A COMPARISON OF YIELD ESTIMATION TECHNIQUES FOR

OLD-FIELD LOBLOLLY PINE PLANTATIONS

by

Mike Robert Strub

Dissertation submitted to the Graduate Faculty of

the Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Statistics

APPROVED:

_____   _____
W. R. Pirie, Co-chairman   H. E. Burkhart, Co-chairman

_____   _____
    J. C. Arnold   M. R. Reynolds

                    _____
                         R. B. Vasey

June, 1976

Blacksburg, Virginia

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

Loblolly pine (*Pinus taeda*) is a fast growing species that is intensively managed for the production of pulpwood, sawlogs, and veneer bolts. High early return on investments (when compared with returns from growth of other species) justifies the large monetary investments required for intensive management. A fundamental requirement of most management systems is yield information. Sophisticated management planning aids such as RETURN (Schweitzer et al., 1967), TEVAP 2 (Myers, 1973) and TIMBER RAM (Navon, 1971) require detailed and accurate yield projections. Yield information is also required for computer simulations such as the Purdue Management Game (Bare, 1970), used to train forest managers.

The topic of this dissertation is the development of quantative models to supply forest managers with yield estimates fundamental to long range planning. Yield is defined as the amount of wood produced by a collection or stand of trees over some time period. The units in which yield is measured are determined by the product which is to be manufactured from the trees. Volume or weight of solid wood contained in the bole of the trees is the common unit of measure when paper is produced from the trees. Larger trees are often measured according to the

1

amount of lumber which can be manufactured from them. The unit of measure for lumber is termed the board foot, defined as 144 cubic inches. Specific models for yield estimation will be discussed later in this chapter and in Chapters 2 and 4. A general description of the derivation of yield models follows.

The yield of a stand of trees depends on the genetic make-up of the trees, the environmental conditions to which the trees are exposed, and the length of time over which the trees are grown. Only a few of these factors have been successfully quantified. For example, genetic make-up of trees is currently under extensive research and an issue of considerable debate; accurate measurement of this factor has not been included in previous growth models. In conventional yield models environmental conditions are commonly described by two quantities -- one a measure of the growing quality of the land upon which the trees are planted, the second an indication of the competition between trees for scarce resources. A third factor that is usually included in yield models is the length of time over which the trees are grown. These three quantifiable variables are termed stand variables. In this study, the measure of site quality is site index, the indicator of competition between trees is the number of trees per acre, and the length of time the trees are grown is stand age.

Site index is a measure of the fertility of the land upon which the trees are planted. It is defined as the average height of the largest trees in the stand at some base age, usually 25 years for loblolly pine plantations. Site index curves can be constructed by assuming the following functional relationship between stand age (A) and the average height of the largest trees (H_d):

$$\log_{10}(H_d) = \beta_0 + \beta_1/A. \qquad (1.1)$$

The expression $\log_{10}(\cdot)$ indicates the base 10 logarithm of the argument. Site index is estimated from a single observation of the age and average height of the largest trees. Yield is often predicted as a function of estimated site index plus other variables. An alternate approach to using site index is to predict yield from H_d, since stand age is also included in both yield and site prediction schemes. The stand variable H_d is often preferred since it is measured, not estimated as is site index.

The second stand variable is the number of trees per acre (N) in the plantation. Increasing N at a fixed age and site index results in increasing competition for scarce resources. Increasing A and holding N fixed also results in increasing competitive stress. Competition is a function of N, A, and site index.

A yield equation is a function used to predict yield from the stand variables H_d, N, and A. Two types of yield estimation techniques will be discussed in this presentation, they will be labeled direct and indirect methods. Direct yield estimation methods relate stand yield directly to stand variables. A common procedure is to use stepwise regression to obtain a functional form and to estimate coefficients. This technique makes little or no use of knowledge of fundamental physical and biological processes which result in the yield of a stand of trees. Indirect yield processes which result in yield production techniques are based on rudimentary processes which result in yield production. The characteristic process of yield formation upon which indirect techniques are based is that the yield of a stand of trees is the summation of the yield of the individual trees in the stand. Other physiological processes may inspire the choice of the functional form used to model individual tree yield.

The purpose of this dissertation is to critically examine direct and indirect yield prediction procedures currently in use, and to suggest new yield prediction procedures which may be superior to those currently in use. Yield prediction models can be compared on several levels. One criterion of the worth of a model is the accuracy and precision of estimates obtained from the model. The manager is most concerned about this

criterion. He is interested in obtaining an accurate
estimate of the yield at the end of some growth period
so that he can pursue a course of action which is
financially rewarding and ecologically sound. A second
criterion, also of concern to the manager, is the amount
of detail contained in the yield information. Direct
yield prediction models supply only a total stand yield
estimate. Indirect yield techniques supply detailed
information related to the size and yield of individual
trees. This information is useful in making management
decisions such as determining optimal harvesting methods.
A third criterion is the biological relevance of the yield
prediction model. This criterion is of lesser importance
to the forest manager. The manager is satisfied if the
yield model conforms to certain pre-conceived notions
of tree growth (e.g. the age at which mean annual increment
should culminate). While these notions are of some
consequence, other biological factors seldom referenced
by forest managers should form the basis of yield predic-
tion models. The importance of biological factors is
supported by the following argument.

Most yield prediction models are functionally
derived from a single data set. Extrapolation of such
models outside the range of observed values in the original
data set is dangerous. Another approach is to base the
functional form of yield equations on the biological

processes which produce yield. This approach should
suggest a model which is less sensitive to pecularities
of the data set.

Objectives

The objectives for this dissertation were: (1) to
compare existing yield prediction techniques in light
of information required for forest management, (2) to
suggest possible improvement of existing techniques
espectially improvements based on biological and physiolog-
ical aspects of yield production, and (3) to derive new
yield prediction techniques based on biological and
physiological aspects of yield production.

Literature Review

Yield of Individual Trees

The data base used to construct direct and indirect
yield equations depends on individual tree yield equations.
Direct determination of the yield of a large number of
trees is prohibitively expensive. Instead individual tree
data are collected and tree yield is predicted from the
easily measured variables, tree diameter at breast height
(D or dbh) and total tree height (H). The prediction
equations developed are used to estimate the yield of
trees by measuring only D and H for each tree. Individual
tree height, H, is more difficult to measure than D and

further reduction in the cost of data collection is
achieved by predicting individual tree volume from
D alone or from D and other stand variables. A
general treatment of individual tree volume equations
is given in Chapter 9 of Hush et al. (1972), and
Chapters 4 to 9 of Spurr (1952).

Yield of individual trees varies according to the
product which is to be manufactured from the tree. In
general tree yield is proportional to the volume of the
bole of the tree. Tree boles assume geometric shapes
similar to cylinders, cones, paraboloids, and neiloids.
In each case listed above the volume of the figure is
proportional to the product of the area of the base and
the height. Analogously total volume of a tree bole
should be approximately proportional to the product of
the square of diameter at breast height and total tree
height. The equation which describes the relationship,

$$V = \beta D^2 H, \qquad (1.2)$$

has been labelled the constant form factor volume equation.

The constant form factor volume equation must be
modified to provide accurate estimation of the yield of
products such as lumber. The top-most section of the
tree bole, often referred to as topwood, cannot be
economically utilized. The point beyond which the bole
of the tree cannot be used for manufacture of a product is

termed the merchantability limit. This merchantability
limit is usually the point on the tree where some minimum
diameter is reached. In this case the yield of the tree
is proportional to the volume of the bole excluding the
topwood, or the volume of the tree up to the merchant-
ability limit. The combined variable volume equation,

$$V = \beta_0 + \beta_1 D^2 H, \qquad (1.3)$$

is commonly used to model tree volume to some merchantability
limit. Equation (1.2) is a special case of (1.3), with
β_0 equal to 0. The intercept, β_0, is negative and
represents loss of volume due to topwood. This loss in
volume should be constant since the area of the base
of the topwood is constant, and, if all trees taper at the
same rate, the height of the topwood will be constant.
Although many other functional forms have been considered,
the combined variable is the most common functional form.
This equation was used by Burkhart et al. (1972) to
predict the yield of individual trees found in loblolly
pine plantations.

Direct Estimation of Yield of Forest Stands

The initial step in the implementation of this procedure is data collection. Some subset of the trees in a forest stand is chosen for measurement, usually all the trees on some fraction of an acre of land. The stand variables are measured for this subset of the stand. Average yield per acre in the stand is estimated by (1) estimating the yield of each tree in the subset according to the technique described in the preceding section, (2) summing individual tree yield to obtain an estimate of the yield of the trees on that fractional part of an acre, and (3) multiplying by some constant to convert the yield estimate to yield per acre.

Multiple regression is then applied to the data described above. The dependent variable is yield per acre or some transformation of yield per acre. The independent variables are the stand variables and functions of the stand variables. A stepwise procedure is often used to aid in the determination of the proper functional form of the yield prediction equation. Yield prediction is described in detail in Chapters 17, 18, 21 and 22 of Spurr (1952), and in Chapter 17 of Hush et al. (1972).

Burkhart et al. (1972) computed yield equations for loblolly pine plantations; multiple regression was used

to estimate average stand yield per acre (Y),

$$\log_{10}(Y) = \beta_0 + \beta_1/A + \beta_2 H_d/A + \beta_3 N + \beta_4 A \log_{10}(N). \quad (1.4)$$

This equation was used to estimate stand yield per acre
for cubic feet of wood contained in the boles of the
trees to various merchantability limits and for board
feet contained in the boles of the trees.

When lumber is cut from the tree bole, the topwood
can be used in the manufacture of paper. Burkhart et
al. (1972) developed a different functional form to
estimate the volume of topwood (CV_2) from the board foot
volume per acre estimate (BF) and the cubic foot volume
per acre estimate (CV_1) obtained from (1.4),

$$\log_{10}(CV_1 - CV_2) = \beta_0 + \beta_1 \log_{10}(BF), \quad (1.5)$$

resulting in

$$CV_2 = CV_1 - 10^{\beta_0} BF^{\beta_1}. \quad (1.6)$$

Indirect Estimation of Yield of Forest Stands

One indirect method of yield estimation is referred
to in the forestry literature as the diameter distribution
method. Stand yield estimates are obtained by first
estimating the frequency distribution of tree diameters
at breast height of the trees in the stand. The estimated
frequency distribution, height prediction equations, and

individual tree yield equations are used to predict stand yield.

The frequency distribution of tree diameters has been studied since the late eighteen hundreds. Meyer (1930) summarized these early attempts to describe diameter distributions. Graphical techniques suggested by Bruce and Reineke (1928) consisted of fitting the normal or log-normal distribution to the data by using special graph paper. Parameters were estimated by drawing a straight line through the center of the data points. Other techniques for estimating parameters suggested by early researchers were based on the method of moments. Meyer (1928) used Charlier curves to describe diameter distributions by equating the coefficient of asymmetry and the coefficient of excess of the data to that of the Charlier curve. Schnur (1934) used Charlier curves and Pearson curves to approximate the diameter distribution of old-field loblolly pine stands. Osborne and Schumacher (1935) used the Pearl-Reed growth curve (a generalized form of the logistic distribution) to describe the diameter distribution of even-aged stands. All of the above techniques are handicapped by computational difficulties. Elaborate techniques and gross assumptions are devoted to reduction of arithmatic operations. The inavailability of mechanical and electronic calculating devices during this early period accounts for the

numerous simplifying assumptions which result in computational ease.

Diameter distributions received little attention during the forties and fifties, but interest was revived in the sixties. Nelson (1964) investigated the correlation between characterizations of the diameter distribution (such as the average diameter, the range of tree diameters, the sample standard deviation, the coefficient of asymmetry, and the coefficient of excess) of loblolly pine and stand growth. He also compared parameter estimates obtained by fitting the gamma distribution and the Pearl-Reed growth curve to the diameter distribution data with stand growth. None of the above characterizations of the diameter distribution were highly correlated with stand growth. Bliss and Reinker (1964) used the three parameter log-normal distribution to describe the diameter distribution of Douglas-fir. They presented graphical and numerical techniques for estimating parameters. The method of moments was used to estimate parameters from data grouped into diameter classes. Chi-square tests indicated that the log-normal distribution did not accurately describe the diameter distribution of Douglas-fir (*Pseudotsuga menziesii*).

The work of Clutter and Bennett (1965) marked the beginning of modern interest in diameter distributions.

This new interest was characterized by the influx of new methods which depended upon the availability of high-speed computing equipment. Clutter and Bennett suggested using the four-parameter beta distribution to approximate the diameter distribution of slash pine (*Pinus elliottii*). Lloyd (1966) used the same distribution to describe diameter distributions for Virginia pine (*Pinus virginia*). Other applications of the four-parameter beta include McGee and Della-Bianca (1967) who described yellow poplar (*Liriodendron tulipifera*), Lenhart and Clutter (1971), Lenhart (1972), and Burkhart and Strub (1974) who worked with old-field loblolly pine plantations. Methods of parameter estimation included moment and maximum likelihood estimation of shape parameters. The smallest and largest diameters observed on a sample plot were used to estimate the parameters that describe the domain of tree diameters.

Bailey and Dell (1973) suggested that the Weibull distribution be used to quantify forest diameter distributions. They described the two and three-parameter Weibull. The three parameter Weibull has a location parameter (λ), a scale parameter (γ), and a shape parameter (δ). The density function suggested by Bailey and Dell to describe frequency of occurance of tree diameters is

$$f(D) = \gamma\delta(D-\lambda)^{\delta-1}e^{-\gamma(D-\lambda)^{\delta}}, \quad \lambda < D. \qquad (1.7)$$

The two-parameter Weibull is a special case of (1.7) with
λ set equal to zero. Bailey and Dell discussed parameter
estimation for only the two-parameter Weibull. They
suggested maximum likelihood and quantile estimation of
the parameters γ and δ. Maximum likelihood estimates
can be obtained from the solutions of

$$[\sum_{i=1}^{n} D_i^{\hat{\delta}} \ln(D_i)] / [\sum_{i=1}^{n} D_i^{\hat{\delta}}] - 1/\hat{\delta} = [\sum_{i=1}^{n} \ln(D_i)]/n \; , \quad (1.8)$$

$$\hat{\gamma} = [\sum_{i=1}^{n} D_i^{\hat{\delta}}/n]^{1/\hat{\delta}} \; . \quad\quad (1.9)$$

The expression $\ln(\cdot)$ indicates the natural logarithm of
the argument. Equation (1.8) must be solved iteratively.
The resulting estimate of δ is used to obtain an estimate
of γ via (1.9). Harter and Moore (1965) discussed a

similar maximum likelihood estimation procedure for the
three-parameter Weibull. Quantile estimators are
obtained by equating two quantiles from the data with
 two quantiles of the Weibull distribution. Dubey (1967)
suggests use of the 93rd and 24th quantiles since this
results in minimum asymptotic generalized variance (the
determinant of the asymptotic variance-covariance matrix)
of the quantile estimates. The quantile estimates (γ and

$\tilde{\delta}$) are calculated from the 93rd quantile (Q_{93}) and the 24th quantile (Q_{24}) by applying

$$\tilde{\delta} = \ln\{\ln(0.07)/\ln(0.76)\}/\ln(Q_{93}/Q_{24}) \, , \qquad (1.10)$$

$$\tilde{\gamma} = -\ln(0.76)/(Q_{24})^{\tilde{\delta}} \, . \qquad (1.11)$$

The shape parameter, δ, is first estimated via (1.10) and then the scale parameter, γ, is estimated from (1.11) and $\tilde{\delta}$.

Once the parameters of the diameter distribution have been estimated for specific sets of stand variables from data sets, the estimates are compared with those stand variables. Schnur (1934) presented the basic technique after which modern efforts are modeled. Schnur plotted parameter estimates against average stand diameter at breast height and fitted smooth free-hand curves through the data points. These curves could be used to predict the parameters of the estimated diameter distribution from average diameter. The modern approach consists of using multiple regression techniques to predict the parameters of the estimated diameter distribution from stand variables by fitting a smooth curve to the data.

Recent efforts to estimate diameter distributions consists of (1) collecting data consisting of measurement

of all tree diameters on a fractional part of an acre
(often 0.1-acre), and measurement of the stand variables
such as number of trees per acre, the average height of
the dominant and co-dominant trees (the largest trees),
and stand age, (2) repeating this process many times
(each sample is called a plot), (3) for each plot fitting
a smooth function such as the Weibull distribution to the
diameter data, thereby obtaining parameter estimates for
each set of stand variables, and (4) using multiple
regression techniques to estimate the diameter distribu-
tion parameters from the stand variables (one regression
equation is required for each diameter distribution
parameter). Once this process is complete, stand yield
is estimated from the estimated diameter distribution.

Bennett and Clutter (1968) first predicted yield
from estimated diameter distributions. They used the
beta distribution to describe the diameter distribution of
slash pine plantations, and to predict stand yields of
pulpwood, sawtimber, gum (used to make turpentine
and other products), and multiple-product yields
for mixtures of these products. An advantage of
this indirect yield estimation procedure is that it
is easily adapted to prediction of multiple-product
yield. Estimation of yield is accomplished by
(1) estimating the expected number of trees in each
diameter class (the range of diameters is divided into

class intervals, usually with a width of one inch) by
multiplying the area under the estimated diameter distribu-
tion curve in the diameter class by the number of trees
per acre, (2) predicting the volume of a tree with dia-
meter equal to the midpoint of the diameter class, this
is accomplished by using an individual tree volume
equation, (3) estimating the volume in each diameter class
as the product of expected number of trees in the diameter
class and the predicted volume of a tree with diameter
equal to the midpoint of the diameter class, and (4)
obtaining an estimate of the average stand yield per
acre as the sum of the estimated yields per acre in each
diameter class. A mathematical representation of this
process is

$$Y = N \sum_{i=1}^{m} P_i g(D_i, A, H_d, N) .$$

(1.12)

The estimated average stand yield per acre is represented
by Y, m is the number of diameter class intervals, and
D_i is the midpoint of the ith diameter class. The
expression for predicting individual tree volume,
$g(D_i, A, H_d, N)$, is a function of the stand variables
as well as tree diameter at breast height. The
relative frequency of occurance of trees in the ith class
interval, P_i, is obtained by integrating the estimated

diameter denstiy over the range of the ith diameter class. This yield estimation procedure has been used by Bennet and Clutter (1968) to predict yields of slash pine plantations, and by Lenhart and Clutter (1971), Lenhart (1972), and Smalley and Bailey (1974a) to predict the yield of loblolly pine plantations. Smalley and Bailey (1974b) used this technique to predict the yield of shortleaf pine (*Pinus echinata*) plantations.

II. PROCEDURES

Data Collection

The data used for this study were from old-field loblolly pine plantations, and from loblolly pine trees grown free from competition. The study area, plot selection procedures, and plot measurements were described by Burkhart et al. (1972).

Study area. Data for this study were collected by field crews from several industrial forestry organizations. Selected loblolly pine plantations were sampled in the Piedmont and Coastal Plain regions of Virginia, and in the Coastal Plain region of Delaware, Maryland and North Carolina. Data from 186 sample plots were used in the analyses reported in this dissertation.

Plot selection. Temporary 0.1-acre, circular sample plots were randomly located in selected stands. To be sampled, plantations were required to be unthinned, contain no interplanting, be free of severe insect or disease damage, be unburned and unpruned, and be relatively free of wildlings.

Plot measurements. On each 0.1-acre plot, dbh was recorded to the nearest 0.1 inch for all trees in the

1-inch dbh class and above. Each tree in the 8-inch
dbh class and above was classed as qualifying or not
qualifying for sawtimber. A sawtimber tree was defined
as being in the 8-inch dbh class or larger and having
at least one 16-foot sawlog to a 6-inch top diameter,
inside bark. Total height was recorded for at least one,
but usually two trees per 1-inch dbh class. Six to eight
dominant and co-dominant trees were selected as site
sample trees and the total age of the stand was determined
from planting records, increment borings, or ring counts
at the stump of the felled specimens.

On each plot two trees (the 10th and 20th trees
measured) were felled and cut into 4-foot sections for
detailed measurements. The following data were recorded
for each felled sample tree:

1. Diameter at breast height to the nearest
   0.1 inch

2. Total tree height to the nearest 0.1 foot

3. Total age of the tree (age at the top of
   each 4-foot bolt was also recorded for
   dominants and codominants)

4. Diameters (inside and outside bark) at
   the stump and at 4-foot intervals up the
   stem to an approximate 2-inch top diameter
   (outside bark).

Additional data not described in Burkhart et al.
(1972) were collected for trees grown free of competition.
Eighty-one sample trees were located in Piedmont and
Coastal Plain Virginia. The trees were located in
abandoned fields, plantations, rights-of-way, and on
private lawns. They were chosen only if free of insect
damage and disease, fire damage, and any sign of competi-
tion as described by Krajicek et al. (1961). Measurements
made on each tree included:

1. Age from seed

2. Diameter at breast height to the
    nearest 0.1 inch

3. Total tree height to the nearest
    foot

4. Length of live tree crown to the
    nearest foot

5. Two crown diameters, at the widest
    crown expanse, measured perpendicular
    to each other to the nearest foot.

Data were collected by the Chesapeake Corporation
of Virginia, Continental Can Company Incorporated,
Glatfelter Pulp Wood Company, Southern Johns-Manville
Products Corporation, and Union Camp Corporation. The
Virginia Division of Forestry contributed some open-grown
tree measurements.

Individual Tree Yield Equations

Individual tree yield equations were required to estimate the yield of trees on each of the 186 sample plots as a preliminary step to derivation and evaluation of the yield estimation procedures. The same individual tree equations were used as an integral part of the indirect yield estimation procedures. The individual tree yield equations used in this dissertation are described by Burkhart et al. (1972) and are modifications of the combined variable equation (1.3); they were developed to simplify data collection. The total height data (usually two trees per 1-inch dbh class) was used to establish the relationship between dbh and total tree height on each plot. The data on each plot were used to estimate β_2, and β_3 in

$$\log_{10}(H_p) = \beta_2 + \beta_3/D. \tag{2.1}$$

This equation was then used to predict total tree height (H_p) from tree diameter at breast height for each tree measured on the same plot. Individual tree volume was then a function of tree diameter at breast height only, described by

$$V = \beta_0 + \beta_1 D^2 H_p = \beta_0 + \beta_1 D^2 10^{\beta_2 + \beta_3/D}. \tag{2.2}$$

The coefficients β_0, and β_1 were estimated from one subsample of trees chosen from trees exposed to a range of stand variables (the felled tree data were used for this purpose) using (1.3); β_2 and β_3 were estimated from the subsample of trees on each plot using (2.1). Estimates of coefficients were substituted into (2.2) to obtain an individual tree yield equation for each sample plot.

These individual tree yield equations were used to predict the volume of each tree on the plot from dbh. The estimates of each tree's yield were summed to obtain an estimate of total plot yield for each of the 186 0.1-acre sample plots; the estimated plot yields were multiplied by ten to obtain an estimate of yield per acre.

A similar approach was used to estimate the topwood contained in individual trees and to estimate topwood per acre from the data collected on each plot. The functional form used to estimate the cubic feet of topwood (TV) contained in individual trees from dbh is described by the following equations:

$$TV = \beta_0 + \beta_1 H/D^2 , \qquad (2.3)$$

$$TV = \beta_0 + \beta_1 D^{-2} 10^{\beta_2 + \beta_3/D} . \qquad (2.4)$$

Equation (2.4) was used to predict topwood volume per
acre from the dbh measurements made on each plot.

New equations, which use H_d in place of H, for
predicting individual tree yield and topwood yield are:

$$V = \beta_0 + \beta_1 D^2 H_d \qquad (2.5)$$

$$TV = \beta_0 + \beta_1 H_d / D^2 \qquad (2.6)$$

These models have some utility when indirect methods of
stand yield estimation are considered. The result of the
individual tree volume estimation effort is summarized by
Table 2.1.

Random Sample

Fifty of the 186 sample plots were chosen via a
random sample stratified on stand age. The fifty plots
were used in the comparison part of this study (observed
and expected yield and diameter distributions were
compared). For this reason, stand yield prediction
procedures are based on the remaining 136 sample plots.

