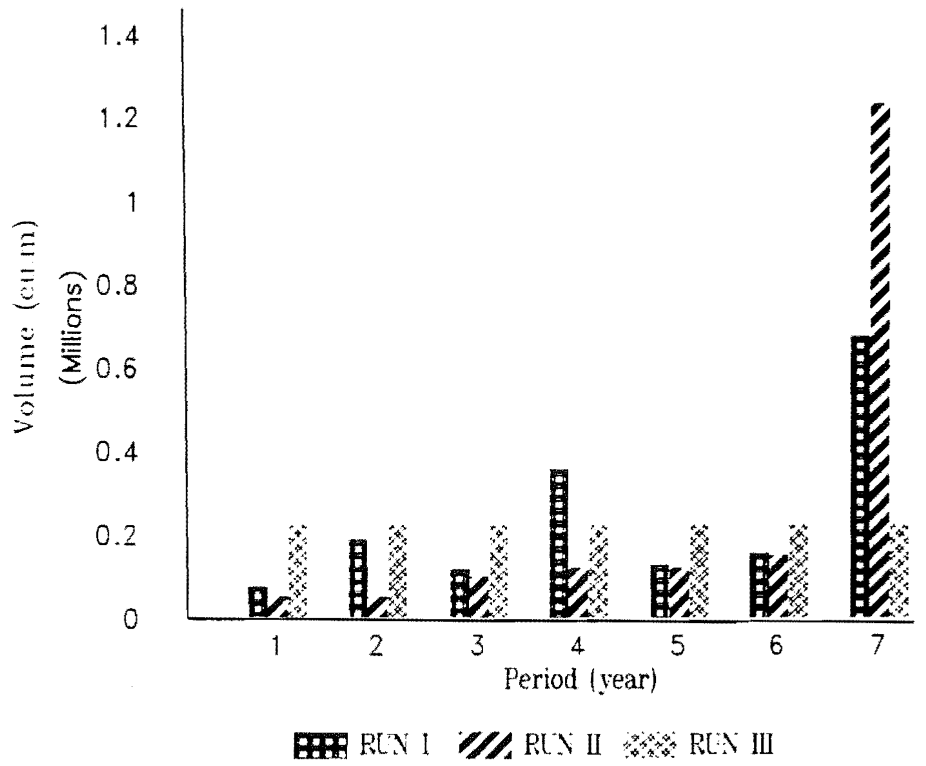


Figure 4.18 Periodic harvest schedules from three sensitivity tests*.

*RUN	VOLUME HARVESTED (m^3)	**NET DISCOUNTED REVENUE ($)**	**OBJECTIVE FUNCTION**
I	1 689 055	12 999 940	Maximize net discounted revenue
II	1 821 473	10 653 788	Maximize volume
III	1 565 550	12 443 886	Max. NDR, non-declining yield

quick re-running of the model. Overall, therefore, the methodology developed in this study provides an improved system of integrated yield forecasting and harvest planning for plantations that improves the quality of decision to the manager and is particularly relevant to plantation management in tropical countries.

CHAPTER 5

SUMMARY,CONCLUSIONS AND RECOMMENDATIONS

5.1 SUMMARY

This two part study to improve medium term planning in pine plantations in the tropics has resulted in the development of a decision support system that has two easily interfaced modular components, one for growth modelling and the other for harvest scheduling. The first part of the study resulted in the development of a growth model with better fits than previous versions through adoption of more recent approaches to modelling and through use of an easier form of interfacing for the user; the second part resulted in the development of a harvest scheduling model using macros and computer spreadsheet packages with which managers are familiar. Finally these two parts were themselves interfaced with each other so that transfers between the two phases of planning could be simply effected.

The study to improve growth and yield prediction produced an integrated growth and yield projection model, providing both a stand and a diameter distribution growth prediction capability for Caribbean pine belonging to Fiji Pine Limited (FPL) in Seaqaqa, Lololo, and Drasa forests. In outlining the methodology, the data sources and validation procedures undertaken to ensure a high degree of reliability in the derived functions are discussed in detail. Implementation of growth and yield modelling in a spreadsheet-driven simulation model is also evaluated.

The stand growth and yield model consisted of equations for projecting stand net basal area per hectare (G/ha), survival (N/ha), maximum diameter outside bark ($Dmax_{ob}$) and standard deviation of diameter outside bark ($Dstd_{ob}$). The equations solved possessed properties desirable for forest growth projection functions, namely that they represent sigmoid shaped curves which have an asymptotic value and which are consistent and path invariant. The mean diameter outside bark ($Dmean_{ob}$) which was also needed to solve the parameters of the distribution was modelled implicitly using the above variables.

The study also assessed the appropriateness of the reverse Weibull probability density function (pdf) as a means of modelling diameter distribution. This pdf was indeed found to be an appropriate distribution to model diameter distribution of the stands for the chosen case study area.

This study, furthermore, demonstrated that the stand equations earlier developed were adequate to solve the parameters of the reverse Weibull probability density function through a combination of the mixed parameter prediction (PPM) and parameter recovery methods (PRM). To solve the location parameter, the study used the PPM technique; to solve the scale and shape parameters, the study used the PRM technique. This approach was found to be effective in ensuring compatibility between projections with the whole stand models and those with the diameter distribution models.

The development of a diameter distribution growth and yield simulation model, called YIELD, on a spreadsheet package with which managers are familiar improved the implementation of the model. Implementing YIELD on a microcomputer spreadsheet provides several advantages and offers unique possibilities not available in other formulations. Simulation runs with YIELD produced output of projected variables that demonstrated the expected mathematical properties and the desired biological relationships that should exist between them. Sensitivity analyses were conducted including simulation runs for various basal areas and levels of stockings. YIELD facilitated the means of making extensive sensitivity analyses on each of the variables that affect yield. For example, an evaluation procedure where stocking and basal area were allowed to vary at an initial age of 5 years, was successfully completed.

Simulation output of the spreadsheet model was also verified against outputs for the same model implemented in FORTRAN for the PC and for the Vax. This step was not essential as the model evaluation, verification and validation conducted should have been enough, but was used nevertheless to provide further reassurance. All versions gave the same outputs for any chosen set of inputs.

Finally for this module, the advantages of spreadsheet simulation which included its interactive capability, its range of simulation options for a variety of purposes and speedy graphical portrayal were all demonstrated.

The second part of the study critically evaluated work undertaken to solve forest regulation problems of harvest scheduling. Harvest scheduling was examined as a complex forest regulation activity involving many data inputs which are difficult to analyze and as a problem that is difficult to solve without computerized optimization techniques. Of the possible tools reviewed, linear programming (LP) was found to be a most appropriate tool to use, but aspects of the previously used harvest scheduling optimization technique needed to be improved, principally matrix generation and report writing.

Thus, the study aimed to develop a harvest scheduling model that attempted to overcome perceived deficiencies in early LP harvest scheduling models or LP in general. The harvest scheduling model developed in this study was easy to prepare, had an output that was easy to interpret and most importantly, was easy to link with other models like the growth and yield model also developed in this study.

The harvest scheduling model was fully validated using a case study area and generated harvest schedules that were nearly identical to the solutions of earlier formulations of the model that used complex matrix generators and report writers. The model also possessed other favourable features such as being interactive and user friendly. The model in its modular form is useful for evaluating various harvesting alternatives based on certain assumptions. The number of experiments conducted with the use of the model showed how it can be a very powerful tool in harvest scheduling analysis.

