PARAMETER ESTIMATION FOR A SHORTLEAF PINE (*Pinus echinata* Mill.) BASAL AREA GROWTH MODEL

By

CHARLES EDWARD ROSE Jr.

Bachelor of Science

Northern Arizona University

Flagstaff, Arizona

1995

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
May, 1998

PARAMETER ESTIMATION FOR A SHORTLEAF

PINE (*Pinus echinata* Mill.) BASAL AREA

GROWTH MODEL

Thesis Approved:

Thomas B. Lynch
Thesis Adviser

Jeanne R. [illegible]

[illegible]

Wayne B. Powell
Dean of the Graduate College

PREFACE

The purpose of this study was to model basal area growth using a system of equations that accounts for tree interdependency within a plot by using seemingly unrelated regression (SUR) to estimate the parameters. A major regression assumption is that the error terms are independent. For Forestry applications, trees within a plot are not independent. If the independent observation assumption is violated the parameter estimates standard errors may be underestimated and the mean square error may be overestimated. Using seemingly unrelated regression to estimate the parameters of a system of equations accounts for the correlation between error terms and tree interdependency within a plot.

I wish to express my sincere gratitude to my major adviser, Dr. Thomas Lynch for his support, understanding, and patience in guiding me through the research and development process. I wish to express my gratitude to my other committee members, Drs. Lawrence Gering and Mark Payton for their support and encouragement. I wish to thank Michael Huebschmann for his guidance and patience with me in using SAS. I would also like to express my gratitude to the Oklahoma State University Department of Statistics for their guidance and support. Finally I would like to express my sincere gratitude to the Oklahoma State University Department of Forestry and the USDA Forest Service for this research opportunity and their financial assistance.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The shortleaf pine *(Pinus echinata* Mill.) forest type is classified by the USDA Forest Service as forests in which pine occupy at least 50% of the stocking of all live trees with shortleaf being the main pine species (USDA Forest Service 1972). Shortleaf pine has a range of over 440,000 square miles in twenty-two states and is the most widely distributed of the southern yellow pines (Willet 1986). Shortleaf pine distribution ranges from Texas to New York and is second to loblolly pine (*Pinus taeda* L.) in terms of total softwood volume for the southern pines. In recent years there has been concern about loblolly pine being planted outside its natural geographic range which has renewed an interest in the shortleaf pine resource (Willet 1986).

Oklahoma has approximately 765,000 acres of shortleaf pine located in the Ouachita Highlands in the southeastern area of the state. Shortleaf pine has a slower juvenile growth rate and regeneration is difficult when compared with the other southern pines, but stem and crown form are generally better (Guldin 1986), and although shade intolerant, shortleaf pine has the ability to tolerate drier upland sites (McWilliams et al. 1986). More than 94% of shortleaf pine stands originate from natural regeneration and these stands outperform other southern pines in areas where there are cold temperatures, ice, and drought conditions (Williston and Balmer 1980).

The primary timber region for shortleaf pine is found in the Ouachita Highlands of eastern Arkansas and western Oklahoma and south into Louisiana and Texas (Braun 1950). Shortleaf pine volume currently accounts for approximately 22% of the southern pine compared with 57% for loblolly pine for the southern pines (McWillams et al. 1986), but in the Ouachita Highlands shortleaf pine is the dominant pine species and represents more than 50% of the softwood volume (van Hees 1980). Shortleaf pine has been steadily declining since the 1960's, mainly due to the replacement of mature stands with other southern pines, predominantly loblolly pine (McWilliams et al. 1986). Because shortleaf pine in the Ouachita Highlands has been characterized as slow growing with below average volume per acre when compared with the other southern pines of the region (Smith 1986), national corporations such as Weyerhaeuser have steadily converted harvested shortleaf pine stands to loblolly plantations. However, non-industrial private individuals own the majority of the timberland throughout the Ouachita Highlands. Because of the high cost of planting, shortleaf pine should continue to be a valuable resource. Loblolly pine has been the dominant southern pine for commercial use in the South and has been extensively managed and planted, but there is renewed interest in the management of shortleaf pine.

Objectives

The purpose of this study is to develop a basal area growth model for natural even-aged shortleaf pine in the Ouachita Highlands of southeastern Oklahoma and eastern Arkansas. Currently the forest projection system for southeastern Oklahoma and eastern Arkansas uses a basal area growth model developed by Hitch (1994). This study seeks to improved upon Hitch's basal area growth model used in the shortleaf pine stand simulator (SLPSS) (Huebschmann et al. 1998) by:

(1) developing a system of equations based on tree diameter rank classes and using seemingly unrelated regression (SUR) to estimate the common parameters;

(2) compromising between a tree and stand-level model and accounting for error correlation between tree diameter classes;

(3) using variables that are biologically reasonable while minimizing the use of highly correlated variables; and

(4) recommending the suitability of the model through validation and evaluation.

CHAPTER II

LITERATURE REVIEW

This literature review focuses on growth and yield models for shortleaf pine and models of other species that pertain to developing a basal area growth model for natural even-aged shortleaf pine of the Ouachita Highlands. Extensive research has been conducted for growth and yield models of the southern pines with most studies concentrating on loblolly and slash pine. Growth and yield models typically use linear and nonlinear regression techniques to estimate the parameters. The classifications for system of equations and parameter estimation for seemingly unrelated regression (SUR) will be discussed. Murphy (1986) compiled a summary of growth and yield studies for shortleaf pine.

Data Classification

Growth and yield studies begin with data collection. Moser and Hall (1969) suggested that forest growth can be considered a time series and that data collection over time approximates a record of forest growth. They suggested three data classifications: "real growth series", "abstract growth series", and "approximated real growth series."

The ideal data source for developing growth and yield equations would be a complete chronological record of several stands from establishment to harvest known as a "real growth series." Because of the time and expense involved in collecting data from establishment to harvest, the "real growth series" is not practical or efficient for data collection. While the "real growth series" has an advantage of following a stand through its entire life span, it is difficult to maintain and record data from stands representing a wide variety of stand conditions for a species study.

One data collection method that has often been employed in yield studies is known as an "abstract growth series." The "abstract growth series" consists of data collected from numerous temporary plots covering a wide range of sites and ages to accurately reflect stand conditions. The "abstract growth series" is desirable for efficient data collection but individual trees can't be monitored for growth over time.

A common method for growth and yield research data collection is a compromise between the two preceding methods and is known as an "approximated real growth series" which consists of permanent plots that are remeasured at fixed intervals to approximate the rate of growth within a geographical location. While an "approximated real growth series" lacks the complete chronological history of the stand, since the remeasurement of permanent plots may approximate the ideal data source for a particular geographical location in a relatively short time it has been widely used by researchers.

The "real growth series" would be the ideal method for data collection for growth and yield equations. However, the most economical and practical method for developing growth models is based on a few repeated measurements of stands representing a variety of ages, site indices, and densities. The "approximated real growth series" gives a good approximation of actual forest growth through remeasurement of numerous plots that incorporate a variety of stand ages and site qualities.

Biological Considerations

Mathematical models for describing growth should be biologically reasonable in light of what is known concerning natural laws and biological processes. Empirical equations may be developed for data sets that accurately describe the data but have no biological basis and therefore may not accurately predict future growth (Vanclay 1994). Among the biological considerations for models of tree growth are that there is an upper asymptote relating to the maximum tree size for a given species, the growth rate after the juvenile stage is inversely related to age, and that growth rate is inversely related to the amount of competition. There are numerous biological processes restricting and limiting the growth and size of a tree that should be considered when developing a model to describe and predict growth. A common technique is to predict the maximum potential growth for a species and then modify the potential growth based on the competition for resources.

Zeide (1989) stated that growth results from cell division and is an inherently exponential process. However, except in the earliest stages of tree development, a simple exponential function may not accurately describe growth because of catabolic processes that restrict the growth of a tree. Some growth equations have two components to account for the anabolic and catabolic interaction of tree growth (e.g. Bertalanffy 1951). Growth equations may differ in structure, but all growth models should conform to reasonable biological behavior.

Bertalanffy (1951) hypothesized that growth of an organism could be thought of as the difference between the anabolic and catabolic rates. The anabolic and catabolic components are opposing forces that are usually described by subtraction or division in a differential equation describing growth. The Chapman-Richards (Chapman 1961. Richards 1959) function is a generalization of Bertalanffy's (1951) growth model, which although empirical has been used extensively in growth and yield equations and has the following form:

$$(1) \qquad \frac{\partial y}{\partial t} = py^{r} - qy$$

where

y = tree size,

t = tree age, and

p, q, and r = constants ($p, q > 0$, $0 < r < 1$).

The Chapman-Richards function's positive or anabolic term describes the cell division while the negative or catabolic components describes environmental and self-regulatory forces that oppose growth (Pienaar and Turnbull 1973). The anabolic and catabolic components conceptually describe the biological processes of tree growth. The Chapman-Richards function is widely used in forest growth and yield studies but is more empirical than theoretical because it is based on a the generalization of Bertalanffy's growth model (Yang et al. 1978). Equations in the differential form usually describe growth as either a linear function or a power function. The Chapman-Richards function is typical of the family of growth equations with the catabolic term ($-qy$) becoming more prominent and restrictive as a tree ages.

Zeide (1989) stated that growth equations are usually combinations of power and exponential functions but the relative growth rate of tree diameter is a power rather than an exponential function of age. Therefore, he proposed the following growth equation form:

$$(2) \qquad \frac{y'}{y} = at^{-b}$$

where

y = tree size,

t = tree age, and

a, b = constants ($a, b > 0$).

Zeide (1989) referred to equation 2 as the power decline since the incremental increase at any given age is proportional to tree size. Because of the correlation between diameter and crown size, the diameter indicates the amount of resources available. The "a" parameter is interpreted as the initial relative growth rate and parameter "b" represents the rate of aging. The power decline equation becomes a Schumacher type equation (Schumacher 1939) when the "b" parameter equals two. Typically the power decline predicts larger growth ratio than exponential equations because power functions decrease more slowly.

There is an extensive body of work relating mathematical models to the biological processes for describing growth processes. Zeide's (1989) study compared conventional equation forms such as the Chapman-Richards function with the power decline. The study results represented by the data set used show that the power decline equation describes growth better than the conventional equations for different species, site qualities, locations, and growth rates. A concern with the power decline was that all predicted diameters were less than actual diameters for all site indices.

Martin and Ek (1984) conducted a study which concluded that empirical equations may be more accurate than theoretical equations for a variety of data, but that theoretical equations are usually more accurate for extrapolating predictions beyond the range of the data. Yang et al. (1978) found that a modified Weibull function was flexible enough to describe most biological growth processes and possesses some desirable theoretical characteristics for modeling growth. Regardless of whether an equation is empirical or theoretical, careful consideration of explanatory variables is needed to provide realistic and robust predictions by a model (Vanclay 1994).

Model Types

In order to meet the need for growth and yield information to effectively manage the shortleaf pine resource, several types of shortleaf pine growth models have been developed. Munro (1974) suggested the following three classifications for growth and yield models: (1) stand-level models, (2) distance-independent tree-level models, and (3) distance-dependent tree-level models. These three classifications can be divided into subclasses (Davis and Johnson 1987).

According to Murphy (1986), regardless of the growth and yield model classification, models are typically either inferential or predictive. Inferential studies are designed statistically to answer specific questions about stand or tree structure. Predictive studies are designed to produce mathematical models that are used to predict growth and yield given certain stand characteristics.

The Davis and Johnson (1987) classification system will be used throughout this paper for growth and yield model classification. Relatively little growth and yield research has been conducted for shortleaf pine in comparison to the other southern pines, particularly loblolly pine.

Stand Models

Stand-level models use stand statistics such as basal area per acre, site index, trees per acre, and volume and are classified by Davis and Johnson (1987) as either density-free or variable-density models. Stand-level models require relatively little information to predict stand growth and yield but only general information is obtained about future stand conditions.

Density-free models

Density-free models use the concept of fully stocked stands to develop "normal" yield tables or average stand density empirical yield tables. The term "normal" refers to ideal fully stocked stands and is based upon the density of a stand that produces the maximum cubic-foot volume. Because "normal" fully stocked stands are subjective, few stands in reality approach the yield of "normal" yield tables. "Normal" yield tables are developed from temporary plots located in the fully stocked portion of a stand representing various ages and site indices. The plot observations are sorted by volume per unit area and site index classes, and volume is then plotted over age to obtained "normal" yield curves. The earliest "normal" yield tables for natural even-aged shortleaf pine are in the USDA Forest Service Miscellaneous Publication 50 (USDA Forest Service 1929) which also provide site index curves. The Miscellaneous Publication 50 data were obtained from 188 temporary plots located throughout the southern United

States and results are presented in tabular and graphical form. Since plots were selected based on "normal" stocking, an adjustment for yield prediction should be made when applying Miscellaneous Publication 50 to stands that are not normally stocked.

Sylvester (1938) constructed yield tables using data collected from 240 repeated measurement plots of loblolly pine in Louisiana and Arkansas and compared anamorphic guide curves and statistical methods for constructing yield tables with the yield tables of Miscellaneous Publication 50. He inferred that the yield tables of Miscellaneous Publication 50 were erroneous. In the comments appearing with Sylvester's (1938) paper, F. X. Schumacher questioned the results because of a lack of information provided in the analysis. Schumacher did note that the difference between the results and Miscellaneous Publication 50 was probably due the subjective nature of what a fully stocked or "normal" stand is and illustrated concerns associated with "normal" yield tables. "Normal" yield tables have been used to predict the growth of a stand by computing the periodic annual increment but are unreliable because the subject stand usually has less density and therefore less growth and stocking than a "normal" stand.

Empirical yield tables are developed from plots having average stand density. A volume versus age relationship is developed for these plots and thus the problem of defining "normal" is eliminated, but the average stocking is still subjective.

The next major growth and yield study for natural even-aged shortleaf pine was by Schumacher and Coile (1960) consisting of density-free growth and yield equations that have been extensively used but have some limitations concerning stocking because density is not a variable in the model. The results are presented in equation form from data collected from 74 temporary plots located in the North Carolina Piedmont district.

Variable-density models

Variable-density models use stand density as an explicit independent variable, usually expressed as basal area per acre or number of trees per acre. Multiple regression techniques are usually used to estimate parameters. Stand-level variable-density models were improvements over density-free models. Schumacher (1939) developed the prototype for the following variable-density model (Clutter et. al. 1983):

$$(3) \qquad \ln(V) = \beta_0 + \beta_1 A^{-1} + \beta_2 f(S) + \beta_3 g(D_s)$$

where

V = an expression of per acre yield,

A = stand age,

$f(S)$ = function of site index,

$g(D_s)$ = function of stand density, and

β_i = parameters.

Buckman (1962) published the first study that directly predicted growth from current stand variables in a way that was compatible with yield using data from Red pine in Minnesota. Compatible growth and yield models are defined as yield models that are derived by the mathematical integration of growth models (Davis and Johnson 1987). Clutter (1963) at about the same time as Buckman (1962) developed a compatible growth and yield model for loblolly pine. The general steps involved in Clutter's (1963) compatible equations are: (1) obtain models for current cubic foot volume and basal area per acre as functions of current age, site index, and basal area, (2) derivatives of the volume and basal area models are taken with respect to age to obtain growth models, (3) coefficients for the growth equations are estimated using linear regression, and (4) growth equations are integrated to obtain the volume and basal area projection models. The major contribution by Buckman (1962) and Clutter (1963) is the compatibility of growth and yield equations.

Murphy and Beltz (1981), using permanent plot data in Arkansas, Louisiana, eastern Oklahoma, and eastern Texas, developed the first variable-density models for natural even-aged shortleaf pine. The volume growth prediction is obtained by using the basal area growth projection equation in conjunction with the stand volume equation. The future basal area per acre is projected as a function of stand density and age and then used to predict volume as a function of site index, age, and stand density (Murphy and Beltz 1981). The following year Murphy (1982) used the same data and basal area projection equation to predict sawtimber volumes for natural even-aged shortleaf pine.

Stand volume variable-density equations for natural even-aged shortleaf pine of eastern Oklahoma and western Arkansas (Lynch et al. 1991) were developed using a stand volume equation that is related to the "Schumacher type yield model" through a logarithmic transformation. The data are from 191 permanent plots that were established in 1985-1987 in a cooperative effort between the USDA Forest Service and Oklahoma State University. These equations may be used to estimate per acre merchantable cubic-foot, sawtimber cubic-foot, boardfoot volumes of natural even-aged shortleaf pine. The volume equations have the following general form:

(4) $$V = \beta_0 \, B^{\beta_1} \, H^{\beta_2}$$

where

V = volume per unit area,

B = basal area per unit area,

H = average total height of dominants and codominants, and

β_i = parameters.

Future volume is predicted by using the projected basal area and predicted height from the site index equation. The basal area is projected using an equation that was developed by Murphy and Beltz (1981) for shortleaf pine. The projected height for the dominant and codominant shortleaf pines for the Ouachita region is obtained using the following equation (Graney and Burkhart 1973):

(5) $$H = [a_0 + a_1 SI][1 - \exp[-(a_2 + a_3 SI) AGE]]^{a_4}$$

where

H = average height of dominants and codominants,

SI = site index,

AGE = stand age, and

a_i = regression coefficients.

Equation 5 is used to predict the average height of the dominant and codominant trees given stand age and site index. An alternative for predicting the average dominant and codominant heights is to use the site index curves in Miscellaneous Publication 50 (USDA Forest Service 1929). Once the height and basal area are projected, equation 4 can be used to predict volume. The basal area projection equation developed by Murphy and Beltz (1981) has a tendency to underestimate future basal area because the model development did not account for ingrowth. A basal area projection equation is presently being developed for natural even-aged shortleaf pine of the Ouachita Highlands which should improve the volume predictions.

Diameter Class Models

Diameter class models provide more information than stand level models since they provide volumes by diameter classes. Tree diameters are placed in diameter classes and volume is computed for each diameter class. The stand volume is calculated by aggregating the diameter class volumes. Diameter class models typically use a probability distribution function such as the Weibull distribution function to allocate trees to diameter classes. Other probability distribution functions have been used with varying success such as the exponential and beta distributions. Bailey and Dell (1973) found the Weibull to be a flexible function that has the capability to assume the full range of unimodal continuous shapes of diameter distributions. The Weibull probability distribution function has the following form:

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

(6)

$$f(X) = 0 \qquad \textit{otherwise}$$

where

X = random variable,

$a \geq 0$, and

b, $c > 0$.

The Weibull probability function has three parameters commonly denoted as *a*, *b*, and *c*. The "*a*" parameter is the location parameter that indicates the lower end of the diameter distribution and must be greater than or equal to zero for forest stands (Clutter et al. 1983). The "*b*" parameter is referred to as the spread parameter and indicates the width of the function and thus indicates the width of the diameter classes. The "*c*" parameter describes the shape of the function (Avery and Burkhart 1994). When the "*a*" and "*c*" parameters are zero and one respectively, the Weibull distribution function is reduced to the exponential distribution function with the inverse "*J*" shape that is characteristic of uneven-aged forests (Clutter et al. 1983). When the "*c*" parameter is approximately 3.6, the distribution is approximately normal (Johnson and Kotz 1970). Once the parameters of the Weibull probability function have been estimated the probabilities associated with each diameter class can be calculated and multiplied by the number of trees in a stand to derive the number of trees for each diameter class.

Smalley and Bailey (1974) used a Weibull distribution function as part of a yield prediction system for shortleaf pine plantations in the Highlands of Tennessee, Alabama, and Georgia. Tree diameter, height, and age were recorded on 104 plots of shortleaf pine plantations. The tree height, survival, and age data were used to estimate the Weibull distribution parameters. Mortality was estimated by the presence of dead or dying trees. The maximum-likelihood estimates of the Weibull function parameters on each plot were related to plot age, density, and site by regression analysis. The Weibull function was then used to estimate the number of trees surviving in each diameter class and a stand table was constructed. The volumes by diameter classes were aggregated to obtain volume per acre.

Individual Tree Models

Individual tree models may have advantages over stand-level models for certain applications since they can simulate the competitive environment of each tree individually by simulating the growth of each individual tree in diameter, height, and crown. These models usually contain equations for prediction of individual tree growth, probability of survival, and volume. Results for all trees in the simulated stand are aggregated for per- acre attributes, volumes, and growth rates (Davis and Johnson 1987). Mortality and growth of an individual tree is dependent upon its relative position and size in comparison to neighboring trees. A major distinction between stand level and individual tree models is that individual tree models aggregate the stand volume after each individual tree's growth and volume is calculated whereas the stand level model aggregates individual tree data into stand characteristics before development of model equations. While individual tree models are data intensive and more time consuming to develop than stand level models, they provide more information about stand and tree dynamics.