Table 2.1. Coefficients of individual tree volume equations

Unit of Volume Measure	Model	Slope	Intercept	Coefficient of Determination	Standard Error of Estimate
Cubic Feet	1.3	0.11691	0.00185	0.970	0.495
Total Stem Inside Bark	2.5	-0.06136	0.00187	0.952	0.614
Cubic Feet Inside Bark	1.3	-0.46236	0.00185	0.959	0.576
To 4" Top Outside Bark	2.5	-0.63594	0.00187	0.945	0.665
Board Feet	1.3 *	-23.67532	0.01102	0.517	0.002
International 0.25-inch	2.5	-37.23422	0.01332	0.856	12.738
Cubic Feet of Topwood	2.3	0.49318	2.7089	0.413	0.564
To 4" Top	2.6	0.67632	2.3883	0.427	0.557

*Weighted regression was used for this model. Least squares was used
to estimate parameters in the model

$$V/D^2H = \beta_0/D^2H + \beta_1.$$

The model was solved for volume and used to estimate individual tree volume
from D and H.

Direct Estimation of Stand Yield

The yield per acre estimates from each of the 136 sample plots and the stand variables measured on each of those sample plots were used to estimate coefficients of the per acre yield equation, (1.4). Results from employing this procedure are summarized by Table 2.2. Estimation of the number of cubic feet of topwood per acre to a 4-inch merchantability limit was accomplished by use of (1.5) and (1.6). Simple linear regression was used to estimate β_0 and β_1 in (1.5). The estimates were respectively -0.16822 and 0.85685. The coefficient of determination was 0.479 and the standard error of estimate was 0.083.

Indirect Estimation of Stand Yield

Maximum likelihood estimation is a statistical technique which is used extensively in indirect estimation of stand yield. This technique and its properties are described in the next three sections of this chapter.

Maximum Likelihood Estimation

Maximum likelihood is a statistical technique used to estimate the parameters of a probability distribution

Table 2.2. Parameter estimates, coefficient of determination and standard error of estimate for per-acre yield prediction via the direct yield estimation approach* for 136 sample plots in loblolly pine plantations

Unit of Measure of Volume	b_0	b_1	b_2	b_3	b_4	Coefficient of Determination (R^2)	Standard Error of Estimate $S_{y \cdot x}$
Cubic Feet Per Acre, Total Stem, Inside Bark	2.45498	−7.21875	0.32198	0.00819	0.00808	0.941	0.048
Cubic Feet Per Acre Inside Bark to a 4-inch Top (OB)	2.62827	−13.38634	0.45609	−0.01633	0.00571	0.934	0.077
Board Feet Per Acre International 0.25-inch to a 6-inch Top (IB)	4.88359	−68.40489	1.22336	−0.14121	.00921	0.812	0.292

*Model: $\text{Log}_{10}Y = b_0 + b_1(1/A) + b_2(H_d/A) + b_3(N/100) + b_4(A)(\text{Log}_{10}N)$

from observations from that probability distribution.
This technique and properties of maximum likelihood
estimates are described in detail in chapter 5 of Zacks
(1971) and chapter 5 of Rao (1973). The method of maximum
likelihood is to find the parameter estimates which
maximize the likelihood function. The likelihood function
is defined as the joint probability density function or .
probability mass function of the observations. Maximizing
this function with respect to choice of parameters results
in choosing the probability distribution with the largest
likelihood function for the data observed. A set of n
independent identically distributed random variables, $\{X_i\}$
each with probability density or mass function $f(X_i, \underline{\theta})$,
with parameter vector $\underline{\theta}$, have the following likelihood
function (L),

$$L = \prod_{i=1}^{n} f(X_i, \underline{\theta}). \tag{2.7}$$

The maximum likelihood estimates of $\underline{\theta}$, $\hat{\underline{\theta}}$, are
obtained by maximizing L. A common technique is to maximize
ln(L). This results in the same estimates since the natural
logarithm is a monotonic transformation, and this often
simplifies the maximization problem. Classical optimiza-
tion techniques are commonly used, the partial derivative

of ln(L) with respect to each parameter is set equal
to zero, resulting in a system of non-linear equations
which must be solved for the parameter vector, $\hat{\underline{\theta}}$. If
the matrix of second partial derivatives with respect
to the parameters, the Hessian matrix, evaluated at
$\hat{\underline{\theta}}$ is negative definite, a local maximum of ln(L) has
been found. Numerical techniques are often used
to find maximum likelihood estimates; these numerical
techniques will be discussed in a later section of
this chapter.

The properties of maximum likelihood estimators
are discussed in Zacks (1971) and Rao (1973). Rao
gives three restrictions on the probability distribu-
tion of X which result in consistency and asymptotic
normality of the maximum likelihood estimators.

> Restriction 1. The first three partial
> derivatives of f(X, $\underline{\theta}$) with
> respect to $\underline{\theta}$ exist, in an
> interval, A, including the
> true value of $\underline{\theta}$.

Restriction 2. At the true value of $\underline{\theta}$,

$$E \left[\frac{\partial f(X,\underline{\theta})}{\partial \theta_i} / f(X, \underline{\theta}) \right] = 0 \qquad \text{for all } i$$

$$E \left[\frac{\partial^2 f(X,\underline{\theta})}{\partial \theta_i \partial \theta_j} / f(X, \underline{\theta}) \right] = 0 \qquad \text{for all } i, j$$

$$E \left[\{\frac{\partial f(X,\underline{\theta})}{\partial \theta_i}\}^2 / f(X, \underline{\theta}) \right] > 0 \qquad \text{for all } i$$

Restriction 3. For every $\underline{\theta}$ in A,

$$\left| \frac{\partial^3 \ln[f(X, \underline{\theta})]}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| < M(X) \qquad \text{for all } i, j, k$$

$$E [M(X)] < K$$

The constant, k, must not depend on $\underline{\theta}$; $M(X)$ is a function of X. These restrictions also guarantee other asymptotic results which will be discussed in the following section.

Discrete Regression

Regression is a well documented statistical technique with wide practical application. There are

numerous references on regression analysis such as the
texts by Graybill (1961) and Draper and Smith (1966).
A basic assumption of most regression models is that the
dependent variable (y) is continuous. Numerous forest
management situations require modeling of discrete
random variables similar to Bernoulli trials. Some
examples of these forestry applications include mortality
of trees, presence or absence of a tree species, and
whether or not a tree is suitable for manufacture of
lumber. The topic of this section is the development of
a discrete regression model. The dependent variable will
be assumed to be a discrete random variable, similar in
nature to a Bernoulli trial. The exact probability mass
function assumed is described by

$$P(y) = [g(\underline{X}, \underline{\theta})]^y [1 - g(\underline{X}, \underline{\theta})]^{1-y} \qquad y = 0, 1. \qquad (2.8)$$

The dependent variable, y, can assume the values zero and
one. Common situations where this model can be applied
are often described as success-failure, or life-death
settings. The probability of success or failure, or
life or death, is a function of a vector of independent
variables, $\underline{X}$, and a vector of parameters, $\underline{\theta}$. No particular
functional form for $g(\underline{X}, \underline{\theta})$ is assumed at this point; how-
ever, the function must be bounded between zero and one.

Reference to this model in the statistical literature is directed to specific functional forms for $g(\underline{X}, \underline{\Theta})$. Walker and Duncan (1967) and Hamilton (1974) address the case where $g(\underline{X}, \underline{\Theta})$ is the logistic distribution function and determine an iterative procedure for finding least squares like estimates of $\underline{\Theta}$. Cox (1970) also restricts discussion to the logistic function; he discusses least squares and maximum likelihood estimation of $\underline{\Theta}$. Probit analysis deals with the case where $g(\underline{X}, \underline{\Theta})$ is the normal or log-normal distribution function. Again iterative procedures are used to determine least squares estimates of the parameters. Garwood (1941) demonstrated that the least squares and maximum likelihood estimates are equivalent in probit analysis. Jennrich and Moore (1975) discussed the use of standard least squares programs in finding maximum likelihood estimates of $\underline{\Theta}$.

The purpose of this section is to develop a general theory concerning estimation of the parameters $\underline{\Theta}$ and inference on those parameters by using the general theory of maximum likelihood. Maximum likelihood has intuitive appeal that the least squares procedure lacks in this instance. The least squares approach is to minimize the sum of the squared distances from $g(\underline{X}, \underline{\Theta})$ to the data points. The value of the dependent variable, y, must be zero or one, hence least squares estimates minimize the

sum of the squared distances from $g(\underline{X}, \underline{\theta})$ to zero or one.
Maximum likelihood estimates of $\underline{\theta}$ are derived from the
likelihood function for n observations of y and $\underline{X}$, namely

$$L = \prod_{i=1}^{n} [g(\underline{X}_i, \underline{\theta})]^{y_i} [1 - g(\underline{X}_i, \underline{\theta})]^{1-y_i}. \qquad (2.9)$$

A more convenient form is the lagarithm of the likelihood
function,

$$\ln(L) = \sum_{i=1}^{n} y_i \ln[g(\underline{X}_i, \underline{\theta})] + \sum_{i=1}^{n} (1-y_i)\ln[1-g(\underline{X}_i, \underline{\theta})]. \qquad (2.10)$$

Maximization of $\ln(L)$ cannot be achieved using classical
techniques except for some elementary forms of $g(\underline{X}, \underline{\theta})$.
An iterative technique for maximizing $\ln(L)$ will be
introduced in the next section.

Inference on the parameter vector, $\underline{\theta}$, is described
in general in Chapter 6 of Rao (1973). The test of
hypothesis $\underline{\theta} = \underline{\theta}_0$ can be performed by using the likeli-
hood ratio statistic proposed by Neyman and Pearson (1928),

$$\Lambda = \frac{\prod\limits_{i=1}^{n} [g(\underline{X}_i, \underline{\theta}_0)]^{y_i} [1 - g(\underline{X}_i, \underline{\theta}_0)]^{1-y_i}}{\prod\limits_{i=1}^{n} [g(\underline{X}_i, \hat{\underline{\theta}})]^{y_i} [1 - g(\underline{X}_i, \hat{\underline{\theta}})]^{1-y_i}} . \qquad (2.11)$$

The distribution of Λ under the null hypothesis can be
determined from (2.8) with $\underline{\theta} = \underline{\theta}_0$, since Λ is a function

of the random variables $\{Y_i\}$ whose probability mass function is given by (2.6). Enumeration of all possible permutations of the $\{Y_i\}$, associated values of $\{\underline{X}_i\}$, $\hat{\underline{\Theta}}$, and the probability of occurance of each permutation as calculated from (2.8) results in the exact distribution of Λ. This information can be used to determine the critical region of the test. The power of the test can be determined for specific alternate values of $\underline{\Theta}$ in a similar fashion.

The Neyman-Pearson likelihood ratio test can also be used to test hypotheses of the form $\underline{\Theta} = \underline{h}(\underline{\gamma})$; where $\underline{h}$ is a vector of functions which depend on the vector of parameters $\underline{\gamma}$. The number of elements in $\underline{\Theta}$ is q, and the number of elements in $\underline{\gamma}$ is S < q. Under the null hypothesis the probability mass function of Y is given by

$$P(y) = \{g[\underline{X}, \underline{h}(\underline{\gamma})]\}^y \{1 - g[\underline{X}, \underline{h}(\underline{\gamma})]\}^{1-y} \quad y = 0, 1. \quad (2.12)$$

Under the null hypothesis, the likelihood function, L, is given by

$$L = \prod_{i=1}^{n} \{g[\underline{X}_i, \underline{h}(\underline{\gamma})]\}^{y_i} \{1 - g[\underline{X}_i, \underline{h}(\underline{\gamma})]\}^{1-y_i}. \quad (2.13)$$

Maximum likelihood estimates of $\underline{\gamma}$ can be found by applying the same iterative scheme as is used to determine maximum likelihood estimates of $\underline{\theta}$. The maximum likelihood estimates of $\underline{\gamma}$ will be denoted $\hat{\underline{\gamma}}$. The test statistic used to test the hypotheses $\underline{\theta} = \underline{h}(\underline{\gamma})$ is given by

$$\Lambda^* = \frac{\prod\limits_{i=1}^{n} \{g[\underline{X}_i, \underline{h}(\hat{\underline{\gamma}})]\}^{y_i} \{1 - g[\underline{X}_i, \underline{h}(\hat{\underline{\gamma}})]\}^{1-y_i}}{\prod\limits_{i=1}^{n} [g(\underline{X}_i, \hat{\underline{\theta}})]^{y_i} [1 - g(\underline{X}_i, \hat{\underline{\theta}})]^{1-y_i}} . \qquad (2.14)$$

An exact test cannot be performed since the true value of $\underline{\theta}$ is not completely specified by the null hypothesis ($\underline{\theta}$ depends on $\underline{\gamma}$ which is not given). An approximate test can be performed by following a procedure similar to the procedure used for the simple test of hypothesis. The approximate distribution of $\{Y_i\}$ under the null hypothesis is given by

$$P(y) = \{g[\underline{X}, h(\hat{\underline{\gamma}})]\}^{y} \{1-g[\underline{X}, h(\hat{\underline{\gamma}})]\}^{1-y} \qquad y = 0,1 . \qquad (2.15)$$

This probability mass function, (2.15), can be used in place of (2.8) to determine the probability mass function of Λ^*. The critical region and power of the test for specific alternative values of $\underline{\theta}$ are determined by following the same procedure as in the simple test. If

the number of observations is large, computational effort to test the hypothesis may become prohibitively large. In this case an asymptotic test can be used.

Rao (1973) demonstrated that the asymptotic distribution of $-2\ln(\Lambda^*)$ is chi-square with q-s degrees of freedom under the null hypothesis if restrictions 1, 2, 3 listed in the previous section hold.

Restriction 1. This restriction is met if the first three partial derivatives of $g(\underline{X}, \underline{\Theta})$ with respect to $\underline{\Theta}$ exist.

Restriction 2. For all values of Θ,

$$E\left[\frac{\partial P(y)}{\partial \Theta_i}/P(y)\right] = E\{y/g(\underline{X},\underline{\Theta}) - (1-y)/[1-g(\underline{X},\underline{\Theta})]\}\frac{\partial g(\underline{X},\underline{\Theta})}{\partial \Theta_i} = 0$$

$$E\left[\frac{\partial^2 P(y)}{\partial \Theta_i \partial \Theta_j}/P(y)\right] = E\{y/g(\underline{X},\underline{\Theta}) - (1-y)/[1-g(\underline{X},\underline{\Theta})]\}\frac{\partial^2 g(\underline{X},\underline{\Theta})}{\partial \Theta_i \partial \Theta_j}$$

$$+ E\left[y(y-1)\{[g(\underline{X},\underline{\Theta})]^{y-2}[1-g(\underline{X},\underline{\Theta})]^{1-y}+2[g(\underline{X},\underline{\Theta})]^{y-1}[1-g(\underline{X},\underline{\Theta})]^{-y}\right.$$

$$+ [g(\underline{X},\Theta)]^{y}[1-g(\underline{X},\underline{\Theta})]^{-1-y}\}\left.\right]\frac{\partial g(\underline{X},\underline{\Theta})}{\partial \Theta_i}\frac{\partial g(\underline{X},\underline{\Theta})}{\partial \Theta_j} = 0$$

$$E\left[\{\frac{\partial P(y)}{\partial \Theta_i}\}^2/P(y)\right] = E\{y^2[g(\underline{X},\underline{\Theta})]^{y-2}[1-g(\underline{X},\underline{\Theta})]^{1-y}$$

$$-y(1-y)[g(\underline{X},\underline{\Theta})]^{y-1}[1-g(\underline{X},\underline{\Theta})]^{-y}$$

$$+ (1-y)^2[g(\underline{X},\underline{\Theta})]^y[1-g(\underline{X},\underline{\Theta})]^{-1-y}\}\left[\frac{\partial g(\underline{X},\underline{\Theta})}{\partial \Theta_1}\right]^2$$

$$= [g(\underline{X},\underline{\Theta})]^{-1}[g(\underline{X},\underline{\Theta})] + 0$$

$$+ [1-g(\underline{X},\underline{\Theta})]^{-1}[1-g(\underline{X},\underline{\Theta})] [\frac{\partial g(\underline{X},\underline{\Theta})}{\partial \Theta_1}]$$

$$= 2 \left[\frac{\partial g(X,\Theta)}{\partial \Theta_i}\right]^2 \neq 0 ,$$

$$\text{if } 0 < g(\underline{X}, \underline{\Theta}) < 1 \text{ and } \frac{\partial g(X,\Theta)}{\partial \Theta_i} \neq 0$$

Restriction 3. For every $\underline{\Theta}$,

$$\frac{\partial^3 \ln[P(y)]}{\partial \Theta_i \partial \Theta_j \partial \Theta_k} = 2 \{\frac{y}{[g(\underline{X},\underline{\Theta})]3} + \frac{1-y}{[1-g(\underline{X},\underline{\Theta})]3}\}$$

$$\{\frac{\partial g(X,\Theta)}{\partial \Theta_i} \quad \frac{\partial g(X,\Theta)}{\partial \Theta_j} \quad \frac{\partial(X,\Theta)}{\partial \Theta_k}\} + \{\frac{1-y}{[1-g(\underline{X},\underline{\Theta})]^2} - \frac{y}{[g(\underline{X},\underline{\Theta})]^2}\}$$

$$\{\frac{\partial^2 g(X,\Theta)}{\partial\Theta_i\partial\Theta_j} \cdot \frac{\partial g(X,\Theta)}{\partial\Theta_k} + \frac{\partial^2 g(X,\Theta)}{\partial\Theta_i\partial\Theta_k} \cdot \frac{\partial g(X,\Theta)}{\partial\Theta_j} + \frac{\partial^2 g(X,\Theta)}{\partial\Theta_j\partial\Theta_k} \cdot \frac{\partial g(X,\Theta)}{\partial\Theta_i}\}$$

$$+ \{\frac{y}{g(\underline{X},\underline{\Theta})} + \frac{1-y}{1-g(\underline{X},\underline{\Theta})}\}\{\frac{\partial^3 g(\underline{X},\underline{\Theta})}{\partial\Theta_i\partial\Theta_j\partial\Theta_k}\}$$

which is bounded if $0<g(\underline{X},\underline{\Theta})<1$ and the first three partial derivatives of $g(\underline{X},\underline{\Theta})$ with respect to $\underline{\Theta}$ are bounded. This theory will be applied to forestry problems in later sections of this chapter.

Optimization Theory

The purpose of this section is to describe general procedures for optimizing functions. These procedures will be used to determine maximum likelihood estimates from data. Optimization schemes for functions of a single variable and functions of more than one variable will be discussed.

Optimization techniques found in the statistical literature are based on the classical method of finding the value of the variable or variables for which the derivative or partial derivatives vanish. These techniques are outlined in Chapter 5 of Zacks (1971) and Chapter 5 of Rao (1973). The techniques described are based on some modification of the Newton-Raphson method for solving non-linear equations.

The Newton-Raphson method is a method for approximating vector of values, $\underline{X}^*$, which maximize or minimize the function $f(\underline{X})$. The gradient vector or vector of partial derivatives with respect to $\underline{X}$ is denoted $\partial f(\underline{X})/\partial\underline{X}$. The matrix of second partial derivatives of $f(\underline{X})$ with respect to $\underline{X}$ is called the Hessian matrix and denoted $\partial^2 f(\underline{X})/\partial\underline{X}^2$. The Newton-Raphson technique is described by

$$\underline{X}_{\nu+1} = \underline{X}_\nu - [\partial^2 f(\underline{X})/\partial\underline{X}|\underline{X}_\nu]^{-1}[\partial f(\underline{X})/\partial\underline{X}|\underline{X}_\nu]. \qquad (2.16)$$

Equation (2.16) describes an iterative procedure, $\underline{X}_{\nu+1}$ is the new "improved" estimate of $\underline{X}^*$, $\underline{X}_\nu$ is the old estimate of $\underline{X}^*$. In practice some initial estimate of $\underline{X}^*$ is obtained, say $\underline{X}_0$. Equation (2.16) is applied repeatedly, each time obtaining a new estimate of $\underline{X}^*$. When the new estimate of $\underline{X}^*$ is close to the old estimate, the procedure has converged and the final estimate of $\underline{X}^*$ should be approximately equal to $\underline{X}^*$. Convergence is defined as agreement of two successive estimates of $\underline{X}^*$ to five significant figures in applications found in this dissertation.

Application of this technique to practical problems results in certain difficulties. Russell (1970) describes these difficulties in Chapter 5.

"Now the method is by no means guaranteed to work in every case. First of all . . . the initial point

$\underline{X}_0$ should be fairly close to $\underline{X}^*$. Also, Newton's method pays no attention to the fact that we are looking for a minimum of f -- it is looking for a point where $\partial f/\partial\underline{X} = 0$. Such a point might or might not be a minimum. It might be a maximum of f or a point where $\partial f/\partial\underline{X} = 0$ but f has neither a maximum nor a minimum. In fact, Newton's method for finding a maxima is exactly the same as Newton's method for finding a minima."

In practice the Newton-Raphson technique works well when optimizing f with respect to a single variable, but often does not work well when $\underline{X}$ represents a vector of variables.

A technique which does work well was suggested by Davidon (1959). A refined version of this technique is described in detail by Fletcher and Powell (1963). They claim that "the method is probably the most powerful general procedure for finding a local minimum which is known at the present time." This technique is described by

$$\underline{X}_{\nu+1} = \underline{X}_\nu - a_\nu H_\nu[\partial f(\underline{X})/\partial\underline{X}|\underline{X}_\nu] \quad (2.17)$$

The positive scalar constant, a_ν, is the number which gives the greatest reduction in f. The matrix H_ν is updated at each iteration. Initially H_1 is set equal to the unit matrix. Thereafter it is updated according to

$$H_{\nu+1} = H_{\nu} + [\underline{X}_{\nu} - \underline{X}_{\nu+1}][\underline{X}_{\nu} - \underline{X}_{\nu+1}]^{T}/[\underline{X}_{\nu} - \underline{X}_{\nu+1}]^{T}\underline{Y}_{\nu}$$

$$-H_{\nu}\underline{Y}_{\nu}\underline{Y}_{\nu}^{T} H_{\nu}/\underline{Y}_{\nu}^{T} H_{\nu}\underline{Y}_{\nu}. \tag{2.18}$$

The superscript T indicates the transpose of the vector. The vector $\underline{Y}_{\nu}$ is defined

$$\underline{Y}_{\nu} = [\partial f(\underline{X})/\partial\underline{X}|\underline{X}_{\nu+1}] - [\partial f(\underline{X})/\partial\underline{X}|\underline{X}_{\nu}]. \tag{2.19}$$

The same convergence criterion can be used as was used for the Newton-Raphson technique. This technique was successfully used to determine maximum likelihood estimates for multiparameter distributions. The application of this technique to estimation of parameters in approximations of diameter distributions is discussed later in this chapter.

The Weibull Approximation to Diameter Distributions

The first step in indirect estimation of yields of forest stands is to approximate the diameter distribution of the stand. One approximation procedure is to assume that tree diameters follow a Weibull distribution and to estimate the parameters of the distribution function for each stand from stand variables. One parameter estimation scheme involved obtaining maximum likelihood estimates for the two-parameter Weibull density for each of the 136 plots. The Weibull density is:

$$f(D) = \gamma\delta D^{\delta-1}e^{-\gamma D^{\delta}} \qquad 0 < D, \gamma, \delta \qquad (2.20)$$

The log-likelihood function associated with this probability density function is

$$\ln(L) = n\ln(\gamma) + n\ln(\delta) + (\delta-1)\sum_{i=1}^{n}\ln(D_i) - \gamma\sum_{i=1}^{n}D_i^{\delta}. \qquad (2.21)$$

The number of trees on a sample plot is n, and D_i represents the dbh of the ith tree on the plot. The two equations which must be solved simultaneously to find maximum likelihood estimates of γ and δ are

$$\left.\frac{\partial\ln(L)}{\partial\gamma}\right|_{\hat{\gamma},\hat{\delta}} = \frac{n}{\gamma} = \sum_{i=1}^{n}D_i^{\delta} = 0, \qquad (2.22)$$

and

$$\left.\frac{\ln(L)}{\partial\delta}\right|_{\hat{\gamma},\hat{\delta}} = \frac{n}{\delta} + \sum_{i=1}^{n}\ln(D_i) - \hat{\gamma}\sum_{i=1}^{n}D_i^{\hat{\delta}}\ln(D_i) = 0. \qquad (2.23)$$

Equation (2.22) can be solved for $\hat{\gamma}$, resulting in

$$\hat{\gamma} = n/\sum_{i=1}^{n}D_i^{\hat{\delta}} \qquad (2.24)$$

This expression for $\hat{\gamma}$ is then substituted into (2.23) to obtain

$$n/\hat{\delta} + \sum_{i=1}^{n} \ln(D_i) - [n/\sum_{i=1}^{n} D_i^{\hat{\delta}}][\sum_{i=1}^{n} D_i^{\hat{\delta}} \ln(D_i)] = 0. \qquad (2.25)$$

This equation can be solved iteratively for $\hat{\delta}$ using the Newton-Raphson method described by (2.16),

$$\hat{\delta}_{\nu+1} = \hat{\delta}_{\nu} - \{n/\hat{\delta}_{\nu} + \sum_{i=1}^{n} \ln(D_i) - [n/\sum_{i=1}^{n} D_i^{\hat{\delta}_{\nu}}][\sum_{i=1}^{n} D_i^{\hat{\delta}_{\nu}} \ln(D_i)]\}/$$

$$\{-n/(\hat{\delta}_{\nu})^2 + [n(\sum_{i=1}^{n} D_i^{\hat{\delta}_{\nu}} \ln(D_i))^2/(\sum_{i=1}^{n} D_i^{\hat{\delta}_{\nu}})^2]$$

$$-[n/\sum_{i=1}^{n} D_i^{\hat{\delta}_{\nu}}][\sum_{i=1}^{n} D_i^{\hat{\delta}_{\nu}} (\ln(D_i))^2]\}. \qquad (2.26)$$

The use of this equation to determine the value of $\hat{\delta}$ which maximizes $\ln(L)$ is described in the optimization theory section of this chapter.