One of the purposes of this research, to develop a methodology whereby growth and harvest scheduling models could be interfaced, was achieved initially through developing the component models individually then in integrated form, called HARVEST. A case study was used to undertake various analyses on and validation of the model, with acceptable results.

5.2 CONCLUSIONS

From the research undertaken, the following can be concluded.

1. The methodology developed in this study has improved growth and yield modelling in the case study area selected, due to:
   a. availability of advanced computer hardware and software which were instrumental in being able to estimate the parameters of the growth models in non-linear form, through specifically adopting algorithms to solve non-linear equations and so improve solutions of the non-linear growth functions without resorting to transformation which had previously resulted in biased estimates;
   b. adoption of a diameter distribution growth and yield model whereby stand and stock tables are able to be predicted with greater sensitivity;
   c. implementation of the diameter distribution model in various computer environments, thus ensuring its portability and above all, implementing it in an environment that can be linked easily to other planning models;
   d. provision of quick sensitivity analysis to evaluate the performance of the model.
2. The average stand projection models, in algebraic differential equation forms have more desirable features than previously. The parameter estimates

were more precise and the equations predicted present and future stand values more accurately than in previous models.

a. net basal area per hectare was best modelled in projection form through modification of the Schumacher equation;

b. maximum diameter was also best modelled through modification of the Schumacher equation;

c. standard deviation of diameter was a modification of the Gompertz equation;

d. mortality proved difficult to project succesfully and was modelled with a limited degree of success through the use of the inverse exponential;

e. mean basal area and dbh_{ob} per tree were implicitly derived;

f. analysis of residuals revealed normally distributed patterns and also indicated that the equations gave unbiased estimates of the stand statistics, with residuals ranging within ± 8 m^2/ha for the net basal area per hectare projection model, ± 7.5 cm for the maximum diameter projection model, ± 1.25 cm for the standard deviation of diameter projection model and ± 40 trees per hectare for the mortality projection equation.

g. variation of growth due to locality was incorporated where needed through use of dummy variables.

3. The diameter distribution growth and yield model provided an efficient means of describing the stands to a degree that was not provided by previous models:
   a. the model gave accurate estimates of diameter class frequencies over the range of stand ages studied because diameter variables were used to estimate the parameters of the pdf;
   b. the model gave compatible stand average estimates and diameter class estimates because the parameter recovery method was used to estimate the parameters of the pdf;
   c. the model gave a more accurate and precise estimate of the diameter distribution, through specifying the location parameter as a function of the maximum diameter;
   d. the model provided estimates of potential log assortments by diameter classes and by summing over all diameter classes, for the whole stand volume;
   e. the model provided a rapid means of conducting sensitivity analysis.
4. Forest regulation problems can be successfully modelled and solved by linear programming algorithms integrated with spreadsheets. Aside from finding that the methodology reported here is easy to apply, this study revealed other advantages it possessed, including:
   a. easy accommodation of new constraints ;
   b. easy substitution of new objectives;

c. easy updating of input data;

d. no sophisticated matrix generator and report writer needed;

e. easily linked with other planning models;

f. easy to construct because the spreadsheets have built-in functions and macros;

g. intensive and quick sensitivity analysis;

5. The growth and yield model can be interfaced with the harvest scheduling model. Trials with the integrated model showed how efficiently it performed.

6. This study emphasized user-friendliness of planning models aimed at encouraging routine use by managers without the disadvantages of coping with technological complexity.

7. There is further room for refining HARVEST's capabilities and properties. For example, improvement in the graphical portrayals could be extended to provide disaggregation of stems into different log classes. Knowledge based programming could also be used in this regard and this study reviewed its application to facilitate formulation and analysis of harvest scheduling problems.

8. Finally, the models developed in this study provided a means for managers to look routinely and on their own at forest regulation problems more deeply. The models were used to compare optimal solutions of various formulations. This was achieved by examining new objectives, adding new constraints and modifying assumptions. All these have been facilitated by the

new and innovative methodology that was developed in this study. The methodology provided an additional tool by providing a complete and exhaustive analysis of all feasible cutting schedules. The interface provided a quick model update when new growth functions were used. Thus, this study demonstrated the efficacy of the modelling methodology that was aimed at providing insights into some of the problems facing planners and decision makers rather than identifying and prescribing strict optimal solutions. The application of the model to a case study in a pine plantation of Fiji led to recommendations on needed data and research which are discussed in the following section.

5.3 RECOMMENDATIONS

The modelling work and the development of the methodology for the improvement of yield forecasting and harvest scheduling in the case study area in a tropical plantation brought forth a list of data and function requirements which are prerequisites for undertaking the task. The implementation of a diameter distribution growth and yield model, for example, required not only raw data on tree diameter variables, but also measurements on other variables like heights and sectional measurements of trees for tree height functions and taper which can be used in projected disaggregation of number and volume in different diameter classes. This study alluded also to the needed investment for data collection to improve yield prediction in tropical plantations.

In harvest scheduling, improvements can be realized by properly interfacing models and making problem formulation and interpretation of results easier. This aspect of the study has still a considerable potential for improvement: for example, development of methodologies whereby the model is able to accommodate any size of the problem. Studies on the application of knowledge based programming should also be conducted for the computer assisted modelling and analysis of harvest scheduling models.

Recent advances in spreadsheets have led to development of the so called three dimensional spreadsheet. The initial disadvantage of three dimensional spreadsheets, as far as their use in optimization models in spreadsheet form is concerned, is apparently the lack of flexibility. The structure of this type of spreadsheet, however, provides an opportunity to improve LP harvest scheduling in a spreadsheet environment.

Finally, this study recommends that intensive plantation forest management should continue to rely on scientific planning models. For planning models to be truly beneficial, it is necessary that they be accurately and effectively created and efficiently adopted. This study attempted to contribute to these ends.

REFERENCES

Abernethy, N. C. 1981. Predicted and Projected Diameter Distributions of Thinned Old-field Slash Pine. M.Sc. Thesis. Sch. For. Res., Univ. of Georgia.

Alston R. M. and D. C. Iverson. 1987. The Road from Timber RAM to FORPLAN: How Far Have We Travelled?. Journal of Forestry, 85(6): 43-49.

Anscombe, F. J. and J. W. Tukey. 1963. The Examination and Analysis of Residuals. Technometrics, 5(2): 141-159.

Araño R. R. and M. L. Bonita. 1977. Log Allocation and Transport to Processing Centrals in Mindanao. Pterocarpus, 3(1): 55-65.

Ashley Software. 1989. Beeline User Manual. New Zealand. 173 pp.

Bailey, R. L. 1972. Development of Unthinned Stands of **Pinus radiata** in New Zealand. Ph.D. Dissertation. University of Georgia. 67 pp.

Bailey , R. L., L. V. Pienaar, B. D. Shiver and J. W. Rheney. 1982. Stand Structure and Yield of Site-prepared Slash Pine Plantation. The Univ. Ga. Agric. Exp. Stn. Bull. 291. 83 pp.

Bailey, R. L., N. C. Abernathy, and E. P. Jones Jr. 1981. Diameter Distribution Models for Repeatedly Thinned Slash Pine Plantations . In Barnett, J. B. (Ed.). Proceedings of the First Biennial Southern Silvicultural Research Conference. U.S.D.A. Forest Service Tech. Rep. SO-34. pp. 115-126.