There are two types of individual tree models, distance-dependent and distance-independent. The primary difference between the two types of individual tree models is competition accountability.

Distance-dependent

Distance-dependent individual tree models have been developed to more accurately simulate competition measures between neighboring trees. According to

Clutter et al. (1983), distance-dependent models may provide more detail concerning tree growth and the relationships between tree biological and ecological interactions but require a tree list with the spatial separation between trees as a major component of the input data. If spatial locations are unknown the simulator must generate a reasonable map of tree locations before beginning the simulation.

The main assumption of distance-dependent models is that better predictions of individual tree growth can be obtained if each neighboring tree size and location is known, but this assumption has not been empirically validated (Clutter et al. 1983). In addition to basic tree measurements, each tree's location must be plotted on X-Y coordinates which gives the location of each tree within a plot. Distance-dependent growth projections usually proceed through the following steps:

(1) competition index is computed for each tree,

(2) mortality probabilities are computed as functions of the competition indices,

(3) periodic growth rate of each tree is predicted over the projection period (usually one year), and

(4) individual tree volumes are predicted from the final projection and aggregated for stand level statistics.

PTAEDA is a distance-dependent model developed for loblolly pine (Daniels and Burkhart 1975) which grows trees individually and assumes that each tree has a theoretical maximum growth potential. PTAEDA was developed for managed loblolly pine plantations to estimate the influence of different sites, spatial patterns, and silvicultural regimes on tree growth. The growth potential is based upon the maximum growth of an open grown tree and tree vigor and competition factors modify the potential growth. The initial stand is determined by the location of each tree in the planting spatial pattern. The PTAEDA model works in two stages: first a Weibull probability distribution function determines the diameter for each individual juvenile tree and the model grows the initial juvenile stand until inter-tree competition begins. Then in the second stage the model accounts for inter-tree competition. The PTAEDA model grows in the juvenile stage until the *CCF*[1] (Crown Competition Factor) (Krajicek et al. 1961) reaches 100 percent, at which time it is assumed the inter-tree competition begins (Davis and Johnson 1987). The distance-dependent competition measure used by the PTAEDA model is a distance-weighted size ratio (DR) index developed by Hegyi (1974). Distance-weighted size ratio indices are defined as the sum of the ratios between the dimensions of each competitor to the subject tree weighted by a function of intertree distance. This definition of competitors has been preferred in recent years and has an advantage of being easy to compute and explaining variation in growth with precision similar to that of other distance-dependent competition indices (Tome and Burkhart 1989). The PTAEDA model has proven to be effective for southern pine plantations but has shortcomings for natural regenerated stands (Davis and Johnson 1987).

In addition to the distance-weighted size ratio competition index used in PTAEDA, numerous competition indices have been developed in an attempt to increase the precision of growth and yield predictions. Most distance-dependent competition indices begin with Staebler's (1951) concept of a circular influence zone around a subject

---

[1]*CCF* is a stand density measure that describes the available area for the average in relation to maximum area of an open grown tree (Avery and Burkhart 1994).

tree in which the competing trees reduce the rate of growth of the subject tree. The overlap from circular influence zones of competing trees measures the amount of competition. Spurr (1962) developed a point density measure in which trees included in a fixed point sampling angle gauge (prism) sweep are considered to be competitors. Brown (1965) developed the area potentially available index as a measure of point density in which the area available for each tree was calculated as the area of the smallest polygon bisecting the intertree lines.

Studies show that distance-dependent competition indices contribute little or no improvement in growth prediction when compared to use of distance-independent measures of competition (Tome and Burkhart 1989).

Distance-independent

Distance-independent models usually project tree growth as a function of diameter and stand level variables, typically having three main components: diameter growth, height growth, and mortality (Avery and Burkhart 1994). Distance-independent models assume that the spatial separation and tree diameters are uniformly distributed throughout the stand (Davis and Johnson 1987). Since the exact location of each tree within a stand is unknown, competition is normally defined by a comparison of a tree's characteristics with that of other trees within a stand. Since distance-independent models do not use spatial information to formulate competition indices, they are less data intensive than distance-dependent models. An assumption of distance-independent models is that if a tree is smaller than the average tree within the stand regarding crown, diameter, and height then the tree lacks competitive vigor in comparison with larger trees.

Two commonly used competition indices are *DD* (ratio of the quadratic mean diameter to an individual tree *DBH*) and *BAL* (the cumulative basal area of trees larger than the subject tree). As the *DD* index decreases the tree is considered to be increasingly vigorous and will grow at rates closer to its maximum theoretical potential. The *BAL* index indicates that as the cumulative basal area of the trees larger than the subject tree decreases, the subject tree is more competitive. The *BAL* index is used in Stage's (1973) PROGNOSIS forest projection system. Krumland (1982) developed a competition index that is based upon a tree's crown and is defined as the percentage of land covered by a live tree crown (measured at a height of 66 percent up the live crown of a subject tree). This competition index has been shown to be effective but is difficult to measure.

Basal area growth is often used as the dependent variable for increase in stem thickness because studies show that the correlation between various competition indices are higher with tree basal area increment than with diameter increment (Bella 1971; Johnson 1973). Since competition among trees for resources is a major component of individual tree models it seems intuitive that basal area growth might yield the better model. A study by West (1979) compared the results of growth equations using both basal area growth and diameter growth as the dependent variable and found that the correlation between tree basal area growth and parameters determining the growth were higher than for tree diameter increments. There was no evidence that the precision of estimates of predicted diameters made with either diameter or basal area growth as the dependent variable differed significantly. West (1979) concluded that the higher correlation of basal area growth equations was probably due to the partial dependence of basal area growth on the initial tree diameter.

The two primary methods for modeling distance-independent tree growth are composite modeling of tree growth as a function of tree, site, and stand characteristics and the potential-modifier growth function. The potential-modifier function models tree growth as a theoretical maximum potential based on an individual's tree characteristics which is multiplied by a modifier to account for stand and tree characteristics as well as competition.

PROGNOSIS is a composite distance-independent individual tree model developed by the USDA Forest Service for the western United States (Stage 1973) that directly predicts the growth of a tree. The PROGNOSIS forest projection system uses the composite growth model approach because the difficulty in obtaining dominant-age site relationships for mixed species stands makes potential growth difficult to estimate. The PROGNOSIS growth model has the following general form (Wykoff 1990):

$$(7) \qquad \ln(dds) = COMP + SITE + \beta_1 \ln(dbh) + \beta_2 \, dbh^2$$

where

$COMP$ = function of competition measures,

$SITE$ = function of site quality,

dbh = diameter at breast height,

dds = 10-year periodic change in squared diameter (inches), and

β_i = regression coefficients.

The PROGNOSIS model predicts the natural logarithm of new growth in square inches (*dds*) as a function of site, tree, stand, and competition characteristics and is easily converted to either basal area or diameter. Since PROGNOSIS was developed for mixed conifers it uses habitat type, geographic location, slope, aspect, and elevation to express site quality rather than site index. Crown ratio, crown competition factor (*CCF*), and the cumulative basal area of trees larger than the subject tree (*BAL*) measure competition. The composite model approach has performed adequately and is used extensively throughout the western United States but is not often applied to southern pines. Wykoff (1990) presents a detailed discussion of the model development and performance.

The potential-modifier growth function used in many forest projection systems consists of the potential component which estimates the theoretical maximum diameter or basal area growth of a tree growing free of competition and a modifier component that reduces the potential based on competition factors. Parameters of potential-modifier growth functions have been fitted in two ways: (1) parameters for a tree's potential growth are fitted based on an individual tree characteristics, then the modifier parameters which are a function of tree and stand characteristics are fitted while holding the potential parameters constant, and (2) parameters for the potential-modifier function are fitted simultaneously.

The STEMS forest projection system, which was developed for the Lake States of the United States, uses a potential-modifier function to estimate individual tree growth (Belcher et al. 1982). The STEMS potential diameter growth function that was developed for 26 species by Hahn and Leary (1979) using data collected from the dominant and codominant trees throughout the Lake States has the following form:

$$(8) \qquad Potential \quad Growth = \beta_1 + \beta_2 D^{\beta_3} + \beta_4 (SI)(CR)(D)^{\beta_5}$$

where

Potential Growth = potential annual dbh growth (inches/year),

D = initial tree dbh (inches),

SI = plot site index (base age = 50),

CR = tree crown ratio, and

β_i = regression coefficients.

The potential diameter growth was derived for each tree species over a ten-year projection period. It uses the initial diameter, crown ratio, and site index as the independent variables. In developing the potential growth function, Hahn and Leary (1979) grouped the independent variables by species, one-inch diameter classes, 10-foot site index classes and 10-percent crown ratio classes. It was assumed that the diameter growth was normally distributed around the mean diameter growth. The potential growth was estimated as the mean growth plus 1.65 standard deviations corresponding to the 95th percentile of dominant and codominant tree diameter growth. The mathematical form of the potential growth model was developed as a generalization of the Richards function (Richards 1959) and has a species-specific intercept along with anabolic and catabolic terms. The catabolic term is an allometric relation of tree diameter and the anabolic term is a product of crown length, site index and an allometric relation of tree diameter. Leary and Holdaway (1979) developed a modifier function for use with Hahn and Leary's (1979) potential growth function which is a function of size, site, and competitive status and has the following form:

$$(9) \quad Competition \quad Modifier = 1 - e^{-f(R)g(AD)[(BA_{\max} - BA)/BA]^{0.5}}$$

where

Competition Modifier = competition index (bounded between 0 and 1),

$f(R)$ = function characterizing the individual tree's relative diameter effect on the modifier,

$g(AD)$ = function characterizing the average stand diameter effect,

BA_{max} = maximum basal area per acre, and

BA = current basal area per acre.

In estimating the parameters of the modifier function, the parameters of the potential growth function are held constant and because the competition modifier adjusts potential growth which is constrained to be between zero and one. As the stand basal area and competition increase the competition modifier approaches zero, and conversely, the competition modifier approaches one as stand density and competition decrease to allow trees growing relatively open and free of competition to approach the theoretical potential growth.

The first distance-independent growth model for shortleaf pine was developed as part of the TWIGS and STEMS forest projection system for the central United States (Shifley 1987). The growth model data came from Forest Inventory and Analysis (FIA) plots in Missouri, Ohio and Indiana and the model uses a potential-modifier growth function. Unlike the Lake States potential growth function, which is the sum of the two components plus an intercept, the Central States potential growth function is the product of two components and was developed as a function of tree size, crown ratio, and site index for each species. The Central States potential growth function has the following form:

$$(10) \quad Potential\ Growth = [\beta_1 BA^{\beta_2} - \beta_3 BA]\,[\beta_4 + \beta_5 SI + \beta_6 CR]$$

where

Potential Growth = potential tree basal area growth (ft^2/year),

BA = tree basal area (ft^2/ac),

SI = plot site index (base age = 50),

CR = individual tree's crown ratio, and

β_i = regression coefficients.

The first bracketed component in the potential growth function above is the growth form of the Chapman-Richards function which is used to model maximum potential growth and the second bracketed component adjusts potential basal area growth by the tree crown ratio and site index, increasing or decreasing maximum potential growth. The potential growth function was constrained to achieve maximum tree diameter within biologically reasonable limits by comparing results from the potential growth model based on the fastest growing five percent of the trees by species and diameter class to the National Register of Big Trees (American Forestry Association 1982).

Shifley and Brand (1984) used a variation of the Chapman-Richards function to account for a maximum tree basal area since potential growth approaches zero as the basal area approaches the biological upper limit. The constraint was implemented by using the first bracketed component of equation 10 and setting the potential growth to zero. Solving for basal area yields the following equation:

(11) $$Max\ \ Basal\ \ Area = \left(\frac{\beta_1}{\beta_3}\right)^{\frac{1}{1 - \beta_2}}$$

where

Max Basal Area = species estimated maximum basal area (ft^2) and

β_i = estimated parameters.

The maximum basal area constrains equation 10 to a pre-specified upper asymptote and was developed to ensure that the model behaves realistically when predicting growth for large diameter trees. Solving equation 11 for β_3 and substituting into equation 10 yields the Central States TWIGS present potential growth equation form.

The Central States modifier component reduces the potential growth in response to competition and expresses growth as a percentage of the estimated potential growth. The modifier has the following form:

(12) $$Modifier = \beta_9 \left[1 - e^{-\left(\frac{\beta_7}{BAL} + \beta_8 D^2\right)\left(1 - \frac{BA}{BA_{max}}\right)^{\frac{1}{2}}} \right]$$

where

Modifier = competition modifier (bounded between 0 and 1),

BAL = basal area (ft^2/ac) of all trees larger than subject tree ,

D = current tree dbh (inches),

BA_{max} = maximum basal area per acre,

BA = stand basal area (ft^2/ac), and

β_i = regression coefficients.

The Central states modifier decreases growth as the basal area of larger trees increases and larger diameter trees receive a smaller proportion of potential growth. Shifley (1987) defined a maximum stand density of 200 square feet of basal area per acre (BA_{MAX}) to constrain the modifier within biological reasonable limits. The shortleaf pine potential modifier function of TWIGS has performed adequately but was fitted to data outside the Ouachita Highlands geographical region.

Hitch (1994) developed the first shortleaf pine distance-independent individual tree for the Ouachita Highlands, which is currently being used in the SLPSS forest projection system for even-aged natural shortleaf pine. The data for this study came from permanent research plots established cooperatively by the USDA Forest Service and the Forestry Department at Oklahoma State University. The plots were established in 1985-87 with remeasurements occurring at four or five year intervals. All shortleaf pine having a diameter of greater than or equal to one inch were recorded and each tree permanently numbered. The following is a summary of basal area growth models Hitch (1994) fitted and evaluated.

A modified PROGNOSIS type model was fitted using a single intercept in place of the location effects and site effects were estimated by site index. PROGNOSIS was developed for uneven-aged stands of mixed conifers in the western United States, consequently Hitch (1994) modified the model to include an independent variable to represent stand age. The model performed adequately with a fit index and *MSE* of 0.593 and 0.000046, respectively. However, the model predicted a maximum tree size of more than 70 inches *DBH*. The large maximum tree diameter is a result of the fact that the research plots do not include stands in which the basal area growth has culminated.

Potential-modifier functions were fitted to the data and evaluated in an effort to better describe growth. Hitch (1994) used three methods to constrain the maximum growth of the potential function. The first was a modified form of Hahn and Leary's (1979) STEMS potential growth model with two-inch diameter classes in which the potential growth was estimated as the mean growth plus 1.65 standard deviations corresponding to the 95th percentile of basal area growth. The second method fitted the fastest growing five-percent of the trees using the dominant and codominant trees in each one inch diameter class with a modified Chapman-Richards function (Shifley and Brand 1984; Shifley 1987) with tree size constrained to a biological reasonable maximum size. The maximum tree size was derived as 36 inches (basal area = 7.068 ft^2) by averaging the maximum diameter found from local records and the largest shortleaf pine recorded in the National Records of Big Trees (American Forestry Association 1992). The third method was suggested by Amateis et al. (1989) who proposed using the data from open grown trees to approximate the potential growth for trees growing without competition. The data for the potential function came from a study of open growing shortleaf pine in the West Gulf Region (Smith et al. 1992). The potential-function was a variation of the Chapman-Richards function that included an intercept component. All three methods appear to constrain the maximum growth within biologically reasonable bounds and were coupled with a modifier to determine which would best fit the data.

Hitch (1994) considered two modifier functions. The first modifier was Shifley's (1987) variation of the STEMS and TWIGS function for the Central States. The second modifier was a modified logistic function developed by Murphy and Shelton (1993). This modifier easily adapts to a variety of stand and tree conditions by including more independent variables to explain more variation in tree growth and has the following form:

$$(13) \qquad Modifier = \frac{1}{1 + \exp\left[\beta_1 BAL + \beta_2 SI + \beta_3 AGE + \beta_4 BA + ...\right]}$$

where

$Modifier$ = competition index (bounded between 0 and 1),

BAL = basal area (ft^2/ac) of all trees larger than subject tree,

SI = plot site index (base age = 50),

AGE = plot age (years),

BA = stand basal area (ft^2), and

β_i = regression coefficients.

The final model Hitch (1994) developed used the modified Chapman-Richards potential growth function (Shifley and Brand 1984) coupled with the modifier developed by Murphy and Shelton (1993). The model predicts growth adequately with a fit index of 0.609 and MSE of 0.000044. Although the model adequately predicts basal area growth, there is concern about model bias because of under-predictions for stands with less than 45 square feet of basal area per acre and a under-representation of young stands with a site index greater than 65 (base age equals 50 years). The model was developed using data from stands that have not reached culmination of basal area per acre and may over or under predict for stands beyond the range of the data.

Bitoki et al. (1997) developed a distance-independent shortleaf pine individual basal area growth model for uneven-aged stands in the Ouachita Highlands. The data came from CFI plots established by the Deltic Farm and Timber Company Inc. in 1965-66 with remeasurement occurring at five-year intervals. Individual trees were measured and recorded for all trees greater than or equal to five-inch dbh. Plots used to develop the model had no harvesting activities or silvicultural treatments during the five-year measurement interval and shortleaf pine comprised at least 70% of the basal area of each plot. The basal area growth model was developed using a potential-modifier basal area growth model for which the parameters were fitted in two steps. First the theoretical potential growth was fitted separately using a variation of the Chapman-Richards function (Shifley and Brand 1984) with one parameter eliminated by using Hitch's (1994) estimate for maximum basal area of an individual shortleaf pine. The other two parameters were fitted using nonlinear regression. The second step consisted of fitting the modifier function parameters to the complete potential modifier model by using nonlinear regression while holding the potential growth constant. This was a variation of the model used in TWIGS and described by Shifley (1987). The model fit appears adequate with a fit index of 0.44, but because the study data came from CFI plots, the data do not represent all stand conditions equally. There is concern about diameter growth predictions for large trees and under-represented site index classes because there are few observations for trees greater than 16 inches *DBH* and for site index classes 40 and 70 (base age equals 50 years).

Potential-modifier models usually estimate the parameters by estimating the maximum potential growth in isolation and then holding potential growth constant while estimating the parameters for the modifier. The potential function is normally fitted separately using either a subset of the data or open grown tree data and the modifier is fitted to the complete data set while holding the potential function constant. Often a series of iterations of the two steps are completed to stabilize the parameters. Murphy and Shelton (1996) proposed fitting the data to the potential modifier function simultaneously. Previously Wensel et al. (1987) fitted a potential modifier function simultaneously for northern California conifers but found that it confounded the potential and modifier effects. Murphy and Shelton's (1996) data are from a study on the growth and development of loblolly pine in Arkansas and Louisiana. They selected a growth function that is biologically reasonable for achieving a maximum growth rate. The potential growth function selected bounded the function to an upper asymptote. Then a variation of the logistic function (equation 13) was selected as the modifier, which is constrained between zero and one. The potential function was fitted to obtain an estimate of the parameters and then the potential-modifier function was fitted simultaneously using nonlinear OLS (Ordinary Least Squares) using the potential growth function estimates as initial values. The results reveal the model achieved a good fit with a fit index of 0.69 and root mean square error of 0.56 square centimeters. The model exhibits logical results when compared with biological processes. The logistic function modifier can be easily adapted to add more variables as needed. However because the components are fitted simultaneously, the potential component cannot be analyzed in isolation and does not necessarily equal the maximum theoretical potential growth.

System of Equations

Growth and yield studies often use a system of equations to describe stand development. Early applications of systems of equations in forestry fitted the parameters of each equation independently using OLS (e.g. Moser 1972). Furnival and Wilson (1971) suggested that fitting the parameters for each equation in a system independently was not satisfactory because a variable may be dependent in one equation and independent in another equation. Therefore, coefficients of one equation may be functionally related to coefficients in another equation, and the residuals of each equation may be correlated. Furnival and Wilson (1971) proposed that parameters for a system of equations describing forest growth and yield could be fitted simultaneously using known econometric techniques. Simultaneously fitting parameters of a system of equations provides an increase in parameter estimation efficiency.