Once maximum likelihood estimates were obtained for each plot, least squares was used to fit a curve to these estimates and the stand variables. Equations obtained for this data set were

$$\ln(\hat{\gamma}) = -9.5549 - 28.9205H_d/N + 0.1081N/H_d, \qquad (2.27)$$

and

$$\ln(\hat{\delta}) = 1.7593 - 0.0080N/H_d. \qquad (2.28)$$

Estimates of γ and δ must be positive according to (2.20). Coefficients of determination are low (0.622 and 0.223, respectively) as has been the case when this technique has been applied in the past (e.g. Smalley and Bailey 1974a and 1974b). The standard errors of estimate were 1.559 and 0.158, respectively.

The coefficient of determination and the standard error of estimate will be reported for each parameter estimation model. The standard error of estimate is an estimate of the standard deviation of observations about the true regression line. A small standard error of estimate indicates a good fit. The coefficient of determination indicates the proportion of variation of the independent variable about its mean which is explained by the regression line. A coefficient of determination close to one indicates a good fit. Comparison of the standard error of estimate and coefficient of determination of different models is valid only when the dependent variable of each model is the same. The coefficients of determination and standard errors of estimate presented in this dissertation are in general only indicators of the goodness of fit of the model they describe and should not be compared between models.

A second procedure for estimating γ and δ from stand variables was based on the 24th and 93rd quantiles.

The 24th and 93rd quantiles were determined for each of the 136 sample plots. Least squares procedures were used to estimate the quantiles, Q, from stand variables. The same functional form was used to estimate each quantile,

$$Q = \beta_0 + \beta_1 H_d + \beta_2 AN/1000 + \beta_3 H_d/N. \qquad (2.29)$$

Parameter estimates, coefficients of determination and standard errors of estimate are contained in Table 2.3. Estimates of γ and δ are obtained from the predicted quantiles and (1.10) and (1.11).

The third estimation procedure, based on the method of moments, is a new technique that uses the first order statistic, $D_{(1)}$, and the average diameter, $\overline{D}$. The expected values of the first order statistic and the mean for the two parameter Weibull distribution are respectively:

$$E(D_{(1)}) = (1/n\gamma)^{1/\delta}\Gamma(1 + 1/\delta), \qquad (2.30)$$

$$E(\overline{D}) = (1/\gamma)^{1/\delta}\Gamma(1 + 1/\delta). \qquad (2.31)$$

The expectation operator is denoted by $E(\cdot)$; $\Gamma(\cdot)$ indicates the gamma function, and n is the sample size. Prediction equations were developed to estimate the average diameter

Table 2.3. Parameter estimates, coefficients of determination, and standard errors of estimate for prediction of quantiles and minimum and average diameters on 0.1- acre plots in loblolly pine stands

Diameter to Be Estimated	b_0	b_1	b_2	b_3	Coefficient of Determination R^2	Standard Error of Estimate $S_{y.x}$
24th Quantile	2.45759	0.05385	-0.05320	13.16225	0.838	0.469
93rd Quantile	3.96878	0.05484	-0.03535	30.41613	0.931	0.454
Minimum Diameter	-0.05729	0.05294	-0.02001	14.7423	0.735	0.554
Average Diameter	2.88209	0.05593	-0.04957	17.8446	0.932	0.336

and the smallest diameter of the trees on the sample plots.
The same functional form as (2.29) was used to estimate
both average and smallest diameter from the stand
variables, namely

$$D = \beta_0 + \beta_1 H_d + \beta_2 AN/1000 + \beta_3 H_d/N . \qquad (2.32)$$

Parameter estimates, coefficients of determination and
standard errors of estimate are also contained in Table 2.3.

Estimates of γ and δ are obtained by equating the
estimated smallest tree diameter, $D_{(1)}$, and $E(D_{(1)})$, and
by equating the estimated average tree diameter, $\bar{D}$, and
$E(\bar{D})$. This resulted in a set of two simultaneous equations
in two unknowns,

$$D_{(1)} = (1/n\gamma')^{1/\delta'} \Gamma(1 + 1/\delta') , \qquad (2.33)$$

and

$$\bar{D} = (1/\gamma')^{1/\delta'} \Gamma(1 + 1/\delta') \qquad (2.34)$$

These equations were solved simultaneously for γ' and δ'

$$\delta' = \frac{\ln(n)}{\ln(\bar{D}') - \ln(D_{(1)}')} \qquad (2.35)$$

$$\gamma' = \left[\frac{\Gamma(1 + 1/\delta')}{\bar{D}'} \right]^{\delta'} \qquad . \quad (2.36)$$

Estimates of γ and δ for a specific set of stand variables are obtained by first predicting $D_{(1)}$ and $\overline{D}$ from prediction equations with the same functional form as (2.32). These estimates are used first to predict δ' from $D_{(1)}$ and $\overline{D}$ via (2.35), and then to predict γ' from δ' and $\overline{D}$ via (2.36). The use of these diameter distribution approximation procedures to indirectly estimate yields will be discussed in a later section of this chapter.

The Normal Approximation to Basal Area Distributions

A new procedure for approximating the diameter distribution of forest stands is to assume that the basal areas (B) of trees in the stand are normally distributed. This approach was inspired by examining frequency histograms of basal areas on the 136 sample plots. The histograms for many of the plots appeared bell shaped, hence it was conjectured that the normal distribution would provide an accurate approximation to the frequency distribution of tree basal areas. The diameter distribution was derived from the basal area distribution, since basal area in square feet and diameter in inches are related according to

$$B = 0.005454 \; D^2.\qquad(2.37)$$

This transformation was used to derive the diameter distribution; the result is that if basal areas are normally

distributed, diameters are distributed as the square root of a normal distribution.

The parameters (μ and σ^2) of the normal distribution,

$$f(B) = \frac{1}{\sqrt{2\pi}\ \sigma}\ e^{-(B-\mu)^2/2\sigma^2} \qquad \infty < B < \infty, \qquad (2.38)$$

were estimated for each of the 136 sample plots. An unbiased estimate of μ, $\bar{B}$, was obtained from

$$\bar{B} = \sum_{i=1}^{n} B_i/n. \qquad (2.39)$$

The basal area of the ith tree on the plot was B_i. An unbiased estimate of σ^2 was obtained from

$$S^2 = \sum_{i=1}^{n} (B_i - \bar{B})^2/(n-1). \qquad (2.40)$$

Once a parameter estimate was obtained for each plot, least squares was used to fit a curve to the estimates and the stand variables. Equations obtained for this data set were

$$\bar{B} = -0.17669 + 0.0054195\ H_d + 101.56/N, \qquad (2.41)$$

and

$$S = 0.042725 + 0.0015267\ AH_d/\sqrt{N}. \qquad (2.42)$$

Coefficients of determination were 0.918 and 0.782, respectively. The standard error of estimates were respectively 0.0275 and 0.0185.

The procedure for approximating a diameter distribution was to first estimate μ and σ from stand variables via (2.41) and (2.42). These estimates were substituted into (2.38) to obtain an approximation of the basal area distribution, from which transformation (2.37) could be used to obtain an approximation to the diameter distribution. This procedure was used in a later section as part of an indirect yield estimation scheme.

<div align="center">

The Transformed Normal Approximation to
Basal Area Distributions

</div>

Frequency historgrams from the 136 sample plots seemed to indicate that some of the diameter distributions were not normally distributed, but somewhat skewed. In an attempt to introduce greater flexability into the assumed basal area distribution, it was conjectured that some power of basal area was normally distributed, namely $B^{1-C}/(1-C)$. This assumption results in a basal area distribution described by

$$f(B) = \frac{B^{-C}}{\sqrt{2\pi A}\ \sigma}\ e^{-[B^{1-C}-A\mu(1-C)]^2/2A\sigma^2(1-C)^2},$$

$$C \neq 1, \qquad -\infty < B < \infty. \qquad (2.43)$$

The parameters of this distribution are μ, σ^2, and C; A represents stand age. Further justification for this assumption will be presented in Chapter 4.

Maximum likelihood estimates of C, μ, and σ^2 were determined for each of the 136 sample plots. The likelihood equation for this density function is:

$$L = \prod_{i=1}^{n} \frac{B_i^{-C}}{\sqrt{2\pi A}\,\sigma} e^{-[B_i^{1-C}-A\mu(1-C)]^2/2A\sigma^2(1-C)^2} \qquad (2.44)$$

The log-likelihood function is more convenient to work with;

$$\ln(L) = -C \sum_{i=1}^{n} \ln(B_i) - (n/2)\ln(2\pi A) - (n/2)\ln(\sigma^2)$$

$$- \sum_{i=1}^{n} [B_i^{1-C}-A\mu(1-C)]^2/2A\sigma^2(1-C)^2 \qquad (2.45)$$

The maximum likelihood estimates of C, μ, and σ^2 are obtained by taking the partial derivatives of $\ln(L)$ with respect to each variable and equating each partial derivative to zero. The resulting three equations are solved simultaneously for C, μ, and σ^2,

$$\frac{\partial \ln(L)}{\partial C}\bigg|_{\hat{C},\hat{\mu},\hat{\sigma}} = -\sum_{i=1}^{n} \ln(B_i) + \{(1-\hat{C})\sum_{i=1}^{n}[B_i^{1-\hat{C}} - A\hat{\mu}(1-\hat{C})]$$

$$[B_i^{1-\hat{C}}\ln(B_i) - A\hat{\mu}] - \sum_{i=1}^{n}[B_i^{1-\hat{C}}-A\hat{\mu}(1-C)]^2\}/[A\hat{\sigma}^2(1-\hat{C})^3]$$

$$= 0, \qquad\qquad\qquad (2.46)$$

$$\frac{\partial \ln(L)}{\partial \mu}\bigg|_{\hat{C},\hat{\mu},\hat{\sigma}^2} = \sum_{i=1}^{n}\left[B_i^{1-\hat{C}} - A\hat{\mu}(1-\hat{C})\right]/\hat{\sigma}^2(1-\hat{C}) = 0, \qquad (2.47)$$

and

$$\frac{\partial \ln(L)}{\partial \sigma^2}\bigg|_{\hat{C},\hat{\mu},\hat{\sigma}^2} = -n/2\hat{\sigma}^2 + \sum_{i=1}^{n}\left[B_i^{1-\hat{C}} - A\hat{\mu}(1-\hat{C})\right]^2/2A\hat{\sigma}^4(1-\hat{C})^2$$

$$= 0. \qquad\qquad\qquad (2.48)$$

Equation (2.38) can be solved directly for $\hat{\mu}$,

$$\hat{\mu} = \sum_{i=1}^{n} B_i^{1-\hat{C}}/An(1-\hat{C}). \qquad\qquad (2.49)$$

Equation (2.39) can be solved directly for $\hat{\sigma}^2$,

$$\hat{\sigma}^2 = \sum_{i=1}^{n}\left[B_i^{1-\hat{C}} - A\hat{\mu}(1-\hat{C})\right]^2/An(1-\hat{C})^2. \qquad (2.50)$$

The value of $\hat{\mu}$ (2.49) can be substituted into (2.50),

$$\hat{\sigma}^2 = \sum_{i=1}^{n}(B_i^{1-\hat{C}} - \sum_{j=1}^{n} B_j^{1-\hat{C}}/n)^2/An(1-\hat{C})^2. \qquad (2.51)$$

The value of $\hat{\mu}$ from (2.49) and the value of $\hat{\sigma}^2$ from (2.51)

can be substituted into (2.46) to obtain an expression which contains only $\hat{C}$,

$$- \sum_{i=1}^{n} \ln(B_i) + \{(1-\hat{C}) \sum_{i=1}^{n} (B_i^{1-\hat{C}} - S_1/n)(B_i^{1-\hat{C}} \ln(B_i))$$

$$- S_1/n(1-\hat{C})) - \sum_{i=1}^{n} (B_i^{1-\hat{C}} - S_1/n)^2\}/$$

$$\{ \sum_{i=1}^{n} (B_i^{1-\hat{C}} - S_1/n)^2(1-\hat{C})/n\} = 0. \qquad (2.52)$$

The expression S_1 is defined by (2.56). This equation must be solved iteratively for $\hat{C}$. The Newton-Raphson method described by (2.16) was used to accomplish this. The expression on the left side of (2.52) corresonds to $\partial f(\underline{X})/\partial \underline{X}$ in (2.16). The iterative procedure used to find $\hat{C}$ is described by

$$\hat{C}_{\nu+1} = \hat{C}_{\nu} - g_{\nu}/g'_{\nu}. \qquad (2.53)$$

The expression on the left side of (2.52) evaluated at $\hat{C}_{\nu}$ is equal to g_{ν}. The derivative of g_{ν} at $\hat{C}_{\nu}$, is equal to g'_{ν}. The expressions g_{ν} and g'_{ν} are described by

$$g_{\nu} = -S_7 + n/(1-C) + (S_1 S_3 - nS_4)/(S_2 - S_1^2/n), \qquad (2.54)$$

$$g'_{\nu} = (2nS_6 - S_3{}^2 - S_1 S_5)/(S_2 - S_1{}^2/n) \, ,$$

$$+ \, 2(2S_1 S_3 S_4 - S_1{}^2 S_3{}^2/n)/(S_2 - S_1{}^2/n)^2 \, + \, n/(1 - C_{\nu})^2 \, , \qquad (2.55)$$

$$S_1 = \sum_{i=1}^{n} B_i{}^{1-\hat{C}_{\nu}} \, , \qquad (2.56)$$

$$S_2 = \sum_{i=1}^{n} B_i{}^{2-2\hat{C}_{\nu}} \, , \qquad (2.57)$$

$$S_3 = \sum_{i=1}^{n} B_i{}^{1-\hat{C}_{\nu}} \ln(B_i) \, , \qquad (2.58)$$

$$S_4 = \sum_{i=1}^{n} B_i{}^{2-2\hat{C}_{\nu}} \ln(B_i) \, , \qquad (2.59)$$

$$S_5 = \sum_{i=1}^{n} B_i{}^{1-\hat{C}_{\nu}} \ln(B_i) \, , \qquad (2.60)$$

$$S_6 = \sum_{i=1}^{n} B_1{}^{2-2C_{\nu}} (\ln(B_i))^2 \, , \qquad (2.61)$$

and

$$S_7 = \sum_{i=1}^{n} \ln(B_i) \, . \qquad (2.62)$$

The algorithm defined by (2.53) through (2.62)
was used as described in the optimization theory section
of this chapter to compute maximum likelihood estimates
of C for each plot. Maximum likelihood estimates of μ
and σ^2 were computed for each plot using (2.49)
and (2.50).

Parameter estimates and measurement of stand
variables for each plot were used to develop equations to
estimate parameter values from stand variables. Equation
(2.63) was used to predict C,

$$C = 0.14198 + 0.66779AH_dN/1000000. \tag{2.63}$$

Simple linear regression was used to estimate coefficients.
The coefficient of determination was 0.271 and the
standard error of estimate was 0.283. Equations (2.63) and
(2.64) were used jointly to estimate μ,

$$A(1-\hat{C})\hat{\mu} = -0.19332 + 0.043100A \tag{2.64}$$

The coefficient of determination was 0.400 and the standard
error of estimate was 0.256. Equations (2.63) and (2.65) were
used jointly to estimate σ^2,

$$\sqrt{A(1-\hat{C})\hat{\sigma}^2} = 0.013473A. \tag{2.65}$$

The coefficient of determination was 0.325 and the standard

error of estimate was 0.107. The use of this method of approximating basal area distributions and the resulting approximate diameter distributions to indirectly estimate yields of forest stands is described later in this chapter.

A Discrete Approximation to Diameter Distribution

A new approach to the estimation of the frequency of occurance of tree diameters based on discrete regression as described earlier in this Chapter is presented in this section. Fundamental to this approach is the diameter class interval system inherent to the measurement of tree diameters at breast height. The roughness of tree bark and irregularity of the shape (departure from circular) of cross-sections of tree boles makes measurement of tree diameters to closer than the nearest 0.1-inch impractical. Measurement of tree diameters is tantamount to grouping the trees into class intervals, each interval 0.1-inch in width. This situation is characteristic of any measurement technique; the width of the class interval is determined by the accuracy of the instruments used to make measurements.

A system of class intervals referenced by the letter j was used to describe the diameter classes. The first diameter class which contained a tree observed on the

136 sample plots was the 1.2-inch diameter class. This class includes diameters greater than 1.15 inches and less than or equal to 1.25 inches. The last

diameter class was the diameter class of the tree with the largest diameter on the 136 sample plots. This was the 156th diameter class which included diameters greater than 16.65 inches and less than or equal to 16.75 inches.

The goal of this new prediction system is to estimate the frequency of occurance of trees in each diameter class from stand variables. This was accomplished by estimating the probability that an observed tree falls in the jth or larger diameter class from stand variables, for all diameter classes which contained observations. A model for such an estimation scheme can be based on the Bernouli trial,

$$Pr(Y_{ijk}) = P_{ij}^{Y_{ijk}}(1-P_{ij})^{1-Y_{ijk}} \qquad Y_{ijk}=0,1. \qquad (2.66)$$

The random variable Y_{ijk} is equal to one if the kth observed tree diameter on the ith plot is in the jth or larger diameter class, and is equal to zero if the diameter is in less than the jth class. The probability P_{ij} depends on j, the class interval which is being

referenced and on the stand variables (of the ith plot)
under which the tree was grown.

This probability might be functionally related to
statistics which describe the diameter distribution such
as the smallest tree diameter on the plot, $D_{(1)}$, the
average tree diameter on the plot, $\overline{D}_i$, and sample quantiles,
Q. A sigmoid relationship between P_{ij} and these statistics
seems reasonable. The sample average can be used to
illustrate this relationship. P_{ij} represents the
probability that a tree is larger than some threshold
value, say t. When $\overline{D}_i$ is much smaller than t a substantial
increase in $\overline{D}_i$ should result in only a small relative
increase in P_{ij} since the number of trees much larger
than $\overline{D}_i$ should be small. When $\overline{D}_i$ and t are not widely
separated, P_{ij} and $\overline{D}_i$ should be approximately linearly
related. When $\overline{D}_i$ is much larger than t, nearly all the
trees are larger than t, therefore P_{ij} is close to one
and again a substantial increase in $\overline{D}_i$ results in only
a small increase in P_{ij}. A model which incorporates
these concepts is the logistic distribution, which can
be written

$$P_{ij} = \frac{1}{1 + e^{\alpha - D_i}} \, . \qquad (2.67)$$

The constant α is positive. The functional form used to describe the relationship between stand variables and each of $D_{(1)}$, $\bar{D}$, and Q in (2.29) and (2.32) can be substituted for $\bar{D}_i$ in (2.67) to obtain

$$P_{ij} = \frac{1}{1 + e^{\beta_{oj} + \beta_{1j}H_{di} + \beta_{2j}A_i N_i + \beta_{3j}H_{di}/N_i}} \, , \qquad (2.68)$$

The coefficients β_{oj}, β_{1j}, β_{2j}, and β_{3j} depend on the jth class diameter. The stand variables on the ith plot are represented by A_i; N_i, H_{di}. Equations (2.66) and (2.68) describe the probability mass function of Y_{ij}.

The likelihood function of Y_{ij} for a fixed value of j is

$$L_j = \prod_{i=1}^{136} \prod_{k=1}^{n_i} Pr(Y_{ijk}) = \prod_{i=1}^{136} \prod_{k=1}^{n_i} P_{ij}^{Y_{ijk}} (1-P_{ij})^{1-Y_{ijk}} \, . \quad (2.69)$$

The likelihood function can be written in a more convenient form by considering the number of observations on the ith plot, n_i and the number of those trees which are in the jth or larger diameter class, m_{ij} ,

$$L_j = \prod_{i=1}^{136} P_{ij}^{m_{ij}} (1-P_{ij})^{n_i - m_{ij}} \, . \qquad (2.70)$$

The log-likelihood function is described by

$$\ln(L_j) = \sum_{i=1}^{136} \left[m_{ij}\ln(P_{ij}) + (n_i - m_{ij})\ln(1-P_{ij}) \right]. \tag{2.71}$$

substituting for P_{ij} from (2.68) results in

$$\ln(L_j) = \sum_{i=1}^{136} \{ -m_{ij}\ln(1 + e^{\beta_{oj} + \beta_{1j}H_{di} + \beta_{2j}A_iN_i + \beta_{3j}H_{di}/N_i})$$

$$+ (n_i - m_{ij})\ln\left[1 - 1/(1 + e^{\beta_{oj} + \beta_{1j}H_{di} + \beta_{2j}A_iN_i + \beta_{3j}H_{di}/N_i}) \right] \}. \tag{2.72}$$

Expression (2.72) is algebraically equivalent to

$$\ln(L_j) = \sum_{i=1}^{136} \{ (n_i - m_{ij})(\beta_{oj} + \beta_{1j}H_{di} + \beta_{2j}A_iN_i + \beta_{3j}H_{di}/N_i)$$

$$- n_i\ln(1 + e^{\beta_{oj} + \beta_{1j}H_{di} + \beta_{2j}A_iN_i + \beta_{3j}H_{di}/N_i}) \}. \tag{2.64}$$

An attempt to maximize this expression with respect to the β's using the Newton-Raphson method described by (2.16) failed. After several iterations the estimated values of the β's became very large or very small, result-ing in machine overflow; the technique diverged.
The method of Fletcher and Powell (1963) described by

(2.17) was then applied. Maximum likelihood estimates

of β_{0j}, β_{1j}, β_{2j}, and β_{3j} were found for each value of

j using this iterative procedure. A subprogram to aid

in the implementation of the scheme suggested by Fletcher

and Powell is described in IBM System/360 Scientific

Subroutine Package Version III Programmer's Manual (1970).

The approximation to diameter distributions is

described by

$$f(D) = 10(\hat{P}_{j+1} - \hat{P}_j), \qquad j/10 + 1.15 < D \le j/10 + 1.25. \quad (2.74)$$

Estimates of P_j, $\hat{P}_j$, are obtained by substituting estimates of

β_{0j}, β_{1j}, β_{2j}, and β_{3j} obtained from the maximization of

(2.73) into

$$\hat{P}_j = \frac{1}{1 + e^{\hat{\beta}_{0j} + \hat{\beta}_{1j}H_d + \hat{\beta}_{2j}AN + \hat{\beta}_{3j}H_d/N}} . \qquad (2.75)$$

The parameter P_1 was set equal to zero, P_{157} was set

equal to one. The use of this method of approximating

diameter distributions in the indirect estimation of

yields of forest stands will be described later in

this chapter.

Prediction of Yield from the Estimated Diameter Distribution

Equation (1.12) has been used successfully to predict
the volume of wood contained in the boles of a stand of
trees. This equation must be generalized to obtain
estimates of the number of board feet of lumber which can
be cut from the boles of a stand of trees. A commonly
used generalization is

$$Y = N \sum_{i=1}^{m} P_i Q_i g(D_i, A, H_d, N).$$ (2.76)

All the symbols except Q_i have been discussed in Chapter
1. The probability that a tree in the ith diameter class
is suitable for the manufacture of lumber (or any other
product under consideration) is represented by Q_i.
Methods of estimating P_i, the probability that a
tree diameter is in the ith diameter class, and
$g(D_i, A, H_d, N)$, an individual tree volume equation,
have been discussed in this chapter. The individual
tree volume equation and an estimation procedure for
Q_i will require further development before estimation
of stand yield will be feasible.