Bailey, R. L. and T. R. Dell 1973. Quantifying Diameter Distributions with the Weibull Functions. Forest Science, 19(2): 97-104.

Baldwin, Jr. V. C., and D. P. Feduccia. 1988. Loblolly Pine Growth and Yield Prediction for Managed West Gulf Plantations. USDA Forest Service Res. Pap. SO-236. 27 pp.

Bates D. M. and D. G. Watts. 1988. Nonlinear Regression Analysis and Its Applications. John Wiley and Sons, Inc. U.S.A. 364 pp.

Bennett, F. A. and J. L. Clutter.1968. Multiple-product Yield Estimates for Unthinned Slash Pine Plantations - Pulpwood, Sawtimber Gum. USDA Forest Service Res. Pap. SE-35. 21 pp.

Bliss, C. I. and K. A. Reinker. 1964. A Lognormal Approach to Diameter Distribution in Even-aged Stands. Forest Science, 10(3): 350-360.

Bodily, S. E. 1986. Spreadsheet Modelling as a Stepping Stone. Interfaces, 16(5): 34-52.

Borders, B. E., R. L. Bailey and M. L. Clutter, 1988. Forest Growth Models: Parameter Estimation Using Real Growth Series. In. Ek, A. R. et al. (Eds.). U.S.D.A. For. Serv., Gen. Tech. Rep. NC-120.

Borders, B. E. 1989. Systems of Equations in Forest Stands Modelling. Forest Science, 35(2): 548-556.

Borders, B. E. and W. D. Patterson. 1990. Projecting Stand Tables: A Comparison of the Weibull Diameter Distribution Method, a Percentile-Based Projection Method, and a Basal Area Growth Projection Method. Forest Science, 36(2): 413-424.

Broad, L. R. 1978. A Revised Yield Forecasting Model for **Pinus Caribaea** in Fiji. M.Sc. Forestry Report. School of Forestry, University of Canterbury. 56 pp.

Broad, L. R. 1985. A Mixed Integer Linear Programming Approach to Forest Utilization Management Problem. Ph.D. Thesis. Univ. of Canterbury. 230 pp.

Bruce, D. and L. C. Wensel. 1987. Modelling Forest Growth: Applications, definitions and Problems. In Ek, A. R., S. R. Stephen and T. E. Burk (Eds.) Proceedings IUFRO Conference on Forest Growth Modelling and Prediction. 23-27 August, 1987. Minneapolis, MN. pp. 1-8.

Buckman, R. E. 1962. Growth and Yield for Red Pine in Minnesota. U.S.D.A. For. Serv. Tech. Bull. 1272. 50 pp.

Burk, T. E. and J. D. Newberry. 1984. A Simple Algorithm for Moment-Based Recovery of Weibull Distribution Parameters. Forest Science, 30(2): 329-332.

Burkhart, H.E. and P.T. Sprinz. 1984. Compatible Cubic Volume and Basal Area Equations for Thinned Old Loblolly Pine Plantations. Forest Science, 30(1): 86-93.

Cao, Q. V. and Burkhart, H. E. 1984. A Segmented Distribution Approach for Modelling Diameter Frequency Data. Forest Science, 30(3): 129-137.

Cao, Q. V., H. E. Burkhart, and R. C. Lemin, Jr. 1982. Diameter Distributions and Yields of Thinned Loblolly Pine Plantations. VPI and SU Sch. For. and Wildl. Resourc. Publ. FWS-1-82. 62 pp.

Clutter, J. L. 1963. Compatible Growth and Yield Models for Loblolly Pine. Forest Science 9(3): 354-371.

Clutter, J. L., J. C. Fortson, L. V. Pienaar, G. H. Brister and R. L. Bailey. 1983. Timber Management: A Quantitative Approach. John Wiley and Sons. 333 pp.

Clutter, J. L. and D. M. Belcher. 1978. Yield of Site Prepared Slash Pine Plantations in the Lower Coastal Plain of Georgia and Florida. pp. 53-70. In Fries, H.E., H.E. Burkhart and T. A. Max (Eds.). Growth Models for Long Term Forecasting of Timber Yields. VPI and SU, School of Forestry and Wildlife Resources Publication. FWS-1-78. 249 pp.

Clutter, J. L. and F. A. Bennett. 1965. Diameter Distribution in Old-field Slash Pine Plantations. Ga. For. Res. Counc. Rep. 13. 9 pp.

Clutter, J. L. and E.P. Jones. 1980. Prediction of Growth After Thinning in old slash pine plantations. USDA For. Serv. Res. Paper SE-127. 14 pp.

Covington, W. W., D. B. Wood, D. L. Young, D. P. Dykstra, and L. D. Garrett. 1988. TEAMS: A Decision Support System for Multiresource Management. Journal of Forestry, 86(8): 25-33.

Coyle, R. G. 1977. Management System Dynamics. John Wiley and Sons, Ltd. 463 pp.

Cornwell, L. W. and D. T. Modianos. 1990. Management Tool: Using Spreadsheets for Simulation Models. Production and Inventory Management Journal, (Third Quarter, 1990): 8-17.

Cromarty, P. L. 1981. Validation of Fiji Pine Commission Yield Forecasting Model. B. Sc. Forestry Dissertation. School of Forestry. Univ. of Canterbury.

Curtis, F. H. 1962. Linear Programming the Management of a Forest Property. Journal of Forestry, 60: 611-616.

D'Agostino, R. B. and M. A. Stephens. 1986. Goodness of Fit Techniques. Marcel Decker Inc. N.Y. 548 pp.

Daellenbach, H. G., J. A. George and D. C. McNickle. 1983. Introduction to Operations Research Technique. 2nd ed. Allyn and Bacon, Inc., 711 pp.

Daniels, R. F., H. E. Burkhart and M. R. Strub. 1979. Yield Estimates for Loblolly Pine Plantations. Journal of Forestry, 77(9): 581-583, 586.

Dargavel, J. B. 1978. A Model for Planning the Development of Industrial Plantations. Australian Forestry, 41(2): 95-107.

Davis, K. P. 1964. The Importance and Uses of Growth Information in Timber Management. Journal of Forestry, 62(7): 490-492.

Davis, L. S. and Johnson, K. N. 1987. Forest Management (3rd ed.). McGraw-Hill Book Co. U.S.A. 790 pp.

de Kluyver, C. A. and A. G. D. Whyte. 1980. Forest Harvest Scheduling in Fiji: A Comparative Case Study of Linear Programming, Heuristic and Decomposition Techniques. NZOR, 8(1): 33-72.

Dell, T. R., D. P. Feduccia, W. F. Mann, Jr., and B. H. Polmer. 1979. Yields of Unthinned slash Pine Plantations on Cut-over Sites in the Westgulf Region. USDA Forest Service Res. Paper SO-147. 84 pp.

Ek, A. R. , J. N. Issos, and R. L. Bailey. 1975. Solving for Weibull Diameter Distribution Parameters to Obtain Specified Mean Diameters. Forest Science, 21(3): 290-292.

Eng, G. 1982. A Methodology for Forest Outturn Assessment and Optimal Tree Bucking. M. Sc. Thesis. University of Canterbury.