The general simultaneous parameter estimation technique for a system of equations may be applied to linear or nonlinear systems with small, large, or unequal sample sizes and with or without imposing constraints between parameters of the system (Reed 1987). The optimal parameter estimation technique is defined by a given system and empirical studies suggest efficiency is gained because of a reduction in the SSE (error sum of squares) (Reed 1987). Parameter estimation for equations within a system is accomplished using either linear or nonlinear regression techniques to minimize the SSE.

There are three steps typically used for estimating parameters of a system of equations. First the parameters for each equation within a system are independently estimated by using linear or nonlinear OLS. Second, the variance-covariance matrix is estimated using the error terms of the independently estimated equations. Third, the generalized least squares is used to estimate the parameters of the system using the residual variance-covariance matrix (parameter constraints may be imposed) (Reed 1987).

Pindyck and Rubinfeld (1981) classified systems of equations as simultaneous equations, recursive equations, or seemingly unrelated equations. Systems of simultaneous equations have variables that are independent in one equation and dependent in another equation and cross-equation error correlation exists (equation 14). Parameter restrictions may occur both within and across equations and independent variables are either endogenous (determined previously in the system of equations) or exogenous (determined independently of the system of equations). Two stage and three stage least squares may be used to estimate biased but consistent parameters (Borders 1989).

$$
\begin{aligned}
Y_1 &= \beta_{10} + \phantom{\beta_{21}Y_1 +{}} \beta_{12}Y_2 + \beta_{13}Y_3 + \beta_{14}X_1 + \beta_{15}X_2 + e_1 \\
Y_2 &= \beta_{20} + \beta_{21}Y_1 \phantom{{}+ \beta_{12}Y_2} + \beta_{23}Y_3 + \beta_{24}X_1 + \beta_{25}X_2 + e_2 \qquad (14) \\
Y_3 &= \beta_{30} + \beta_{31}Y_1 + \beta_{32}Y_2 \phantom{{}+ \beta_{13}Y_3} + \beta_{34}X_1 + \beta_{35}X_2 + e_3
\end{aligned}
$$

where

Y_i = endogenous variables,
X_i = exogenous variables, and
e_i = error term.

Recursive systems of equations have sequential relationships between endogenous variables and OLS can be used for parameter estimation if no cross equation correlation between the error components exists. A recursive system of equations by definition has no correlation between right-hand side endogenous variables and the error components of the left-hand side endogenous variables. The OLS estimation for recursive system of equations requires that error components for the system of equations be pairwise uncorrelated (Borders 1989). Because recursive systems of equations are not cross correlated there is no gain in efficiency for simultaneous parameter estimation unless constraints are imposed between equation parameters (Reed 1987).

Seemingly unrelated equations have no analytical relationship between equations but are linked because the error terms across equations are correlated (Pindyck and Rubinfeld 1981). If the cov $(e_i, e_j) = 0$ for all combinations of i and j then SUR is inappropriate, but if the cov $(e_i, e_j) \neq 0$ then a correlation between the errors of the equations exists and SUR may provide a gain in parameter estimation efficiency. The three step procedure for estimating parameters for a system of equations discussed previously is appropriate and parameter estimation efficiency is improved by accounting for the cross equation correlation. Zellner (1962) suggests that efficiency in parameter estimation may be gained if the system is viewed as a single equation and is accomplished by using the generalized least squares estimation.

Lynch and Murphy (1995) used SUR for parameter estimation in developing a compatible height prediction and projection system for natural even-aged shortleaf pine of the Ouachita Highlands. The study data are from 208 permanent growth and yield research plots established from 1985-89 and the data include individual tree heights, diameters, plot densities, ages, and site indices. The two compatible height equations developed are: (1) height prediction for time one and (2) height prediction/projection for time two given the height at time one. The SUR parameter estimation technique was used because error correlation was expected from the common parameters and because heights measured at different times are correlated. SUR was used for efficient parameter estimates for both equations by providing consistency across equations with parameter restrictions placed across equations. OLS estimates for each individual equation were presented for comparison to SUR estimates but the OLS individually estimated parameters differed by as much as 20% from the SUR parameter estimates. Since the equations are compatible and interrelated using the first equation with OLS parameters to project height in the second equation may be inaccurate. Studies have indicated that parameter variances estimated from large samples using SUR may be less than parameter variances obtained using OLS (Judge et al. 1988). The advantage of SUR in the compatible height prediction and projection system is that the parameters may be used by either equation depending upon available information for height prediction. SUR provided a good fit for both equations while accounting for error correlation between the two equations.

Hasenauer et al. (1998) recently fitted a system of three equations separately using OLS, and simultaneously using two- and three-stage least squares. The three equations are a basal area growth increment model, height increment model. and crown ratio model. The data are from the Austrian National Forest Inventory which consists of over 7,500 Norway Spruce (*Picea abies* L. Karst). Since the results indicated high cross-equation correlation, the three-stage least squares was the most efficient technique.

CHAPTER III

DATA

The data are primarily from the cooperative study being conducted by the USDA Forest Service and Oklahoma State University Department of Forestry to develop growth and yield models for natural even-aged shortleaf pine stands of the Ouachita and Ozark National Forests of southeastern Oklahoma and eastern Arkansas. These data constitute an approximated real growth series. Original plot installation was during the dormant season of 1985-1987 when basic forest measurements were recorded, subsequently remeasurements were recorded on a four or five year interval for each plot. For a detailed discussion of plot reconnaissance, installation, and location see the USDA Forest Service establishment and progress report (Murphy 1988a).

A total of 191 plots located in the Ouachita and Ozark National Forests were installed which covered a wide spectrum of site, age, and density classes. There are four site index, age, and density classes for a total of 64 combinations (Table 1).

Each combination of the site-age-density classes originally was to have 3 replicates for a total of 192 plots. However only two plots were located for the age 20 years, site index > 75 feet, and residual basal area 30 square feet combination. Therefore, only 191 plots were actually installed (Murphy 1988a).

Table 1. Attributes and class ranges for the USDA Forest Service-Oklahoma State University cooperative research plots for the natural even-aged shortleaf pine growth and yield study.

Attribute	Class Range	Class Midpoint
Basal area per acre (sq.ft.)	16-45	30
	46-75	60
	76-105	90
	106-135	120
Site index (base age = 50 years)	< 56	
	56-65	60
	66-75	70
	> 75	
Age (years)	11-30	20
	31-50	40
	51-70	60
	71-90	80

The original plots were selected based upon age-site-density classes and the following stand criteria:

(1) naturally regenerated stands containing at least 70 percent shortleaf pine in terms of basal area for trees 0.6 inches *DBH* and larger;

(2) maximum age range of dominant and codominant trees was 10 years or less;

(3) less than 10-foot variation for site index within a stand;

(4) even-aged forest distribution with no obvious holes or clumping and no more than two age classes per plot; and

(5) no significant insect, disease, or fire damage and no harvesting during the previous five years.

For plot reconnaissance, the shortleaf pine and other species plot basal area were tallied separately using a 10-factor prism. Five dominant or codominant shortleaf pines were selected at each plot for height and age measurement. The initial stand information was used for plot site-age-density classifications and to assign silviculture prescriptions.

The study design called for 0.2 acre circular plots, surrounded by a 33- foot isolation buffer for each site-age-density combination. Silviculture prescriptions were required to control any existing hardwoods greater than or equal to one-inch diameter at ground level and shortleaf pine were thinned when necessary to achieve the desired basal area for both the plot and buffer. The residual shortleaf pines on the 0.2 acre plot were numbered, measured, located from plot center, and tallied for all trees greater than or equal to one inch *DBH*. In addition, the crown class of each tree was recorded.

At each plot representative shortleaf pines for each diameter class were selected for measuring total height and the height to live crown. The age for the representative dominant and codominant sample trees was determined using an increment borer. The annual rings of each increment core were counted and five years added to derive tree age. The plot site index was calculated using a site index equation developed by Graney and Burkhart (1973) for shortleaf pine of the Ouachita Highlands based on the average total height and age of the representative dominant and codominant trees. The *DBH* was measured to the nearest tenth of an inch for all shortleaf pine greater than or equal to one inch. Stand basal area was derived by summing the basal area of individual trees on a plot and expanding to a per acre basis. The crown ratio was computed as the ratio of the crown length to the total height for trees that were measured for height. For trees not measured for height, the crown ratio was predicted using the model form developed by Dyer and Burkhart (1987) with the parameters fitted using nonlinear regression in SAS (SAS Institute Inc. 1989). The following distance-independent individual tree competition measures were computed for each plot. The basal area of all trees as large or larger than the subject tree (*BAL*) was calculated and expanded to a per acre basis. The *CCF* was computed using a technique by Rogers and Sander (1984) for shortleaf pine and the ratio of the quadratic mean diameter to individual tree dbh (*DD*) was calculated. For individual growth models discussed previously additional attributes were computed when appropriate. The midpoint of an attribute was calculated when appropriate by using the observed values attained by an attribute at time one and two and averaging the two observations. If mortality occurred for a given tree between measurement periods then half of the initial measurement was used, a technique described by Bolton and Meldahl (1990).

In addition, data from 25 plots comprising an "approximated real growth" series were available from a study initiated by Frank Freese in thinned stands of even-aged natural shortleaf pine in the Ouachita Highlands. The study was initiated in 1963-64 and 25 of the 35 0.2-acre plots that were installed in the Ouacita National Forest still exist. The initial 35 plots were installed with the following stand residual basal area per acre: 45, 65, 85, 105, and 125 square feet. The remaining 25 plots were assigned to these residual basal area levels in 1988. Remeasurements of these plots during 1988 were consistent with the methods discussed previously. In addition, the other 10 plots were available for developing the growth model by utilizing the historical data on these plots. One of these 10 plots was removed from consideration because of fire damage. This left a total of 34 plots for developing the growth model. Murphy (1988b) provides further details for this study.

Eight plots from the cooperative growth and yield study were decommissioned because of stand damage or failure to follow silvicultural prescriptions. The remaining 183 plots from this study, together with the 34 plots from Freese's study, constitute a total of 217 plots that are available for developing growth and yield equations for natural even-aged shortleaf pine of the Ouachita Highlands. The data set consists of 8928 individual tree observations for which the summary statistics (Table 2) were computed.

Table 2. Summary statistics for the complete data set of the Ouachita Highlands natural even-aged shortleaf pine study for developing a basal area growth model (N = 8928).

Attribute	Minimum	Maximum	Mean	Standard Error
Plot Age (years)	21.0	96.0	45.1	18.71
Plot Site Index (base age=50)	38.9	87.1	57.3	9.70
Stand Basal Area (sq. ft.)	22.53	177.12	106.80	30.93
DBH (inches)	1.2	24.9	7.93	3.74
Crown Ratio	0.1310	0.7636	0.3799	0.063
CCF	13.72	255.44	132.32	48.86
Avg. Annual Individual Tree Basal Area Growth	-0.010	0.0718	0.0124	0.0102
Individual Tree Basal Area	0.00723	3.383	0.4196	0.3795
Quadratic Mean Diameter to Individual Tree *DBH* (*DD*)	0.4424	4.241	1.0875	0.3205
Basal Area of all Trees as Large or Larger than the Subject Tree (*BAL*)	0	171.37	66.00	37.80

The number of trees by *DBH* class was calculated for the complete data set (Figure 1). As illustrated by the graph, there are fewer observations in the 2, 14, 16, and 18-inch *DBH* classes in comparison with the other *DBH* classes. The number of trees in the 14, 16, and 18 *DBH* classes should increase over time as trees grow into these classes. Since the study entails thinning from below, the future number of trees in the 2-inch *DBH* class may not increase and could decrease.

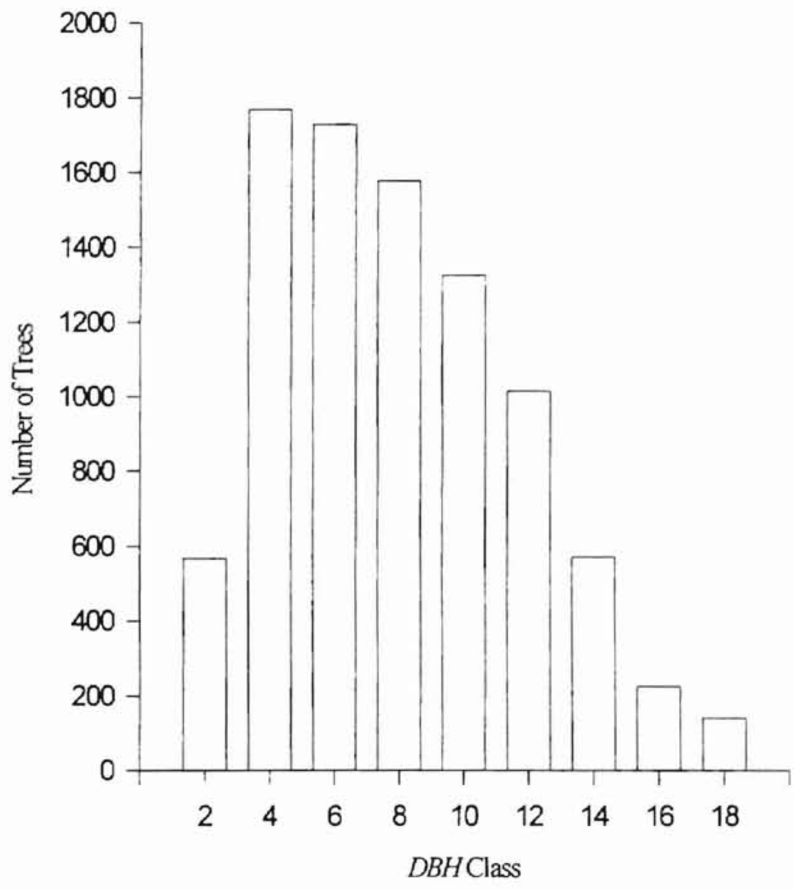

Figure 1. Number of trees by diameter class for the complete data set (N = 8928).

CHAPTER IV

METHODS

Model Considerations

Among the major linear regression assumptions is that the errors are independent and identically normally distributed. Nonlinear and linear regression models differ in that the nonlinear model least squares estimators of their parameters are not unbiased, normally distributed, or minimum variance estimators. Ratkowsky (1990) stated that the regression assumptions for nonlinear regression need only be correct approximately because the least squares criterion tends to be robust in minor departures from the assumptions. Major departures from the regression assumptions such as a dependency between the error terms can lead to significant estimate errors. Except for the independence assumption, nonlinear regression models tend to conform to the linear regression assumptions asymptotically as the sample size approaches infinity. Trees within a plot have some interdependency because of the competition for resources on the plot. Also trees within a plot share a similar microenvironment, which may be above or below average for tree growth. When the assumption of independent errors is violated, the *MSE* and standard error of the parameter estimates may seriously underestimate the variance of the error terms and the standard error of the parameter estimates when calculated according to the OLS procedures (Neter et al. 1996). A system of equations was developed to account for tree interdependency within a plot and SUR was used for parameter estimation. The SUR parameter estimation is an appropriate technique because the error terms from trees within a plot are correlated. To determine the feasibility of using a system of equations to model even-aged natural stands of shortleaf pine some of the model forms discussed previously were examined.

Since the basal area growth model may be applied outside the data geographical range, the model should behave within known biological limitations when predicting the maximum basal area or diameter for shortleaf pine. The model should follow a realistic growth pattern and behave logically with respect to the independent variables. As growth approaches zero the individual tree basal area should approach an upper asymptote that corresponds to a biologically reasonable diameter.

Individual tree growth may be predicted as basal area increment or diameter increment or as a function of either basal area or diameter. As discussed previously, West (1979) concluded that for the species in his growth study, there was no significant difference between the predictions obtained using either basal area or diameter increment to model growth. Individual tree diameter growth culminates before basal area growth and thus may have an advantage over basal area growth for predicting growth in young stands. The fit index is normally higher for basal area growth than for diameter growth using the same data because the range of diameter growth is smaller in comparison to the range of basal area growth (Shifley 1987). Consequently, the denominator of the fit index (corrected total sum of squares) is smaller for diameter growth than for basal area growth. Individual tree volumes are normally proportional to the product of tree height and basal area and consequently it is logical to predict basal area growth directly. Also, because the initial diameter is known for this study, diameter growth can be computed from basal area growth. Whether to model growth using basal area or diameter is normally determined by study objectives. Since the growth model developed may be used in the shortleaf pine simulator for the Ouachita Highlands, it is logical to use a basal area growth model. The average annual basal area growth (*AABAG*) was used to model growth since remeasurements occurred over different intervals. The *AABAG* facilitates using the model over any projection period.

The dependent variables in model development focused on variables that are currently obtained during the inventory and variables that can be derived from basic forest measurements. Some measurements typically taken during an inventory include *DBH*, site index, stand age, and basal area per acre. Competition measures such as *CCF* can be computed from the inventory measurements. The model will be distance-independent and attributes discussed earlier for distance-independent models will be considered. Independent variables used in the model development will be examined for significance using an alpha level of 0.05. Independent variables deemed insignificant will be removed from the model.

Revised Data Set

The data set was revised for modeling basal area growth using a system of equations. The current data set contains one record for each individual tree. This data set was revised to create four classes corresponding to the individual tree diameters within each plot. The revised data set for modeling basal area growth using a system of four equations was developed through the following steps. Individual tree diameters within each plot were ranked in ascending order. The four *DBH* rank classes were computed by dividing the ranked *DBH* tree list by four. If the ranked *DBH* tree list was evenly divisible by four then each *DBH* rank class had an equal number of trees. The *DBH* rank class one corresponds to the *DBH* of the smallest tree(s). A subroutine program was written in SAS to place trees in the correct *DBH* rank classes if the number of trees on a plot was not evenly divisible by four (Appendix A). For example (Table 3), if a plot has six trees the first and second class would have two trees and the remaining two classes would have one tree.

Table 3. Example of ranking and placing trees in correct *DBH* rank class by plot for use in a system of equations.

Plot Number	*DBH* (inches)	Rank	*DBH* Rank Class
134	6.3	1	1
134	7.1	2	1
134	8.2	3	2
134	10.3	4	2
134	10.9	5	3
134	13.2	6	4

A visual inspection of the revised data was conducted to insure that trees were placed in the proper class. The revised data set has a total of 217 plot records. Plot 261 was removed when fitting the system of equations models to the complete data set because it has only two trees and could not be used for parameter estimation in the system of equations.

Development of a System of Equations

The revised data set was used for modeling the system of four equations. The system of four equations for each model has the following general form:

$$\overline{AABAG_1} = \frac{1}{n_1}\sum_{j=1}^{n_1} f(X_1) + e_1$$

$$\overline{AABAG_2} = \frac{1}{n_2}\sum_{j=1}^{n_2} f(X_2) + e_2$$

$$\overline{AABAG_3} = \frac{1}{n_3}\sum_{j=1}^{n_3} f(X_3) + e_3$$

$$\overline{AABAG_4} = \frac{1}{n_4}\sum_{j=1}^{n_4} f(X_4) + e_4$$

where

$AABAG_{ij}$ = average annual basal area growth for tree j of class i within a plot,

f = a function of stand and individual tree characteristics,

X_i = a vector of stand and class i individual tree characteristics on a plot,

n_i = number of trees in class i, and

e_i = error component associated with each class within a plot.

Seemingly unrelated regression is an appropriate technique for estimating parameters in this system of equations because a correlation is expected between the error components within each plot for the four classes. The correlation is expected because trees within a plot are ecologically interdependent and competing for finite resources.

The SUR parameter estimation technique has been proven to provide a gain in efficiency when the error components are correlated and the structure of each equation differs (Zellner 1962). When the generalized system of equations uses linear functions (f), independent variables are the means of each attribute by class within each plot. Class means could not be used for nonlinear functions (f) because a function evaluated at the mean does not equal the mean of the function for nonlinear equations. Therefore, a program in SAS PROC MODEL (SAS Institute Inc. 1989) was written to estimate the parameters using an iterative process. The program used 217 records corresponding to the plots. Each individual tree within a plot and its respective attributes were arrayed along one record. The program evaluated the function for each plot between iterations and computed the mean of the function for each class within a plot (Appendix B).

Models

Three basal area growth models were developed for trial use in a system of equations. Model 1 (Hitch 1994) is the current model being used in the shortleaf pine simulator and was used as a basis for comparison. The numerator of Model 1 is the potential function developed by Shifley (1987) and is constrained for a biologically reasonable tree size (M). A biologically reasonable tree size for shortleaf pine of this geographical location was derived as having a 36-inch *DBH* and the equivalent basal area is 7.068384 square feet. The potential function initial parameters were estimated by fitting the potential function separately using the five-percent fastest growing trees by one-inch diameter class. If a diameter class had less than 21 observations then all trees for that diameter class were used to estimate the initial potential function parameters.