Individual tree yield equations such as (2.2) and (2.4)
cannot be used to estimate individual tree yield since the

coefficients β_2 and β_3 in each equation were estimated for a specific set of stand variables using the model described by (2.1). An alternate method of individual tree yield prediction is based on a tree height prediction equation presented by Burkhart and Strub (1974),

$$\log_{10}H_p = 0.53815 + 0.77975 \log_{10}(H_d) - 1.17713/A$$
$$+ 0.35468 \log_{10}(N)/D + 4.11014/(AD) - 2.10285/D .$$

$$(2.77)$$

The coefficient of determination for this least squares fit was 0.911 and the standard error of estimate was 0.040. The data set described earlier in this Chapter was used to estimate coefficients. Estimates of tree height, H_p, based on stand variables, A, H_d, and N, and tree diameter at breast height, D, can be substituted into (1.3) and (2.3) to obtain the following individual tree equations,

$$V = \beta_0 + \beta_1 D^2 10^{(0.53815+0.77975 \log_{10}(H_d)-1.17713/A}$$
$$+0.35468 \log_{10}(N)/D + 4.11014/(AD)-2.10285/D),$$

$$(2.78)$$

$$TV = \beta_0+\beta_1 D^{-2} 10^{(0.53815+0.77975 \log_{10}(H_d)-1.17713/A}$$
$$+0.35468 \log_{10}(N)/D+4.11014/(AD)-2.10285/D).$$

$$(2.79)$$

Estimates of β_0 and β_1 are contained in Table 2.1. These individual tree yield equations can be used in (2.76) to predict stand yield. Alternatively, two new yield equations, (2.5), and (2.6), developed earlier in this chapter can be used. These new equations offer greater simplicity, and are well suited to use in a new indirect yield projection procedure that will be developed later in this chapter.

The probabilities, Q_i, were estimated from individual tree data. When each tree diameter was measured, the person conducting the inventory classified the tree as suitable or not suitable for manufacture of lumber. The estimates of Q_i, $\hat{Q}_i$, were determined by dividing the number of trees measured which fell in the ith diameter class and were suitable for the manufacture of lumber by the total number of trees measured which fell in the ith diameter class.

Diameter Class (i)	$\hat{Q}_i$
Less Than 8	0
8	0.3246
9	0.5322
10	0.9385
11	0.9851
Greater Than 11	1

The class interval width used in the above estimation procedure was one inch. The ith diameter class included trees greater than i-0.45 inches, but less than or equal to i+0.55 inches. This completes the information required to estimate yield via (2.76). The results of a comparison of yields estimated with (2.76) and other yield estimation techniques is contained in Chapter 3.

A new class-interval-free approach to estimation of yield from diameter distributions will be developed from the procedure described by (2.76). A similar approach has been published by Strub and Burkhart (1975). The class interval system inherent in (2.76) is artificially constructed with the class interval limits being arbitrarily chosen. A class-interval-free system for yield prediction can be obtained by examining the limiting form of (2.76) as the number of class intervals, m, increases without bound, the width of individual class intervals approaches zero, and the product of the number of class intervals and class interval width remains constant. This procedure is represented symbolically by

$$Y = \lim_{m \to \infty} N \sum_{i=1}^{m} P_i Q_i g(D_i, A, H_d, N). \qquad (2.80)$$

The probability P_i is obtained by integrating the estimated diameter distribution, say f(D), over the range of the ith diameter class. This is represented symbolically by

$$P_i = \int_{D_i - w/2}^{D_i + w/2} f(D) dD. \qquad (2.81)$$

The symbol w represents the width of a class interval. Substituting this expression into (2.80) results in

$$Y = \lim_{m \to \infty} N \sum_{i=1}^{m} \int_{D_i - w/2}^{D_i + w/2} f(D) dD Q_i g(D_i, A, H_d, N). \qquad (2.82)$$

The restriction that mw be constant dictates that m and w are functionally related as described in

$$w = (U - L)/m. \qquad (2.83)$$

The symbol L represents the lower limit of the 1st class interval and U represents the upper limit of the mth class interval. Substituting this expression into (2.82) results in

$$Y = \lim_{m \to \infty} N \sum_{i=1}^{m} \int_{D_i - (U-L)/(2m)}^{D_i + (U-L)/(2m)} f(D) dD Q_i g(D_i, A, H_d, N). \qquad (2.84)$$

It is now necessary to assume that Q_i is a continuous function of D_i, say $q(D_i)$,

$$Y = \lim_{m \to \infty} \sum_{i=1}^{m} \int_{D_i-(U-L)/(2m)}^{D_i+(U-L)/(2m)} f(D)dD q(D_i) g(D_i,A,H_d,N). \qquad (2.85)$$

This assumption will be discussed later in this section. The definition of the Riemann integral ensures that (2.85) has the same limit as

$$Y = \lim_{m \to \infty} N \sum_{i=1}^{m} f(D_i) \Delta D_i q(D_i) g(D_i,A,H_d;N). \qquad (2.86)$$

The expression ΔD_i is equal to the class interval width. Another application of the definition of the Riemann integral results in

$$Y = N \int_{L}^{U} f(D) q(D) g(D,A,H_d,N)dD. \qquad (2.87)$$

The estimation of stand yield in cubic feet of solid wood in the tree boles when the Weibull approximation to diameter distributions is utilized will be used as an example of the application of this new technique. In this case $q(D) = 1$ for all values of D. Substitution of (2.5) for $g(D, A, H_d, N)$ and (2.20) for $f(D)$ results in

$$Y = N \int_{0}^{\infty} (\gamma \delta D^{\delta-1} e^{-\gamma D^{\delta}}) \quad (\beta_0 + \beta_1 D^2 H_d)dD. \qquad (2.88)$$

Analytical evaluation of the integral results in

$$Y = N\beta_0 + N\beta_1 H_d (1/\gamma)^{2/\delta} \Gamma(1 + 2/\delta). \qquad (2.89)$$

The appropriate estimates of β_0 and β_1 are listed in Table 2.1. They are respectively -0.06136 and 0.00187. Equations (2.27) and (2.28) or (1.10), (1.11) and (2.29) or (2.32), (2.33) and (2.34) can be used to estimate γ and δ. These estimates are substituted into (2.89). The resulting value of Y is an indirect estimate of the average yield per acre of the volume of solid wood contained in the boles of trees in the stand.

When the trees in a stand are to be used for the production of pulpwood, logs smaller than 4-inches in diameter cannot be economically utilized. Trees with a diameter at breast height smaller than 4.55-inches cannot be economically utilized for production of paper. Estimation of stand yield is accomplished in this case by choosing an individual tree volume equation which predicts cubic foot volume to a 4-inch merchantability limit, such equations are described earlier in this chapter, (2.5) or (2.78), and coefficient estimates are contained in Table 2.1. The function q(D) takes on the following values,

$$q(D) = \begin{cases} 0 & D \leq 4.55 \\ 1 & D > 4.55. \end{cases} \qquad (2.90)$$

The diameter distribution, f(D), can be estimated via one of the three schemes described earlier in this Chapter. When all of these functions are substituted into (2.87), analytical evaluation of equation (2.87) was not possible. The approximate value of the integral was obtained in that case with the trapazoid rule. This numerical technique is described by (2.82)

$$\int_{L}^{U} h(x)dx = \frac{w}{2}[h(L)+2h(L+W)+.\ .\ .+2h(U-W)+h(W)]. \quad (2.91)$$

The symbol h(x) represents the function being integrated, w is the class interval width. The class interval width used in the numerical approximations found in this dissertation is 0.01 inches. A smaller class interval results in costly increases in computer time required to calculate yield estimates, a larger class interval might result in large errors due to the approximation.

A third indirect estimation procedure for predicting board foot volume per acre (the average number of board feet which can be cut from the boles of the trees on an acre of land in a stand) was developed. The individual tree volume equations, (2.5) or (2.78), were developed earlier in this chapter. They are also summarized in Table 2.1. Any of the three diameter distribution estimation procedures can be used. Choice

of a suitable functional form for q(D) is more complex

than the previous two cases. Trees with a diameter at

breast height less than or equal to 7.55 inches are not

suitable for manufacture of lumber. The function q(D)

assumes the value zero over this range. Trees larger

than 7.55 inches may or may not be suitable for manufacture

of lumber. Plotting individual tree dbh versus relative

frequency of trees with the same dbh which were suitable

for manufacture of lumber indicates that an appropriate

choice for q(D) is

$$q(D) = \begin{cases} 0 & D \leq 7.55 \\ 1 - e^{\beta(7.55-D)} & D > 7.55, \end{cases} \qquad (2.92)$$

This continuous function of D is bounded between zero

and one. The function is monotone increasing and has

an upper asymptote of one. The derivative of q(D) is

defined for all values of D except D equal to 7.55. The

technique of discrete regression described earlier in

this chapter was used to estimate β. Data used to

estimate β were described earlier in this chapter

and consisted of the tree diameter at breast

height of each tree over 7.55 inches measured to the

nearest 0.1-inch plus an estimation of whether or not

the tree is suitable for production of lumber in the opinion of the person measuring the tree.

The theory developed in the discrete regression section of this chapter can be applied to the estimation of β in

$$P(Y) = \left[1 - e^{\beta(7.55-D)}\right]^{Y}\left[e^{\beta(7.55-D)}\right]^{1-Y} \qquad Y=0,1. \qquad (2.93)$$

The random variable Y assumes the value one if a tree D inches in diameter is suitable for the manufacture of lumber and Y assumes the value zero if the tree is not suitable for the manufacture of lumber. This equation was obtained by substituting the function describing q(D) in (2.92) for $g(\underline{X}, \underline{\theta})$ in (2.8). The log-likelihood function can be obtained in a similar fashion from (2.10),

$$\ln(L) = \sum_{i=1}^{n} Y_i \ln\left[1 - e^{\beta(7.55-D_i)}\right] + \sum_{i=1}^{n} (1-Y_i)\beta(7.55-D_i). \qquad (2.94)$$

The diameter of the ith tree measured is represented by D_i, $Y_i = 0$ if the ith tree measured was not suitable for manufacture of lumber and $Y_i = 1$ if the ith tree was suitable for manufacture of lumber. An equation equivalent to (2.94) which has greater computational efficiency is

$$\ln(L) = \sum_{j=1}^{k} m_j \ln\left[1 - e^{\beta(0.05-j/10)}\right] + \sum_{j=1}^{k} n_j \beta(0.05-j/10). \quad (2.95)$$

The symbol m_j represents the number of trees between 7.45 + j/10 and 7.55 + j/10 inches in diameter which were suitable for production of sawtimber, n_j represents the number which were not suitable for production of sawtimber, and k represents the number of diameter classes observed. The Newton-Raphson method of optimization can be used to find the value of β which maximizes $\ln(L)$. This is the maximum likelihood estimate of β and was found to be 1.2286 for the data collected on 186 sample plots.

It has been previously mentioned that data for this study were contributed by several forest industries. The personnel that gathered data, and therefore estimated whether or not a tree is suitable for the production of lumber, vary from company to company. It was hypothesized that since personnel vary from company to company, perhaps β also varies from company to company. The data collected by Glatfelter Pulp Wood Company was divided into data collected on the mainland and data collected on the eastern shore of Maryland and Virginia, and treated separately. The theory presented in the discrete regression section of this chapter was used to test this hypothesis.

Maximum likelihood estimates of β for the data collected by each company, and the value of the log-likelihood equation for each of those estimates were obtained by maximizing (2.95) for the data collected by that company. Six estimates were obtained for the five companies (two estimates for Glatfelter) which contributed data. This information was used to test the following hypothesis,

$$H_0 : \beta_i = \beta \qquad i = 1. \ . \ . \ ., 6$$

$$H_1 : \beta_i \neq \beta_j \qquad \text{for some } i, j; i \neq j.$$

The estimates of β_i, the parameter value for the ith company, and the optimal value of $\ln(L)$, $\ln[L_i(\beta_i)]$, are listed below for each company and for the case when β is constant over all companies (pooled).

Company	$\hat{\beta}$	$\ln[L(\hat{\beta})]$
1	0.97141	2834.8
2	0.49181	1182.3
3	1.79125	3450.7
4	1.96346	852.0
5	0.06035	38.6
6	0.68680	133.0
All Companies	1.22861	8857.1

The test statistic suggested, following (2.14), in the discrete regression section of this chapter is

$$-2\ln(\Lambda^*) = \ln[L_p(\hat{\beta}_p)] - \sum_{i=1}^{6} \ln[L_i(\hat{\beta}_i)]. \qquad (2.96)$$

where $\ln[L_i(\hat{\beta}_i)]$ indicates the maximum value of the log-likelihood function for the data from the ith company, and $\ln[L_p(\hat{\beta}_p)]$ represents the maximum of the log-likelihood function for the pooled data. Evaluation of (2.96) resulted in a test statistic of 365.7.

A sample size of 2200 necessitates the use of an asymptotic test. It has been demonstrated in the discrete regression section that the asymptotic distribution of the test statistic is chi-square with five degrees of freedom. The hypothesis can be rejected at the 0.001 level of significance. For this reason a different function, q(D) was used for prediction of board feet yield for each company. The function used for each company was (2.92) with a different estimate of β for each company.

After q(D) was determined board feet yields were estimated via (2.87) and (2.93). In addition topwood yield per acre was estimated. Topwood included wood found between the 6-inch top diameter merchantability limit for sawtimber and the 4-inch top diameter merchantability

limit for pulpwood plus wood found in trees not suitable
for the production of lumber. Each of the two components
of topwood was estimated separately.

The component of topwood consisting of wood found
between the 6-inch merchantability limit for sawtimber
and the 4-inch merchantability limit for pulpwood was
estimated first. Individual tree volume equations of the
form (2.6) or (2.79) were used. Equation (2.92) was used to
estimate $q(D)$, using a different value of β for yield
estimates for each company. Each of the four diameter
distribution approximations were used. These functions
were substituted into (2.87) and final evaluation of
yield was accomplished via (2.91).

The second component of topwood yield was contained
in the trees which were not suitable for production of
lumber. Individual tree volume equations used to indirectly
estimate topwood yield were of the form of (2.5) and (2.78).
The proportion of the stand not suitable for production of
sawtimber can be estimated as one minus the proportion of
the stand which is suitable for sawtimber.

$$q(D) = \begin{cases} 0 & D \leq 4.55 \\ 1 & 4.55 < D \leq 7.55 \\ e^{\beta(7.55-D)} & D > 7.55 \quad (2.97) \end{cases}$$

The approximation to the diameter distribution can be any one of the four developed earlier in this chapter. Equations (2.86) and (2.91) were used to obtain estimates of this component. The topwood yield estimate was the sum of the estimates for each of the components. A comparison of these indirect techniques for estimating topwood yield and the direct method of estimating topwood yield is found in chapter 3. This comparison is based on the 50 randomly selected plots.

Comparison Techniques for Evaluating Yield Prediction

Kolmogorov-Smirnov Tests

Bailey (1972) suggested use of the Kolmogorov-Smirnov test to check the validity of an assumed diameter distribution. This test is described in chapter 4 of Gibbons (1971). The test is used to compare the empirical distribution function (in this case calculated from the diameter measurements made on a plot) with some hypothesized distribution function. The null hypothesis tested is that the hypothesized distribution is the true underlying probability distribution of the population

which was sampled. When applied to testing diameter
distributions the null hypothesis becomes the hypothesized
diameter distribution (such as the Weibull or the trans-
formed normal) is in fact the true diameter distribution.
One assumption of this test is that the hypothesized
distribution be completely specified, including specifica-
tion of all parameters. When Bailey applied this test,
he violated this assumption by estimating parameters of
the hypothesized diameter distribution (the Weibull)
from the data used to calculate the empirical distribu-
tion function. The violation of this assumption results
in an extremely conservative test, that is the probability
of a type I error is much less than specified.

Bailey used this procedure on 63 plots to perform
tests comparing two and three-parameter Weibull distribu-
tion to the observed diameter distribution. Several
estimation schemes including maximum likelihood were used
to estimate parameters of the hypothesized distribution
from the sample data. The proportion of the 63 plots on
which the null hypothesis was rejected was compared with
the probability of a type I error specified for all the
tests. If this proportion was less than the probability
of a type I error it was concluded that the null distribu-
tion was the correct approximation to the observed

diameter distribution. There is some danger in following this procedure. It is known that in this case the test is conservative, but the true probability of a type I error is not known.

The danger of applying the Kolmogorov-Smirnov test as Bailey suggests was examined via computer simulation. Simulation techniques described by Box and Muller (1958) were used to generate 1,000 samples, each containing 80 standard normal pseudo random numbers. Three different methods were used to test the hypothesis: the variates are normally distributed. The first test was the Kolmogorov-Smirnov test described by Gibbons (1971), for this test the hypothesized distribution was the standard normal, that is a mean of zero and a variance of one were specified. The second test was the test suggested by Bailey (1972). The hypothesized distribution was the normal, however $\overline{X}$ and S^2 were calculated from each sample and used to estimate the mean and variance of each of the 50 populations sampled. A third test suggested by Stephens (1974) was applied to the same data. Stephens presents modifications for the critical region of the Kolmogorov-Smirnov test when the null hypothesis is that the random variables are normally distributed, and the population mean and variance are estimated by $\overline{X}$ and S^2. Further discussion of the properties of this test was presented by Green and Hegazy (1976).

The number of samples for which the null hypothesis was rejected (hence it was concluded that the variates were not normally distributed) was observed for each of the three tests when the probability of a type I error was 0.1, 0.05, and 0.01. The results of these tests are listed in Table 2.4. If the test is operating correctly, we would expect the proportion of samples for which the null hypothesis was rejected to be approximately equal to the probability of a type I error. Table 2.4 indicates that the test suggested by Bailey is extremely conservative, and that the other two tests are not conservative.

The normal approximation to basal areas introduced earlier in this chapter and described by (2.38) was used to illustrate the danger of using the Kolmogorov-Smirnov test as suggested by Bailey. The data collected on each of the 186 sample plots were used to test the hypothesis: tree basal areas are normally distributed. The mean and variance of the basal areas was estimated for each plot by $\bar{X}$ and S^2. The estimated mean and variance were used to specify the null distribution. The results of the tests suggested by Bailey and described in Table 2.5 seem to indicate that basal areas are normally distributed since the proportion of plots on which the null hypothesis was rejected is less than the probability of a type I error. However, as we have seen in the previous example,

Table 2.4. A comparison of three tests for normality from the simulation of 1000 standard normal samples of size 80.

| Test | Probability of Type I Error | | |
| | 0.1 | 0.05 | 0.01 |
	Number of Samples Rejected		
Kolmogorov-Smirnov (normal distribution with mean = 0 and variance = 1 specified)	92	53	8
Test suggested by Bailey (1972) (normal distribution specified; mean and variance estimated by $\overline{X}$ and S^2)	0	0	0
Stephens (1974) modification of the Kolmogorov-Smirnov test (normal distribution specified, mean and variance estimated by $\overline{X}$ and S^2)	100	48	5

Table 2.5. A comparison of tests for normality of tree basal areas on 186 sample plots from old-field loblolly pine plantations.

| | Probability of a Type I Error | | |
| | 0.10 | 0.05 | 0.01 |
	Number (and proportion) of Plots Rejected		
Test suggested by Bailey (1972 (normal distribution specified, mean and variance estimated by $\overline{X}$ and S^2)	11(0.052)	5(0.027)	1(0.005)
Stephens (1974) modification of the Kolmogorov-Smirnov test (normal distribution specified, mean and variance estimated by $\overline{X}$ and S^2)	52(0.280)	36(0.194)	20(0.108)

the test described by Bailey is extremely conservative.
When Stephens' modification of the Kolmogorov-Smirnov
test was applied to the data it was concluded that
basal areas are not normally distributed. The results
of these tests, given in Table 2.5, are that the null
hypothesis was rejected on a larger proportion of the
plots than the probability of a type I error, hence the
conclusion that tree basal areas are not normally
distributed. The conservativeness of the test suggested
by Bailey resulted in an incorrect conclusion.

Stephens (1974) presents modifications of the
Kolmogorov-Smirnov test for the cases of normal and
exponential null distributions. The modified tests are
not distribution-free since a new critical region must
be derived for each null distribution. Tests appropriate
for the Weibull distribution function, (2.20), the trans-
formed normal distribution function, (2.43), and the
discrete approximation described by (2.74) were not
available. A second problem encountered with the Stephens
modification of the Kolmogorov-Smirnov test is that this
procedure does not detect inaccuracy in prediction of
diameter distributions resulting from estimation of
parameters from stand variables. Such inaccuracy
effects the accuracy of approximations of diameter
distributions and yield estimates. A scheme

for making such a comparison between estimated and observed diameter distributions should take into account error introduced from estimation of parameters from stand variables.

Such a comparison scheme is based on the 50 plots removed at random from the total sample. The remaining data, 136 plots, were used to develop schemes for estimating parameters from stand variables. One of the diameter distribution prediction schemes developed earlier in this chapter can be used to estimate the diameter distribution on each of the 50 randomly selected plots from stand variables. In this manner the hypothesized diameter distribution is completely specified for each plot without using the tree diameter measurements from the 50 plots. Observed and estimated diameter distributions were compared via the Kolmogorov-Smirnov test in this manner. The results of these tests indicate that perhaps none of the diameter distribution schemes developed in this chapter are completely adequate; these results are described in detail in the following chapter.

Comparison of Observed and Estimated Yields

Each of the yield prediction techniques presented in this chapter were implemented on the 136 sample plots and the results were used to estimate the yield on the 50 randomly selected sample plots from the associated stand

variables. The average differences between observed and expected yields as well as the variances of the differences are presented in the next chapter.

Comparison of Culmination of Mean Annual Increment

The mean annual increment of a forest stand is the yield per acre of the stand at a particular age divided by that age. The point at which the mean annual increment culminates (the age at which mean annual increment is maximized) has been used as a criterion to determine rotation length (the age at which the trees should be harvested). Davis (1966, on pages 226 and 227) describes the behavior of mean annual increment.

> "Several items of significance hold in general regardless of the source and kind of data. The first is that on any given site the indicated rotation for maximum board foot production is longer than that for cubic foot production. This is because of the nature of the board foot unit; large trees contain proportionately more board feet than do small trees in comparison to cubic foot measure. Also, board foot measure does not contain all trees, only those above some specified minimum diameter limit. The second is that the rotation of maximum mean annual increment increases with decreasing site quality. As would be expected, this trend is much more strongly evident in board foot than in cubic foot growth. In addition to producing much less timber, the poorer sites take longer to reach their point of maximum mean annual increment."

Davis also indicates that loblolly pine stand yield in cubic feet culminates between 30 and 35 years whereas board foot yield culminates between 40 and 50 years. Davis does not indicate whether these ages are for planted or natural stands. The age of culmination of mean annual increment for a variety of stand variables was calculated with each of the more promising yield prediction techniques. These results are compared with the notions described by Davis.

A direct and several indirect approaches to stand yield estimation have been developed in this chapter. A method for comparing the accuracy and precision of each of the yield estimation techniques has been suggested. The results of this comparison are described in chapter 3. A second criterion for comparing yield estimation techniques is to examine the age of culmination of mean annual increment for various combinations of stand variables. The behavior of the age of culmination of mean annual increment should behave in the manner described by Davis (1966). The results of this comparison are presented in chapter 3. A third method of comparison has been suggested for the indirect yield estimation techniques, which supply the forest manager with addition information, namely an approximation to the frequency distribution of tree diameters. A method for comparing

the accuracy of the diameter distribution approximations
has been developed in this chapter and the results of
this comparison are also presented in chapter 3.

III. RESULTS AND DISCUSSION

Kolmogorov-Smirnov Tests

Six diameter distribution estimation schemes were
presented in Chapter II. The first three diameter
distribution estimation schemes involved use of the
Weibull density density function, (2.20). The first
of these involved estimating the parameters of
the Weibull density function with (2.27) and (2.28). The
second technique involved estimating the 24th and 93rd
quantiles of the diameter distribution from stand variables,
and then estimating the parameters of the Weibull distribu-
tion function from the quantiles, (1.10) and (1.11). The
third scheme involved estimating the smallest and average
diameter from stand variables, and then estimating Weibull
parameters from (2.35) and (2.36). The fourth scheme
involved assuming that tree basal areas follow a normal
distribution, estimating parameters from stand variables
via (2.41), and (2.42), and obtaining the diameter
distribution from transformation (2.37). The fifth method
was to assume tree basal areas follow a transformed normal
distribution, estimate parameters from (2.63), (2.64), and
(2.65), and obtain the diameter distribution from trans-
formation (2.37). The last scheme was to estimate the
diameter distribution with the discrete approach described
by (2.74), and (2.75).