Feduccia, D. P. , T. R. Dell , W. F. Mann, Jr. and B. H. Polmer. 1979. Yields of Loblolly Pine Plantations on Cutover Sites in the Westgulf Region. U.S.D.A. Forest Service Res. Pap. SO-148. 88 pp.

Field, R. C. 1984. National Forest Planning is Promoting U.S. Forest Service Acceptance of OR. Interfaces, 14(5): 67-76.

Fisher, M. L. 1985. Interactive Optimization. Annals of Operations Research. 5(1-4): 541-556.

Frazier, J. R. 1981. Compatible Whole-Stand and Diameter Distribution Models for Loblolly Pine Plantations. Ph.D. Thesis. Virginia Polytechnic Institute and State Univ. 125 pp.

Garcia, O. M. 1981a. IFS, An Interactive Forest Simulator for Long Range Planning. New Zealand Journal of Forestry Science, 11: 8-22.

Garcia, O. M. 1981b. Simplified Method-of-Moment Estimation for the Weibull Distribution. New Zealand Journal of Forestry Science, 11(3): 304-306.

Garcia, O. M. 1984. FOLPI, A Forestry-Oriented Linear Programming Interpreter. In Nagumo, H. et al. (eds.). Proceedings IUFRO Symposium on Forest Management Planning and Managerial Economics. University of Tokyo. pp. 293-305.

Geiser, M. 1977. A Yield Forecasting Model for Pinus caribaea in Fiji. M. For. Sci. Report. University of Canterbury. 66 pp.

Greenberg, H. J. 1987. ANALYZE: A Computer-Assisted Analysis System for Linear Programming Models. Operations Research Letters, 6(5): 349-255.

Greenberg, H. J., C. Lucas and G. Mitra. 1987. Computer Assisted Modelling and Analysis of Linear Programming Problems: Towards a Unified Framework. IMA Journal of Mathematics in Management, 1987(1): 251-266.

Hafley, W. L. and H. T. Schreuder. 1977. Statistical Distributions for Fitting Diameter and Height Data in Even Aged Stands. Canadian Journal of Forestry Research, 7(3): 481-487.

Ho, J. K. 1987. OPTIMACROS, Optimization with Spreadsheet Macros. Operations Research Letters, 6(2): 99-103.

Hoganson, H. M. and D. W. Rose. 1984. A Simulation Approach for Optimal Timber Management Scheduling. Forest Science, 30(1): 220-238.

Hyink, D. M. 1980. Diameter Distribution Approaches to Growth and Yield Modelling. In Brown, K. M. and F. D. Clarke (Eds.). Forecasting Forest Stand Dynamics. School of Forestry, Lakehead University, Thunderbay, Ontario. 138-163.

Hyink, D. M. and J. W. Moser, Jr. 1983. A Generalized Framework for Projecting Forest Yield and Stand Structure Using Diameter Distributions. Forest Science, 29(1): 85-95.

Jamnick, M. S. 1990. A Comparison of FORMAN and Linear Programming Approaches to Timber Harvest Scheduling. Canadian Journal of Forestry Research, 20(9): 1351-1360.

Johnson, K. N. and P. L. Tedder. 1983. Linear Programming vs. Binary Search in Periodic Harvest Level Calculation. Forest Science, 29(3): 569-581.

Jones, C. and J. Carmona. 1987. A New Species of Modelling System. IMA Journal of Mathematics in Management. 1987(1): 225-236.

Kent, B. M. 1989. Forest Service Land Management Planners Introduction to Linear Programming. U.S.D.A. For. Serv. Gen. Tech. Rep. RM-173. 36 pp.

Kent, B., B. B. Bare, R. C. Field and G. Bradley. 1991. Natural Resource Land Management Planning Using Large-Scale Linear Programs: The U.S.D.A. Forest Service Experience with FORPLAN. Operations Research, 39 (1): 13-27.

Kidd, W. E. Jr, E. F. Thompson, and P. H. Hoepner. 1966. Forest Regulation by Linear Programming - A Case Study. Journal of Forestry, 64(9): 611-613.

Kowero, G. S. and Dykstra, D. P. 1988. Improving Long-Term Management Plans for a Forest Plantation in Tanzania, Using Linear Programming. Forest Ecology and Management, 24(1988): 203-216.

Kuru, G. A. 1989. A Diameter Distribution Growth and Yield Model for Naturally Regenerated and Seeded Stands of Radiata Pine in Kaingaroa Forest. M.Sc. Thesis. University of Canterbury. 102 pp.

Lamb, A. F. A. 1973. **Pinus Caribaea** Fast Growing Trees in Lowland Tropics No. 6. Department of Forestry. Oxford.

Leak, W. B. 1964. Estimating Maximum Allowable Timber Yields by Linear Programming. U.S. Forest Service Res. Pap. NE-17, 9pp. NE Forest Exp. Sta. Upper Dorby, Penn.

Leefers, L. A. 1991. Incorporating Linear Programming and Monte Carlo Simulation in a Spreadsheet-Based Harvest Scheduling Model. In Buford, M. A. (Compiler). Proceedings of the 1991 Symposium on Systems Analysis in Forest Resources. March 3-6, 1991. Charleston, South Carolina, U.S.A.

Lohrey, R.E. and R. L. Bailey. 1976. Yield Tables and Stand Structures for Unthinned Longleaf Plantations in Louisiana and Texas. USDA Forest Service Res. Pap. SO-133. 53 pp.

Loucks, D. P. 1964. The Development of an Optimal Program for Sustained Yield Management. Journal of Forestry, 62(7): 485-490.

Ma, P., F. H. Murphy and E. A. Stohr. 1989. Representing Knowledge About Linear Programming Formulation. Annals of Operations Research, 21(1-4): 149-172.

MacKinney, A. L. and L. E. Chaiken. 1939. Volume, Yield and Growth of Loblolly Pine in the Mid-Atlantic Coastal Region. U.S. For. Serv. Appalachian For. Exp. Stn. Tech. Note 33. 30 pp.

Manley, B. R. 1977. Yield Projection of **Pinus caribaea var. Hondurensis** at Drasa, Fiji Using the Weibull Function. FRI Report No. 52. Unpublished.

Manley, B. R. and J. A. Threadgill. 1991. LP Used for Valuation and Planning of New Zealand Plantation Forests. Interfaces, 21(6): 66-79.

Matney, T. G. and A. D. Sullivan. 1982. Compatible Stand and Stock Tables for Thinned and Unthinned Loblolly Pine Stands. Forest Science, 28(28): 161-171.

McGee, C. E. and L. Della-Bianca. 1967. Diameter Distributions in Natural Yellow-Poplar Stands. USDA Forest Service Res. Pap. SE-25. 7 pp.

McGuigan, B. J. 1992. REGRAM: A Flexible Resource Management Optimizer. In Whyte, A. G. D. (Ed.). Proceedings IUFRO S3.04.01 Conference. 27-31, January, 1992. Christchurch.

McTague, J. P. and R. L. Bailey. 1987. Compatible Basal Area and Diameter Distribution for Loblolly Pine Plantations in Santa Catarina, Brazil. Forest Science, 33(1): 43-51.

Meyer, H. A. 1952. Structure, Growth and Strain in Balanced Uneven-aged Forests. Journal of Forestry, 50(2): 85-92.