The denominator of Model 1 is the competition modifier suggested by Murphy and Shelton (1996) in which the parameters were estimated while holding the potential function parameters constant. To reduce bias exhibited by Model 1 with respect to some *DBH* classes, one iteration was completed to re-estimate the parameters of both the potential and competition functions as suggested by Wensel et al. (1987). The Model 1 basal area growth model has the following form:

$$AABAG = \frac{\beta_1 BA^{\beta_2} - \left(\dfrac{\beta_1 BA}{M^{(1-\beta_2)}} \right)}{1 + e^{(\beta_3 + \beta_4 SBA + \beta_5 AGE + \beta_6 DD + \beta_7 BA)}}$$

where

$AABAG$ = average annual basal area growth,

BA = individual tree basal area,

M = maximum basal area (M = 7.068384),

SBA = stand basal area,

AGE = stand age,

DD = diameter of subject tree divided by quadratic mean diameter, and

β_i = parameters.

Model 2 uses an individual tree equation that is mathematically identical to Model 1 but to estimate parameters using SUR, a system of four equations was formed corresponding to the four diameter rank classes within each plot. Model 2 has the following form:

$$\overline{AABAG_1} = \frac{1}{n_1}\sum_{j=1}^{n_1} \frac{\beta_1 BA_j^{\beta_2} - \left(\frac{\beta_1 BA_j}{M^{(1-\beta_2)}}\right)}{1+e^{(\beta_3+\beta_4 SBA+\beta_5 AGE+\beta_6 DD_j+\beta_7 BA_j)}}$$

$$\overline{AABAG_2} = \frac{1}{n_2}\sum_{j=1}^{n_2} \frac{\beta_1 BA_j^{\beta_2} - \left(\frac{\beta_1 BA_j}{M^{(1-\beta_2)}}\right)}{1+e^{(\beta_3+\beta_4 SBA+\beta_5 AGE+\beta_6 DD_j+\beta_7 BA_j)}}$$

$$\overline{AABAG_3} = \frac{1}{n_3}\sum_{j=1}^{n_3} \frac{\beta_1 BA_j^{\beta_2} - \left(\frac{\beta_1 BA_j}{M^{(1-\beta_2)}}\right)}{1+e^{(\beta_3+\beta_4 SBA+\beta_5 AGE+\beta_6 DD_i+\beta_7 BA_j)}}$$

$$\overline{AABAG_4} = \frac{1}{n_4}\sum_{j=1}^{n_4} \frac{\beta_1 BA_j^{\beta_2} - \left(\frac{\beta_1 BA_j}{M^{(1-\beta_2)}}\right)}{1+e^{(\beta_3+\beta_4 SBA+\beta_5 AGE+\beta_6 DD_j+\beta_7 BA_j)}}$$

where

j = individual tree observation(s) within a class (class = 1, 2, 3, 4) on a plot,

AABAG, BA, M, SBA, AGE, β_i, and *DD* defined previously, and

n_i = observations in class *i* on a plot.

The parameters for Model 2 were estimated using the technique described for Model 1 with the following modifications. The parameters were estimated using the SUR option in the SAS PROC MODEL procedure with parameter restrictions placed across the four equations.

Model 3 uses a modified Weibull probability function as the potential and the same competition modifier used for Models 1 and 2.

$$\overline{AABAG_1} = \frac{1}{n_1} \sum_{j=1}^{n_1} \frac{\beta_0 \left(\frac{\beta_2}{\beta_1}\right)\left(\frac{BA_j}{\beta_1}\right)^{(\beta_2 - 1)} e^{-\left(\frac{BA_j}{\beta_1}\right)^{\beta_2}}}{1 + e^{(\beta_3 + \beta_4 SBA + \beta_5 AGE + \beta_6 DD_j + \beta_7 BA_j)}}$$

$$\overline{AABAG_2} = \frac{1}{n_2} \sum_{j=1}^{n_2} \frac{\beta_0 \left(\frac{\beta_2}{\beta_1}\right)\left(\frac{BA_j}{\beta_1}\right)^{(\beta_2 - 1)} e^{-\left(\frac{BA_j}{\beta_1}\right)^{\beta_2}}}{1 + e^{(\beta_3 + \beta_4 SBA + \beta_5 AGE + \beta_6 DD_j + \beta_7 BA_j)}}$$

$$\overline{AABAG_3} = \frac{1}{n_3} \sum_{j=1}^{n_3} \frac{\beta_0 \left(\frac{\beta_2}{\beta_1}\right)\left(\frac{BA_j}{\beta_1}\right)^{(\beta_2 - 1)} e^{-\left(\frac{BA_j}{\beta_1}\right)^{\beta_2}}}{1 + e^{(\beta_3 + \beta_4 SBA + \beta_5 AGE + \beta_6 DD_j + \beta_7 BA_j)}}$$

$$\overline{AABAG_4} = \frac{1}{n_4} \sum_{j=1}^{n_4} \frac{\beta_0 \left(\frac{\beta_2}{\beta_1}\right)\left(\frac{BA_j}{\beta_1}\right)^{(\beta_2 - 1)} e^{-\left(\frac{BA_j}{\beta_1}\right)^{\beta_2}}}{1 + e^{(\beta_3 + \beta_4 SBA + \beta_5 AGE + \beta_6 DD_j + \beta_7 BA_j)}}$$

where

AABAG, BA, SBA, AGE, DD, β_i, n_i, and j are as defined previously.

The Weibull probability density function was modified by adding a parameter (β_0) that scaled the function. The modified Weibull probability density function was selected as the potential function because of its flexibility and to determine the viability of using the function to model basal area growth. The potential function was fitted to the fastest growing five percent of the trees using the method described previously. Then the competition modifier was fitted while holding the potential function constant. One iteration of refitting the potential and modifier functions was completed as described for Models 1 and 2 to re-estimate the parameters and remove some bias with respect to certain *DBH* classes.

The composite function (Model 4) selected is a variation of the PROGNOSIS model (equation 7). Model 4 predicts the natural log of basal area growth as a linear function of stand and tree characteristics. The system of four equations has the following form:

$$\overline{\ln(\text{AABAG}_1)} = \beta_0 + \beta_1 \overline{\ln(\text{dbh}_1)} + \beta_2 \overline{\text{dbh}_1^2} + \beta_3 \overline{\text{BAL}_1} + \beta_4 \overline{\text{CR}_1} + \beta_5 \overline{\text{CR}_1^2} + \beta_6 \overline{\text{CCF}_1} + \beta_7 \overline{\text{SI}} + \beta_8 \overline{\text{AGE}}$$

$$\overline{\ln(\text{AABAG}_2)} = \beta_0 + \beta_1 \overline{\ln(\text{dbh}_2)} + \beta_2 \overline{\text{dbh}_2^2} + \beta_3 \overline{\text{BAL}_2} + \beta_4 \overline{\text{CR}_2} + \beta_5 \overline{\text{CR}_2^2} + \beta_6 \overline{\text{CCF}_2} + \beta_7 \overline{\text{SI}} + \beta_8 \overline{\text{AGE}}$$

$$\overline{\ln(\text{AABAG}_3)} = \beta_0 + \beta_1 \overline{\ln(\text{dbh}_3)} + \beta_2 \overline{\text{dbh}_3^2} + \beta_3 \overline{\text{BAL}_3} + \beta_4 \overline{\text{CR}_3} + \beta_5 \overline{\text{CR}_3^2} + \beta_6 \overline{\text{CCF}_3} + \beta_7 \overline{\text{SI}} + \beta_8 \overline{\text{AGE}}$$

$$\overline{\ln(\text{AABAG}_4)} = \beta_0 + \beta_1 \overline{\ln(\text{dbh}_4)} + \beta_2 \overline{\text{dbh}_4^2} + \beta_3 \overline{\text{BAL}_4} + \beta_4 \overline{\text{CR}_4} + \beta_5 \overline{\text{CR}_4^2} + \beta_6 \overline{\text{CCF}_4} + \beta_7 \overline{\text{SI}} + \beta_8 \overline{\text{AGE}}$$

where

$AABAG_{ij}$ = average annual basal area growth of tree j in class i (square feet),

DBH = tree diameter (inches),

BAL = cumulative basal area of all trees larger than subjective tree,

CR = crown ratio,

CCF = crown competition factor,

SI = site index (base age = 50),

AGE = stand age (years), and

β_i = estimated parameters.

The dependent variable in this model requires computing the natural log which is a problem because of the negative or zero growth observations. The data set consisted of approximately five-percent of the average annual basal area growth measurements having zero or negative growth. These observations may be attributed to logical reasons such as peeling bark, which results in a smaller diameter being recorded during the second measurement than at the first measurement. Because Model 4 required computing the natural log of basal area growth, the observations that were less than or equal to zero were constrained to be number that was less than the smallest observable basal area growth. The smallest observable growth during a measurement period corresponds to a tree growing from 1.0 to 1.1 inches with a basal area growth of 0.00145 square feet. All *AABAG* with less than or equal to zero growth were constrained to be 0.00125 square feet.

Calibration and Validation

To calibrate and evaluate the performance of the growth models, the complete data set was divided into a calibration and validation data set. The plots were stratified by the site indices, stand ages, and basal area per acre combinations. Approximately 1/3rd of the plots within each combination of age, site index, and basal area classes were selected randomly to form a validation data set. The calibration and validation data sets contain 149 and 68 plots, respectively. The summary statistics for the calibration and validation data sets are similar with no substantial differences. The summary statistics for the calibration and validation data sets are presented in Tables 4 and 5, respectively.

The calibration data set was used to estimate parameters for the models. The calibration of the models was focused on the statistical fitting process using either linear or nonlinear regression techniques to assure that the models perform satisfactorily with respect to the fit index and mean square error and to examine the residuals for heterogeneity. The computer program was verified to ensure that the trees were placed in the correct classes and that the program was operating correctly.

The validation process was used to determine model performance using the estimated parameters from the calibration data set to determine the fit and performance of the models on an independent data set. Validation tests were both subjective and objective. The subjective tests were used to observe the predicted basal area growth and determine if the projections were realistic with respect to current knowledge of shortleaf pine growth. Model predictions should follow known principles of tree and stand development and should have biological integrity at the stand extremes. Models were examined under extreme conditions. such as for young and small diameter trees to evaluate model performance. The objective tests are outlined under model evaluation criteria.

Table 4. Summary statistics for the calibration data set of the Ouachita Highlands natural even-aged shortleaf pine study for developing a basal area growth model (N = 6099).

Attribute	Minimum	Maximum	Mean	Standard Error
Plot Age (years)	21.0	94.0	45.1	18.53
Plot Site Index (base age=50)	38.9	85.9	57.6	9.50
Stand Basal Area (sq. ft.)	29.57	177.12	108.22	31.66
DBH (inches)	1.5	24.9	8.04	3.69
Crown Ratio	0.1310	0.7636	0.3785	0.062
CCF	26.46	255.44	133.05	50.90
Avg. Annual Individual Tree Basal Area Growth	-0.008	0.0718	0.0127	0.0102
Individual Tree Basal Area	0.0123	3.383	0.4268	0.3757
Quadratic Mean Diameter to Individual Tree *DBH* (*DD*)	0.4424	4.2410	1.0875	0.3191
Basal Area of all Trees as Large or Larger than the Subject Tree (*BAL*)	0	171.37	66.74	38.47

Table 5. Summary statistics for the validation data set of the Ouachita Highlands natural even-aged shortleaf pine study for developing a basal area growth model (N = 2829).

Attribute	Minimum	Maximum	Mean	Standard Error
Plot Age (years)	22.0	96.0	43.4	18.98
Plot Site Index (base age=50)	40.0	87.1	56.8	10.1
Stand Basal Area (sq. ft.)	22.53	140.37	103.74	29.06
DBH (inches)	1.2	23.2	7.70	3.86
Crown Ratio	0.1557	0.6619	0.3824	0.058
CCF	13.72	203.69	130.74	44.11
Avg. Annual Individual Tree Basal Area Growth	-0.010	0.0700	0.0119	0.0100
Individual Tree Basal Area	0.00723	2.924	0.4041	0.3871
Quadratic Mean Diameter to Individual Tree *DBH* (*DD*)	0.4520	3.725	1.0934	0.3235
Basal Area of all Trees as Large or Larger than the Subject Tree (*BAL*)	0	139.24	64.42	36.28

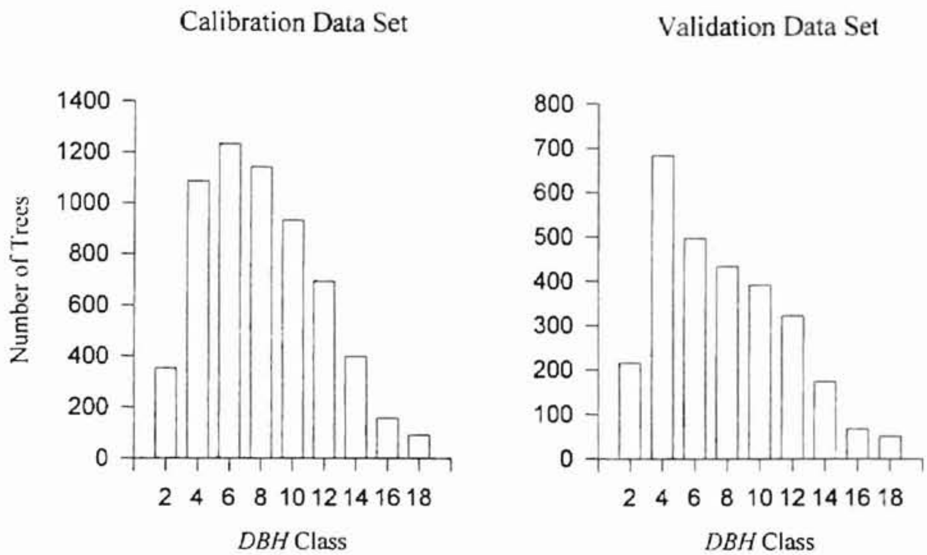

Figure 2. Number of trees by *DBH* class for the calibration (N = 6099) and validation (N = 2829) data sets.

The number of trees by *DBH* class graphs for the calibration and validation data sets exhibit the same general pattern. The 4-inch *DBH* class has the largest deviation from the complete data set number of trees by *DBH* class for both the calibration and validation data sets. The calibration and validation data sets are under- and over-represented for the 4-inch *DBH* class, respectively.

Model Evaluation Criteria

Model evaluation was conducted by computing the fit index and *MSE* and examining model performance across the data range with respect to *DBH*, site index, basal area per acre, and age classes. The goal was to determine model performance and to examine the models developed by using a system of equations to evaluate the validity of this method.

The fit index for each model was computed to determine how well the model fits the data. The fit index has the following form:

$$FIT\ INDEX = 1 - \frac{\sum_{i=1}^{n} e_i^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}$$

where

n = number of observations in validation, calibration, or complete data set,

e_i = residual of the ith observation, and

Y_i = average annual basal area growth for tree i.

In addition to the fit index, several other statistics were computed to evaluate the performance of the models for the calibration and validation data sets. The mean square error (*MSE*) was computed as the sum of the squared difference between the predicted basal area growth and the actual basal area growth divided by the degrees of freedom and is a measure of the dispersion.

$$MSE = \frac{\sum_{i=1}^{n} e_i^2}{n - p}$$

where

p = number of parameters, and

e_i and n are defined previously.

The *MSE* was computed for the complete calibration and validation data sets. In addition, the *MSE* was calculated using the validation data set for the *DBH*, age, site index, and basal area classes. The average deviation has the following form.

$$Average\ Deviation = \frac{\sum_{i=1}^{n} e_i}{n}$$

where

e_i and n are defined previously.

The average deviation by *DBH*, age, and basal area class was computed as the sum of the errors divided by the total number of observations and is used to detect bias.

The mean absolute deviation was calculated as the sum of the absolute difference between the predicted average annual basal area growth and the actual average annual basal area growth divided by the number of observations. The mean absolute deviation indicates the average absolute deviation from the mean *AABAG* and has the following form:

$$Average\ Absolute\ Deviation = \frac{\sum_{i=1}^{n} | e_i |}{n}$$

where

e_i and n are defined previously.

The average absolute error as a percentage of the mean *AABAG* by attribute was computed as the sum of the absolute deviations divided by the sum of the actual basal area growth.

$$Average\ Percent\ Absolute\ Deviation = \frac{\sum_{i=1}^{n} | e_i |}{\sum_{i=1}^{n} Y_i} (100)$$

where

e_i, Y_i, and n are defined previously.

The above model evaluation criteria are presented in tabular form. Box plots of the residuals were constructed by *DBH*, site index, age, and basal area classes to illustrate model biases and departures from normality.

The model(s) that performed best were then fitted to the complete data set of 217 plots. Further analysis was then conducted by computing the fit index, *MSE*, and average deviations by attribute classes for the complete data set. Bar charts of the average deviation by attribute classes and *AABAG* were constructed using the complete data set to detect biases and compare model performance.

CHAPTER V

RESULTS

Calibration

All four models used the same calibration data set, which consists of 149 plots and 6099 individual tree observations. Model 1 parameters were fitted using nonlinear regression using the 6099 individual tree records. Models 2 and 3 used a system of four nonlinear equations to fit the parameters using the plot mean of the function for each *DBH* rank class. Model 4 used the plot means of each *DBH* rank class to fit the parameters to a system of four linear equations. Models 2, 3, and 4 used the *DBH* rank classes by plot with 149 plot observations used for fitting the parameters. The parameter estimates, standard errors, and descriptions for Models 1 and 2, Model 3, and Model 4 are presented in Tables 6, 7, and 8, respectively.

Table 6. Parameter estimates, standard errors, and descriptions for Models 1 and 2 when fitted to the calibration data set.

	Model 1		Model 2		
Parameter	Estimate	Standard Error	Estimate	Standard Error	Description
β_1	0.081555	0.001626	0.081248	0.005159	*BA*
β_2	0.573986	0.010216	0.572551	0.037990	*BA* Power
β_3	-3.453924	0.095805	-2.816660	0.323730	Intercept
β_4	0.015943	0.000401	0.015596	0.001332	*SBA*
β_5	0.029879	0.001049	0.023381	0.003361	*AGE*
β_6	1.191026	0.075979	0.895971	0.209410	*DD*
β_7	-1.065994	0.049297	-0.982284	0.144010	*BA*

Table 7. Parameter estimates, standard errors, and descriptions for Model 3 when fitted to the calibration data set.

Parameter	Estimate	Standard Error	Description
β_0	0.237509	0.029240	Weibull Multiplier
β_1	3.432234	0.383970	Weibull Spread Parameter
β_2	1.482893	0.039310	Weibull Shape Parameter
β_3	-2.662908	0.312800	Logistic Intercept
β_4	0.014920	0.001283	*SBA*
β_5	0.024019	0.003245	*AGE*
β_6	0.837411	0.202140	*DD*
β_7	-0.977300	0.138470	*BA*

Models 1, 2, and 3 used the same competition modifier and consequently the signs of parameters β_3-β_7 are logically identical. The logical properties of the signs for the estimated parameters were examined while holding the other estimated parameters constant. Although illogical coefficient signs are possible in a valid predictive model due to multicolinearity, logical signs are desirable. For Models 1, 2, and 3, the signs for the estimated parameters β_4 (stand basal area), β_5 (stand age), and β_6 (quadratic mean diameter to individual tree *DBH*) are logically positive. This implies that as these attributes increase, such as stand age, the rate of growth decreases, and conversely, parameter β_7 (individual tree basal area) is negative and indicates that as the individual tree basal area increases its growth rate increases.

Table 8. Parameter estimates, standard errors, and descriptions for Model 4 when fitted to the calibration data set.

Parameter	Estimate	Standard Error	Description
β_1	-6.567868	0.232430	Intercept
β_2	1.457768	0.088260	Natural Log *DBH*
β_3	-0.003002	0.000727	*BAL*
β_4	1.388672	0.344310	*CR*
β_5	-0.016536	0.001858	*CCF*
β_6	-0.004905	0.000577	*AGE*

Model 4 is a variation of the PROGNOSIS model, with crown ratio squared, *DBH* squared, and site index excluded from the final model because they were insignificant (alpha level of 0.05). The signs of Model 4 attributes are logical and consistent with known forest growth patterns. Model 4 estimated parameters β_3 (basal area of all trees as large or larger than the subject tree), β_5 (plot age), and β_6 (*CCF*) have negative signs which are consistent with growth patterns. For instance, as *CCF* increases, tree competition increases and hence the rate of growth decreases. The signs for the estimated parameters β_2 (*DBH*) and β_4 (crown ratio) are logically positive.