The six techniques were used to estimate the diameter distribution on each of the 50 randomly selected plots. Diameter measurements for the 50 plots were not used in the estimation procedures. These estimated diameter distributions were compared with the observed diameter distributions by applying the Kolmogorov-Smirnov test to the data on each of the 50 plots. The results of this comparison are summarized in Table 3.1.

Examination of Table 3.1 indicates that among the three Weibull schemes the estimators based on the smallest and average tree diameters perform the worst. The other two Weibull schemes are consistently better. No clear advantage is evident in the two remaining Weibull schemes. The normal approximation to basal area distributions performed as well or better than any of the Weibull schemes. The transformed normal approach is by far the worst of all the approximations tested. Further discussion of this result will appear in chapter 4. Performance of the discrete approximation to diameter distributions was approximately the same as the normal basal area approach. No one scheme exhibits a clear advantage over all the other schemes. These tests also indicate that perhaps none of the diameter distribution estimation techniques considered thus far adequately describes diameter distributions. An alternate method of estimating diameter distributions will be developed in Chapter 4.

Table 3.1. Frequency of rejection for Kolmogorov-Smirnov tests comparing estimated and observed diameter distributions on the 50 randomly selected plots

Estimation Procedure	Significance Level (α)				
	0.15	0.10	0.05	0.025	0.01
Weibull Distribution Function -- Parameter Estimation Equations					
(2.27) and (2.28)	25(0.50)*	22(0.44)	20(0.40)	17(0.34)	15(0.30)
(1.10), (1.11) and (2.29)	26(0.52)	25(0.50)	23(0.46)	16(0.32)	11(0.22)
(2.32), (2.35) and (2.36)	32(0.64)	27(0.54)	23(0.46)	22(0.44)	16(0.32)
Normal Distribution Function -- Parameter Estimation Equations					
(2.41) and (2.42)	27(0.45)	24(0.48)	23(0.46)	12(0.24)	16(0.32)
Transformed Normal Distribution Function -- Parameter Estimation Equations					
(2.63), (2.64) and (2.65)	39(0.78)	37(0.74)	35(0.70)	33(0.66)	31(0.62)
Discrete Approximation to Diameter Distributions	25(0.50)	24(0.48)	18(0.36)	17(0.34)	14(0.28)

*The first entry is the number of plots on which the null hypothesis was rejected; the second entry in parenthesis is the first entry divided by 50, or the proportion of the plots upon which the null hypothesis was rejected.

Comparison of Observed and Estimated Yields

The direct and indirect yield prediction techniques
presented in chapter 2 were used to estimate yield from
stand variables on each of the 50 randomly selected plots.
The Weibull distribution functions and each of the three
associated parameter estimation schemes were used to
estimate yield. The transformed normal distribution
function and the discrete approximation to diameter
distributions were both used to estimate yield. Table
3.2 contains the results of this yield comparison.

The first row of Table 3.2 contains the comparison
of observed yields and yields estimated directly from
stand variables by use of multiple linear regression.
Equations (1.4) and (1.6) were used to obtain the yield
estimates. Estimates of coefficients in (1.4) are found
in Table 2.2. The estimates of β_0 and β_1 in (1.6) used in
yield estimation are, respectively, -0.16822 and 0.85685.

The next three rows of Table 3.2 contains a compari-
son based on the indirect method of yield estimation
utilizing the Weibull distribution and based on the class
interval method of yield estimation with 1-inch diameter
class intervals. The yield prediction procedure is
described by (2.76). Individual tree yield equations
(2.78) and (2.79) were substituted into (2.76). Each of
the three methods of parameter estimation was used to
obtain the three comparisons.

Table 3.2. Means and variances of differences between observed and estimated yields on 50 randomly selected plots

Estimation Procedure	Total Cubic Foot Volume Inside Bark	Cubic Volume Inside Bark to 4-inch Top	Board Foot Volume	Topwood Volume to a 4-inch Top
Direct Methods				
Yield Equations (1.4) and (1.6)	12.1* 297.7	53.9 318.9	77.0 1007.2	39.6 280.4
Indirect Methods				
Weibull Distribution Function Parameter Estimation and Yield Prediction Equations				
(2.27),(2.28),(2.76),(2.78) and (2.79)	-60.3 493.9	-69.6 501.6	-560.2 2565.8	34.1 262.1
(1.10),(1.11),(2.36),(2.76), (2.78) and (2.79)	-23.2 305.5	-36.5 327.7	56.9 1062.8	-51.7 265.3
(2.32),(2.34),(2.36),(2.76), (2.78) and (2.79)	105.5 297.4	86.0 319.2	673.9 1160.4	-60.3 272.7
(2.5),(2.6),(2.27),(2.28), and (2.87)	-122.3 489.9	-125.9 499.7	-740.4 2341.1	-26.8 248.6
(2.5),(2.6),(1.10),(1.11), (2.29) and (2.87)	-97.1 326.2	-105.7 343.7	-203.4 1035.4	-60.3 254.6
(2.5),(2.6),(2.32),(2.35), (2.36) and (2.87)	18.1 306.7	-0.5 328.2	498.4 1026.7	109.8 287.4

Table 3.2 (Cont)

Estimation Procedure	Total Cubic Foot Volume Inside Bark	Cubic Volume Inside Bark to 4-inch Top	Board Foot Volume	Topwood Volume to a 4-inch Top
Normal Distribution Function Parameter Estimation and Yield Prediction Equations				
(2.5),(2.6),(2.41),(2.42), and (2.87)	-67.3 382.4	-98.3 395.9	-42.6 1046.3	-87.8 307.5
Transformed Normal Distribution Function -- Parameter Estimation and Yield Prediction Equations				
(2.5),(2.6),(2.63),(2.64), (2.65) and (2.87)	-953.9 1038.4	-1016.4 1018.6	-6575.8 6281.8	348.1 517.0
Discrete Approximation to Diameter Distributions Parameter Estimation and Yield Prediction Equations				
(2.76),(2.78) and (2.79) (0.1-inch class intervals)	-27.0 306.4	-32.4 324.6	-66.3 1126.7	-14.5 249.5
(2.78),(2.79) and (2.87)	-27.0 306.4	-32.4 324.6	-60.0 1106.6	-14.2 234.7

*The first entry is the average difference (observed plot yield minus estimated plot yield); the second estimate is the variance of the differences.

The fifth through seventh rows of Table 3.2 are based on estimates calculated from the Weibull distribution function and the class-interval-free approach to yield estimation. The yield prediction procedure is described by (2.87). Individual tree yield equations (2.5) and (2.6) were substituted into (2.87). Again each of the three methods of parameter estimation was used to obtain the three comparisons.

The normal basal area approach was used to estimate yields described in row eight of Table 3.2. Parameters were estimated from (2.41) and (2.42), then substituted into (2.38). The estimated diameter distribution was derived from this estimated basal area distribution, (2.38), and transformation (2.37). The estimated diameter distribution and individual tree volume equations (2.5), and (2.6) were substituted into (2.87) to obtain indirect yield estimates.

The ninth row of Table 3.2 contains a comparison based on the transformed normal distribution function and the class-interval-free approach to yield prediction. Yield estimation equation (2.87) and individual tree yield equations (2.5) and (2.6) were used to calculate yield estimates.

The last two rows of Table 3.2 contain a comparison based on the discrete approximation to diameter

distributions. The comparison contained in the tenth row is based on the class interval approach to yield prediction. The class interval width used in estimation was 1/10-inch. Equation (2.76) describes the procedure used. Individual tree volume equations (2.78) and (2.79) were substituted into (2.76). The average number of trees suitable for production of lumber by diameter class on the 186 sample plots was substituted for Q_i. The 11th row of Table 3.2 is a comparison based on the class-interval-free method of yield prediction. This procedure is described by equation (2.87). Individual tree yield equations (2.78) and (2.79) were used. The only difference between yield estimates used in rows nine and ten is the method of estimating the probability that a tree is suitable for production of lumber. Equation (2.92) was used in row ten. A comparison of rows nine and ten indicates that (2.92) is preferred over class averages for estimating the probability that a tree is suitable for production of lumber from tree dbh.

The direct method of yield estimation performed nearly as well as any of the indirect yield estimation techniques for estimation of all products considered. The Weibull distribution function, quantile estimation of parameters, and the class interval yield estimation procedure with one inch class intervals was also

quite accurate in predicting total cubic foot volume
and cubic foot volume to a 4-inch top. The normal basal
area approach provided accurate and precise estimates
of board foot yield, however, total cubic volume,
and cubic volume to a 4-inch top estimates were less
accurate, and topwood volume estimates were not nearly
as accurate as estimates obtained via direct estimation
of topwood yield. The discrete diameter distribution
approach provided accurate yield estimates of all
products. The transformed normal distribution proved
a poor choice for estimating yield of loblolly pine stands.

Comparison of Age of Culmination of Mean Annual Increment

Four of the most accurate (as determined in the
previous section) yield estimation procedures were
compared. The first estimation procedure was the direct
method of estimating yield from stand variables,
described by (1.4). The second and third estimation
procedures were indirect techniques based on the Weibull
distribution function. The second procedure made use of
quantile parameter estimates, from equations (1.10), (1.11),
and (2.29), and the 1-inch-class-interval method of yield
estimation. The third technique made use of parameter
estimates based on the smallest and largest diameters,

from equations (2.32), (2.35), and (2.36), and the class-interval-free method of yield estimation. All of these yield estimation procedures appear in the comparison described by Table 3.1.

Each of these yield estimation methods requires input of the stand variables age (A), average height of the dominant and codominant stand (H_d), and existing number of trees per acre (N). Yield at various ages must be calculated to determine the age at which mean annual increment culminates, hence H_d and N must also be estimated for each age for which a yield estimate is calculated. Estimation of H_d is accomplished by assuming a site index (S) and then using site index curves to estimate H_d from A. Estimation of N is accomplished by assuming an initial number of trees per acre at age zero (N_0) and then using survival curves to estimate N from A.

The site index curves used in this part of the study were developed by Burkhart et al. (1972). They can be used to estimate H_d from A and S,

$$H_d = S(10)^{0.23461 - 5.86537/A} \tag{3.1}$$

Site index, S, is a measure of the growing quality of the land upon which the trees are planted, in particular

S is the average height of the dominant and codominant

trees at age 25.

The survival curves used in this study were developed

by Lenhart and Clutter (1971). These survival curves are

$$N = N_0 \, \Phi(9.3745-0.67637 \, \log_{10}(A)-0.96269 \, \log_{10}(N_0)) \quad (3.2)$$

The symbol $\Phi(\cdot)$ represents the standard normal distribu-

tion function.

The results of the comparison of the age of

culmination of mean annual increment for the different

yield estimation techniques appear in Tables 3.3, 3.4, 3.5,

3.6, and 3. 7. None of the yield estimation systems behaves

exactly as described by Davis (1966).

The direct method of yield estimation of cubic

foot yield does not culminate before age 50 on poor

sites. Culmination is early on the better sites. The

board foot volumes culminate too early on all sites. The

relationship between site index and age of culmination

of mean annual increment is monotone decreasing as

described by Davis.

The indirect method utilizing the Weibull distribu-

tion function, quantile estimators, and 1-inch class

intervals culminated at approximately the correct age

for cubic foot and board foot yield. This method does

Table 3.3. Age at which mean annual increment culminates for various site index and initial density combinations when yields are estimated directly from stand variables via multiple regression equation (1.4).

Site Index	Initial Number of Trees Per Acre			
	600	800	1000	1200
	Cubic Volume to a Four-Inch Top			
40	50+(72.3)*	50+(74.7)	50+(76.1)	50+(76.8)
50	50+(94.9)	50+(98.1)	50+(99.9)	50+(100.9)
60	25(124.0)	26(123.1)	50+(131.2)	50+(132.4)
70	20(190.5)	23(187.8)	23(184.5)	23(180.8)
80	20(297.2)	21(292.4)	21(286.4)	21(279.9)
90	19(468.2)	20(459.7)	20(449.6)	20(438.8)
	Board Foot Volume			
40	36(22.3)	37(15.3)	38(10.8)	39(7.9)
50	32(56.9)	32(38.7)	33(27.2)	33(19.6)
60	29(155.3)	29(105.0)	29(73.2)	30(52.4)
70	26(446.9)	27(300.4)	27(208.3)	27(148.0)
80	25(1336.9)	25(895.4)	25(617.8)	25(436.5)
90	23(4123.1)	24(2746.8)	24(1889.7)	24(1331.2)

*The first entry is the age at which mean annual increment culminates, the second entry is the mean annual increment at this age. An entry of 50+ indicates that mean annual increment culminates after age 50, in this case the second entry is the annual increment at age 50.

Table 3.4. Age at which mean annual increment culminates for
for various site index and initial density
combinations when yields are estimated
indirectly with the Weibull distribution
function, quantile approach to parameter
estimation, and 1-inch class interval yield
estimation procedure

Site	Initial Number of Trees Per Acre			
Index	600	800	1000	1200
	Cubic Volume to a Four-Inch Top			
40	30(45.5)*	29(45.2)	28(44.5)	27(43.4)
50	30(81.5)	29(82.7)	28(83.5)	27(83.6)
60	31(130.3)	29(133.2)	28(135.9)	27(138.0)
70	32(193.5)	30(198.5)	29(203.6)	28(208.1)
80	32(272.9)	31(280.3)	30(288.3)	29(295.7)
90	33(370.1)	31(380.5)	30(392.0)	29(403.0)
	Board Foot Volume			
40	50+(54.8)	50+(28.8)	50+(16.0)	50+(9.4)
50	50+(200.7)	50+(137.0)	50+(93.1)	50+(64.0)
60	47(456.7)	49(367.4)	50+(291.0)	50+(228.1)
70	43(823.9)	44(727.4)	45(634.7)	46(547.2)
80	40(1301.0)	41(1214.5)	42(1126.0)	43(1033.5)
90	38(1886.0)	39(1824.7)	39(1756.8)	40(1679.4)

*The first entry is the age at which mean annual increment
culminates, the second entry is the mean annual increment
at this age. An entry of 50+ indicates that mean annual
increment culminates after age 50, in this case the
second entry is the annual increment at age 50.

Table 3.5.   Age at which mean annual increment culminates
             for various site index and initial density
             combinations when yields are estimated
             indirectly with the Weibull distribution
             function, minimum and average diameter
             approach to parameter estimation, and the
             class-interval-free yield estimation
             procedure

Site Index	Initial Number of Trees Per Acre			
	600	800	1000	1200
	Cubic Volume to a Four-Inch Top			
40	30(41.9)*	30(41.7)	29(40.4)	29(38.2)
50	29(77.4)	29(80.0)	28(81.3)	28(81.3)
60	25(120.1)	26(129.2)	27(135.0)	27(138.7)
70	21(163.0)	22(181.4)	23(195.4)	24(205.9)
80	18(203.5)	19(231.8)	20(254.8)	21(273.4)
90	17(240.1)	17(279.3)	18(310.8)	18(337.9)
	Board Foot Volume			
40	50+(21.9)	47(3.6)	38(0.3)	31(0.02)
50	50+(149.3)	50+(72.0)	49(22.7)	42(4.0)
60	47(410.4)	50+(311.6)	50+(205.0)	50+(108.0)
70	42(782.1)	45(701.0)	46(599.1)	48(478.4)
80	38(1257.8)	40(1212.8)	41(1144.8)	43(1051.7)
90	37(1841.2)	37(1840.4)	37(1818.3)	38(1771.2)

*The first entry is the age at which mean annual increment
culminates, the second entry is the mean annual increment
at this age. An entry of 50+ indicates that mean annual
increment culminates after age 50, in this case the
second entry is the annual increment at age 50.

Table 3.6. Age at which mean annual increment culminates
for various site index and initial density
combinations when yields are estimated
indirectly with the normal basal areas
approach to diameter distribution estimation
and the class-interval-free yield estimation
procedure

Site Index	Initial Number of Trees Per Acre			
	600	800	1000	1200
	Cubic Volume to a Four-Inch Top			
40	30(54.9)	33(56.2)	36(57.9)	38(59.8)
50	28(88.2)	30(94.18)	32(100.1)	33(105.8)
60	27(128.8)	28(141.5)	29(153.6)	30(165.1)
70	26(176.4)	27(197.9)	28(218.2)	29(237.3)
80	26(230.9)	27(263.3)	27(293.6)	28(322.0)
90	25(292.5)	26(337.5)	27(379.5)	27(419.0)
	Board Foot Volume			
40	50+(125.9)	50+(113.6)	50+(104.2)	50+(96.7)
50	50+(261.8)	50+(261.4)	50+(260.8)	50+(260.2)
60	48(438.3)	50+(462.8)	50+(483.8)	50+(502.0)
70	41(663.0)	45(716.7)	47(767.1)	49(813.7)
80	37(939.6)	39(1036.1)	41(1125.4)	42(1208.0)
90	34(1266.5)	36(1421.2)	37(1564.1)	38(1696.4)

*The first entry is the age at which mean annual increment
culminates, the second entry is the mean annual increment
at this age. An entry of 50+ indicates that the mean
annual increment culminates after age 50, in this case
the second entry is the annual increment at age 50.

Table 3.7. Age at which mean annual increment culminates
            for various site index and initial density
            combinations when yields are estimated
            indirectly with the discrete approximation to
            diameter distributions and the class-interval-
            free yield estimation procedure.

Site Index	Initial Number of Trees Per Acre			
	600	800	1000	1200
Cubic Volume to a Four-Inch Top				
40	25(41.3)*	25(42.1)	25(41.7)	25(40.8)
50	25(72.5)	24(75.0)	24(75.8)	24(75.7)
60	27(122.0)	25(125.2)	24(125.5)	25(125.6)
70	32(197.7)	39(718.7)	36(864.4)	35(984.5)
80	29(625.2)	27(1068.0)	25(1258.6)	25(1435.9)
90	23(395.3)	21(1347.7)	21(1614.7)	20(1830.8)
Board Foot Volume				
40	50+(14.1)	50+(8.1)	50+(6.1)	50+(5.5)
50	50+(105.7)	50+(56.2)	50+(37.3)	50+(30.1)
60	50+(510.7)	50+(337.4)	50+(227.0)	50+(170.4)
70	46(2804.1)	39(3558.3)	36(4143.7)	35(4634.4)
80	29(3157.9)	27(5358.9)	26(6186.7)	25(6919.5)
90	27(1853.2)	21(6714.97)	21(7987.3)	21(8890.6)

*The first entry is the age at which mean annual increment
culminates, the second entry is the mean annual increment
at this age. An entry of 50+ indicates that mean annual
increment culminates after age 50, in this case the
second entry is the annual increment at age 50.

not agree with the relation between site index and
culmination as described by Davis. Culmination of mean
annual increment of cubic foot yield is monotone increasing
with site index.

The indirect method utilizing the Weibull distribu-
tion function, estimators based on the smallest and
average dbh, and the class-interval-free approach to
yield estimation culminated at too early an age when
cubic foot yield was estimated. Culmination of board foot
yield did not occur before age 50 on poorer sites, and low
densities, and occured at an early age on poor sites at
high densities. Board foot volumes culminated as described
by Davis on the better sites.

The normal basal area approach to indirect yield
estimation produced cubic volume yield estimates which
culminated earlier than described by Davis on higher site
indexes. The board foot yields culminated too late on
high site indexes.

The discrete approximation to diameter distribu-
tions also produced results at odds with the description
of Davis. The relationship between culmination of cubic
foot yield and site index was not monotone. Board foot
yields did not culminate before age 50 on poor sites,
and culminated too early on good sites.

Although none of these methods produced results which agreed with the notions described by Davis, some of the differences may be due to the fact that Davis was referring to natural stands and not plantations. Mean annual increment generally culminates earlier in plantations than in natural stands. Since the yield prediction schemes described in this dissertation describe plantations, the age of culmination of mean annual increment should be somewhat earlier than described by Davis. This would explain many of the descrepencies observed in Tables 3.3 thru 3.7.

When this argument is taken into account, the normal basal area method of yield prediction produces ages of culmination of mean annual increment which agree with the notions of Davis better than the other methods compared. The age of culmination of mean annual increment was monotone decreasing with site index, which is the proper behavior according to Davis. The age at which mean annual increment culminates is about five years less than anticipated by Davis, however this can be attributed to the early culmination of mean annual increment in plantations.

When the three comparisons described in this chapter are considered jointly, the normal basal area

approach consistently performed as well or better than
the other methods. Diameter distributions were as
accurately estimated by this technique as any other
considered. Cubic volume yield estimates derived from
the normal basal area approach were nearly as
accurate and precise as other methods of yield
estimation, and board foot yield estimates were more
accurate than other methods considered. When culmination
of mean annual increment is considered, the normal basal
area approach shows a clear advantage over other
techniques described. These arguments indicate that the
normal basal area approach to diameter distribution and
yield estimation is to be preferred over other methods
compared in this chapter.

The general lack of accuracy and precision in
estimating diameter distributions and yields of forest
stands indicates that, perhaps none of the methods described
in this chapter are completely adequate. A new approach
to development of diameter distribution and yield estima-
tion procedures will be presented in the next chapter.
This new approach is an attempt to systematically develop
an estimation technique based on the physiological
processes of tree growth. Perhaps this new system will
provide estimates which are more accurate and precise
than estimates described in this chapter.

IV. A UNIFIED APPROACH TO YIELD PREDICTION

The growth of biological organisms in general and
trees in particular is a phenomenon which has been
modeled in many different ways. The models described in
Chapters I and II are only a small part of the entire
modeling effort. A more general model (in the sense
that it would be applicable to a wider range of species
and stand conditions) might be derived by incorporating
the methodology of several of the more successful models
into the development of a single growth model. A yield
model could be derived from the newly developed growth
model by accumulating growth over many time periods.

The Differential Equation Approach

A variety of growth models have been reviewed by
Brister (1972); a large proportion of the models are
based on growth or growth rates. One common approach is
to equate growth or growth rate and some function of
tree size and stand variables. Growth rate is by
definition the instantaneous rate of change in tree size,
growth is the change in tree size over some fixed period
of time. Thus differential equations (when modeling
growth rate) and difference equations (when modeling
growth) are appropriate mathematical approaches to tree

growth modeling. An estimate of tree size is obtained
via the solution of the differential or difference
equation. A simple example of this technique is the
exponential growth curve commonly used to describe growth
of fish and wildlife populations. This model may also
have application to tree growth, in some limited situations,
to describe the number of cells in a tree. Derivation
of the exponential model begins by assuming the instanteous
growth rate is proportional to the number of individuals in
the population, or cells in the tree in our case. If cell
size is constant the model is

$$\frac{dS(t)}{dt} = kS(t),$$ (4.1)

where

$S(t)$ = some measure of the size of the tree
at time t (number of cells x average
cell size)

t = time measured from some starting
point, t = 0

k = a constant of proportionality.

In addition, it is assumed that the tree size at the
starting point, t = 0, is S_0. The solution of (4.1)
with this added boundary condition is

$$S(t) = S_0 e^{kt}.$$ (4.2)

According to this simple model, tree size at time t
is a function of initial tree size, the time which has
elapsed and some constant commonly described as the
individual or cellular growth rate.

Ledig (1969) used a differential equation approach
to model growth of loblolly pine seedlings over a short
period of time. His model is

$$\frac{dY(t)}{dt} = kP(t)L(t) \; , \qquad\qquad (4.3)$$

where

  $Y(t)$ = total dry weight of pine seedling

  $P(t)$ = net assimilation rate of CO_2 per

     unit needle weight (a function of

     time and an indication of the

     rate of production of photo-

     synthate)

  k = a constant to convert net assimilation

     rate of CO_2 to growth rate of dry

     weight of the pine seedlings

  $L(t)$ = needle dry weight at time t .

The important difference between (4.3) and (4.1) is
that (4.3) takes into account certain physiological
notions of tree growth, namely that the rate of produc-
tion of the photosynthate which is utilized in production

of new needles can vary with time and genetic structure of the seedlings. Other stand variables that would also affect P(t) and L(t) but which were controlled in Ledig's experiments are season, weather, site index, and competition.

The model developed by Ledig was expanded by Promnitz (1972) to describe the growth of individual parts of trees such as crown, bole and roots. Promnitz uses a system of differential equations, each similar in nature to the model developed by Ledig, to describe the dry weight growth and yield of each plant part. Both Ledig and Promnitz modeled tree seedling growth only over a short period of time (less than six months).

Ledig and Promnitz indicated that their models might be applied to tree growth for more than one growing season. However, certain difficulties arise when attempting to apply these models to plantation grown trees for up to thirty year periods. Direct measurement of model components such as net assimilation rate of CO_2 for large trees is difficult or impossible. Control of conditions such as the effect of weather, site and genetic stock is prohibitively expensive. However, the incorporation of tree physiology considerations into growth and yield models is desirable.