Munro, D. D. 1974. Forest Growth Models- A Prognosis. In Fries, J. (Ed.). Growth Models for Tree and Stand Simulation. Royal College of Forestry, Stockholm, Sweden.

Nautiyal, J. C. and P. H. Pearse. 1967. Optimizing the Conversion to Sustained Yield - A Programming Solution. Forest Science, 13(2): 131-139.

Navon, D. I. 1971. TimberRAM, A Long Range Planning Method for Commercial Timberlands Under Multiple Use Management. U.S.D.A. For. Serv. Res. Pap. PSW-70.

Nelson, T. C. 1964. Diameter Distribution and Growth of Loblolly Pine. Forest Science, 10(1): 105-115.

Newby, M. J. 1980. The properties of Moment Estimators for the Weibull Distribution Based on Sample Coefficient of Variation. Technometrics, 22(2): 187-194.

Ogweno, D. C. O. 1988. Optimum Life of Production Assets in Short and Medium Term Timber Harvesting Projects. M.Sc. Thesis. University of British Columbia. 98 pp.

Parekh, R. 1990. Capacity/Inventory Planning Using a Spreadsheet. Production and Inventory Management Journal, (First Quarter, 1990): 1-3.

Patel, N. 1985. Forest Inventory System Manual. Fiji Pine Commission Manual. Lautoka, Fiji. 108 pp.

Pinder, J. E., J. G. Weiner and M. H. Smith. 1978. The Weibull Distribution : A New Method of Summarizing Survivorship Data. Ecology, 59: 175-179.

Pienaar, L. V. 1979. An Approximation of Basal Area Growth after Thinning Based on Growth in Unthinned Plantations. Forest Science, 25(2): 223-232.

Reid, S. B. 1986. Revisions of Equations to Forecast Yields of Caribbean pine in Lololo Forest, Fiji. B. For. Sci. Dissertation. School of Forestry. Univ. of Canterbury. 72 pp.

Reynolds, M. R. Jr., T. E. Burk and W. Huang. 1988. Goodness-of-Fit Tests and Model Selection Procedures for Diameter Distribution Models. Forest Science, 34(2): 373-399.

SAS Institute, Inc. 1985. SAS Procedures Guide for Personal Computers, Version 6 edition. Cary, NC.

Sassaman, R. W., E. Hold, and K. Bergsvick. 1972. User's Manual for a Computer Program for Simulating Intensively Managed Allowable Cut. U.S.D.A. For. Serv. Gen. Tech. Rep. PNW-1, Pacific Northwest Forest and Range Exp. Stn. Portland. 50 pp.

Schreuder, H. T. , W. L. Hafley and F. A. Bennett. 1979. Yield Prediction for Unthinned Natural Slash Pine Stands. Forest Science, 25(1): 25-30.

Schreuder, H. T. and W. T. Swank. 1974. Coniferous Stands Characterized with the Weibull Distribution. Canadian Journal of Forest Research, 4(4): 518-523.

Schmidt, J. S. and P. L. Tedder. 1981. A Comprehensive Examination of Economic Optimization Simulation Methods. Forest Science, 27(3): 523-536.

Sharda, R. 1984. Linear Programming on Microcomputers: A Survey. Interfaces, 14(6): 27-38.

Sharda, R. 1986. Optimization Using Spreadsheets on a Microcomputer. Annals of Operations Research, 5(1-4): 599-612.

Sharda, 1988. The State of the Art of Linear Programming on Personal Computers. Interfaces, 18(4): 49-58.

Smalley , G. W. and R. L. Bailey. 1974a. Yield Tables and Stand Structures for Loblolly Pine Plantations in Tennessee, Alabama and Georgia Highlands. U.S.D.A. Forest Service Res. Pap. SO-96. 81 pp.

Smalley, G. W. and R. L. Bailey. 1974b. Yield Tables and Stand Structures for Shortleaf Pine Plantations in Tennessee, Alabama and Georgia Highlands. USDA Forest Service Res. Pap. SO-97. 57 pp.

Somers, G. L. , R. G. Oderwald, W. R. Harms and O. G. Langdon 1980. Predicting Mortality with a Weibull Distribution. Forest Science, 26(2): 291-300.

Strub M. R. and H. E. Burkhart. 1975. A Class-Free Method for Obtaining Expected Yields from Diameter Distributions. Forest Science, 21(1): 67-69.

Strub, M. R., D. P. Feduccia and V. C. Baldwin, Jr. 1981. A Diameter Distribution Method Useful in Compatible Growth and Yield Modelling of Thinned Stands. In Barnett, J. D. (Ed.). Proceedings of the First Biennial Southern Silvicultural Research Conference. U.S.D.A. Forest Service Tech. Rep. SO-34. 375 pp.

Sullivan, A. D. and J. L. Clutter. 1972. A Simultaneous Growth and Yield Model for Loblolly Pine. Forest Science 18(1): 76-86.

Sumichrast, R. T. 1990. Using a Microcomputer Decision Support System for Production Planning. Production and Inventory Management Journal, (First Quarter, 1990): 59-61.

Teeguarden, D. E. and H. L. von Sperber. 1968. Scheduling Douglas fir Reforestation Investments: A Comparison of Methods. Forest Science, 14(4): 354-367.

Turner, B. J., R. W. Bernarz and J. B. Dargavel. 1977. A Model to Generate Stand Strategies for Intensively Managed Radiata Pine Plantations. Australian Forestry. 40(4): 255-267.

Villanueva, T. R. and A. G. D. Whyte. 1992. Integrated Yield Forecasting and Harvest Scheduling for Tropical Plantations. In Whyte, A. G. D. (Ed.). Proceedings IUFRO S3.04.01 Conference. 27-31 January. Christchurch.

Wardle, P. A. 1965. Forest Management and Operational Research: A Linear Programming Study. Management Science, 11(10): B260-B270.

Ware, G. O. and J. L. Clutter. 1971. A Mathematical Programming System for the Management of Industrial Forests. Forest Science, 17(4): 428-445.

Wasil, E.. B. Golden, and R. Sharda. 1989. Mathematical Programming Software for the Microcomputer: Recent Advances, Comparisons and Trends. Impacts of Recent Computer Advances on Operations Research. pp. 263-272.

Weintraub, A. and D. Navon. 1976. A Forest Management Planning Model Integrating Silvicultural and Transportation Activities. Management Science, 22(12): 1299-1309.

Welch, J. S. Jr. 1987. The Data Management Needs of Mathematical Programming Applications. IMA Journal of Mathematics in Management, 1987(1): 237-250.

Whyte, A. G. D. 1978. Inventory and Yield Forecasting Systems for Caribbean pine in Fiji. In Elliot, D. A. (compiler). Proceedings, FRI Symposium 20, Mensuration for Management Planning of Exotic Forest Plantations. pp. 41-47.

Whyte, A. G. D. 1983. Development of the Fiji Pine Commission: An Account of the Emergence of a Major New Industry in a Developing Country. New Zealand Journal of Forestry, 28(4): 412-422.

Whyte, A. G. D. 1987. A Resource Planning and Management System for Pine Plantations in Fiji. Paper Presented to Sessions on Operations Research in Forestry. 13th International Federation for Information Processing. Conference on System Modelling and Optimization. Tokyo, Japan.