Models 1 and 2 use the same potential function, which was initially fitted separately from the competition modifier. The potential function has a fit index of 0.7593 and *MSE* of 0.000052. Model 3 uses a modified Weibull probability density function as the potential function which has a fit index of 0.7749 and a *MSE* of 0.000048 when fitted to the fastest growing five percent of the trees by one-inch diameter class. The individual tree growth function used in Models 1 and 2 are identical mathematically but the degrees of freedom for the standard errors differ. Model 1 estimated parameters are based upon 6099 individual tree observations whereas Model 2 uses a system of equations that are based upon 149 plot observations.

The fit index and *MSE* for all four models are presented in Table 9. Model 1 had the highest fit index (0.6270) and Model 4 the lowest (0.5740). Models 2 and 3 fit indices and *MSE*'s are similar to those of Model 1. All models provide a reasonable fit with a small variance of the error terms.

Table 9. Fit index and mean square error for all models using the calibration data set.

Model	Fit Index	*MSE*
1	0.6270	0.0000391
2	0.6206	0.0000397
3	0.6180	0.0000400
4	0.5740	0.0000450

All model residuals were plotted against their respective model attributes and *DBH* to detect any trends. Except for *DBH*, there was no evidence of trends. The plots of residuals versus the *DBH* classes for models 1, 2, and 3 revealed some bias and slight heterogeneity of variance. As discussed in Methods (Chapter 4), one iteration for the potential and modifier functions was performed to re-estimate the parameters. The iteration removed some of the bias with respect to *DBH* and provided a better fit.

To determine whether SUR was an appropriate technique, the correlation among residuals was examined for the system of equations models. The residual correlation matrix is presented below for Model 2 with Models 3 and 4 illustrating similar results. The *DBH* rank classes are 1-4 (left to right, top to bottom). As illustrated by the matrix, there is a moderate amount of correlation between the plot *DBH* rank classes. Consequently SUR may provide a gain in parameter estimation efficiency.

$$\begin{bmatrix} 1.0 & 0.5684 & 0.6341 & 0.5099 \\ & 1.0 & 0.6532 & 0.5165 \\ & & 1.0 & 0.6440 \\ & & & 1.0 \end{bmatrix}$$

The pairwise correlation for *DBH* rank classes might be positive due to the fact that the study entails thinning from below and consequently the suppressed trees are removed. Since the microenvironment on a plot is similar for all *DBH* rank classes, if growth in the first rank class is above the mean, growth in the second rank class tends to be above the mean. Therefore, the residuals tend to be positively correlated.

The parameter correlation matrices were examined to detect multicollinearity between independent attributes. The correlation matrix for Model 1 is presented below with the other models showing similar results. There is a strong inverse correlation between the intercept (β_3) and both stand basal area (β_4) and quadratic mean diameter to individual tree *DBH* (β_6). There is also a significant inverse correlation between stand age (β_5) and individual tree basal area (β_7). The asymptotic correlation of parameters matrix for Model 1 β_3-β_7 (left to right, top to bottom) is

$$\begin{bmatrix} 1.0 & -0.64510426 & 0.04746676 & -0.81875629 & -0.33982952 \\ & 1.0 & -0.11238271 & 0.29022445 & 0.18386998 \\ & & 1.0 & -0.40950827 & -0.80374116 \\ & & & 1.0 & 0.51353762 \\ & & & & 1.0 \end{bmatrix}$$

Multicollinearity may cause the parameter estimates to differ substantially when using different techniques for parameter estimation or when the data are updated.

Validation

The validation data set consists of 68 plots and 2829 individual tree observations. All models were evaluated using the individual tree observations in the validation data set to determine model performance. The validation summary statistics for all models are presented Table 10.

Table 10. Summary statistics for all models using the validation data set.

Model	Fit Index	*MSE*	Average Error	Error Percentage	Mean Absolute Error
1	0.6074	0.0000393	0.0007998	37.11	0.0044087
2	0.6036	0.0000397	0.0005774	37.47	0.0044513
3	0.6016	0.0000399	0.0005439	37.62	0.0044693
4	0.5821	0.0000418	-0.00103900	38.84	0.0046148

Note: Error Percentage = average absolute error as a percentage of mean *AABAG*

The fit index and *MSE* results for the validation data set when compared to the calibration data set for Models 1, 2, and 3 are inferior, while Model 4 improved. Model 3 illustrated the smallest absolute value of average error (0.0005439) while Model 4 has the largest absolute value of average error (0.001039). All models except Model 4 exhibit an overprediction for average error. The variance among the average absolute error as a percentage of the mean *AABAG* for Models 1, 2, and 3 is relatively small while Model 4 has the highest error percentage. The mean absolute error among the four models is similar with Model 4 having the highest absolute error.

The validation data set was used to detect any trends by *DBH*, site index, basal area, and age classes for average deviation, *MSE*, average absolute error, and average absolute error as a percentage of the mean *AABAG* by attribute class. The results for the average deviation by *DBH* class are presented in Table 11.

Table 11. Average deviation for all models by *DBH* class using the validation data set.

DBH Class	N	Model 1	Model 2	Model 3	Model 4
2	215	0.0009295	0.0010123	0.0011487	-0.0000192
4	683	0.0009268	0.0005659	0.0005602	-0.0017151
6	496	-0.0002946	-0.0008933	-0.0009947	-0.0028879
8	433	0.0006813	0.0003399	0.0002422	-0.0011852
10	391	0.0009385	0.0008093	0.0007911	-0.0002121
12	321	0.0020258	0.0021786	0.0022184	0.0010329
14	173	0.0025150	0.0027801	0.0028366	0.0011660
16	67	-0.0008632	-0.0007556	-0.0007553	-0.0014970
18	50	-0.0022730	-0.0024204	-0.0032815	-0.0033731

Models 1, 2, and 3 have identical signs by *DBH* class for average deviation. Model 4 has the smallest average deviation for *DBH* class 2. Model 2 underpredicts for the 2-inch *DBH* class while the other models overpredict. Model 4 performs better for *DBH* classes 10, 12, and 14. Model 3 performs better for the 4 and 8-inch *DBH* classes. Model 1 is superior for the 6-inch *DBH* class and for the 16 and 18-inch *DBH* classes all models underpredict. The average deviation by site index, basal area, and age classes were computed and are presented in Table 12.

Table 12. Average deviations by site index, basal area per acre, and age classes for all models using the validation data set.

	N	Model 1	Model 2	Model 3	Model 4
Site Index					
<56	1485	0.0004838	0.0001059	0.0000623	-0.0018454
60	622	0.0006002	0.0005599	0.0005223	-0.0005920
70	568	0.0019515	0.0019576	0.0019641	0.0005699
>75	154	0.0004048	0.0001030	0.0000358	-0.0010047
Ba/ac					
30	229	-0.0020835	-0.0025523	-0.0026437	-0.0038334
60	239	0.0027902	0.0030303	0.0028908	0.0009413
90	608	0.0007943	0.0005074	0.0004648	-0.0017232
120	1753	0.0009070	0.0006760	0.0006677	-0.0007069
Age					
20	1024	-0.0002492	-0.0008470	-0.0008633	-0.0029609
40	797	0.0017839	0.0012287	0.0012309	-0.0000471
60	722	0.0022831	0.0024280	0.0024225	0.0014040
80	286	-0.0019318	-0.0008095	-0.0010751	-0.0030910

Models 1, 2, and 3 have similar trends for site index, basal area, and age classes. For site index classes, Models 1, 2, and 3 all overpredict for average deviation. Model 4 has underpredictions for site indices <56, 60, and >75 and performs best for the site index class 70. Models 1, 2, and 3 results are similar for all four basal area classes. They underpredict for basal area of 30 square feet per acre and overpredict for the other three basal area classes. Model 4 has more variation within attribute classes when compared with the other models. Model 4 performs the best for basal area class 60 and is inferior for the basal area class 90. Models 1, 2, and 3 underpredict for age classes 20 and 80, and overpredict for the other two age classes. Model 4 performs best for age class 40. Model 1 performs best for the age class 20, and Model 2 is superior for the older age class (80). The *MSE* for all models was computed by *DBH* class and presented in Table 13.

Table 13. Mean square error for all models by *DBH* class using the validation data set.

Class	N	Model 1	Model 2	Model 3	Model 4
2	215	0.0000023	0.0000024	0.0000027	0.0000019
4	683	0.0000077	0.0000073	0.0000073	0.0000113
6	496	0.0000266	0.0000295	0.0000297	0.0000385
8	433	0.0000382	0.0000412	0.0000409	0.0000401
10	391	0.0000522	0.0000519	0.0000517	0.0000463
12	321	0.0000626	0.0000593	0.0000600	0.0000581
14	173	0.0000921	0.0000913	0.0000917	0.0000882
16	67	0.0001854	0.0001835	0.0001826	0.0002026
18	50	0.0002300	0.0002309	0.0002380	0.0002652

With the exception of 4-inch *DBH* class where Model 4 performs best, all models perform similarly by *DBH* class for mean square error. The mean square error by site index, basal area per acre, and age classes are presented in Table 14.

Table 14. Mean square error for all models by site index, basal area per acre, and age classes using the validation data set.

	N	Model 1	Model 2	Model 3	Model 4
Site Index					
<56	1485	0.0000275	0.0000278	0.0000279	0.0000310
60	622	0.0000402	0.0000408	0.0000408	0.0000432
70	568	0.0000572	0.0000579	0.0000587	0.0000560
>75	154	0.0000920	0.0000914	0.0000926	0.0000974
Ba/ac					
30	229	0.0000710	0.0000766	0.0000766	0.0000759
60	239	0.0000967	0.0000937	0.0000933	0.0000982
90	608	0.0000355	0.0000348	0.0000350	0.0000382
120	1753	0.0000295	0.0000300	0.0000303	0.0000318
Age					
20	1024	0.0000147	0.0000169	0.0000173	0.0000251
40	797	0.0000438	0.0000431	0.0000430	0.0000415
60	722	0.0000510	0.0000513	0.0000519	0.0000466
80	286	0.0000895	0.0000864	0.0000866	0.0000951

The mean square error for all models reveals little variation among models by site index, basal area per acre, or age class. The validation data set results for mean absolute deviations by *DBH*, site index, basal area per acre, and age classes for all models are presented in Tables 15 and 16.

Table 15. Mean absolute deviation for all models by *DBH* class using the validation data set.

DBH Class	N	Model 1	Model 2	Model 3	Model 4
2	215	0.0011973	0.0012668	0.0013658	0.0010710
4	683	0.0022436	0.0021662	0.0021877	0.0026637
6	496	0.0039337	0.0041680	0.0041841	0.0049294
8	433	0.0046312	0.0048731	0.0048551	0.0048393
10	391	0.0054917	0.0054545	0.0054410	0.0051348
12	321	0.0062618	0.0061451	0.0061720	0.0058953
14	173	0.0077675	0.0077006	0.0077255	0.0074574
16	67	0.0101988	0.0102113	0.0101765	0.0100904
18	50	0.0108322	0.0108360	0.0110230	0.0119798

The mean absolute deviation by *DBH* class reveals that all models behave similarly and there is not a substantial amount of variation among models for the *DBH* classes. All models exhibit the trend of mean absolute deviation increasing as the *DBH* class increases with the 16 and 18-inch *DBH* classes having the highest mean absolute deviation for all models.

Table 16. Mean absolute deviation for all models by site index, basal area per acre, and age classes using the validation data set.

	N	Model 1	Model 2	Model 3	Model 4
Site Index					
<56	1485	0.0034833	0.0035020	0.0035275	0.0038789
60	622	0.0048433	0.0049299	0.0049177	0.0049857
70	568	0.0057531	0.0058382	0.0058661	0.0055378
>75	154	0.0066185	0.0065563	0.0065871	0.0068090
Ba/ac					
30	229	0.0060362	0.0062881	0.0063330	0.0067268
60	239	0.0077023	0.0076989	0.0076447	0.0074454
90	608	0.0040877	0.0040347	0.0040476	0.0045127
120	1753	0.0038584	0.0039130	0.0039391	0.0039884
Age					
20	1024	0.0027089	0.0028279	0.0028731	0.0035761
40	797	0.0048471	0.0048167	0.0048185	0.0047709
60	722	0.0053553	0.0054140	0.0054262	0.0049689
80	286	0.0069086	0.0068150	0.0067949	0.0070047

All models behave similarly for mean absolute deviation by site index, basal area per acre, and age classes. For both site index and age, all models exhibit higher mean absolute error as the site index and age class increases. All models exhibit the same trend for basal area per acre mean absolute deviation. Their smallest respective absolute deviations occur in the basal area classes 90 and 120, and their largest respective absolute deviations occurring in the basal area per acre 30 and 60 classes.

The average absolute error as a percentage of the mean *AABAG* by *DBH*, site index, basal area per acre, and age classes for all models was computed and are presented in Tables 17 and 18.

Table 17. Average absolute error as a percentage of mean average annual basal area growth by *DBH* class for all models using the validation data set.

DBH Class	N	Model 1	Model 2	Model 3	Model 4
2	215	62.53	66.16	71.32	55.93
4	683	34.88	33.68	34.01	41.41
6	496	36.48	38.65	38.80	45.71
8	433	36.39	38.29	38.15	38.03
10	391	37.54	37.28	37.19	35.10
12	321	38.94	38.21	38.38	36.66
14	173	39.11	38.77	38.90	37.55
16	67	37.10	37.14	37.02	36.74
18	50	30.31	30.32	30.84	33.52

All models exhibit their respective highest average absolute error as a percentage of average *AABAG* by *DBH* class for *DBH* class 2. Model 4 performs the best (55.93%) and Model 3 performs the worst (71.32%) for the 2-inch *DBH* class. The average absolute error as a percentage of average *AABAG* by *DBH* class for the other *DBH* classes reveals that all models perform similarly with no substantial difference between the respective models. The average absolute error as a percentage of the average *AABAG* by site index, basal area per acre, and age classes are presented in Table 18.

Table 18. Average absolute error as a percentage of mean *AABAG* by site index, basal area per acre, and age classes for all models using the validation data set.

	N	Model 1	Model 2	Model 3	Model 4
Site Index					
<56	1485	35.93	36.12	36.38	40.01
60	622	37.29	37.96	37.86	38.39
70	568	42.79	43.42	43.63	41.19
>75	154	29.14	28.87	29.00	29.98
Ba/ac					
30	229	26.59	27.69	27.90	29.63
60	239	39.23	39.21	38.94	37.92
90	608	31.10	30.69	30.79	34.33
120	1753	43.01	43.61	43.90	44.45
Age					
20	1024	29.04	30.31	30.80	38.33
40	797	39.16	38.91	38.92	38.54
60	722	43.27	43.82	43.92	40.22
80	286	37.48	36.97	36.86	38.00

Models 1, 2, and 3 behave similarly for average absolute error as a percentage of mean *AABAG* for site index, basal area per acre, and age classes. Model 4 behaves similarly to the other models with a few exceptions. Model 4 behaves worse for site index class <56 and slightly better for site index class 70. Model 4 average absolute error as a percentage of the mean *AABAG* is higher with respect to basal area per acre classes 30, 90, and 120 and lower for basal area class 60. All models have their respective highest average absolute error as a percentage of the mean *AABAG* in the following attribute classes: site index 70, basal area per acre 120, and age 60. Models 1, 2, and 3 have their respective lowest average absolute error as a percentage of the mean *AABAG* in the following attribute classes: site index >75, basal area per acre 30, and age 20.

Figure 3 illustrates the residuals box plots by *DBH* classes. Models 1, 2, and 3 illustrate similar trends with respect to bias by *DBH* class with Model 1 and 2 performing about equally and Model 3 illustrating slightly higher bias for the lower and higher *DBH* classes. Model 4 exhibits the worst bias for the largest *DBH* class and least bias for the smallest *DBH* class. All models demonstrate an underprediction bias for the largest *DBH* class (18^{+}). All four models have their highest skewness in the 18-inch *DBH* class.

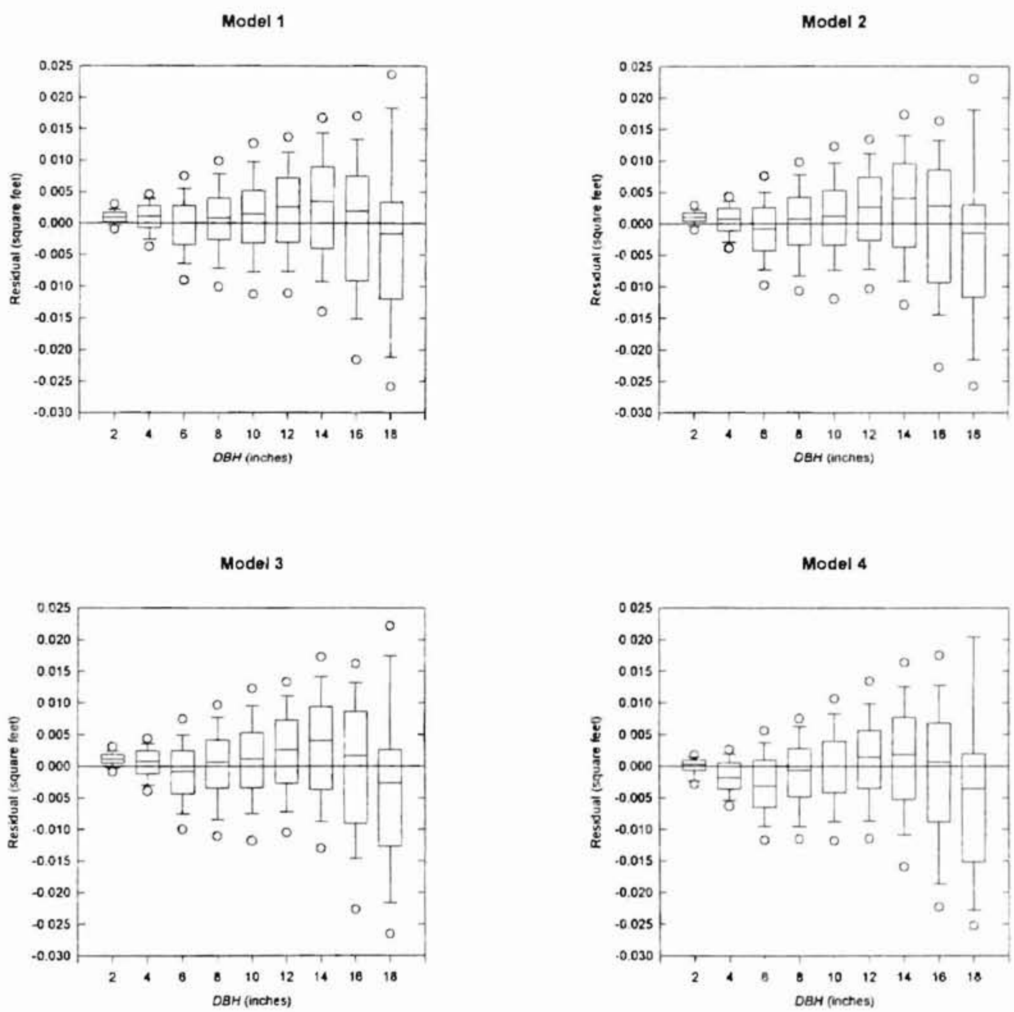

Figure 3. Box plots of residuals by *DBH* class for all models using the validation data set.

Figure 4 illustrates the box plots for residuals by site index class. Models 1, 2, and 3 perform similarly with the difference between Models 1 and 2 negligible. Models 2 and 3 illustrates slightly less bias in the plot site index class >75 and slightly more in site index class 70. Model 4 behavior is substantially different from the other models with overprediction bias only in site index class 70. Model 4 does exhibit the least bias with respect to site index classes 60 and 70.