Past Application of Differential Equation Models to Growth and Yield Prediction

Clutter (1963) applied the concept of a differential-integral relationship between growth and yield to the direct multiple linear regression approach of yield prediction. Similar work has been presented by Schaegel and Kulow (1969), Moser and Hall (1969), and Moser (1972, 1974). When considering growth and yield models Clutter defined compatible as: "Such models are here defined as compatible when the yield model can be obtained by summation of the predicted growth through the appropriate growth periods, or more precisely, when the algebraic form of the yield model can be derived by mathematical integration of the growth model." The relationship Clutter described is embodied in (4.1) and (4.2)((4.1) describes growth rate and (4.2) is the yield model).

To avoid pitfalls, however, growth rate and growth must be rigorously defined. In this presentation, growth rate is defined as the first derivative of yield, that is the instantaneous rate of change of the size of the tree. Growth is the change in size of a tree over some fixed time period, usually one year begining in the spring and ending in late summer or early fall. These definitions can be applied to a stand of trees as well as an individual tree. Therefore growth rate can be modeled with

differential equations, whereas growth is best modeled
with difference equations.

Clutter describes growth rate and yield on a
stand basis. The model for growth rate is

$$\frac{dV(t)}{dt} = \frac{\beta_2 V(t)}{B(t)} \left(\frac{dB(t)}{dt}\right) - \frac{\beta_3 V(t)}{t^2}. \qquad (4.4)$$

Although Clutter describes (4.4) as a growth equation
he introduces it as a differential equation and it will
be labeled growth rate here. The model for stand yield
is

$$\ln V(t) = \beta_0 + \beta_1 SI + \beta_2 (\ln B(t)) + \frac{\beta_3}{t}. \qquad (4.5)$$

where

   $V(t)$ = stand volume per acre at time t

   $B(t)$ = stand basal area per acre at time t

   SI = site index of the stand.

Clutter estimated the coefficients (β_i) in (4.4) and (4.5)
from data by applying multiple linear regression techniques.
Although not explicitly described in his paper, Clutter
made the following assumptions to apply least squares,

$$V(t) = \frac{V(t + 2.5) + V(t - 2.5)}{2} \qquad (4.6)$$

$$\frac{dV(t)}{dt} = \frac{V(t + 2.5) - V(t - 2.5)}{5} \qquad (4.7)$$

$$B(t) = \frac{B(t + 2.5) + B(t - 2.5)}{2} \qquad (4.8)$$

$$\frac{dB(t)}{dt} = \frac{B(t + 2.5) - B(t - 2.5)}{5} \qquad (4.9)$$

These assumptions were made to accomodate the data which consisted of permanent plots remeasured every five years. Clutter justified this procedure with the statement: "with growth periods of only five years, however, it is felt that the errors introduced by this approximation are inconsequential." It is interesting to note that Ledig detected large differences in growth rate within a single growing season; surely great differences in growth rate occur within a single year as the growth rate is zero during the winter months.

A more logical approach would be to obtain a growth model from the first difference of yield equation (4.5). Such an approach is complicated by the log transform used in the yield equation, since the first difference of $\log(X)$ with respect to X is not as simply evaluated as is the second derivative, but this slight increase in the effort of obtaining the model should be off-set by less stringent assumptions when fitting the data. Such an approach was suggested by Burkman (1962). He defines growth and yield in the following manner. "Unless otherwise specified, growth, for the purposes of this paper, is periodic net annual increment. Growth can be in basal

area, cubic feet, cordwood or board feet. Yield is defined as the summation of a number of net annual growth increments."

Another problem encountered by Clutter is that although the functional forms of his yield and growth rate equations are compatible, the estimates of the coefficients in the equations are not. That is, β_2 in (4.4) may not be equal to β_2 in (4.5). Sullivan and Clutter (1972) suggested an estimation procedure which overcomes this problem. Although Sullivan and Clutter add a random error term to each equation, this and all preceding models in this chapter do not make reasonable and systematic allowance for the stochastic nature of growth and yield of forest stands.

A Stochastic Approach to Modeling of Growth and Yield

A basic problem encountered with the above approaches is their deterministic nature. Tree growth is dependent upon weather which must be considered a random effect as it certainly cannot be accurately forecasted. The site quality of the land on which each tree is planted may vary widely over small areas, especially in the Piedmont region. These variations, termed microsite, cannot be easily detected and quantified, and so must be treated as a random element. The genetic make-up of planting stock is generally unknown, although in some cases only a few

improved genetic types are planted. All these effects, their interactions, and interactions between the trees themselves make a strong case for the use of a stochastic model for tree growth and yield. The approach of stochastic difference equations will be developed in this paper.

Although it is impractical to determine the age of a tree closer than to the nearest year, it is practical to measure yearly growth of the bole and larger limbs. Trees grown in temperate regions contain annual rings, which are the result of differences in cell size. Wood cells produced in the spring are larger than those produced in late summer. The result is a pattern of concentric circles when the tree is cut along the horizontal plane, which can be used to determine yearly diameter growth. When the tree is cut along the vertical plane a series of converging lines result, which can be used to determine yearly height growth. The study of these tree rings is called dendrochronology and is often concerned with the use of annual rings as a record of weather conditions. Tree rings will be used here as justification for a model based on yearly tree growth.

A simple model can be developed from the assumptions implicit in (4.1), namely that growth is proportional to size. The corresponding difference equation would be

$$\Delta S(t) = S(t) - S(t-1) = kS(t-1) \ . \qquad (4.10)$$

The solution to the difference equation 4.10 with the boundary condition $S(0) = S_0$ is

$$S(t) = S_0(1+k)^t. \qquad (4.11)$$

A stochastic element can be added to (4.10) by allowing k, the cellular growth, to vary from year to year,

$$\Delta S(t) = X_t S(t-1). \qquad (4.12)$$

The solution to (4.12) depends on the distribution of X_t. If the X_t are independent and identically distributed with mean μ and variance σ^2, the limiting distribution of S can be determined. This assumption might be valid if the variability of X_t is due largely to the weather the tth year. Cramér(1946) determines a solution to (4.12) by first solving for X_t ,

$$X_t = \frac{\Delta S(t)}{S(t-1)} = \frac{S(t) - S(t-1)}{S(t-1)} \ , \qquad (4.13)$$

then summing both sides of the equality with respect to t,

$$\sum_{t=0}^{n} X_t = \sum_{t=0}^{n} \frac{S(t) - S(t-1)}{S(t-1)} \quad . \qquad (4.14)$$

If the yearly growth is small in comparison to the size of the tree in year n, a useful approximation is

$$\sum_{t=0}^{n} X_t \simeq \int_{s(0)}^{s(n)} \frac{dX}{X} \qquad (4.15)$$

The solution of (4.15) is

$$\sum_{t=0}^{n} X_t \simeq \ln[S(n)] - \ln(S_0) \qquad (4.16)$$

The assumptions concerning X_t guarantee that the sum of the X_t is asympotically normally distributed. The limiting distribution of $S(n)$ is therefore the log normal distribution,

$$\frac{\ln[S(n)] - \ln(S_0) - n\mu}{\sqrt{n} \ \sigma} \sim AN(0, 1). \qquad (4.17)$$

Derivation of the Transformed Normal Diameter Distribution

The transformed normal distribution mentioned in earlier chapters is a generalization of the model described by (4.17). The growth in a particular year might not be proportional to the size of the tree, but to the size raised to some power, say C. The difference equation describing this phenomenon is

$$\Delta S(t) = X_t [S(t - 1)]^C \qquad (4.18)$$

The power, C, could be interpreted as the effect of competition on tree growth. A tree grown with no competition from surrounding trees would not be as limited in water, nutrient, CO_2 and sunlight intake as a tree grown in close proximity to neighboring trees. The intake of the open grown tree would be roughly proportional to the size of the tree. A larger tree would have a larger crown and could utilize more sunlight. A similar argument could be applied to the other components of growth. Increases in competition would give even greater advantage to the larger trees. Not only is the smaller tree limited by the size of its crown and root system, but it is further limited by competition with surrounding trees. Therefore we would expect C to be small for open-grown trees, and to increase as competition increases. The solution to (4.18) is obtained in the same manner as (4.12)

Solving for X_t and summing results in

$$\sum_{t=0}^{n} X_t = \sum_{t=0}^{n} \frac{S(t) - S(t-1)}{[S(t)]^c} . \qquad (4.19)$$

An approximation to (4.19) is

$$\sum_{t=0}^{n} X_t \simeq \int_{S(0)}^{S(n)} \frac{dX}{X^c} . \qquad (4.20)$$

and for $c \neq 1$ we obtain

$$\sum_{t=0}^{n} X_t \simeq \frac{[S(n)]^{1-c} - (S_0)^{1-c}}{1-c} . \qquad (4.21)$$

Finally, the approximate distribution of $S(n)$ is determined from

$$\frac{[S(n)]^{1-c} - (S_0)^{1-c} - n\mu(1-c)}{\sqrt{n} \ \sigma(1-c)} \sim AN(0,1). \qquad (4.22)$$

The results of Chapter 3 indicate that diameter distribution and yield estimation schemes derived by substituting basal area for S in (4.22) performed poorly, indicating that the model is inadequate. One possibility is that error was introduced when (4.22) was expanded from an individual tree model to a stand model. This was accomplished by assuming that all basal areas in a

stand are independent and identically distributed
according to (4.22). This may not be true since all
the trees in a stand are exposed to the same weather
conditions. At least part of the randomness explained
by X_t is due to weather effects. Since the basal areas
of trees in a stand each depend on a set of random
variables $\{X_t\}$ and sets of random variables are correlated
due to the common weather effects, perhaps the basal
areas of trees in a forest stand are also correlated,
violating the assumption made to develop the diameter
distribution and yield estimation schemes. An approach
which circumvents this weakness is developed in the
remainder of this chapter.

A Stochastic Model for Open-Grown Loblolly Pine

Competition for growing space in stand-grown trees
is a little understood phenomenon which complicates the
modeling effort. Competition effects will not be
considered initially to simplify derivation of the model.
The effect of competition on the open-grown tree model
will be considered in a later section. A model pertaining
to the growth of the 81 open-grown trees described in
Chapter II is derived in this section.

An implicit assumption of the stochastic models
derived in the previous section is that tree growth is

a Markov process, i.e., given the present size of
the tree, future growth is statistically independent of
past growth. The independence of the X_t and the form
of the difference equations imply that tree growth as
modeled is a Markov process. Consideration of this and
certain physiological aspects indicate that a good measure
of tree size is the ordered pair consisting of the tree
height and basal area.

Prediction of bole volume is the primary objective of
this modeling effort, and it was demonstrated in
Chapters 1 and 2 that bole volume is highly correlated
with the product of height and basal area. Growth of
tree boles is the result of activity of two meristems.
Height growth results from activity of the primary
meristems at the tip of the highest shoot of the tree.
Basal area growth is the result of activity of the
secondary meristem often referred to as the cambium,
located just under the bark of the tree. The processes
involved in growth of each of these areas is different
although the activity of the two meristems is correlated.
Finally, correlation between tree height and basal area,
and crown size (an important factor in the growth model
since photosynthesis takes place in the tree crown)
is high, as can be seen from Table 4.1. The difference
equations will be constructed in a manner which ensures

Table 4.1. Simple correlations (R) between state
variables (tree height and basal area) and
important tree dimensions (crown length,
crown diameter, crown area, and crown volume)
for 81 open-grown loblolly pine trees

	Crown Length	Crown Diameter	Crown Area	Crown Volume
Height	0.966	0.922	0.901	0.813
Basal Area	0.876	0.921	0.975	0.974

that the ordered pair (height, basal area) is a Markov process.

The open-grown tree growth model is further simplified by considering height alone as a Markov process. Figure 4.1 indicates that over long periods of time height growth is nearly linear. This assumption also has physiological support. The primary meristem is located at the tip of the crown and so the proximity to the photosynthate producing crown does not change as is the case with the other meristems. This would indicate that height growth is relatively constant over time and only depends on random fluxuations due to weather. Over longer periods of time height growth is reduced, perhaps due to the increased effort required to move water and nutrients to the tip of the tree. The concentration of photosynthate in the tree, particularly at the apical meristem, also influences height growth. This concentration should be proportional to the crown ratio, the proportion of the bole of the tree supporting foliage. This ratio is constant over a wide range of open grown tree sizes as illustrated by Figure 4.2. The relationship between tree height and crown length is a straight line through the origin, with a small amount of residual
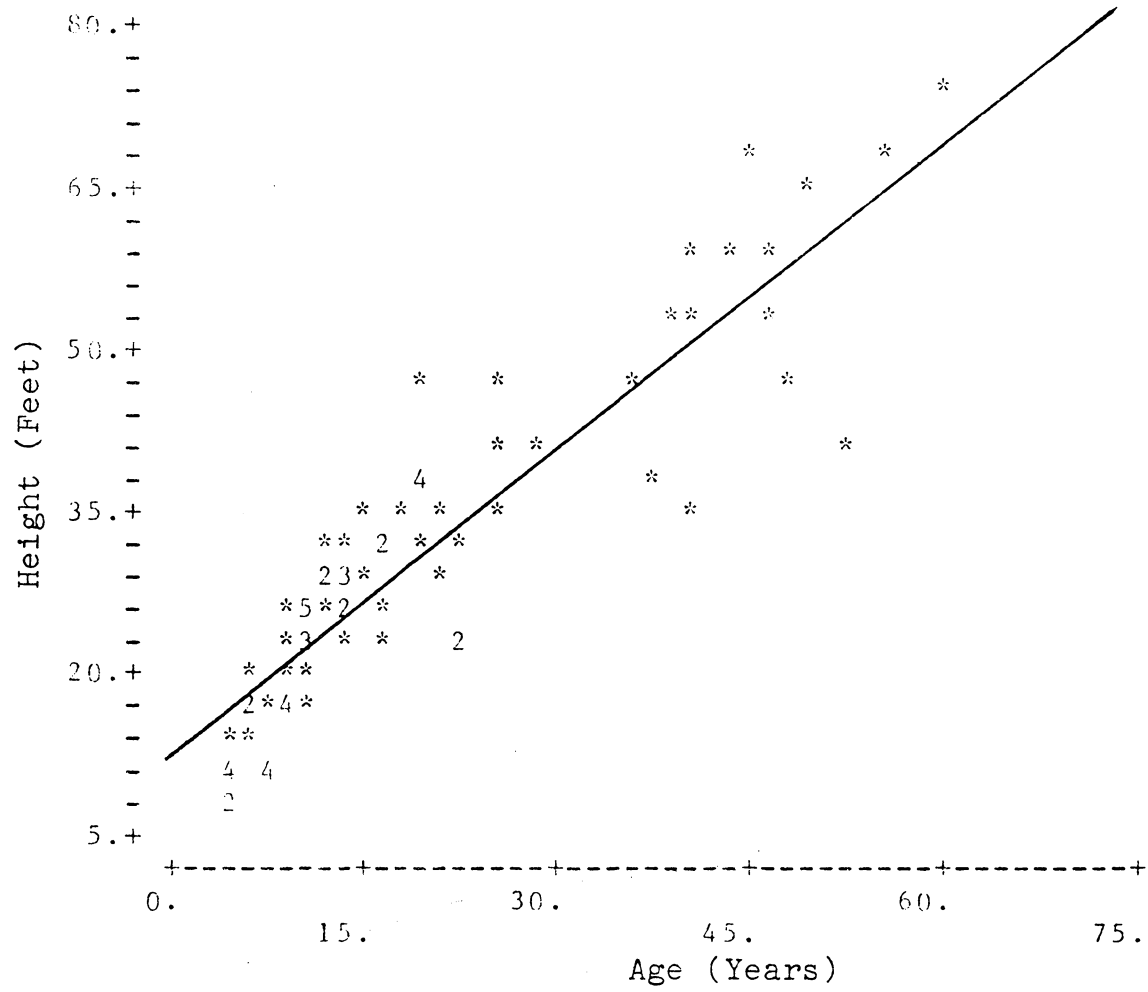
Figure 4.1. The relationship between height and age, and least squares fit for 81 open-grown loblolly pine trees.

Figure 4.2. The relationship between crown length and height, and least squares fit for 81 open-grown loblolly pine trees.

variation. This can be contrasted with Figure 4.3 which illustrates the same phenomenon in stand grown trees. The scattering of data points indicates that crown ratio is not constant in stand-grown trees. One of the most important effects of competition is the resulting decrease in crown length, which results in less photosynthate being produced, and the accompanying decrease in tree growth.

The arguments listed above support the use of the following model for height growth of open-grown trees,

$$\Delta H(t) = H(t) - H(t - 1) = X_t. \qquad (4.23)$$

This stochastic difference equation describes annual height growth as a single random variable, X_t. Year to year randomness is a result of weather conditions. The distribution of X_t is determined by the genetic make-up of the tree and the site upon which the tree is planted. The solution to (4.23) is that the current height of the tree is the summation of the yearly growth of the tree,

$$H(A) = \sum_{t=1}^{A} X_t \qquad (4.24)$$

```
42.0+
         -
         -                                                                              *
         -
35.0+                                                        *
         -
         -                                          *              *
         -                                    *           *  *
         -                               *              *  *
20.0+                              *      *  *  2   22         *
         -                      *    2  *****  5**
         -                  *   **    **3*2    *3*      *  **  *
         -              *   2*  2*2*2*  42  3*****2  2
21.0+              3*2222*23*22  *2**22**              *
         -          **24*53322*2*3*53524*22*      *     *
         -          *3**2*42242* *32327?*2**2  *
         -          3435*52353 6 *34*34  *        *      *
         -          54432633*23*53*435 3 ****  *       *
14.0+          65352 *3*224*4*42423 *
         -        *34*****43 *** 2 ** 2        * *
         -      *  32 3    * * * *
         -    *22  *  2***              *
         -
7.0+ *
         +--------+--------+--------+--------+--------+--------+Cl
      15.       45.       60.       75.
         30.               60.               90.
```

Crown Length (Feet)

Height (Feet)
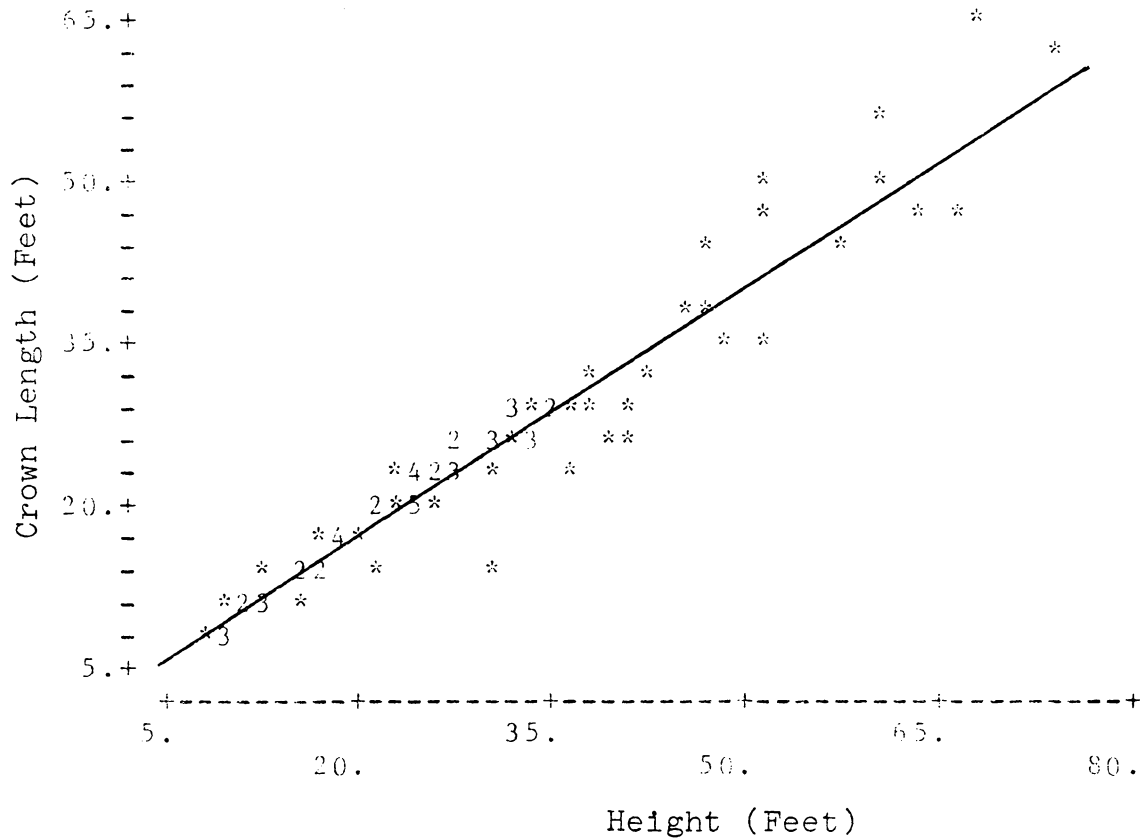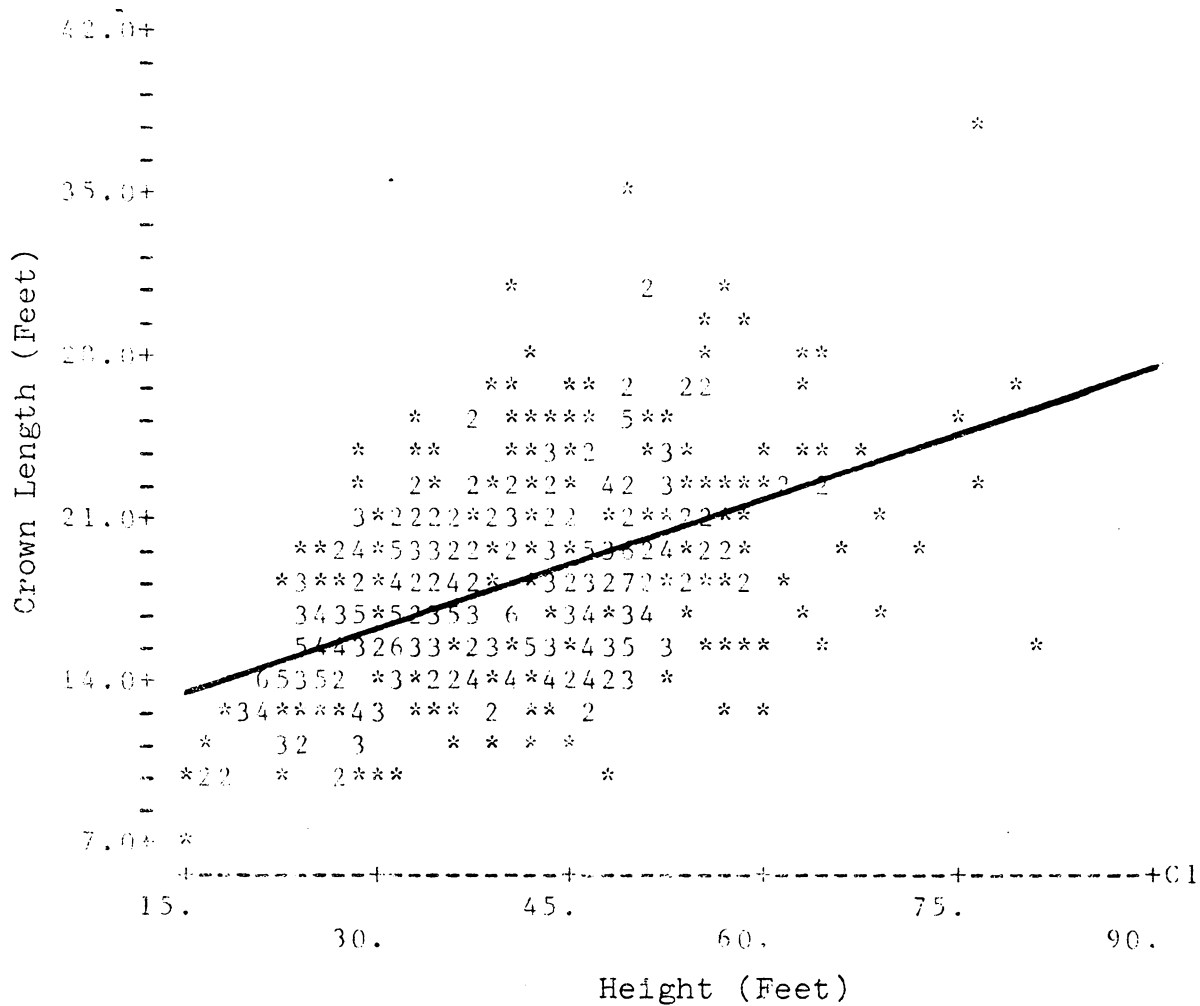
Figure 4.3.   The relationship between crown length and height, and
              least squares fit for 477 stand-grown loblolly pine trees.