Whyte, A. G. D. 1989. Decision Support Systems for Tropical Pine Plantation in Fiji. In Proceedings, Recent Developments in Tree Plantations in Humid/Sub-humid Tropics of Asia. 5-9 June 1989. Universiti Pertanian Malaysia, Kuala Lumpur. pp. 81-92.

Winter, F. W. 1989. market Segmentation Using Modelling and Lotus 1-2-3. Interfaces, 19(6): 83-94.

Woollons, R. C. 1989. Advanced Growth and Yield Modelling. Lecture Notes. University of Canterbury. (Unpublished).

Wybourne, L. J. 1982. A Review of Spacing and Thinning in **Pinus caribaea var. Hondurensis** in Fiji (Analysis of Spacing and Thinning Trials at Seaqaqa). B. For. Sci. Dissertation. Univ. of Canterbury. 117 pp.

Xu, L. 1990. Growth and Yield of Douglas Fir Plantations in the Central North Island of New Zealand. Ph.D. Thesis. Univ. of Canterbury. 244 pp.

Zarnoch, S. J. and T. R. Dell. 1985. An Evaluation of Percentile and Maximum Likelihood Estimators of Weibull Parameters. Forest Science, 31(1): 260-268.

APPENDIX A

PRACTICAL RULES IN MODELLING GROWTH

This appendix contains a copy of a report on the experiences in modelling growth and yield. They are reported here as practical rules which helped in the modelling work for this project.

PRACTICAL RULES IN MODELLING GROWTH

Teodoro R. Villanueva

School of Forestry

University of Canterbury

May 1990

Practical Rules in Modelling Growth

Given a set of data on growth, modelling forest growth starts with the decision on which functional models to begin with. One usually starts with the basic models, in the form, $Y_2 = f(Y_1, T_1, T_2)$, where the response variable Y_2 measured at time T_2 is expressed as a function of the same variable measured at initial time T_1 and a measure of elapsed time as a function of T_1 and T_2. The variable Y can be basal area, height, stems per hectare or any diameter variables.

It is also however equally valid to start with a functional model that contains all possible variables and discard systematically. In both cases, scientific reasons should be given considerations.

In forest growth modelling, one usually begins with simple previously used and published by previous researches. The more commonly adopted models are the Schumacher, Gompertz and Chapman-Richards growth functions. Growth modelling can therefore consider the steps below.

Step 1. Before fitting any model, check the validity and correctness of the data. This can be done by plotting Y_2 vs. T_1, Y_2 vs. T_2, and Y_2 vs Y_1 making sure that Y_2 behaves the way it should with respect to each of the predictor variables. Checks to ensure correctness of entry or coding of data can include:

a) Check that $Y_2 > Y_1$, $T_2 > T_1$, N_1-N_2 which is the change in stocking is less than a set value, for example, 200 stems per hectare or as low as 1 stem per hectare.This value depends upon the data being analyzed, with which the researcher should be familiar.

b) Countercheck coded data against original plot data.

c) Obtain descriptive statistics of the variables, for example means, variance, minimum and maximum values among others.

d) Check that assumptions on error terms are met. Initially, one can only check the constancy of variances by:

- Plotting the data and visually assesing if the spread of the data on Y_2 increases or decreases with respect to any of the predictor variable, or
- Obtaining and plotting averages versus estimated variances for the replicated Y_2.
- There should be no systematic relationship if there is constancy of variances.

Step 2. Specify the model to be fitted. Forest growth functions like the ones mentioned above as starting models are usually correctly specified models since they have been tested to consider biological and theoretical principles. They have also been previously fitted to some growth data. Their specific characteristics include sigmoid shaped curve, an asymptotic value when age approaches infinity, consistency and path invariance.

Step 3. Run the model.

a) Obtain good starting values for the parameters - values from which convergence is quickly obtained. Related researches can be useful sources of information on starting values. Initial estimates of parameters can also be solved by simultaneous solution of **p** equations resulting from substitution of **p** set of observations, where **p** is the number of parameters to be estimated.

b) Choose the iterative technique, for example DUD or the Gauss-Newton. The choice of technique affects the speed with which convergence is achieved. The process is much quicker with the Gauss- Newton but it initially requires that partial derivatives of the model be specified.

c) Obtain convergence and ensure that the:

model is correctly specified;

model is correctly coded;

data are correctly entered;

observations are valid;

starting values are correct; and,

starting values correspond to the correct parameters.

Step 4. Examine and analyze the residuals.

a) Obtain summary statistics of residuals, including mean, normality, standard deviation, skewness and extreme values.

b) Obtain and analyze plots of residuals.

. Residual against predicted values for model inadequacies and inequality of variance.

. Residual against observed variable Y_1 for randomness of residuals.

Residuals against observed variable T_1 for randomness and independence of residuals.

- Residuals against other predictor variables to test adequacy of fit.

c) Detect outliers among residuals.

- Set a rule for declaring residuals. This can be done by setting a value beyond which a residual can be declared an outlier, for example 3 or 4 times the standard deviation.
- Examine the data indicated by the detected outlier. Do not reject automatically, but examine the data for explicable errors.

Step 5. Compare the models. Choose the model with the smallest residual mean square and the most random looking spread of residuals.

Step 6. Assess the fit of the model chosen and modify the parameters by including other independent variables. Guidelines for assessment are:

a) are parameter values sensible?

b) is there convergence at a local minimum?

c) do parameter estimates make sense biologically?

d) check the significance of the **t** values but be aware, however, that because of correlated data, the standard error is underestimated. The standard rules in assessing the significance of the **t** values may not apply. If values are quite critical, for example where the probability of rejection is very low, the whims of the researcher can come in, using his or her knowledge of the data being analyzed.

e) Check the asymptotic correlation matrix of parameters. High correlation may indicate overparametrization.

f) Study the following plots, charts and output data.

- Plot of fitted values of Y_2 overlaid with observed Y_2.
- Plot of residuals versus predicted Y_2 to check homoscedascity in the error term.
- Plot of residuals versus T_1 to check randomness in the error term.
- Plot of residuals versus Y_2 to check randomness in the error term.
- Probability plots of residuals to check normality and error mean equal to zero, i.e. bias in estimate.
- Frequency distribution to verify normality of residuals.
- Data on normality and skewness including spread and mean of residuals, overall and by specified classes.

Step 7. Add in new terms. Repeat steps 3 to 6. Stop when there is no further improvement in fit, for example if the error sum of squares is reduced by only 5%. This value is derived by the formula $((RSS_1 - RSS_2)/RSS_2)*100$. While this is arbitrary, others can be less strict. Again the knowledge of the data is necessary for a sound decision to adopt or reject a new model. Also consider simplicity and purpose of the data analysis to explain or account for the behaviour of the data and not simply to get the best fit. Also, consider biologically sensible parameter estimates in terms of sign and magnitude.

Step 8. Present results, including the final model and estimates of the parameters, standard errors and correlation. Discuss the model in the context of the problem being modelled.

REFERENCES

BATES, D. M. and D. G. WATTS. 1988. Nonlinear Regression Analysis and its Applications. John Wiley and Sons. 365pp.

DRAPER, N. R. and H. SMITH. 1981. Applied Regression Analysis. 2nd ed. John Wiley and Sons. 709pp.

WOOLLONS, R. C. 1989. Advanced Growth and Yield Modelling. Lecture Notes. School of Forestry, University of Canterbury (unpublished).