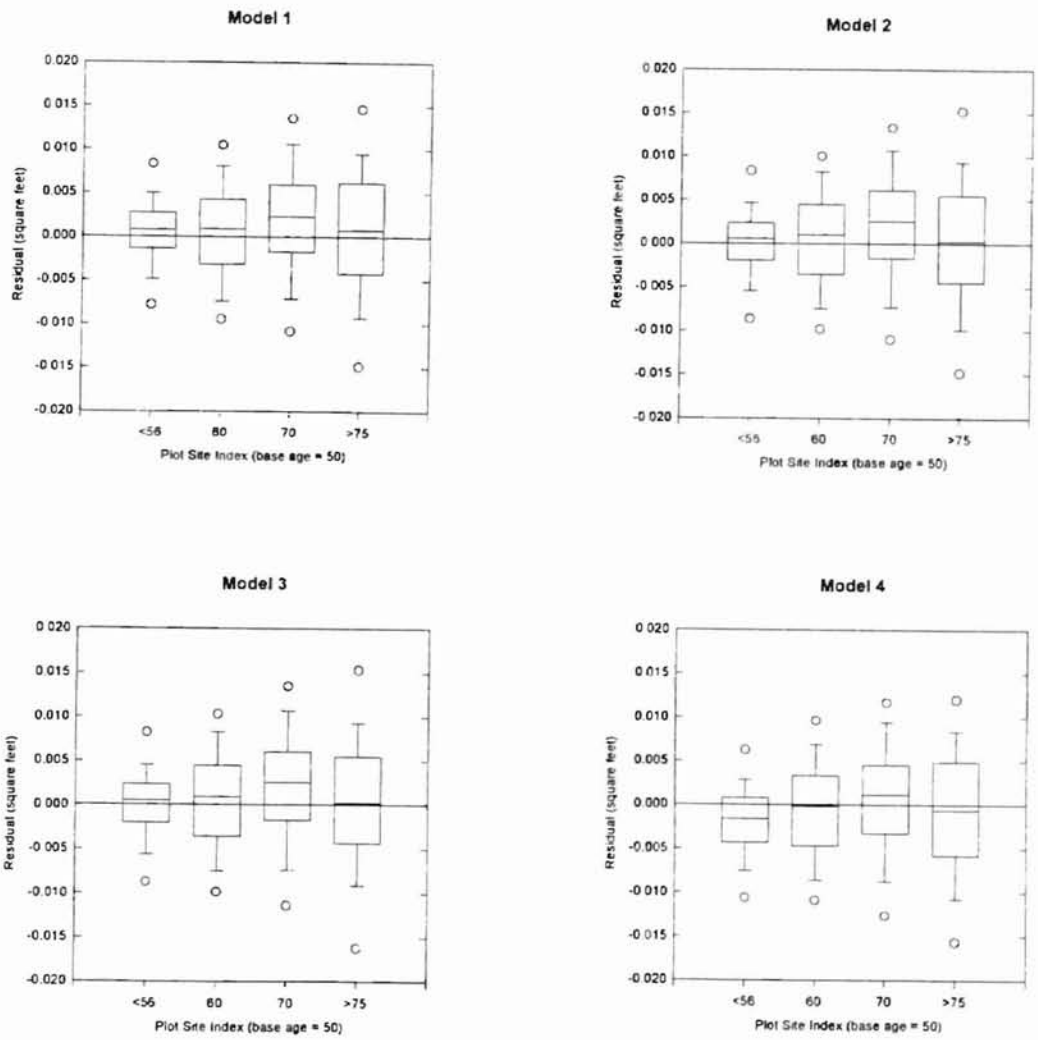

Figure 4. Box plots of residuals by site index class for all models using the validation data set.

Figure 5 illustrates the box plots of residuals by plot basal area class. Models 1, 2, and 3 behave similarly with Models 2 and 3 performing slightly better for plot basal area 90 and Model 1 better for plot basal area 30. Model 4 has the least bias for basal area class 120 and 60, but has substantially more underprediction bias for basal area class 30.

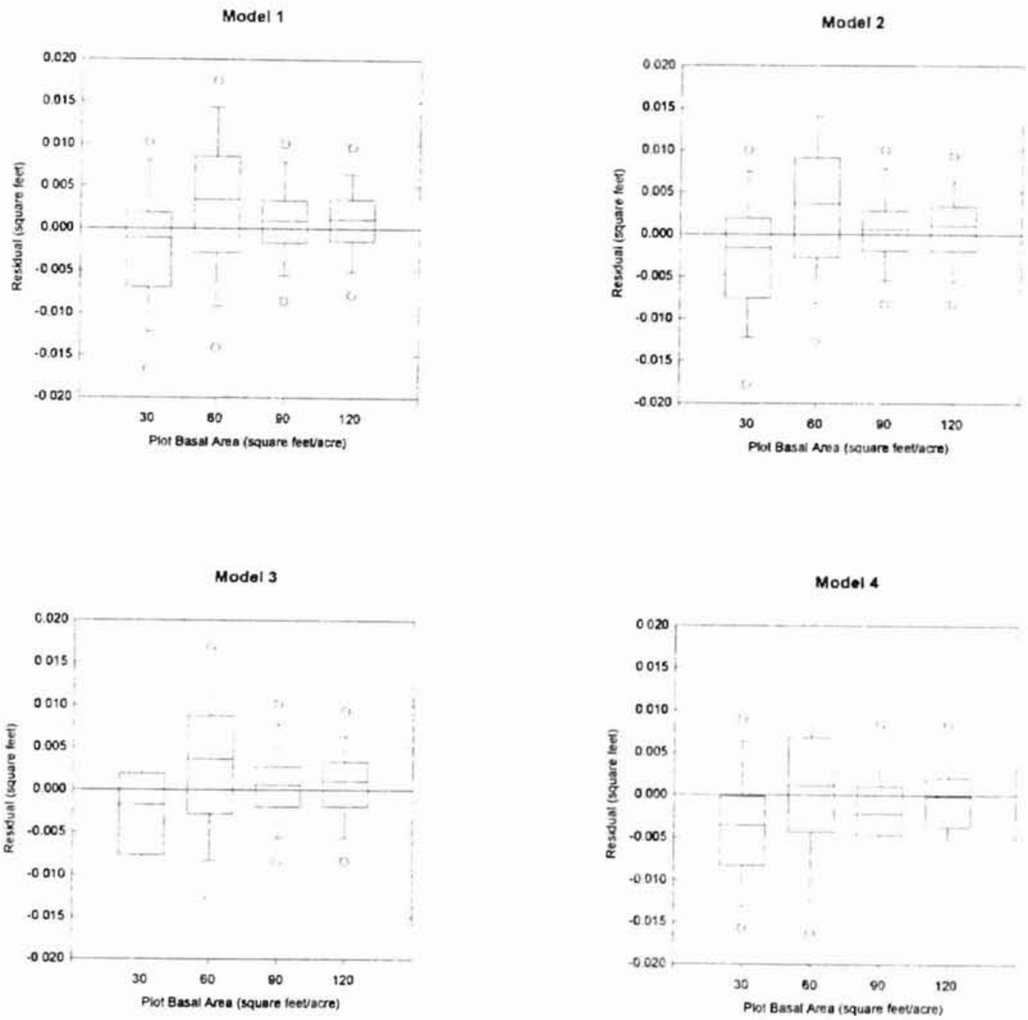

Figure 5. Box plots of the residuals by plot basal area class for all models using the validation data set.

The box plots of the residuals by age class are presented in Figure 6. Models 1, 2, and 3 behave similarly for all age classes. Models 1 and 2 perform the best for the age classes 20 and 80. Model 4 has strong departures in biases with respect to the other models. Model 4 exhibits the least bias for age class 40 but has significantly more underprediction bias in age class 20 with respect to the other models.

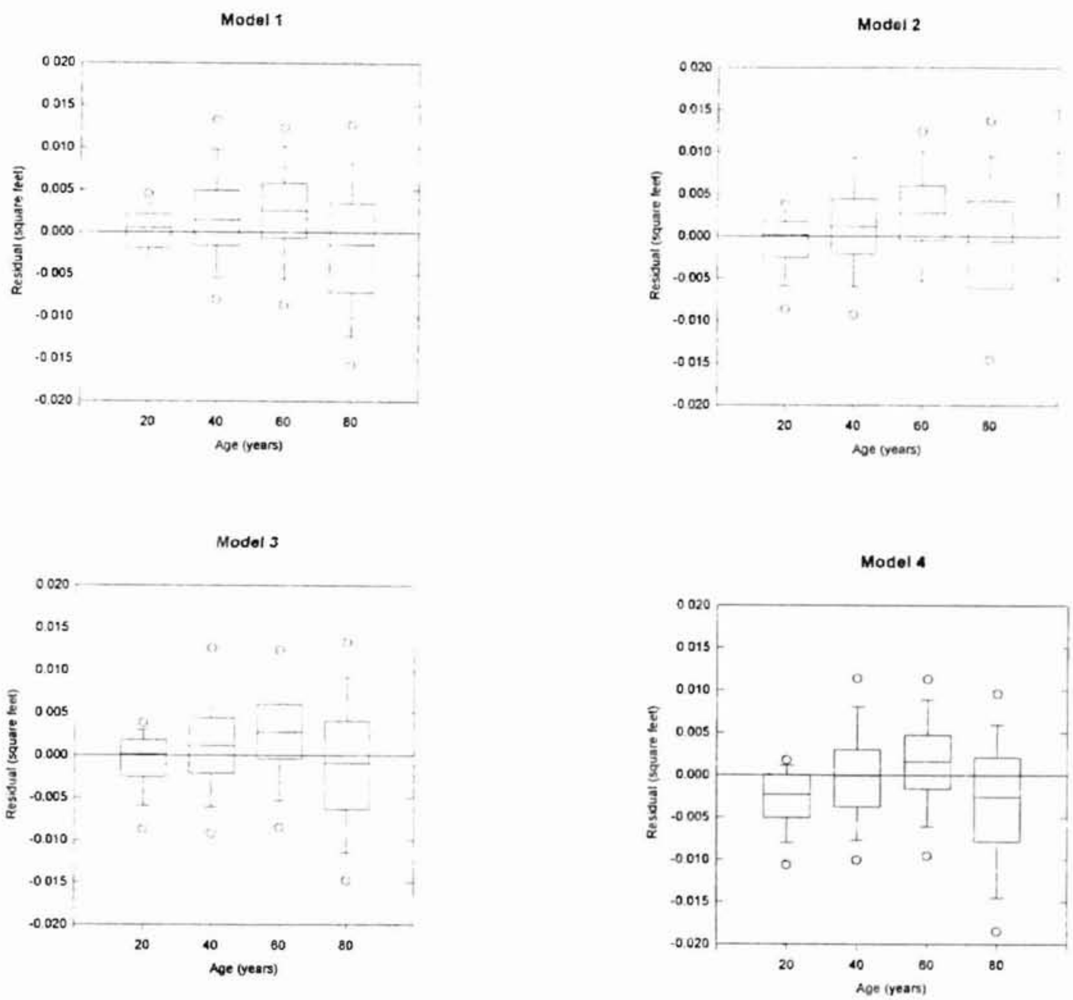

Figure 6. Box plots of residuals by plot age class for all models using the validation data set.

CHAPTER VI

DISCUSSION

The calibration and validation results reveal that Models 1, 2, and 3 perform well with no substantial differences among the models. Model 4 performs adequately but its biological integrity is compromised because it predicts a diameter larger than 70 inches. There could be some concern about using Model 4 to simulate the future forest structure for old stands. Model 4 performs best for the 2-inch diameter class for average error, *MSE*, average absolute error, and average absolute error as a percentage of the mean *AABAG*. Although Model 4 performs best for the 2-inch *DBH* class, it still performs poorly for this diameter class. Model 4 was excluded from further consideration because for the model calibration and validation results it has the least favorable fit index, *MSE*, average error, average absolute error as a percentage of the mean *AABAG*, and mean absolute error.

Models 1, 2, and 3 perform similarly for the calibration and validation data sets. Although the Model 1 fit index, *MSE*, average absolute error as a percentage of the mean *AABAG*, and mean absolute error are better, the differences among the three models is negligible. For example, the respective fit indexes using the validation data set for Models 1, 2, and 3 are 0.6074, 0.6036, and 0.6016, respectively. Model 3 has the smallest average error of 0.0005439, but the other models perform well with Models 1 and 2 having average errors of 0.0007998 and 0.0005774, respectively.

Model 1 does perform best using the validation data set for the average deviation criterion for *DBH* classes 2, 6, 12, 14, and 18, but with the exception of *DBH* class 6 the difference between models is negligible. The difference between Models 1, 2, and 3 for the 6-inch *DBH* class is -0.0002946, -0.0008933, and –0.0009947, respectively. Model 2 performs best overall for the average deviation by site index, basal area per acre and age class. Although Model 1 performs best for some attribute classes, the difference between Model 1 and Models 2 and 3 predictions is smaller than when Model 2 performs best.

There is little distinguishable difference in *MSE* and mean absolute deviation for each of the attributes. Model 3 does perform worst overall for mean absolute deviation but the difference is negligible among the models.

A concern with the shortleaf pine simulator has been the poor performance of the 2-inch diameter class. The 2-inch diameter class has the highest average absolute error as a percentage of the mean *AABAG*. For the 2-inch diameter class, Models 1, 2, and 3 have average absolute error as a percentage of the mean *AABAG* of 62.53, 66.16, and 71.32, respectively. The 2-inch diameter class is more sensitive than the other diameter classes because if all diameter classes have the same amount of measurable diameter growth the 2-inch diameter class will have the highest relative amount of basal area growth. For example, a tree growing from 2.0 to 2.1-inches has approximately 10% increase in basal area growth, whereas a tree growing from 10.0 to 10.1-inches has approximately 2% increase in basal area growth. The average absolute error as a percentage of the mean

AABAG for Models 1, 2, and 3 by site index, age, and basal area per acre classes provides little basis for differentiation among respective models.

The box plots of Models 1, 2, and 3 reveal no substantial differences among the three models with most attribute classes having a fairly symmetric distribution. Models 1, 2, and 3 have their highest *DBH* class skewness (negative) for the 16 and 18-inch diameter classes. This skewness was expected because there are few observations in the 16 and 18-inch *DBH* classes. All three models exhibit their highest plot basal area skewness (negative) for the plot basal area class 30. This was expected because there are relatively few observations in this plot basal area class. The 20-year age class for all three models has the greatest skewness (negative skewed) among age classes. Although the 20-year age class has the most observations among age classes there has been a tendency for the young age class to perform poorly. The box plots illustrate no substantial differences for the overall bias pattern for Models 1, 2, and 3.

Complete Data Set

Models 1, 2, and 3 were fitted to the complete data set for further evaluation. Plot 261 was removed from the data set for fitting the parameters for Models 2 and 3. There were only two trees on plot 261 and it could not be used for estimating the parameters for a system of equations. Models 1 and 2 and Model 3 parameter estimates, standard errors, and descriptions are presented in Tables 19 and 20, respectively.

Table 19. Parameter estimates, standard errors, and descriptions for Models 1 and 2 when fitted to complete data set.

	Model 1		Model 2		
Parameter	Estimate	Standard Error	Estimate	Standard Error	Description
β_1	0.0815392	0.0013261	0.083670	0.004831	*BA*
β_2	0.5727113	0.0083359	0.582591	0.033420	*BA* Power
β_3	-3.4613420	0.0787408	-2.768013	0.274390	Intercept
β_4	0.0160711	0.0003341	0.015550	0.001077	*SBA*
β_5	0.0295606	0.0008705	0.023695	0.002699	*AGE*
β_6	1.2477951	0.0629028	0.916957	0.186520	*DD*
β_7	-1.0603063	0.0401967	-0.994703	0.120200	*BA*

Table 20. Parameter estimates, standard errors, and descriptions for Model 3 when fitted to complete data set.

Parameter	Estimate	Standard Error	Description
β_0	0.279316	0.03920	Weibull Multiplier
β_1	3.930101	0.49676	Weibull Spread Parameter
β_2	1.459578	0.03460	Weibull Shape Parameter
β_3	-2.667933	0.26603	Logistic Intercept
β_4	0.015163	0.00105	*SBA*
β_5	0.023272	0.00261	*AGE*
β_6	0.907143	0.18089	*DD*
β_7	-0.959767	0.11546	*BA*

SUR provides a gain in efficiency for large data sets but a comparison of the standard errors for the estimated parameters for Model 1 and Models 2 and 3 could not be conducted. Model 1 estimated parameters are based upon 8928 individual tree observations whereas Models 2 and 3 parameter estimates are based upon 216 plot observations. While Model 1 appears to have a lower standard error for the estimates, when the independence assumption of regression is violated there may be a significant underestimation of the *MSE* and standard errors of the estimated parameters. Models 2 and 3 were fitted using OLS with no parameter restrictions placed across the equations for comparison with the SUR fit. There was a large gain in efficiency when Models 2 and 3 were fitted using SUR with parameter restrictions placed across equations. However, it cannot be concluded that Models 2 and 3 provide a gain in parameter estimation efficiency when compared with Model 1. The fit index, *MSE*, and *SSE* for Models 1, 2, and 3 when fitted to the complete data set are presented in Table 21.

Table 21. Models 1, 2, and 3 fit index, mean square error, and error sum of squares when fitted to the entire data set.

Model	Fit Index	*MSE*	*SSE*
1	0.6224	0.0000390	0.3482911
2	0.6154	0.0000397	0.3547762
3	0.6144	0.0000398	0.3556626

Model 1 has the highest fit index (0.6224) and lowest *MSE* (0.000039) but as expressed previously, when the independence assumption is violated the *MSE* may be significantly underestimated. If the *MSE* is underestimated then the SSE would be underestimated and the fit index would be overestimated. There is not a substantial difference between the three models, with Models 2 and 3 having fit indices of 0.6154 and 0.6144, and *MSE*'s of 0.0000397 and 0.0000398, respectively. Although *MSE* and fit indices for Models 2 and 3 are less favorable than Model 1, these models account for the interdependency among trees within a plot.

Horizontal bar charts of the average deviation by class attributes and mean *AABAG* are presented in Figures 7-10. Figure 7 illustrates the average deviation by *DBH* class and mean *AABAG* for Models 1, 2, and 3. This illustrates that the 2-inch *DBH* class has the largest bias relative to basal area growth. Models 1 and 2 are relatively equal for the 2-inch diameter class with Model 3 having a larger bias than growth. All three models overpredict for the 2-inch diameter class. Model 2 has virtually no bias for the 4-inch *DBH* class with Models 1 and 3 illustrating a small overprediction bias. Model 1 has the least bias for the 6, 8, and 10-inch *DBH* classes with all three models underpredicting on average. Model 1 has the least bias for the 16-inch *DBH* class. Model 2 is superior for the 18-inch *DBH* class with all three models having their largest bias for the 18-inch *DBH* class. The 18-inch *DBH* class has the largest absolute bias (underprediction) for all models which was expected because this *DBH* class contains few observations.

The average deviation by site index class and mean *AABAG* is presented in Figure 8. All models exhibit their respective smallest biases in the extreme site index classes (<56 and >75). For site index class <56, Model 1 has a slight overprediction while Models 2 and 3 have underpredictions. All three models perform about equally for all site index classes.

The plot basal area average deviation and mean *AABAG* graph is presented in Figure 9. All models have their respective smallest and largest bias for plot basal area class 90 and 30, respectively. All models underpredict for plot basal area on average except for plot basal area class 90. Model 1 performs best for plot basal area class 30 but there is little substantial difference among the models.

The graph of average deviation by age class and mean *AABAG* is presented in Figure 9. All models have underprediction bias for age classes 20, 40, and 80 with Model 1 performing best for age class 20 and worst for age class 80. There is little differentiation between Models 2 and 3 for all age classes. The *AABAG* decreased from age 40 to 60 before increasing substantially from age 60 to 80 and may be the result of age class 60 having substantially more basal area per acre than age classes 40 and 80.