An estimate of the expected value of X_t is the sample average,

$$\bar{X}_t = \sum_{t=1}^{A} X_t / A = H(A)/A. \qquad (4.25)$$

Empirical estimates of higher moments and the distribution of X_t is prohibited by the nature of the open-grown data. Only a single height-age measurement was made for each tree (a unique combination of genetic make-up and growing site). Therefore an estimate of the variance of X_t must be based on a single observation. This problem could be overcome if annual growth data were available. Such data would consist of the yearly height of single trees. This type of data could be obtained by sectioning older trees and measuring growth as recorded by the annual rings.

A model which takes into account the additional work required to transport water from the roots to the apical meristem as tree height increases is described by the following simple linear expression,

$$\Delta H(t) = X_t - \beta_0 H(t - 1). \qquad (4.26)$$

A linear approximation implies that the work required to transport the water increases linearly with tree height. The solution to the stochastic difference (4.26) is

$$H(A) = (1 - \beta_0)^A \sum_{t=1}^{A} X_t / (1 - \beta_0)^t. \qquad (4.27)$$

The value of β_0 may vary from tree to tree due to differences in genetic make-up and microsite; hence, the estimation of β_0 and the moments and distribution of X_t is again precluded by the nature of the open-grown data. Individual tree height growth data would yield such estimates. This model would be applicable to height growth of open-grown trees over longer period of time (greater than thirty years).

Basal area growth is largely dependent on the amount of photosynthate produced by the tree and the proximity of the meristem producing the basal area growth to the site of photosynthate production. The site of this production is the tree crown, and so the amount of photosynthate produced should be proportional to crown size. Sunlight is one of the primary requirements of the process, and so

production might be proportional to the amount of crown surface directly exposed to sunlight. Crown shape is approximately conical, so crown area (the surface area of the tree crown exposed to direct sunlight) is nearly proportional to the product of crown width and crown length. Figures 4.2 and 4.4 illustrate that both crown length and crown width are proportional to height in open-grown trees. The crown ratio of open-grown trees is large (over eighty percent of the height of the trees contains branches) resulting in close proximity of the meristem producing basal area growth to the tree crown. These arguments indicate that the relationship between basal area growth and height squared is linear,

$$\Delta B(t) = \beta_0 + \beta_1 H(t)^2 . \qquad (4.28)$$

Solving for basal area results in

$$B(A) = \sum_{t=1}^{A} [\beta_0 + \beta_1 H(t)^2] = A\beta_0 + \beta_1 \sum_{t=1}^{A} H(t)^2 . \qquad (4.29)$$

The nature of the data preclude estimation of β_0 and β_1 which may vary from tree to tree due to differences in genetic make-up and microsite. The following assumptions permit evaluation of this basal area growth model. Adopting the height growth philosophy of (4.24) and treating X_t as a constant rather than a random variable
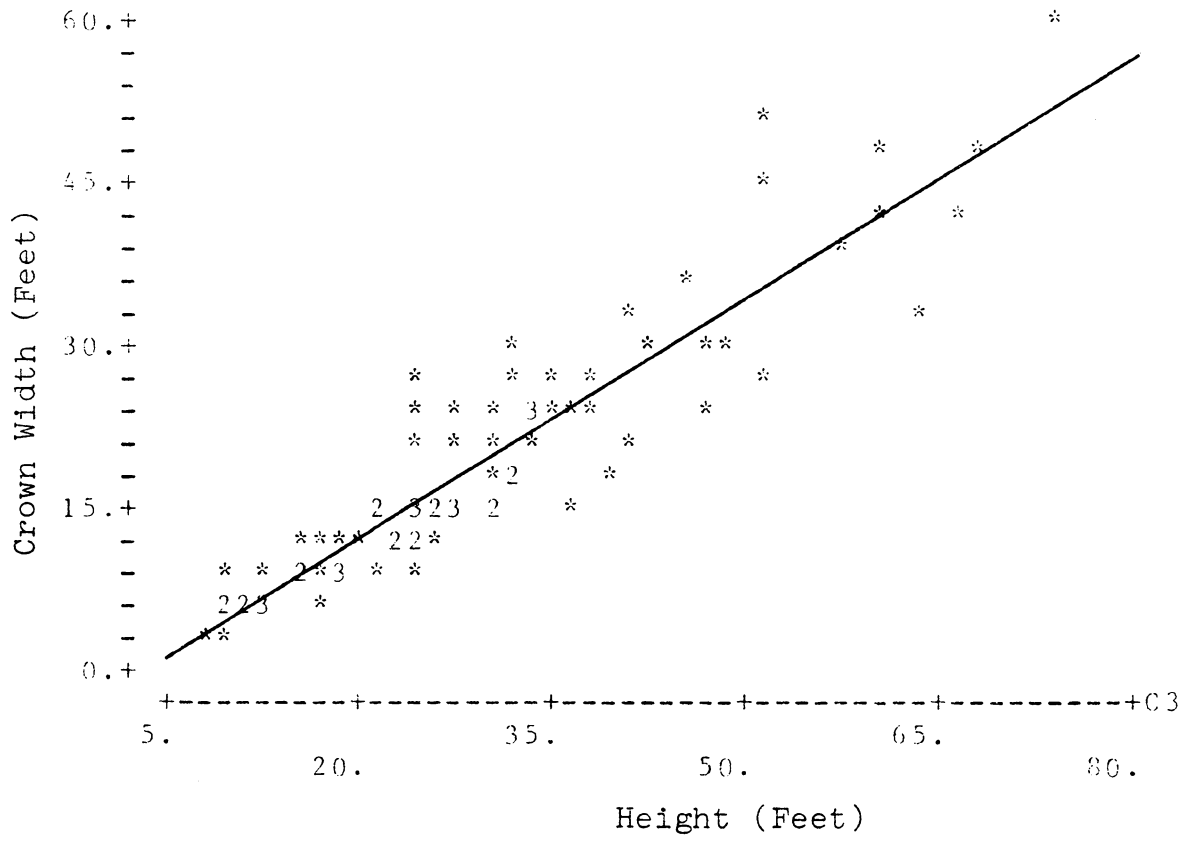
Figure 4.4. The relationship between crown width and height, and least squares fit for 81 open-grown loblolly pine trees.

results in the following approximations,

$$H(t) = \sum_{i=1}^{t} X_i \simeq \sum_{i=1}^{t} \frac{H(A)}{A} = t\left[\frac{H(A)}{A}\right] \qquad (4.30)$$

$$B(A) \simeq A\beta_0 + \beta_1 \sum_{t=1}^{A} \left\{t\left[\frac{H(A)}{A}\right]\right\}^2 = A\beta_0 + \beta_1 \left[\frac{H(A)}{A}\right]^2 \sum_{t=1}^{A} t^2$$

$$= A\beta_0 + \beta_1\left[\frac{H(A)}{A}\right]^2\left[\frac{(A)(A+1)(2A+1)}{6}\right]$$

$$= A\beta_0 + \beta_1\left[\frac{(H(A))^2(2A^2+3A+1)}{6A}\right]. \qquad (4.31)$$

Strub et al. (1975) presented a regression equation for estimating open-grown tree dbh as a function of tree age and height,

$$D(A) = \beta_0 + \beta_1 H(A) + \beta_2 A. \qquad (4.32)$$

A major difference between (4.31) and (4.32) is that (4.31) estimates tree basal area (in square feet) and (4.32) estimates tree diameter at breast height in inches. Tree diameter is directly proportional to the square root of basal area permitting the use of (4.31) in a fashion similar to (4.32). Table 4.2 is a comparison of the two regression

Table 4.2. Evaluation of two equations for estimating
            open-grown tree diameter at breast height
            and basal area from tree age and height.

Dependent Variable	Prediction* Equation	Standard Error of Estimate	Coefficient of Determination
Dbh	1	-2.28	0.892
Dbh	2	2.14	0.913
Basal Area	1	0.417	0.881
Basal Area	2	0.448	0.846

*Equation 1

$$B(A) = 0.013245A + 0.48527\left[\frac{(H(A))^2(2A^2+3A+1)}{60000A}\right] \quad (4.31)$$

Equation 2

$$D(A) = -2.4223 + 0.28658H(A) + 0.20947A$$

equations, listing a measure of the amount of variation
unexplained by each regression equation (the standard
error of estimate) and the percentage of the total
variation of the dependent variable explained by the
regression equation (the coefficient of determination).
This comparison is performed when tree diameter and tree
basal area are each the variable of interest.

The first two rows of Table 4.2 represent a
comparison of the two regression equations when diameter
at breast height is predicted from tree height and age.
Tree dbh is predicted from (4.31) by first estima-
ting β_0 and β_1 by using least squares and the data
collected from the 81 open-grown trees. These estimates
were substituted for β_0 and β_1 in equation (4.31). This
equation is used to estimate tree basal area from tree
height and age. Estimated tree diameter is obtained by
multiplying the square root of estimated basal area by the
constant 13.54. Least squares estimates of β_0, β_1,
and β_2 substituted into (4.32) result in an equation
which can be used to estimate tree diameter directly from
tree height and age. The sum of the squared differences
between observed tree diameter and estimated tree diameter
was calculated for each prediction equation. The
coefficient of determination and standard error of
estimate was calculated from the sum of squared differences.

A similar approach was used to obtain rows three
and four of Table 4.2. The dependent variable was basal
area for this part of the comparison. Equation (4.31) was
used to estimate tree basal area directly from tree height
and age. Equation (4.32) was used to estimate tree
diameter from height and age. Estimates of basal area
were obtained by multiplying the square of diameter by the
constant 0.005454. The sum of squared differences between
observed and estimated tree basal area was calculated for
each prediction equation. These sums of squared differences
were used to calculate the coefficients of determination
and standard errors of estimate contained in rows three
and four of Table 4.2.

Table 4.2 indicates that (4.32) is the
preferred equation for estimating tree dbh, and (4.31)
is the preferred equation for estimating tree basal
area. The slight advantage shown by each equation is
probably due to the least squares estimation procedure
used to determine the coefficients, one equation is the
least squares fit for dbh, the other for basal area. This
suggests that the assumptions made to develop (4.24)
and (4.28) are reasonable, since an adequate estimation
model resulted from these assumptions.

Equations (4.24) and (4.28) could be used to model
growth of open-grown trees. Use of this model depends on
assuming some probability distribution for X_t and

estimation of the parameters β_0 and β_1. If a random sample of open-grown trees is considered, β_0 and β_1 are also random variables. Hence, the value of β_0 and β_1 vary with each measurement of tree basal area, height, and age. Estimation of parameters and probability distributions is again prohibited by the nature of the data. Data with multiple observations over time would permit estimation and subsequent application of these types of models.

Stochastic Modeling of Stand-Grown Trees

Many of the assumptions valid to the development of the open-grown tree model are not applicable to competition-grown trees. The complicating effect of competition results in scatter diagrams similar to Figure 4.3 when data from a broad range of stand variables are examined. Data from the sectioned trees (two on each plot) described in chapter 2 were available for use in modeling height growth of stand-grown trees. These data consisted of age height pairs for each tree. Due to the method of collection of the data height measurements are accurate only to the nearest four feet. When these crude data were examined, a linear relationship between height and age was noted. Therefore (4.24) is applicable to competition grown trees, however the

distribution of X_t is further modified by competitive stress. Strub et al. (1975) presented arguments which indicate that height growth of open grown trees is generally less than height growth of trees grown under competitive stress when all other variables are constant. The crude nature of the height data hinders estimation of the distribution of X_t under different stand conditions.

Modeling of basal area growth in competitive-grown trees is even more limited by the nature of the data. The relationship between height and crown size is weak. The distance from the area of growth to the crown of the tree varies with stand conditions. These arguments indicate that (4.29) must be modified to add flexibility needed to model the growth of trees grown under competitive stress. The modeling efforts of Ledig (1969) described earlier in this chapter suggest an appropriate solution to this problem might be

$$\Delta B(t) = \beta_0 + \beta_1 P(t)L(t) \qquad (4.33)$$

The functions $L(t)$ and $P(t)$ will be discussed in detail in the next two sections.

Effective Crown Area

In (4.33) $L(t)$ is the area of the crown exposed to direct sunlight. The effective crown area of open-grown

trees can be determined from the crown width and crown
length if a conical crown is assumed, since the entire
surface of the open-grown tree crown is exposed to
direct sunlight. This is not true of trees grown under
competitive stress. The top of the crown is exposed to
direct sunlight and the bottom of the crown is shaded by
neighboring trees. The result of this shading is that
the needles on the lower branches have a much lower rate
of photosynthate production than the needles located
near the top of the tree. In fact, the rate of photo-
synthate production in lower branches is often so low
that more photosynthate is used in respiration than is
produced in photosynthesis, hence the net rate of
production is negative. The effective crown area, that
is the area of the crown which is actively engaged in net
positive production of photosynthate is located in the
upper regions of the tree. A method of estimating the
effective crown area can be derived from the results of
Strub, et al. (1975). This method was based on compari-
sons of growth of open-grown and stand-grown trees.

Diameter growth of open-grown and stand-grown lob-
lolly pine trees was compared by Strub, et al. They
determined that the diameter growth of stand-grown trees
begins to differ from that of open-grown trees at approxima-
tely the same time that the crown competition factor reaches
100. A detailed description of the crown competition

factor (CCF) can be found in Krajicek et al. (1961).
The crown competition factor is based on the relationship
between crown diameter and dbh of open-grown trees. This
relationship is linear and the crown diameter of open-
grown trees can be accurately estimated from dbh by
means of simple linear regression. The vertical
projection of the crown onto the ground is simply π
multiplied by the crown radius squared. These two
relationships can be used to estimate the projected
area occupied by the crown of any open-grown tree of
specified dbh. CCF is determined by associating with
each tree in the stand the projected crown area (termed
maximum crown area) of an open-grown tree with the same
dbh as the competition-grown tree. The CCF is the ratio
of the sum of the maximum crown areas to the area of the
land occupied by the stand, expressed as a percentage. A
CCF of 100 would indicate that if the crowns of all the
trees in a stand occupied an area equal to the area
occupied by open-grown trees of the same dbh, the land
would be exactly covered by tree crowns. The results of
Strub et al. indicate that after this point in time the
diameter growth of stand-grown trees differs from that
of open-grown trees.

This change in growth patterns might indicate that
when CCF reaches 100 the effective crown area of the

stand-grown trees reaches a maximum. Up to this time
(the time at which CCF reaches 100) the growth of the
trees would follow the same pattern as open-grown
trees, described by (4.28). Once CCF surpasses 100,
growth is the result of activity in the effective crown,
the area of which remains approximately equal to crown
area when CCF reaches 100. Effective crown area then
depends on the age at which the CCF reaches 100.

Figure 4.4 indicates that crown diameter is
proportional to height in open-grown trees. Projected
crown area should then be proportional to the square of
height. This phenomenon is illustrated by Figure 4.5.
Regression was used to estimate the average projected
crown area from tree height,

$$PCA(t) = 0.39401 \ H(t)^2 . \qquad (4.34)$$

The coefficient of determination is 80.7 percent and the
standard error of estimate 233.8 square feet. The product
of average projected crown area and the number of trees
per acre   results in the land area occupied by tree
crowns when all trees are open-grown. The CCF of N open-
grown trees is estimated by

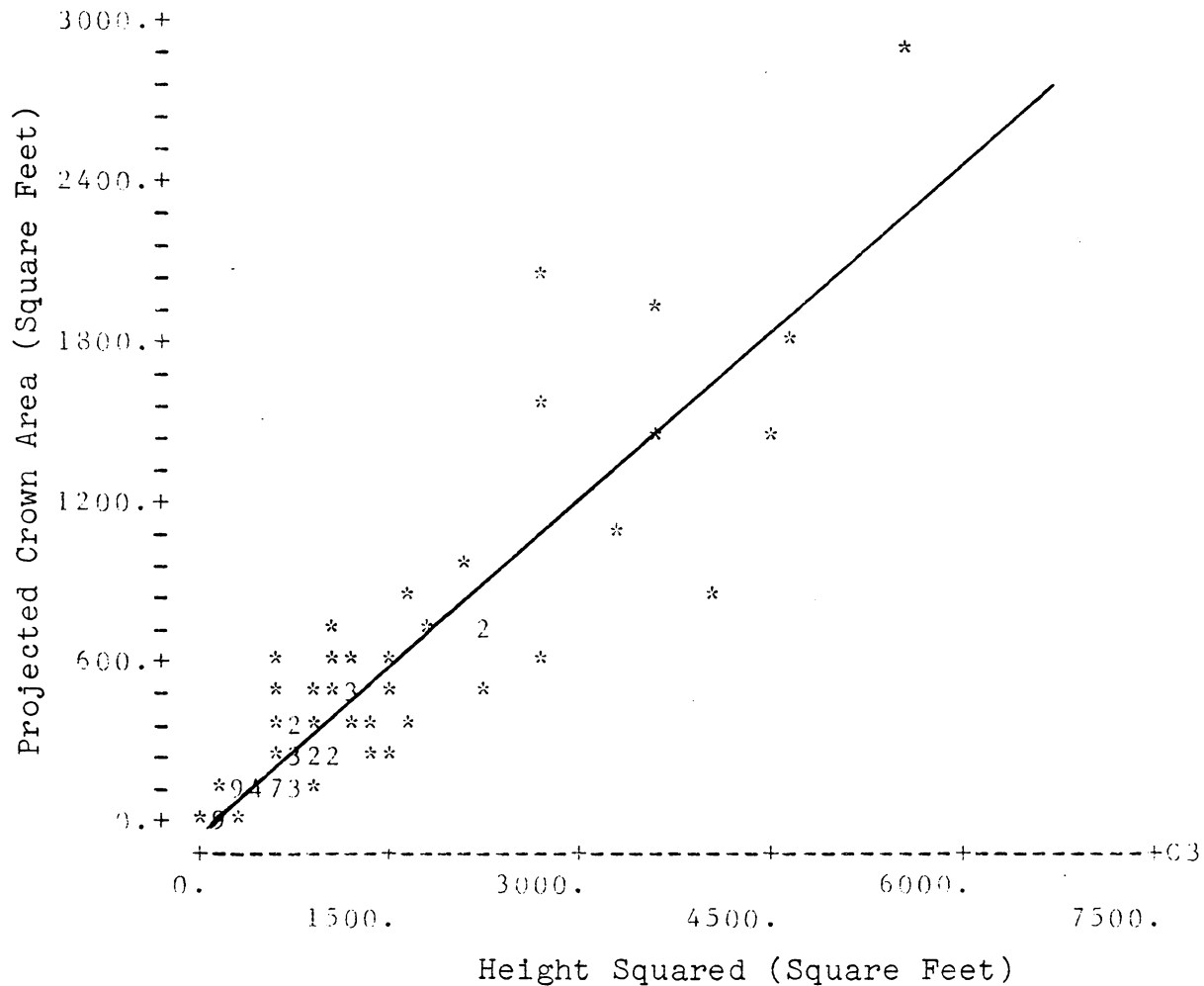$$CCF(t) = 0.39401 \ N \ H(t)^2/43560 . \qquad (4.35)$$

Figure 4.5.   The relationship between crown area and the square of
height, and least squares fit for 81 open-grown loblolly
pine trees.

Occupied land area must be divided by the number of
square feet in an acre (43560) since projected crown
area is measured in square feet. The age at which CCF
reaches 100 is determined by substituting 100 and solving
4.35 for t. A solution is obtained by approximating
H(t) with tH_d/A, an assumption similar to (4.30),

$$t = \frac{332.5\ A}{H_d\sqrt{N}} .$$ \qquad (4.36)

The effective crown area of stand-grown trees is
equivalent to that of open-grown trees up to age A_c, and
remains fixed at the same value after age A_c,

$$A_c = \left[\!\left[\frac{332.5\ A}{H_d\sqrt{N}} \right]\!\right] .$$ \qquad (4.37)

Where $[\![\cdot]\!]$ indicates the greatest integer function.

As previously discussed, crown area is proportional
to the square of height in open-grown trees,

$$CA(t) = \beta_2 H(t)^2 .$$ \qquad (4.38)

The effective crown area of competition grown trees is
described by

$$ECA(t) = \begin{cases} \beta_2 H(t)^2 & t \leq A_c \\ \beta_2 H(A_c)^2 & t > A_c . \end{cases}$$ \qquad (4.39)

The area of the crown exposed to direct sunlight, L(t),
is proportional to effective crown area, ECA(t),

$$L(t) = \gamma ECA(t). \tag{4.40}$$

Photosynthate Allocation

The term P(t) in (4.33) represents the proportion
of the net photosynthate produced which results in basal
area growth. The amount of photosynthate utilized in
basal area growth is a function of the amount of photo-
synthate which reaches the secondary meristem at breast
height. Transport of photosynthate is then of primary
importance to the determination of a proper expression
for P(t).

Photosynthate is transported by long thin strings
of hollow straw-like cells called phloem, which run from
the tip of the tree to the roots and along tree branches.
Phloem is located between the cambium and the bark of the
tree. The phloem cells are water filled and transport
materials such as photosynthate from the crown to the
roots. Cambial cells located adjacent to the phloem
intercept photosynthate on its downward route in the
phloem. If the overall concentration of photosynthate in
a tree is low, a large proportion of the photosynthate
produced would be intercepted by cambial cells located
in or close to the crown of the tree. As the concentra-
tion photosynthate in the tree is increased, the cambial
tissue near the top of the tree reaches some saturation

point, that is some limit to the amount of photosynthate which can be utilized by the cell is reached. A larger amount of photosynthate then reaches cambial cells located near the base of the tree.

This argument indicates that the proportion of photosynthate allocated to basal area growth is proportional to the concentration of photosynthate in the tree bole. This concentration should be approximately proportional to the ratio of the length of the effective crown to the height of the tree. Figure 4.3 indicates that the crown length of open-grown trees is proportional to height. The preceding discussion indicates that effective crown length (the length of the portion of the crown which has a positive net effect on the production of photosynthate) of stand-grown trees would be equal to total crown length until CCF reaches 100. The effective crown length after CCF surpasses 100 would be somewhat less than total crown length and approximately equal to crown length at the time CCF reaches 100. If α is the ratio of crown length to height in open-grown trees, $P(t)$ can be modeled as

$$P(t) = \begin{cases} \alpha \, \dfrac{CL(t)}{H(t)} = \beta_3 & t \leq A_c \\[4mm] \alpha \, \dfrac{CL(A_c)}{H(t)} = \dfrac{\beta_3 \, H(A_c)}{H(t)} & t > A_c \, . \end{cases} \qquad (4.41)$$

The symbol β_3 represents the product of the crown ratio of the tree if it were open-grown and some constant to convert effective crown ratio to the proportion of photosynthate which reaches the cambium at breast height.

A Description of Basal Area Growth in Stand-Grown Trees

Substituting (4.39), (4.40), and (4.41) into (4.33) an expression for basal area growth is obtained,

$$
\Delta B(t) = \begin{cases} \beta_0 + \gamma \beta_1 \beta_3 \beta_2 H(t)^2 & t \leq A_c \\ \\ \beta_0 + \gamma \beta_1 \left[\dfrac{\beta_3 H(A_c)}{H(t)}\right] \beta_2 H(A_c)^2 & t > A_c. \end{cases} \qquad (4.42)
$$

Combining parameters γ, β_1, β_2 and β_3 results in

$$
\Delta B(t) = \begin{cases} \beta_0 + \beta_1{'} H(t)^2 & t \leq A_c \\ \\ \beta_0 + \beta_1{'} \dfrac{H(A_c)^3}{H(t)} & t > A_c. \end{cases} \qquad (4.43)
$$

Equations (4.23), (4.37) and (4.43) can be used jointly to model growth of individual stand grown trees.