APPENDIX B

DATA FOR HARVEST SCHEDULING CASE STUDY

The input tables which were used in the construction of **HARVEST** can be viewed by retrieving the worksheet of the model using a spreadsheet package, preferably VPPlanner 3D.

1. Set up the appropriate spreadsheet package.
2. Retrieve the worksheet by using the File Retrieve commands on the menu. A prompt for the name of the file appears.
3. Type the file name including the correct specific path name, i.e. A:\VPP3D\HARVEST.WKS, if the distribution diskette containing the file is in drive A:.
4. Once the spreadsheet has been retrieved, tables can be directly accessed by invoking the appropriate {Go to} command of the spreadsheet package for example by pressing F5 in VPP3D, and then responding with the range names below when prompted for a cell address.

   IMPORTANT:The cells in this worksheet are unprotected and the user is advised not to modify the formula cells in them in order to avoid inadvertent changes to the program. In an instance that this happens, and the original formula cannot be written again, exit the worksheet without saving it by the Worksheet Erase Yes commands on the menu. Then retrieve the original worksheet again. It is also advisable that backups of the files are made.

The tables and their corresponding rangenames are:

APPENDIX C

TABLES CONTAINED IN HARVEST CONSTRAINTS SECTION

\VPP3D\HARV_OUT.WK1 is a worksheet file that contains the results of a sample run of **HARVEST**. These results can be viewed by retrieving the worksheet of the model using a spreadsheet package, preferably VPPlanner 3D.

1. Set up the appropriate spreadsheet package.
2. Retrieve the worksheet by using the File Retrieve commands on the menu. A prompt for the name of the file appears.
3. Type the file name including the correct specific path name, i.e. A:\VPP3D\HARV_OUT.WK1, if the distribution diskette containing the file is in drive A:.
4. Once the spreadsheet has been retrieved, tables can be directly accessed by invoking the appropriate {Go to} command of the spreadsheet package for example by pressing F5 in VPP3D, and then responding with CONSCRN when prompted for a cell address.

   IMPORTANT: The cells in this worksheet are unprotected and the user is advised not to modify the formula cells in them in order to avoid inadvertent changes to the program. In an instance that this happens, and the original formula cannot be written again, exit the worksheet without saving it by the Worksheet Erase Yes commands on the menu. Then retrieve the original worksheet again. It is also advisable that backups of the files are made.

The constraints as they appear in this section of the worksheet are:

Table	Descriptive Title
Table C.1	Area Accounting Constraints
Table C.2	Cutting Method Constraints
	Maximum Use of Method 2
	Minimum Use of Method 4
	Overall maximum Use of Machine Intensive Method
Table C.3	Port Requirement Constraints
	Port 1 Constraints
	Port 2 Constraints
	Combined Annual Port Requirements
Table C.4	Periodic Harvest Regulation Constraints
	Periodic Volume Harvested
	Periodic Area Harvested
Table C.5	Equivalent Annual Applications of Methods
Table C.6	Summary Report

APPENDIX D

Diskettes containing BAREA.DAT and DIAMETER.DAT

\SAS\BAREA.DAT and **\SAS\DIAMETER.DAT** are two SAS files that contain the basal area and diameter projection data sets respectively. These files can be viewed on a terminal or printed on a printer by using the DOS Type or Print commands and specifying the appropriate pathname and filename. The column labels in the yield data format are as defined in Table 3.1 and the column labels in the projection data set are as defined in Table 3.2. Refer to these table for the description of the different variables that appear in these files.

APPENDIX E

Diskettes containing YIELD

The different versions of **YIELD** are contained in the diskettes.

1. **\DIADIS\DIDISTR1.FOR** is the Vax Fortran version source code. This file can be viewed on a terminal or printed on a printer by using the DOS Type or Print commands, for example the command

   Type A:\Didistr1.For | More

   gives a view of the contents of the file in screenful.

2. \DIADIS\DIDISTR.BAT is the Pc Fortran version. The program has been included in a batch program which can be invoked by typing Didistr while still in the DIADIS subdirectory. It should be ensured however before running this batch program that an output file named MODEL.OUT has not been created. Type Dir to ensure this. The files in the directory are as below, no more no less.

   DIDISTR.BAT
   DIADIS.EXE
   BROWSE.COM
   KEYPRESS.COM
   WAIT.COM
   PRINT.COM
   DIDISTR1.FOR.

   If the file MODEL.OUT is listed, delete this by typing Delete Model.out before typing Didistr.

   The user is then prompted for response to provide data inputs related to the initial conditions of the stand. The program is not foolproof and may just hang when the initial inputs are very highly unrealistic. The program can be restarted by typing 8888. When the inputting of data is finished, the simulation is started and an output is created. From here, a screen is shown where the user is asked on how he/she would like to view the output of the simulation.

3. \VPP3D\YIELD.WKS is the spreadsheet version. The macros which were used to construct the spreadsheet version of YIELD are not unique to VPPlanner 3D, thus can be run to compatible spreadsheet packages. Start up the spreadsheet package and retrieve the worksheet using the File Retrieve commands on the menu. A prompt for the file name appears. Type a:\VPP3D\Yield.wks press Enter. The program's opening screen appears and a pull down menu of the different tasks the model can do. IMPORTANT: User can always go back to the main menu by holding the Alt key and pressing 0 [Alt-0]. The cells in this worksheet are unprotected and the user is advised not to modify the formula cells and the macros in it in order to avoid inadvertent changes to the program. In an instance that this happens, and the original formula and macros cannot be written again, exit the worksheet without saving it by the Worksheet Erase Yes commands on the menu. Then retrieve the original worksheet again. It is also advisable that backups of the files are made.

APPENDIX F

Stand and stock tables generated by YIELD

Sensitivity Test I: LOLOLO/DRASA FOREST

PROJECTED STAND STATISTICS
LOLOLO/DRASA FOREST

AGE (yrs)	BASAL AREA (sq.m./ha)	STOCKING (/ha)	MEAN DIAMETER (cm)	MAXIMUM DIAMETER (cm.)	DIAMETER STANDARD DEVIATION (cm.)	MEAN BASAL AREA (sq.m./ha)
5.	15.0	1180.0	12.1	20.0	4.0	.0127
15.	39.4	1173.1	19.7	33.2	6.5	.0336
25.	48.6	1166.2	21.5	37.7	8.3	.0417

LOG GRADES STAND/STOCK PROJECTION
LOLOLO/DRASA FOREST

STAND AGE = 5. WEIBULL PARAMETERS A =20.00 B = 8.94 C = 1.98

DIAMETER CLASS (cm)	CLASS FREQUENCY (/ha)	AVERAGE HEIGHT (m)	CLASS VOLUME (cu.m./ha)	P1	P2	S1	S2	S3	CHIP	WASTE
					-L---O---G		G---R---A---D---E---S- (cu.m./ha)			
7.5	265.5	6.0	7.237	.00	.00	.00	.00	.00	5.04	2.19
12.5	520.5	9.6	33.299	.00	.00	.00	.00	.00	23.75	9.55
17.5	320.9	11.7	41.798	.00	.00	22.70	.00	.00	13.68	5.42
22.5	.0	13.1	.000	.00	.00	.00	.00	.00	.00	.00
TOTAL	1106.9		82.334	.00	.00	22.70	.00	.00	42.48	17.16