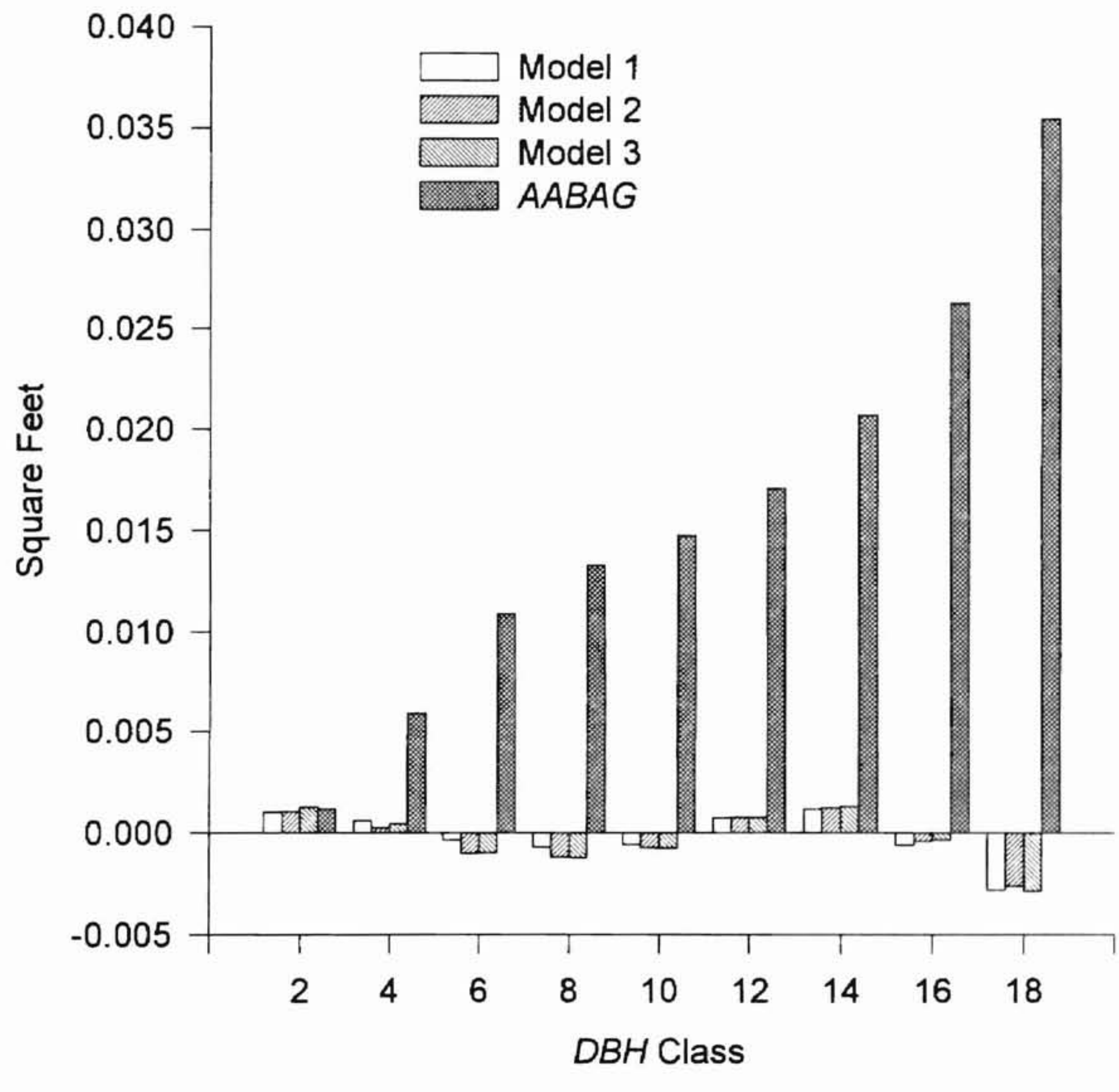

Figure 7. Average deviation by *DBH* class and mean average annual basal area growth for Models 1, 2, and 3 when fitted to the complete data set.

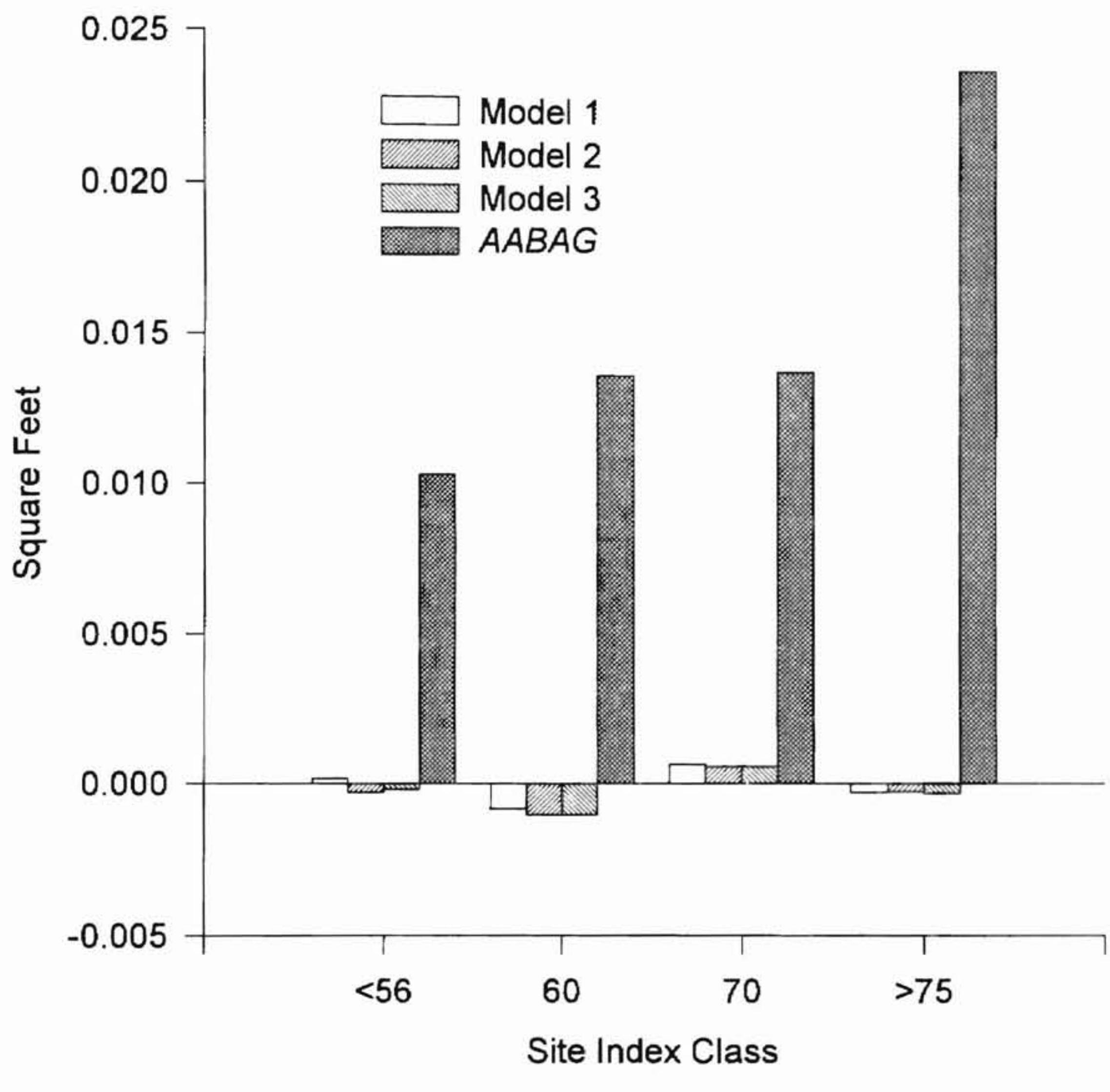

Figure 8. Average deviation by site index class and mean average annual basal area growth for Models 1, 2, and 3 when fitted to the complete data set.

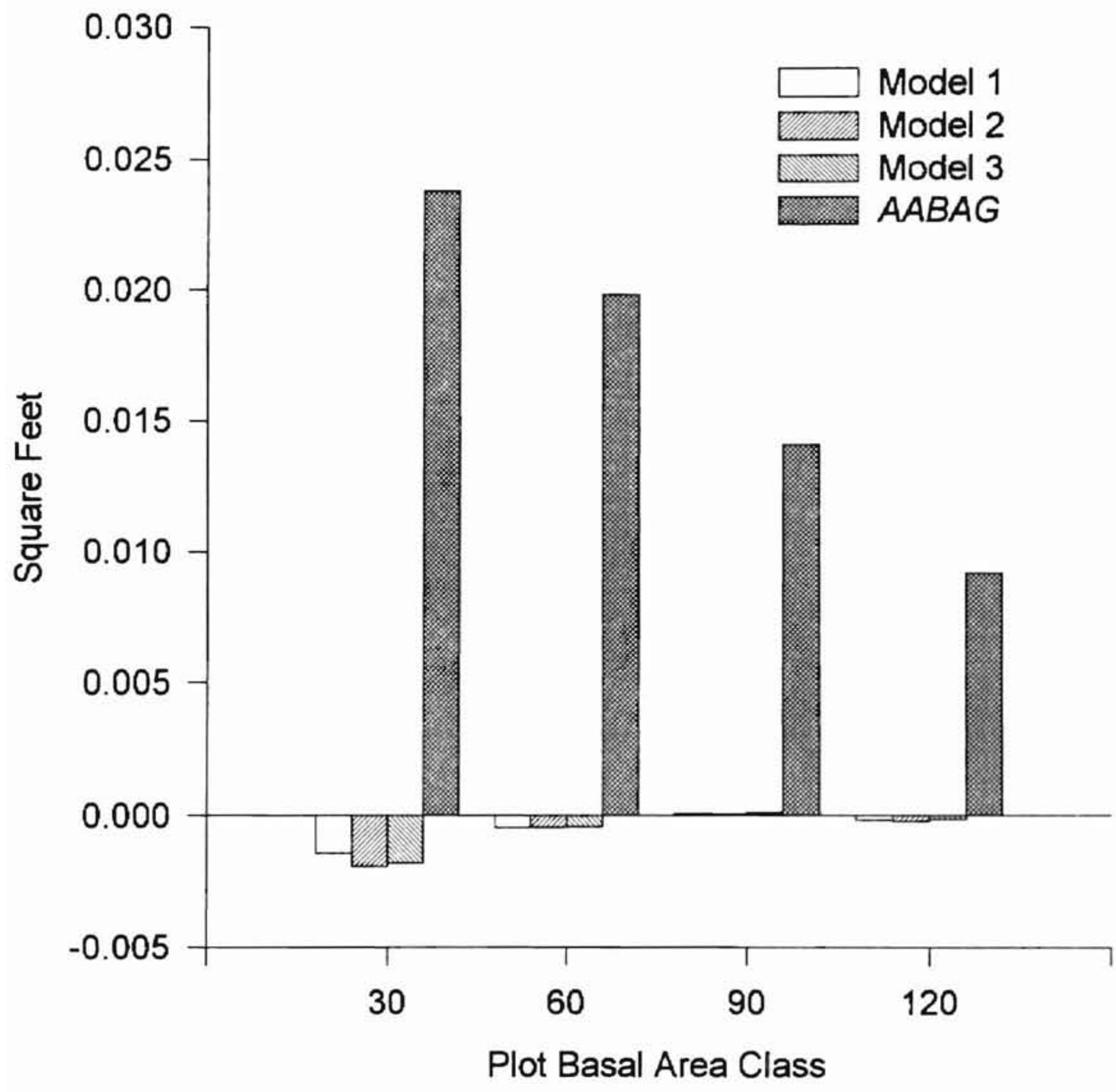

Figure 9. Average deviation by plot basal area class and mean average annual basal area growth for Models 1, 2, and 3 when fitted to the complete data set.

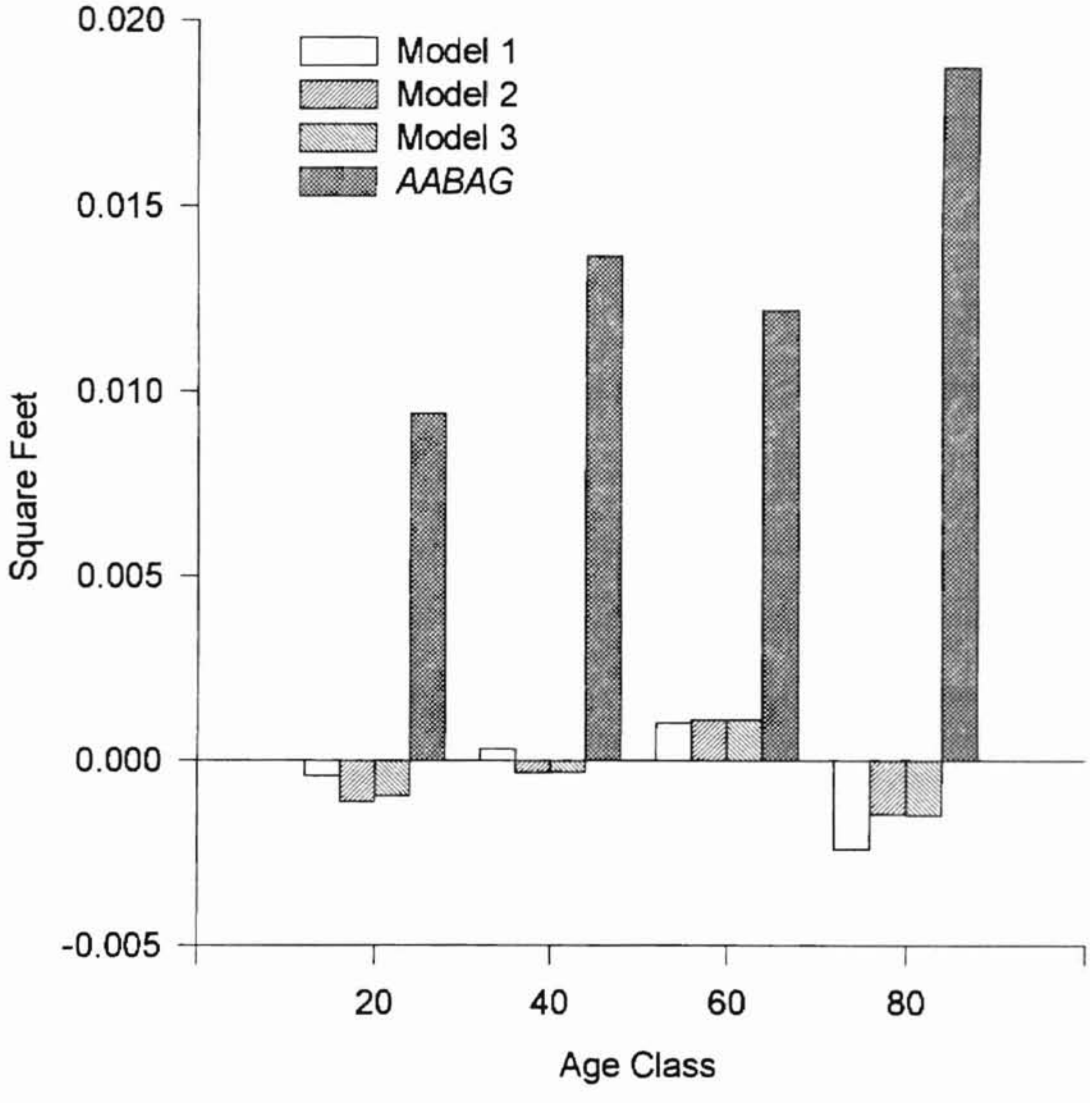

Figure 10. Average deviation by age class and mean average annual basal area growth for Models 1, 2, and 3 when fitted to the complete data set.

Conclusion

Model 4 was excluded as a candidate for use in the shortleaf pine simulator because for the calibration and validation tests it has the least favorable fit index, *MSE*, average error, average absolute error, and average absolute error as a percentage of the mean *AABAG*. In addition, the biological integrity of Model 4 is compromise because it predicts a *DBH* in excess of 70-inches.

The calibration, validation, and complete data set results reveal that neither Model 1, 2, or 3 is clearly superior. In practical terms, any of these models are viable candidates for use in the shortleaf pine simulator. It appears that Models 1 and 2 are slightly superior to Model 3, but Model 3 has better overall consistency. Model 3 uses a modified Weibull potential function, which has demonstrated its viability as a potential function, but further investigation is warranted. Models 2 and 3 account for the interdependency among trees within a plot and are theoretically more correct. These models may be enhanced by increasing the number of equations used in the system of equations.

As discussed earlier, major departures from the assumption of uncorrelated errors may result in a significant underestimation for the standard error(s) of the estimated parameter(s) and *MSE*. If the *MSE* is underestimated then the t-value for testing if $\beta_i = 0$ will be overestimated and consequently β_i may be rejected when it should not be rejected. Since all three models behave similarly, it is recommended that Model 2 be used as the basal area growth model for the shortleaf pine simulator because it has the same mathematical form as the current basal area growth model (Model 1) and accounts for tree interdependency within a plot.

LITERATURE CITED

Amateis, R.R., H.E. Burkhart, and T.A. Walsh. 1989. Diameter increment and survival equations for loblolly pine trees growing in thinned and unthinned plantations on cutover site-prepared lands. South. J. Appl. For. 13:170-174.

American Forestry Association. 1982. National register of big trees. American Forests. 88(4):18-47.

Avery, T.E., and H.E. Burkhart. 1994. Forest Measurements, 4th ed. New York NY: McGraw-Hill inc.. 408 p.

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

Belcher, D.W., M.R. Holdaway, and G.J. Brand. 1982. A description of STEMS. North Central Experiment Station. USDA For. Ser. Res. Bull. SO-135.

Bella, I.E. 1971. A new competition model for individual trees. For. Sci. 17:364-372.

Bertalanffy, L. 1951. Theoretische Biologe. (Band II). Franke, Bern. 403 p.

Bitoki, O., L.R. Gering, T.B. Lynch, and P.A. Murphy. 1997. An individual tree basal area growth model for uneven-aged stands of shortleaf pine (*Pinus echinata* Mill.) in the Ouachita Mountains in Arkansas and Oklahoma. Presented at the Ninth Biennial Southern Silvicultural Research Conference; February, 1997; Clemson, SC. (in press).

Bolton, R.K., and R.S. Meldahl. 1990. Design and development of a multipurpose forest projection system for southern forests. Auburn University, Alabama, Bull. 603. 51 p.

Borders, B.E. 1989. Systems of equations in forest stand modeling. For. Sci. 35:548-556.

Braun, E.L. 1950. Deciduous forests of eastern North America. The Blakiston Co., Philadelphia, PA. 596 p.

Brown, G.S. 1965. Point density in stems per acre. New Zealand For. Res. Note No. 38.

Buckman, R.E. 1962. Growth and yield of red pine in Minnesota. USDA Forest Serv. Tech. Bull. 1272. 50 p.

Chapman, D.G. 1961. Statistical problems in population dynamics. In: Proc. Fourth Berkeley Symp. On Mathematical Statistics and Probability. Univ. Calif. Press, Berkeley.

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

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

Daniels, R.F., and H.E. Burkhart. 1975. Simulation of individual tree growth and stand development in managed loblolly pine plantations. Blacksburg, Virginia: School of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University. FWS-5-75. 69 p.

Davis, L.S., and K.N. Johnson. 1987. Forest Management 3rd ed. New York, NY: McGraw-Hill inc. 790 p.

Dyer, M.E., and H.E. Burkhart. 1987. Compatible crown ratio and crown height models. Can. J. For. Res. 17:572-574.

Furnival, G.M., and R.W. Wilson Jr. 1971. Systems of equations for predicting forest growth and yield. Stat. Eco. 3:43-55.

Graney, D.L., and H.E. Burkhart. 1973. Polymorphic site index curves for shortleaf pine in the Ouachita Mountains. USDA For. Ser. Res. Pap. SO-85. 12 p.

Guldin, J.M. 1986. Ecology of shortleaf pine. In: Proc. Symp. on the Shortleaf Pine Ecosystem. Arkansas Coop. Ext. Ser., University of Arkansas, AR, p. 25-40.

Hahn, J.T., and R.A. Leary. 1979. Potential diameter growth functions. In: A generalization forest growth projection system applied to the Lake States region. USDA For. Serv. Gen. Tech. Rep. NC-49. Pp. 22-26.

Hasenauer. H., R.A. Monserud, and T.G. Gregoire. 1998. Using simultaneous regression techniques with individual-tree growth models. For. Sci. 44:87-95.

Hegyi, F. 1974. A simulation model for managing jack pine stands. Growth Models for Tree and Stand Simulation, IUFRO, Proceedings of Working Party. S.4.01-4:74-90.

Hitch, K.L. 1994. A distance-independent individual tree basal area growth model for natural even-aged stands of shortleaf pine in Eastern Oklahoma and Western Arkansas. M.S. Thesis, Oklahoma State University, OK, 78 p.

Huebschmann, M.M., T.B. Lynch, and P.A. Murphy. 1998. Shortleaf pine stand simulator: An even-aged natural shortleaf pine growth-and-yield model. In press.