These three equations contain only a single random element, X_t, the random variable which represents the effect of weather on tree height growth. When an entire stand of trees is considered this interpretation changes;

β_0 and β_1 are fixed parameters only for a single tree, and vary from tree to tree in a stand due to changes in genetic structure and microsite variability. These sources of variability can be included in the model by substituting the random variables Y and Z for the parameters β_0 and β_1,

$$\Delta B(t) = \begin{cases} Y + ZH(t)^2 & t \le A_c \\[2ex] Y + ZH(A_c)^3/H(t) & t > A_c. \end{cases} \qquad (4.44)$$

All trees in the stand are exposed to the same weather conditions, hence differences between trees and the resulting height and basal area distributions are due to differences in the genetic structure of the trees and microsite changes, not exposure to different weather conditions. Weather would cause variation in diameter distributions between stands of trees, since different stands are separated geographically as well as in time. This argument indicates that variation between height and basal areas of trees in a stand does not result from the random variable X_t. Forest managers would like to estimate the joint height and basal area distribution for average weather conditions. The expected value (with respect to the random variable X_t) of tree height and

basal area are of prime interest to the forest manager.

Evaluation of expected tree height from equation 4.23 is straight forward,

$$E_X\left[H(A)\right] = E_X\left[\sum_{t=1}^{A} X(t)\right] = A\mu_X \ . \qquad (4.45)$$

The expected value of X_t is represented by μ_X. The distribution of X_t may vary from tree to tree due to differences in genetic make-up of the tree and differences in microsite. This indicates that for a stand of trees μ_X is a random variable, represented here by the symbol U, and

$$E_X[H(A)] = AU \ . \qquad (4.46)$$

A similar approach can be applied to tree basal areas,

$$E_X[B(A)] = \sum_{t=1}^{A} E_X[\Delta B(t)] = \sum_{t=1}^{A} E_X[Y + ZH(t)^2] \qquad A \leq A_c. \quad (4.47)$$

The random variables Y and Z are independent of X_t since choice of genetic make-up and microsite is independent of weather, resulting in

$$E_X[B(A)] = \sum_{t=1}^{A} \{Y + Z\, E_X[H(t)^2]\} = AY + Z\sum_{t=1}^{A} E_X[H(t)^2]$$

$$= AY + Z\sum_{t=1}^{A} E_X\left[\left(\sum_{i=1}^{t} X_i\right)^2\right] = AY + Z\sum_{t=1}^{A} [t\sigma_X^2 + (t\mu_X)^2]$$

$$= AY + Z\left[\sigma_X^2\,\frac{A(A+1)}{2} + \mu_X^2\,\frac{A(A+1)(2A+1)}{6}\right]$$

$$A \leq A_c . \tag{4.48}$$

Equation (4.48) describes tree basal areas of open-grown trees and stand-grown trees in which the CCF has not yet reached 100. Again μ_X and σ_X^2 (the mean and variance of X_t) may vary from tree to tree and will be represented by the random variables U and V,

$$E_X[B(A)] = AY + Z\left[V\frac{A(A+1)}{2} + U^2\,\frac{A(A+1)(2A+1)}{6}\right] \quad A \leq A_c. \tag{4.49}$$

The expression for expected basal area of stand-grown trees after CCF surpasses 100 is more difficult to evaluate,

$$E_X\left[B(A)\right] = E_X\left[B(A_c) + \sum_{t=A_c+1}^{A} \Delta B(t)\right] = E_X\, B(A_c)$$

$$+ \sum_{t=A_c+1}^{A} E_X\left[\Delta B(t)\right]$$

$$= E_X\left[B(A_c)\right] + \sum_{t=A_c+1}^{A} E_X\left[Y+ZH(A_c)^3/H(t)\right]$$

$$= A_c Y + Z\left[V\frac{A_c(A_c+1)}{2} + U^2\frac{A_c(A_c+1)(2A_c+1)}{6}\right]$$

$$+ (A - A_c)Y + Z\sum_{t=A_c+1}^{A} E_X\left[H(A_c)^3/H(t)\right]$$

$$= AY + Z\left[V\frac{A_c(A_c+1)}{2} + U^2\frac{A_c(A_c+1)(2A_c+1)}{6}\right.$$

$$+ \left.\sum_{t=A_c+1}^{A} E_X\left[H(A_c)^3/H(t)\right]\right] \qquad A > A_c. \qquad (4.50)$$

The expected value in the final expression of (4.50) is
a function of the parameters of the distribution of X_t and
of t. The parameters of the distribution of X_t vary from
tree to tree, hence when the basal area distribution of
a stand of trees is considered (versus individual tree
model (4.43)) the expected value (with respect to X) in
(4.50) may be a random variable. Specification of the
distribution of X_t may permit evaluation of the expected

value in (4.50). This evaluation would facilitate further development of this model.

Estimation of Distributions and Parameters in the Basal Area - Height Model

Yield can be estimated if the joint probability function of tree basal area and height is specified. This joint probability density function could be estimated from (4.46), (4.49), and (4.50). Data requirements for such an estimation procedure are in excess of the data available for this study. The required data could be obtained by sectioning trees.

Data could be collected for all trees on some fractional part of many stands, just as the trees on the 0.1-acre plots were measured for this study. Stand age, height of the dominant and codominant trees, and the number of trees per acre should be determined for each plot. Each tree should then be felled and sectioned to determine height and basal area at the end of each year growth from the annual rings. Basal area and height growth for each of the trees could be measured in this manner.

Average height growth, μ_X, and the variance of height growth, σ^2_X, over time could be estimated for each tree from this data. The expected value in

(4.49) could be estimated for each tree by substituting height growth in the ith year for X_i. The frequency distribution of the estimates for all the trees on a plot could be used to determine estimates of the probability distributions of U, V, and the expected value in the last expression of (4.50). This data could also be used to determine the relationship between these probability distributions and the stand parameters.

The coefficients describing individual tree basal area growth, β_0 and β_1', in (4.43) could be estimated for individual trees by applying the least squares criterion. The distribution of β_0 and β_1' within a stand, that is the distribution of Y and Z, would be estimated by the frequency distribution of the estimates of β_0 and β_1' for individual trees. The relationship between stand parameters and the distribution of Y and Z could also be examined.

These techniques could be used to derive a procedure for estimating the expected (with respect to weather) distribution of basal area and height. The data set described above would also serve as a check on the models derived in the previous section.

Application of the Basal Area – Height Model to Estimation of Yields of Forest Stands

The procedure described in the previous section might be used to estimate the joint distribution of tree heights and basal areas for average weather conditions. This joint distribution could be used to estimate stand yield from stand variables. Individual tree yield is often modeled as a function of tree height and diameter at breast height squared (which is directly proportional to basal area). Incorporation of such an individual tree yield equation into a generalized version of indirect stand yield equation (2.67) results in a stand yield estimation procedure based on the joint distribution of tree height and basal area.

A generalized version of (2.76) requires use of two sets of class intervals. The variable i indexes a set of tree basal area classes with class midpoints B_i. The variable j indexes a set of height classes with class midpoints H_j. The probability P_{ij} represents the probability that for a specified set of stand variables, a tree is in the ith diameter class and the jth height class. The proportion of the trees in the ith diameter class and the jth height class which are suitable for

the manufacture of a product such as lumber is represented by Q_{ij}. An individual tree volume equation such as (1.3) is represented by $g(B_i, H_j)$. Since basal area is directly proportional to the square of tree diameter at breast height, an equation which is equivalent to (1.3) is

$$V = g(B_i, H_j) = \beta_0 + \beta_1 BH. \qquad (4.51)$$

Stand yield could be predicted by using the m basal area classes and the n height classes in the following fashion,

$$Y = N \sum_{i=1}^{m} \sum_{j=1}^{n} P_{ij} Q_{ij} g(B_i, H_j). \qquad (4.52)$$

Prediction of P_{ij} would result from approximation of the joint density function of tree height and basal area as a function of stand variables just as the probability density function for tree diameter was predicted as a function of stand variables. The approximation of the joint density function ($f(B, H)$) is integrated over the range of the ith diameter class and the jth height class to obtain P_{ij},

$$P_{ij} = \int_{H_j - w_h/2}^{H_j + w_h/2} \int_{B_i - w_b/2}^{B_i + w_b/2} f(B, H) dB dH. \qquad (4.53)$$

The width of a basal area class is w_b, the width of a height class is w_h. The probability that a tree in the ith basal area class and the jth height class is suitable for the production of lumber, Q_{ij}, could be estimated from individual tree data or predicted as a function of basal area and height by using the same techniques used to estimate Q_j.

The class-interval-free analog to (2.87) can be developed by evaluating the limit of (4.52) as the number of basal area classes (m) and the number of height classes (n) increase without bound,

$$Y = \lim_{\substack{m \to \infty \\ n \to \infty}} N \sum_{i=1}^{m} \sum_{j=1}^{n} P_{ij} Q_{ij} g(B_i, H_j). \qquad (4.54)$$

The range of the tree basal area classes (mw_b) and the range of the height classes (nw_h) are held constant. The limits of the range of basal area classes are represented by L_b and U_b, and the limits of the range of height classes are L_h and U_h. The interval widths can be described in the following manner,

$$w_b = (U_b - L_b)/m, \qquad (4.55)$$

$$w_h = (U_h - L_h)/n. \qquad (4.56)$$

Substituting (4.52), (4.54), and (4.55) into (4.53) results in

$$Y = \lim_{\substack{m \to \infty \\ n \to \infty}} N \sum_{i=1}^{m} \sum_{j=1}^{n} \int_{H_j - (U_h - L_h)/(2n)}^{H_j + (U_h - L_h)/(2n)} \int_{B_i - (U_b - L_b)/(2m)}^{B_i + (U_b - L_b)/(2m)}$$

$$f(B, H)dBdHQ_{ij}g(B_i, H_j). \qquad (4.57)$$

Replacing Q_{ij} with a continuous function of B and H, q(B, H) results in

$$Y = \lim_{\substack{m \to \infty \\ n \to \infty}} N \sum_{i=1}^{m} \sum_{j=1}^{n} \int_{H_j - (U_h - L_h)/(2n)}^{H_j + (U_h - L_h)/(2n)} \int_{B_i - (U_b - L_b)/(2m)}^{B_i + (U_b - L_b)/(2m)}$$

$$f(B, H)dBdHq(B_i, H_j)g(B_i, H_j). \qquad (4.58)$$

The definition of the multiple integral ensures that (4.59) has the same limit as (4.58),

$$Y = \lim_{\substack{m \to \infty \\ n \to \infty}} N \sum_{i=1}^{m} \sum_{j=1}^{n} f(B_i, H_j)\Delta B_i \Delta H_j q(B_i, H_j)g(B_i, H_j). \qquad (4.59)$$

The class interval width w_b is represented by ΔB_i and w_h is represented by ΔH_j. If q and g are continuous functions application of the definition of multiple integrals results in

$$Y = N \int_{L_h}^{U_h} \int_{L_b}^{U_b} f(B, H)q(B, H)g(B, H)dBdH . \qquad (4.60)$$

This class-interval free stand yield equation can be used to estimate yield from the joint density of basal area and height.

V. CONCLUSIONS AND RECOMMENDATIONS

The comparisons and tests presented in chapter 3 indicate that no proposed diameter distribution estimation technique adequately described the diameter distribution of loblolly pine plantations for the age, site and density combinations observed. A logical approach to the development of a diameter distribution estimation technique is presented in chapter 4. Such an approach requires incremental data not currently available. Consideration of the physiology of individual tree growth could lead to accurate and biologically sound prediction equations.

The comparisons presented in chapter 3 also indicate that indirect yield estimation can be as accurate as direct yield estimation. The Weibull distribution function, (2.20), estimation of parameters based on quantiles, (2.29), (1.10), and (1.11), and yield estimation based on 1-inch class intervals, (2.76), provides a simple yet accurate scheme for yield estimation. A more complex technique which requires extensive use of high-speed computers is the discrete approximation to diamter distributions, (2.74) and (2.76). This technique also provided accurate results. When only total cubic foot volume and cubic foot volume to a 4-inch merchant-ability limit were estimated, the Weibull distribution

156

function, (2.20), estimators based on the smallest
and average diameter, (2.32), (2.35), and (2.36), and
the class-interval-free yield estimation procedures,
(2.87), provide accurate board foot yield estimates.

Consideration of culmination mean annual increment
also supported use of estimates based on the normal basal
area approach to diameter distribution estimation and the
class-interval-free yield estimation procedures. None
of the other yield estimation schemes conformed exactly
to preconceived notions of the behavior of mean annual
increment culmination. These results are also discussed
in chapter 3.

The various methods of yield estimation can
be compared by considering the allowance each technique
makes for the stochastic nature of growth of forest
stands. Sources of variation in growth can be due to
weather effects, microsite and genetic effects, and
differences in stand variables which result in variation
between stands of trees. The manager cannot control
weather, hence, he is usually interested in yield for
some average set of weather conditions. Microsite and
genetic variation results in differences in tree size
and yield within a stand. Stand variables are the only
sources of variation which have been successfully
quantified. Consequently the stand variable variation

can be modeled deterministically, while the weather,
microsite, and genetic effects have not been successfully
quantified, hence they contribute to errors in yield
estimation.

The direct method of yield estimation treats the
stand variables in a deterministic fashion. The weather
and other factors are modeled as a single additive
error term which has constant variance over the range
of stand variables. This is an over-simplification of
the many physiological processes which result in varia-
tion in individual tree and stand yield.

The indirect method of yield estimation treats
the stand variables in a deterministic fashion. Microsite
and genetic variation is accounted for by estimating the
diameter distribution. This approach does not make
systematic allowance for weather fluxuations.

The unified method of yield estimation described
in chapter 4 was developed by considering variation due
to weather, microsite and genetic effects, and stand
characteristics. Such a model requires more detailed
input data. If variation in growth between and within
stands, is to be separated, the stand model must be
based on growth data which measures between and within
stand variation. Such data were not available in this

study which precluded completion and exact specifica-
tion of the unified method of yield estimation.

The previous argument and comparisons presented
in chapter 3 indicate that the choice of procedure for
estimation of yields of forest stands depends on the
nature of the data available for development of the
procedure. If yield measurements are available for
only a single point in time, as was the case for data
used in this dissertation, the normal basal area approach
to yield estimation presented in chapter 2 should be
used. Equations (2.37), (2.38), (2.41), and (2.42) can
be used to approximate the diameter distribution of
forest stands. The approximate diameter distribution
and individual tree volume equation (2.5) can be
substituted into (2.87) to obtain stand yield estimates.
This procedure provides an accurate estimate of yield
and supplies the forest manager with additional informa-
tion in the form of an approximate diameter distribution.

The procedures developed in chapter 4 should be
used to estimate stand yield if data consisting of
repeated measurements of yield of the same trees at
different ages were available. The biologically based
yield estimation procedure should be applicable to a
wide range of stand variables, and supplies the manager

with detailed information about individual tree and
stand growth and yield.

LITERATURE CITED

Bailey, R. L. 1972. Development of unthinned stands of *Pinus radiata* in New Zealand. Unpublished dissertation. University of Georgia. 67 pp.

Bailey, R. L. and T. R. Dell. 1973. Quantifying diameter distributions with the Weibull function. Forest Science 19:97-104.

Bare, B. B. 1970. User's Manual for the Purdue forest management game. Center for Quantitative Studies paper 38. University of Washington. 72 pp.

Bennett, F. A. and J. L. Clutter. 1968. Multiple-product yield estimates for unthinned slash pine plantations -- pulpwood, sawtimber, gum. U.S.D.A. Forest Service Research Paper SE-35. 21 pp.

Bliss, C. I. and K. A. Reinker. 1964. A log-normal approach to diameter distributions in even-aged stands. Forest Science 10:350-360.

Box, G. E. P., and M. E. Muller. 1958. A note on the generation of random normal deviates. Annal of Mathematical Statistics 24:610-611.

Brister, G. H. 1972. A biometrical study of the development of secondary xylem in Douglas fir. Unpublished dissertation, University of Washington. pp. 70-104.

Bruce, D. and L. H. Reineke. 1928. The use of alinement charts in constructing forest stand tables. Journal of Agricultural Research 38:289-309.

Buckman, R. E. 1962. Growth and yield of red pine in Minnesota. U.S.D.A. Technical Bulletin 1272. 50 pp.

Burkhart, H. E., R. C. Parker, M. R. Strub, and R. G. Oderwald. 1972. Yields of old-field loblolly pine plantations. FWS-3-72. Division of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University, Blacksburg, VA. 51 pp.,

Burkhart, H. E., and M. R. Strub. 1974. A model for simulation of planted loblolly pine stands. IN Growth models for tree and stand simulation. Royal College of Forestry, Stockholm, Sweden. pp. 128-135.

Clutter, J. L. 1963. Compatible growth and yield models for loblolly pine. Forest Science 9:354-371.

Clutter, J. L. and F. A. Bennett. 1965. Diameter distributions in old-field slash pine plantations. Georgia Forest Research Council Report No. 13. 9 pp.

Cox, D. R. 1970. The Analysis of Binary Data. Methuen and Company, London, England. 142 pp.

Cramér, H. 1946. Mathematical methods of statistics. Princeton Mathematical Series, No. 9, Princeton University Press. pp. 218-220.

Davidon, W. C. 1959. Variable metric method for minimization. A.E.C. Research and Development Report, ANL-5990 (Rev.). 21 pp.

Davis, K. P. 1966. Forest management: regulation and valuation. McGraw-Hill Book Company. New York, NY. 519 pp.

Draper, N. R. and H. Smith. 1966. Applied regression analysis. John Wiley and Sons, Inc., New York, NY. 407 pp.

Dubey, S. D. 1967. Some percentile estimators for Weibull parameters. Technometrics 9, 119-129.

Fletcher, R. and M. J. D. Powell. 1963. A rapidly convergent descent method for minimization. Computer Journal 6:163-168.

Garwood, F. 1941. The application of maximum likelihood to dosage mortality curves. Biometrika 32:46-58.

Gibbons, J. D. 1971. Nonparametric statistical inference. McGraw-Hill Book Company, New York, NY. 306 pp.

Graybill, F. A. 1961. An introduction to linear statistical models, volume I. McGraw-Hill Book Co., Inc., New York, NY. 463 pp.

Green, J. R., and Y. A. S. Hegazy. 1976. Powerful modified-EDF goodness-of-fit tests. Journal of the American Statistical Association 71:204-209.

Hamilton, David A., Jr. 1974. Event probabilities estimated by regression. U.S.D.A. Forest Service Research Paper INT-152. 18 pp.

Harter, H. L. and A. H. Moore. 1965. Maximum likelihood estimation of the parameters of gamma and Weibull populations from complete and from censored samples. Technometrics 7:639-643.

Hush, B., C. I. Miller and T. W. Beers. 1972. Forest Mensuration. Ronald Press Company, New York, NY. 410 pp.

International Business Machine Corporation. 1970. IBM System/360 Scientific Subroutine Package Version III Programmer's Manual. pp. 221-225.

Jennrich, R. I., and R. H. Moore. 1976. Maximum likelihood estimation by means of nonlinear least squares IN American Statistical Association 1975 Proceedings of the Statistical Computing Section American Statistical Association, Washington, DC. pp. 57-65.

Krajicek, J. E., K. A. Binkman and S. F. Gingrich. 1961. Crown competition -- a measure of density. Forest Science 7:35-42.

Ledig, F. T. 1969. A growth model for tree seedlings based on the rate of photosynthesis and the distribution of photosynthate. Photosynthetica 3:263-275.

Lenhart, J. D. 1972. Cubic-foot yields for unthinned old-field loblolly pine plantations in the interior West Gulf Coastal Plain. Stephen F. Austin State University, Texas Forest Paper No. 14, 46 pp.

Lenhart, J. D. and J. L. Clutter. 1971. Cubic-foot yield loblolly pine plantations in the Georgia Piedmont. Georgia Forest Research Council Report 22 - Series 3. 13 pp.

Lloyd, F. T. 1966. A description of Virginia pine diameter distributions. Unpublished thesis, Virginia Polytechnic Institute and State University. 46 pp.

McGee, C. E. and L. Della-Bianca. 1967. Diameter distributions in natural yellow-popular stands. U.S.D.A. Forest Service Research Paper SE-25. 7 pp.

Meyer, W. H. 1928. Rates of growth of immature douglas
    fir as shown by periodic remeasurements on permanent
    plots. Journal of Agricultural Research 36:193-215.

Meyer, W. H. 1930. Diameter distribution series in even-
    aged forest stands. Yale University, School of
    Forestry Bulletin 28. 105 p.

Moser, J. W. 1972. Dynamics of an uneven-aged forest
    stand. Forest Science 18-184-191.

Moser, J. W. 1974. A system of equations for the
    components of forest growth. IN Growth Models for
    Tree and Stand Simulation. Royal College of
    Forestry, Stockholm, Sweden. pp. 260-287.

Moser, J. W. and O. F. Hall. 1969. Deriving growth and
    yield functions for uneven-aged forest stands.
    Forest Science 15:183-188.

Myers, C. A. 1973. Computerized preparation of timber
    management plans: TEVAP 2. U.S.D.A. Forest
    Service Research Paper RM-115. 7 pp.

Navon, D. I. 1971. Timber ram . . . a long-range plan-
    ning method for commercial timber lands under
    multiple-use management. U.S.D.A., Forest Service
    Research Paper PSW-70. 22 pp.

Nelson, T. C. 1964. Diameter distribution and growth
    of loblolly pine. Forest Science 10:105-115.

Neyman, J. and E. S. Pearson. 1928. On the use and
    interpretation of certain test criteria for
    purposes of statistical inference. Biometrika
    20A:175-240 and 263-294.

Osborne, J. G. and F. X. Schumacher. 1935. The
     construction of normal-yield and stand tables
     for even-age timber stands. Journal of
     Agricultural Research 51:547-564.

Promnitz, L. C. 1972. A photosynthate allocation model
     for tree growth. Unpublished dissertation. Iowa
     State University. 67 pp.

Rao, C. R. 1973. Linear statistical inference and its
     applications. John Wiley and Sons, Inc., New
     York, NY. 625 pp.

Russell, D. 1970. Optimization Theory. W. A. Benjamin,
     Inc., New York, NY. 405 pp.

Schlaegel, B. E. and D. L. Kulow. 1969. Compatible
     growth and yield equations for West Virginia yellow
     poplar. West Virginia Agricultural Experiment
     Station Bulletin 573T. 20 pp.

Schnur, G. L. 1934. Diameter distributions for old-
     field loblolly pine stands in Maryland. Journal
     of Agricultural Research 49:731-743.

Schweitzer, D. L., A. L. Lundgren and R. F. Wambach.
     1967. A computer program for evaluating long-
     term forestry investments. U.S.D.A. Forest
     Service Research Paper NC-10, St. Paul, MN. 34 pp.

Smalley, G. W. and R. L. Bailey. 1974a. Yield tables
     and stand structure for loblolly pine plantations
     in Tennessee, Alabama and Georgia Highlands.
     U.S.D.A., Forest Service Research Paper SO-96.
     81 pp.

Smalley, G. W. and R. L. Bailey. 1974b. Yield tables and stand structure for shortleaf pine plantations in Tennessee, Alabama and Georgia Highlands. U.S.D.A., Forest Service Research Paper SO-97. 58 pp.

Spurr, S. H. 1952. Forest Inventory. Ronald Press Company, New York, NY. 476 pp.

Stephens, M. A. 1974. EDF statistics for goodness of fit and some comparisons. Journal of the American Statistical Association 69:730-737.

Strub, M. R. and H. E. Burkhart. 1975. A class-interval-free method for obtaining expected yields from diameter distributions. Forest Science 21:67-69.

Strub, M. R., R. B. Vasey and H. E. Burkhart. 1975. A comparison of diameter growth characteristics and crown competition factors. Forest Science 21:427-431.

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

Walker, S. H. and D. B. Duncan. 1967. Estimation of the probability of an event as a function of several independent variables. Biometrika 54:167-179.

Zacks, S. 1971. The theory of statistical inference. John Wiley and Sons, Inc., New York, NY. 609 pp.

The vita has been removed from
the scanned document

A COMPARISON OF YIELD ESTIMATION TECHNIQUES FOR OLD-FIELD LOBLOLLY PINE PLANTATIONS

by

Mike Robert Strub

(ABSTRACT)

Efficient management of loblolly pine (*Pinus taeda*) plantations requires accurate estimation of stand yield at rotation. Methods of estimating yield from stand variables, existing number of trees per acre, height of the dominant and co-dominant stand, and stand age, are presented. These methods include a multiple regression model for estimating stand yield directly from stand variables, and indirect approaches for first approximating the frequency distribution of tree diameters, and then estimating stand yield from individual tree volume equations and the approximate diameter distribution. The accuracy and precision of the stand yield estimates and the diameter distribution approximations were compared. The age at which mean annual increment culminated for various combinations of stand variables was also compared. A new stochastic model for individual tree and stand growth and yield was presented. This model was based on the physiological processes which result in tree growth.

Data required to fully develop this model were not available. Recommendations as to which methods should be used to estimate yields and approximate diameter distributions were made.