LOG GRADES STAND/STOCK PROJECTION
LOLOLO/DRASA FOREST

STAND AGE =15. WEIBULL PARAMETERS A =33.20 B =15.24 C = 2.08

DIAMETER CLASS (cm)	CLASS FREQUENCY (/ha)	AVERAGE HEIGHT (m)	CLASS VOLUME (cu.m./ha)	P1	P2	S1	S2	S3	CHIP	WASTE
					-L---O---G		G---R---A---D---E---S- (cu.m./ha)			
7.5	74.5	9.5	2.494	.00	.00	.00	.00	.00	1.25	1.25
12.5	169.3	15.2	15.512	.00	.00	.00	.00	.00	10.31	5.20
17.5	282.7	18.6	55.584	.00	.00	28.23	.00	.00	23.44	3.91
22.5	331.9	20.8	116.019	.00	64.13	.00	.00	23.87	20.24	7.77
27.5	237.8	22.3	130.975	.00	68.57	.00	39.47	.00	18.22	4.72
32.5	44.8	23.5	35.895	.00	18.09	.00	10.77	4.20	2.09	.75
TOTAL	1141.0		356.480	.00	150.79	28.23	50.24	28.07	75.54	23.60

Stand and stock tables generated by YIELD

Sensitivity Test I: LOLOLO/DRASA FOREST ... p. 2

LOG GRADES STAND/STOCK PROJECTION
LOLOLO/DRASA FOREST

STAND AGE =25.   WEIBULL PARAMETERS A =37.71 B =18.22 C = 1.94

DIAMETER CLASS (cm)	CLASS FREQUENCY (/ha)	AVERAGE HEIGHT (m)	CLASS VOLUME (cu.m./ha)	P1	P2	S1	S2	S3	CHIP	WASTE
				-L---O---G			G---R---A---D---E---S- (cu.m./ha)			
7.5	70.4	15.0	3.048	.00	.00	.00	.00	.00	.00	3.05
12.5	129.4	24.0	17.527	.00	.00	.00	.00	.00	15.23	2.29
17.5	200.7	29.4	60.546	.00	25.43	.00	.00	.00	26.19	8.93
22.5	256.3	32.9	139.439	.00	53.30	.00	37.79	13.73	24.99	9.63
27.5	256.4	35.4	221.255	.00	79.57	.00	97.79	14.00	23.44	5.81
32.5	172.1	37.2	216.591	.00	74.64	.00	119.69	.00	18.01	4.01
37.5	28.7	38.5	49.715	.00	16.60	.00	27.34	3.24	1.88	.64
TOTAL	1114.0		708.121	.00	249.53	.00	282.61	30.98	109.74	34.37

APPENDIX F (con't.)

Stand and stock tables generated by YIELD

Sensitivity Test I: SEAQAQA FOREST

PROJECTED STAND STATISTICS
SEAQAQA FOREST

AGE (yrs)	BASAL AREA (sq.m./ha)	STOCKING (/ha)	MEAN DIAMETER (cm)	MAXIMUM DIAMETER (cm.)	DIAMETER STANDARD DEVIATION (cm.)	MEAN BASAL AREA (sq.m./ha)
5.	15.0	1180.0	12.1	20.0	4.0	.0127
15.	41.8	1173.1	20.4	40.0	6.5	.0356
25.	51.3	1166.2	22.2	46.4	8.3	.0440

LOG GRADES STAND/STOCK PROJECTION
SEAQAQA FOREST

STAND AGE = 5. WEIBULL PARAMETERS A =20.00 B = 8.94 C = 1.98

DIAMETER CLASS (cm)	CLASS FREQUENCY (/ha)	AVERAGE HEIGHT (m)	CLASS VOLUME (cu.m./ha)	P1	-L---O---G P2	S1	G---R---A---D---E---S- S2 (cu.m./ha)	S3	CHIP	WASTE
7.5	265.5	6.0	7.237	.00	.00	.00	.00	.00	.00	7.24
12.5	520.5	9.6	33.299	.00	.00	.00	.00	.00	29.28	4.01
17.5	320.9	11.7	41.798	.00	.00	23.04	.00	.00	13.07	5.69
22.5	.0	13.1	.000	.00	.00	.00	.00	.00	.00	.00
TOTAL	1106.9		82.334	.00	.00	23.04	.00	.00	42.35	16.94

LOG GRADES STAND/STOCK PROJECTION
SEAQAQA FOREST

STAND AGE =15. WEIBULL PARAMETERS A =40.02 B =22.01 C = 3.04

DIAMETER CLASS (cm)	CLASS FREQUENCY (/ha)	AVERAGE HEIGHT (m)	CLASS VOLUME (cu.m./ha)	P1	-L---O---G P2	S1	G---R---A---D---E---S- S2 (cu.m./ha)	S3	CHIP	WASTE
7.5	70.7	9.5	2.365	.00	.00	.00	.00	.00	.00	2.37
12.5	178.1	15.2	16.323	.00	.00	.00	.00	.00	11.28	5.04
17.5	286.1	18.6	56.246	.00	.00	29.10	.00	.00	22.81	4.34
22.5	303.3	20.8	106.015	.00	59.43	.00	.00	20.77	18.01	7.80
27.5	212.7	22.3	117.179	.00	62.40	.00	33.68	.00	16.27	4.83
32.5	89.6	23.5	71.708	.00	36.82	.00	20.52	8.20	4.39	1.78
37.5	13.1	24.3	14.390	.00	7.19	.00	6.27	.00	.70	.23
42.5	.0	25.0	.000	.00	.00	.00	.00	.00	.00	.00
TOTAL	1153.5		384.226	.00	165.84	29.10	60.47	28.97	73.46	26.38

Stand and stock tables generated by YIELD

Sensitivity Test I: SEAQAQA FOREST ... p. 2

LOG GRADES STAND/STOCK PROJECTION
SEAQAQA FOREST

STAND AGE =25. WEIBULL PARAMETERS A =46.45 B =27.17 C = 2.91

DIAMETER CLASS (cm)	CLASS FREQUENCY (/ha)	AVERAGE HEIGHT (m)	CLASS VOLUME (cu.m./ha)	P1	P2	S1	S2	S3	CHIP	WASTE
				-L---O---G			G---R---A---D---E---S- (cu.m./ha)			
7.5	72.9	15.0	3.155	.00	.00	.00	.00	.00	.00	3.15
12.5	141.4	24.0	19.154	.00	.00	.00	.00	.00	16.53	2.63
17.5	210.4	29.4	63.471	.00	27.41	.00	.00	.00	26.33	9.73
22.5	243.1	32.9	132.251	.00	52.20	.00	34.62	12.37	23.05	10.01
27.5	219.1	35.4	189.030	.00	70.37	.00	80.40	11.54	20.33	5.69
32.5	150.9	37.2	189.871	.00	67.85	.00	101.22	.00	16.38	4.15
37.5	72.8	38.5	126.085	.00	43.71	.00	66.89	8.30	5.17	1.97
42.5	17.3	39.6	39.487	13.38	.00	.00	23.83	.00	2.02	.00
47.5	.2	40.5	.663	.22	.00	.16	.24	.02	.01	.00
TOTAL	1128.0		763.166	13.60	261.54	.16	307.20	32.23	109.81	37.34