Johnson, E.W. 1973. Relationships between point density and growth. Agriculture Experiment Station, Auburn University, Alabama, Bull. 447.

Johnson, N.L., and S. Kotz. 1970. Continuous Univariate Distributions-1. John Wiley and Sons, New York. 300 p.

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lutkepohl, and T.C. Lee. 1988. Introduction to the theory and practice of econometrics. Wiley, New York. 1024 p.

Krajicek, J.E., K.A. Brinkman, and S.F. Gingrich. 1961. Crown competition - A measure of density. For. Sci. 7:35-42.

Krumland, B.E. 1982. A tree-based forest yield projection system for the North Coast Region of California. Ph.D. dissertation, Department of Forestry and Resource Management , University of California, Berkeley.

Leary, R.A., and M.R. Holdaway. 1979. Modifier function. In: A generalization forest growth projection system applied to the Lake States region. USDA For. Serv. Gen. Tech.. Rep. NC-49. pp. 31-38.

Lynch, T.B., and P.A. Murphy. 1995. A compatible height prediction and projection system for individual trees in natural, even-aged shortleaf pine stands. For. Sci. 41:194-209.

Lynch, T.B., P.A. Murphy and E.R. Lawson 1991. Stand volume equations for managed natural even-aged shortleaf pine in eastern Oklahoma and Western Arkansas. Stillwater, Oklahoma: Agricultural Experiment Station Division of Agricultural Oklahoma State University. P-921. 12 p.

Martin, G.L., and A.L. Ek. 1984. A comparison of competition measures and growth models for predicting plantation red pine diameter and height growth. For. Sci. 30:731-743.

McWilliams, W.H., R.M. Sheffield, M.H. Hansen, and T.W. Birch 1986. The shortleaf resource. In: Proc. Symp. on the Shortleaf Pine Ecosystem. Arkansas Coop. Ext. Ser., University of Arkansas, Monticello, AR, p. 9-25.

Moser, J.W., Jr. 1972. Dynamic of an uneven-aged forest stand. For. Sci. 18:184-191.

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

Munro, D.D. 1974. Forest growth models-a prognosis. In: Growth models for tree and stand simulation. Royal Coll. For., Res. Notes No. 30, Stockholm.

Murphy, P.A. 1982. Sawtimber growth and yield for natural even-aged stands of shortleaf pine in the West Gulf. USDA For. Ser. Res. Pap. SO-181. 13 p.

Murphy, P.A. 1986. Growth and yield of shortleaf pine. In: Proc. Symp. on the Shortleaf Pine Ecosystem. Arkansas Coop. Ext. Ser., University of Arkansas, AR, p. 159-177.

Murphy, P.A. 1988a. Establishment and progress report: growth and yield of thinned natural shortleaf pine stands in the Ozark and Ouachita National Forests. So. For. Exp. Sta., For. Sci. Lab., Monticello AR. 52 p.

Murphy, P.A. 1988b. Study plan and progress report: Growth relationships for even-aged stands of shortleaf pine. So. For. Exp. Sta., For. Sci. Lab., Monticello AR. 22 p.

Murphy, P.A. and R.C. Beltz. 1981. Growth and yield of shortleaf pine in the West Gulf region. USDA For. Ser. Res. Pap. SO-169. 15 p.

Murphy, P.A. And M.G. Shelton. 1993. Progress report: Effects of density, site quality, and diameter on growth and yield of uneven-aged loblolly pine stands. Monticello, AR:FS-SO-4106-36 (Problem 4).

Murphy, P.A. And M.G. Shelton. 1996. An individual-tree basal area growth model for loblolly pine stands. Can. J. For. Res. 26:327-331.

Neter, J., M.H. Kutner, C.J. Nachtsheim, and W. Wasserman. 1996. Applied Linear Regression Models 3rd ed. Burr Ridge, Il. Third Ed. Irwin. 720 p.

Pienaar, L.V., and K.J. Turnbull. 1973. The Chapman-Richards generalization of Von Bertalanffy's growth model for basal area growth and yield in even-aged stands. For. Sci. 19:2-22.

Pindyck, R.S., and D.L. Rubinfeld. 1981. Econometric models and economic forecasts 2nd ed. McGraw-Hill, New York. 630 p.

Ratkowsky, D.A. 1990. Handbook of Nonlinear Regression Models. Marcel Dekker, Inc., New York 241 p.

Reed, D.D. 1987. Estimation procedures for analytically related growth and yield models. In Proc. 1986 Soc. Am. For. Nat. Conv. Soc. Am. For., Bethesda, MD. 352 p.

Richards, F.J. 1959. A flexible growth function for empirical use. J. Exp. Bot. 10(29):290-300.

Rogers, R., and I.L. Sander. 1984. Intermediate thinnings in a Missouri shortleaf pine stand: a 30-year history. Ed. E. Shoulders. pp 216-219 In: Proc. Third Biennial Southern Silvicultural Research Conference. Atlanta, Georgia. New Orleans, Louisiana: Southern Forest Experiment Station. USDA For. Serv. Gen. Tech. Rep. SO-54.

SAS Institute Inc. SAS/STAT™ User's Guide, Release 6.03 ed. Cary, NC: SAS Institute Inc., 1988. 1028 p.

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

Schumacher, F.X., and T.S. Coile. 1960. Growth and yield of natural stands of the southern pines. T. S. Coile, Inc., Durham, N.C. 115 p.

Shifley, S.R. 1987. A generalized system of models forecasting central states tree growth. USDA For. Res. Pap. NC-279 10 p.

Shifley, S.R., and G.J. Brand. 1984. Chapman-Richards growth function constrained for maximum tree size. For. Sci. 30:1066-1070.

Smalley, G.W., and R.L. Bailey. 1974. Yield tables and stand structure for shortleaf pine plantations in Tennessee, Alabama, and Georgia highlands. Southern Forest Experiment Station. USDA For. Ser. Res. Pap. SO-97. 58 p.

Smith, K.L. 1986. Historical perspective. In: Proc. Symp. on the Shortleaf Pine Ecosystem. Arkansas Coop. Ext. Ser., University of Arkansas, AR, p. 1-8.

Smith, W.R., R.M. Farrar, P.A. Murphy, J.L. Yeiser, N.N. Meldahl, J.L. Kush. 1992. Crown and basal area relationships of open-grown southern pines for modeling competition and growth. Can. J. For. Res. 22:341-347.

Spurr, S.H. 1962. A measure of point density. For. Sci. 8:85-96.

Staebler, G.R. 1951. Growth and spacing in an even-aged stand of Douglas-fir. M.F. Thesis, School of Natural Resources, University of Michigan, Ann Arbor.

Stage, A.R. 1973. Prognosis model for stand development. USDA Forest Service Research Paper INT-137.

Sylvester, W.A. 1938. A comparison of two methods of yield table construction. J. For., 38(681-686).

Tome, M., and H.E. Burkhart. 1989. Diatance-dependent competition measures for predicting growth of individual trees. For. Sci. 35(3): 816-831.

Wensel, L.C., W.J. Meerschaert, and G.S. Biging. 1987. Tree height and diameter growth models for northern California conifers. Hilgardia, 55(8):1-20.

West, P.W. 1979. Use of diameter increment and basal area increment in tree growth studies. Can. J. Forest Res. 10:71-77.

Willett, R.L. 1986. Forward to symposium on the shortleaf pine ecosystem. In:Proc.Symp. On the Shortleaf Pine Ecosystem. Arkansas Coop Ext. Ser., University of Arkansas, Monticello, AR. P. iii-iv.

Williston, H.L. and W.E. Balmer. 1980. Shortleaf pine management. USDA For. Ser. Forestry report SA-FR6. 9 p.

Wykoff, W.R. 1990. A basal area increment model for individual conifers in the northern Rocky Mountains. For. Sci. 36:1077-1104.

Wykoff, W.R., N.L. Crookston, and A.R. Stage. 1982. User's guide to the stand PROGNOSIS model. U.S. Forest Service General Technical Report INT-133.

USDA Forest Service. 1929 (rev. 1976). Volume, yield, and stand tables for second-growth southern pines. USDA Misc. Publ. 50. 202 p.

USDA Forest Service. 1972. Forest survey handbook. USDA For. Ser. FS 4809.11, Washington DC.

van Hees, W.S. 1980. Arkansas forests: trends and perspectives. Monticello, Arkansas: Southern Forest Experiment Station. USDA For. Serv. Res. Bull. SO-77. 32 p.

Vanclay, J.K. 1994. Modelling Forest Growth and Yield. Wallingford UK CAG Int. 312 p.

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

Zeide, B. 1989. Accuracy of equations describing diameter growth. Can. J. For. Res. 19:1283-1286.

Zellner, A. 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. J. Am. Stat. Assoc. 57:348-368.

APPENDICES

APPENDIX A

SAS® SUBROUTINE PROGRAM THAT RANKS AND CREATES FOUR CLASSES WITHIN EACH PLOT CORRESPONDING TO THE INDIVIDUAL TREE *DBH* WITHIN A PLOT

```
/*RANKING DATA SET BY DBH WITHIN EACH PLOT IN ASCENDING ORDER*/
/*IDENTICAL DBH'S WITHIN A PLOT ARE PLACED IN THE LOWER RANK*/
PROC RANK DATA=SET3 TIES=LOW OUT=SET4;
BY PLOT;
RANKS RDBH_MID;
VAR DBH_MID;

        /*CREATING DATA SET5 FROM THE RANKED SET4*/
DATA SET5;
SET SET4;

        /*DO LOOP FOR COMPARING TREES WITHIN A PLOT FOR CREATING
       FOUR CLASSES, EQUAL DBH RANKS ARE PLACED INTO THE LOWER CLASS*/
DO;
      /*CREATING VARIABLES CORRESPONDING TO QUARTILES
                                    A=25%, B=50%, C=75%*/
A=NUM/4;
B=NUM/2;
C=NUM*(3/4);

        /*USING MOD FUNCTION TO DETERMINE HOW MANY REMAINING TREES
          THERE ARE PER PLOT AND WHAT CLASS THEY BELONG IN*/

        /*FOR EXAMPLE, IF PLOT HAS 5 TREES, MOD(NUM,4) WILL RETURN 1,
          THIS ONE TREE WILL BE PUT IN THE LOWER CLASS, IF MOD RETURNS
          2, THEN THE TWO TREES WILL BE SPLIT BETWEEN CLASS ONE AND
          TWO*/

        /*USED A MOD FUNCTION FOR EACH OF THE FOUR POSSIBILITIES, COULD
          HAVE A REMAINDER OF 0, 1,2, OR 3: ADDED ONE TO A, B, C,
          TO INSURE PROPER PLACEMENT OF TREES WHEN MOD = 1, 2, OR 3*/

IF MOD(NUM,4)=0 AND RDBH_MID<=A THEN CLASS=1;
IF MOD(NUM,4)=0 AND RDBH_MID>A AND RDBH_MID<=B THEN CLASS=2;
IF MOD(NUM,4)=0 AND RDBH_MID>B AND RDBH_MID<=C THEN CLASS=3;
IF MOD(NUM,4)=0 AND RDBH_MID>C THEN CLASS=4;
```

```
IF MOD(NUM,4)=1 AND RDBH_MID<=(A+1) THEN CLASS=1;
IF MOD(NUM,4)=1 AND RDBH_MID>(A+1) AND RDBH_MID<=(B+1) THEN CLASS=2;
IF MOD(NUM,4)=1 AND RDBH_MID>(B+1) AND RDBH_MID<=(C+1) THEN CLASS=3;
IF MOD(NUM,4)=1 AND RDBH_MID>(C+1) THEN CLASS=4;

IF MOD(NUM,4)=2 AND RDBH_MID<=(A+1) THEN CLASS=1;
IF MOD(NUM,4)=2 AND RDBH_MID>(A+1) AND RDBH_MID<=(B+1) THEN CLASS=2;
IF MOD(NUM,4)=2 AND RDBH_MID>(B+1) AND RDBH_MID<=(C+1) THEN CLASS=3;
IF MOD(NUM,4)=2 AND RDBH_MID>(C+1) THEN CLASS=4;

IF MOD(NUM,4)=3 AND RDBH_MID<=(A+1) THEN CLASS=1;
IF MOD(NUM,4)=3 AND RDBH_MID>(A+1) AND RDBH_MID<=(B+1) THEN CLASS=2;
IF MOD(NUM,4)=3 AND RDBH_MID>(B+1) AND RDBH_MID<=(C+1) THEN CLASS=3;
IF MOD(NUM,4)=3 AND RDBH_MID>(C+1) THEN CLASS=4;
END;

/*ONCE TREES ARE PLACED INTO THE FOUR CLASSES, THE ATTRIBUTE MEANS
WITHIN A PLOT BY CLASS CAN BE COMPUTED FOR USE IN THE LINEAR SYSTEM OF
EQUATIONS (MODEL 4).  THE DATA SET CAN ALSO NOW BE ARRAYED FOR USE IN
THE NONLINEAR SYSTEM OF EQUATIONS (MODELS 2 AND 3) (SEE APPENDIX B FOR
THE NONLINEAR PROGRAM)*/
```

APPENDIX B

SAS® SUBROUTINE PROGRAM THAT COMPUTES THE MEANS FOR A NONLINEAR FUNCTION BY *DBH* RANK CLASS WITHIN EACH PLOT BETWEEN EACH ITERATION FOR FITTING A NONLINEAR SYSTEM OF EQUATIONS

```
     /*THE DATA SET ARRAY IS USED WITHIN THE SAS PROC MODEL PROCEDURE
          TO COMPUTE THE NONLINEAR FUNCTION MEANS BY CLASS WITHIN
          EACH PLOT*/

PROC MODEL DATA=PATH.ARRAY;

     /*PARAMETER ESTIMATES FOR MODEL 2, UNCOVER THE POTENTIAL OR
          COMPETITION PARAMETERS THAT ARE NOT ESTIMATED DURING
          FITTING PROCESS. FOR EXAMPLE, WHEN FITTING THE COMPETITION
          MODIFIER (DENOMINATOR) THE POTENTIAL PARAMETERS ARE NOT
          ESTIMATED. POTENTIAL PARAMETERS=B1,B2, MODIFIER
          PARAMETERS=B3-B7*/

/*B3=-2.820689; B4=.015676; B5=.023993; B6=.91717; B7=-.990005;*/
  B1=.08367; B2=.582591;

     /*ARRAYS FOR ATTRIBUTES BY DBH CLASS, NUMBER OF ELEMENTS MUST BE
          LARGER THAN THE MAXIMUM NUMBER OF TREES IN A CLASS WITHIN A
          PLOT, SET AT 150*/
  array dbh1{150}; array dbh2{150}; array dbh3{150}; array dbh4{150};
  array pba1{150}; array pba2{150}; array pba3{150}; array pba4{150};
  array dd1{150}; array dd2{150}; array dd3{150}; array dd4{150};
  array ba1{150}; array ba2{150}; array ba3{150}; array ba4{150};
  array pa1{150}; array pa2{150}; array pa3{150}; array pa4{150};

     /*DBH=INDIVIDUAL TREE DBH, PBA=PLOT BASAL AREA, DD=INDIVIDUAL
          TREE DBH TO QUADRATIC MEAN DIAMETER, BA=INDIVIDUAL TREE
          BASAL AREA, PA=PLOT AGE*/
```

```
      /*DO LOOP FOR COMPUTING THE NONLINEAR FUNCTION MEANS BY PLOT AND
            CLASS BETWEEN EACH ITERATION*/

            /*COMPUTING FUNCTION MEANS OF CLASS 1 BY PLOT*/
DO I=1 TO 150;
  IF I=1 THEN N1=0;
        IF I=1 THEN S1=0;
        IF BA1{I}^=-1 THEN DO;
        s1=s1+(((b1*ba1{I}**b2)-(b1*ba1{I}/7.068384**(1-b2)))
                      /(1+exp(b3+b4*PBA1+b5*PA1+b6*DD1{I}+b7*BA1{I})));
               N1=N1+1;

        END;
END;
        EST_G1=S1/N1;

      /*COMPUTING FUNCTION MEANS OF CLASS 2 BY PLOT*/
DO I=1 TO 150;
        IF I=1 THEN N2=0;
               IF I=1 THEN S2=0;
        IF BA2{I}^=-1 THEN DO;
        s2=s2+(((b1*ba2{I}**b2)-(b1*ba2{I}/7.068384**(1-b2)))
                      /(1+exp(b3+b4*PBA2+b5*PA2+b6*DD2{I}+b7*BA2{I})));
               N2=N2+1;
        END;
END;
        EST_G2=S2/N2;

      /*COMPUTING FUNCTION MEANS OF CLASS 3 BY PLOT*/
DO I=1 TO 150;
        IF I=1 THEN N3=0;
               IF I=1 THEN S3=0;
        IF BA3{I}^=-1 THEN DO;
        s3=s3+(((b1*ba3{I}**b2)-(b1*ba3{I}/7.068384**(1-b2)))
                      /(1+exp(b3+b4*PBA3+b5*PA3+b6*DD3{I}+b7*BA3{I})));
               N3=N3+1;
        END;
END;
        EST_G3=S3/N3;
```

```
      /*COMPUTING FUNCTION MEANS FOR CLASS 4 BY PLOT*/
DO I=1 TO 150;
        IF I=1 THEN N4=0;
                IF I=1 THEN S4=0;
        IF BA4{I}^=-1 THEN DO;
        s4=s4+(((b1*ba4{I}**b2)-(b1*ba4{I}/7.068384**(1-b2)))
                      /(1+exp(b3+B4*PBA4+b5*PA4+b6*DD4{I}+b7*BA4{I})));
                N4=N4+1;
        END;
END;
        EST_G4=S4/N4;

      /*FITTING THE MEAN AVERAGE ANNUAL BASAL AREA GROWTH TO THE
            ESTIMATED AABAG, PARAMETER RESTRICTIONS PLACED ACROSS
            EQUATIONS*/
BAG1=EST_G1;
BAG2=EST_G2;
BAG3=EST_G3;
BAG4=EST_G4;

      /*BAG_i=MEAN OF THE AVERAGE ANNUAL BASAL AREA GROWTH FOR EACH DBH
            CLASS WITHIN A PLOT, EST_G_i=ESTIMATED AVERAGE ANNUAL BASAL
            AREA GROWTH FROM THE FUNCTION ABOVE USING THE STARTING
            PARAMETER VALUES DURING THE FIRST ITERATION*/


      /*ESTIMATED STARTING VALUES FOR PARAMETERS FOR FITTING EQUATIONS,
            COVER PARAMETERS VALUES THAT NOT ESTIMATED FOR THIS RUN*/
PARMS /*b1=.84 b2=.593*/ B3=-2.827049 B4=.015562 B5=.023576 B6=.876575
            B7=-.951213;

      /*USING ITERATIVE SUR TO FIT SYSTEM OF EQUATIONS,
FIT BAG1 BAG2 BAG3 BAG4/ITSUR METHOD=MARQUARDT CORR;

RUN;
```

VITA

Charles Edward Rose Jr.

Candidate for the Degree of

Master of Science

Thesis: PARAMETER ESTIMATION FOR A SHORTLEAF PINE (*Pinus echinata* Mill.) BASAL AREA GROWTH MODEL

Major Field: Forest Resources

Biographical:

Education: Received Associate of Science degree in Electronics from Condie Junior College, Campbell, California in January 1988. Received Bachelor of Science Degree in Forest Resources from Northern Arizona University, Flagstaff, Arizona in December 1995. Completed the requirements for the Master of Science degree with a major in Forest Resources at Oklahoma State University in May, 1998.

Experience: Employed as an Instructional Aide at Coconino Community College; assisted students in math and science. Worked for SEC Inc.; Sedona, Arizona performing stand examination and inventory, summer 1996. Employed by Boise Cascade Corporation; Boise, Idaho as a Forest Analyst, summer 1997. Employed by Oklahoma State University Department of Forestry as a graduate research assistant; Oklahoma State University, Department of Forestry, August 1996 to present.

Academic Achievements: Forestry Alumni, Wesley Clark Memorial, Arizona Resources Committee, Flagstaff Cabinet Forestry, and Arizona Community College Transfer Scholarships; Northern Arizona University. Afanasiev Distinguished, College of Agriculture Sciences and Natural Resources, and Presidential Graduate Fellowships, Dean's Tuition Waiver; Oklahoma State University.

Professional Memberships: Phi Kappa Phi, Xi Sigma Pi, Society of American Foresters.