Louisiana State University

LSU Digital Commons

Evaluation of methods to predict Weibull parameters for characterizing diameter distributions

Krishna Prasad Poudel
Louisiana State University and Agricultural and Mechanical College

Recommended Citation

EVALUATION OF METHODS TO PREDICT WEIBULL PARAMETERS FOR CHARACTERIZING DIAMETER DISTRIBUTIONS

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science

in

The School of Renewable Natural Resources

by
Krishna Prasad Poudel
B.S., Tribhuvan University, 2007
August 2011

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Compared to other distribution functions, the Weibull distribution has been more widely used in describing diameter distributions because of its flexibility and relative simplicity. Parameters of the Weibull distribution are generally predicted either by the parameter prediction method or by the parameter recovery method. The coefficients of the regression equations for predicting Weibull parameters, moments, or percentiles are often estimated by use of different approaches such as ordinary least squares (OLS), seemingly unrelated regression (SUR) or cumulative distribution function regression (CDFR). However, there is no strong rationale for preferring one method over the other. We developed and evaluated different methods of predicting parameters of Weibull distribution to characterize diameter distribution using data from the Southwide Seed Source Study.

The SUR and the CDFR approaches were applied to ten different parameter prediction and parameter recovery methods. A modified CDFR approach was developed by modifying the CDFR technique such that the CDF is computed using information from diameter classes instead of individual trees as in the CDFR approach. These methods were evaluated based on four goodness-of-fit statistics (Anderson-Darling, Kolmogorov-Smirnov, negative Log-Likelihood, and Error Index). The CDFR approach provided better results than the SUR approach for all methods. The Modified CDFR approach consistently provided better results than the SUR approach, and was superior to the CDFR approach in all evaluation statistics but the Anderson-Darling statistic.

CHAPTER 1. INTRODUCTION

Forest management decisions are based on information about both current and future resource conditions which require accurate predictions of growth and yield. Because volume, value, conversion costs, and product specifications are dependent on diameter of the tree, stand properties can be well characterized with diameter distributions (Bailey and Dell 1973). Various distribution functions, such as normal, gamma, Johnson's S_B, beta, and Weibull have been used in the past in describing diameter distributions of forest stands.

The Weibull function has been the most widely used distribution function for describing diameter distributions because of its flexibility and relative simplicity. Parameters of Weibull distribution are generally predicted either by parameter prediction method (PPM) or by parameter recovery method (PRM). The parameter prediction method relates the parameters of a distribution function with stand variables using regression equations, whereas the parameter recovery method predicts diameter percentiles or moments, from which the distribution parameters are recovered (Gorgoso et al. 2007).

The coefficients of the regression equations for predicting Weibull parameters (in the parameter prediction method) and moments or percentiles (in the parameter recovery methods) are often estimated by use of ordinary least squares or seemingly unrelated regression. Cao (2004) obtained the regression coefficients in the PPM by minimizing the sum of squared differences between the observed and predicted cumulative probability. He termed this new approach the CDF regression (CDFR) method, which produced better goodness-of-fit statistics than other methods.

A successful diameter-distribution model requires good prediction of its parameters. The objective of this study was to develop and evaluate new methods for predicting parameters of the Weibull probability density function for characterizing diameter distributions.

CHAPTER 2. LITERATURE REVIEW

2.1. Growth and Yield Models

Growth and yield models are functions of current stand conditions to predict future growth and yield. Yield predictions in the United States began with the use normal yield tables for natural even-aged stands of a given species (Knoebel et al. 1986). Growth and yield models can be classified into three broad categories: whole-stand models, size-class models, and individual-tree models.

2.1.1. Whole-Stand Models

Whole-stand models require few details to simulate growth, but provide rather general information about the future stand. They predict future yields as a function of stand-level attributes, such as stand age, site index, and stand density. MacKinney et al. (1937) and MacKinney and Chaiken (1939) used multiple regression techniques to construct variable-density yield equations for loblolly pine stands. Since then, a great number of growth and yield models have been developed using regression techniques (Burkhart et al. 1972; Beck and Della-Bianca 1972; Murphy 1983; Pienaar and Rheney 1993; Lenhart 1996; Coble 2009).

Buckman (1962) and Clutter (1963) obtained compatibility between growth and yield by developing models in which yield was obtained through mathematical integration of the growth equation over time. Sullivan and Clutter (1972) refined Clutter's (1963) equations to develop a simultaneous growth and yield model for loblolly pine that provided not only analytically, but also numerically consistent growth and yield predictions. Beck and Della-Bianca (1972) have applied this method to yellow-poplar data.

3

Burkhart and Sprinz (1984) fitted both basal area and volume projection functions simultaneously by minimizing the squared-error loss functions. A similar approach was used by Knoebel et al. (1986) to develop a simultaneous growth and yield model for thinned stands of yellow-poplar. Van Deusen (1988) showed that the minimization of the squared error loss function is equivalent to using seemingly unrelated regression (SUR). To account for correlation across the equations, researchers have used special regression procedures such as three-stage least squares (Borders and Bailey 1986, Pienaar and Harrison 1989) and SUR (Coble 2009).

Ochi and Cao (2003) developed an annual growth and yield model and suggested that annual growth models provided better predictions of stand survival, basal area, and volume than compatible growth models. The annual growth approach also offered flexibility and path invariance property which ensures that projections to a future age remain the same regardless of different paths (or intermediate future ages) it takes.

2.1.2. Size-Class Models

Size-class models are a compromise between whole-stand models and individual-tree models because if the class size is infinitely large and only one class exists, the method is the whole-stand approach and when the class width is infinitely small, single tree as a class, then the method is the individual-tree approach (Vanclay 1991). Similar to whole-stand models, stand-level attributes, such as age, site index, and stand density are used as inputs to these models, but the technique provides detail on stand structure.

2.1.2.1. Stand-Table Projection Methods

Stand tables give number of trees per diameter class. Stand-table projection methods predict future stand tables based on current stand tables. Ek (1974) introduced a set of nonlinear

equation models for three growth components (ingrowth, mortality, and survivor growth) for individual diameter classes. Clutter and Jones (1980) gave an algorithm for projecting stand tables in old-field slash pine plantations. This projection algorithm differed from conventional stand table projection methods in that it did not move trees from one class to another, but instead moved the entire class forward in time. Pienaar and Harrison (1988) applied this method for unthinned even-aged stands based on long-term remeasurement data from a slash pine stand. Borders and Patterson (1990) evaluated a Weibull diameter distribution model, a percentile-based projection model, and a basal area growth projection model for projecting stand and stock tables for loblolly pine. Nepal and Somers (1992) proposed an algorithm to project current stand table and then adjust the future stand table to match the estimate of future basal area and survival. Corral-Rivas et al. (2009), based on relatively small experimental dataset, found satisfactory result of application of this algorithm to *Eucalyptus grandis* in South Africa. Cao and Baldwin (1999) modified Nepal and Somers (1992) algorithm by applying the constrained-least-squares method for constraining future stand tables. An individual-tree model was later incorporated by Cao (2007) into the algorithm to predict mortality and diameter growth of each diameter class.

2.1.2.2. Diameter-Distribution Methods

Stand yields have also been predicted based on the assumption that diameter distribution of a stand can be characterized by a probability density function. Clutter and Bennett (1965) applied the beta distribution to describe diameter distributions of data from old-field slash pine plantations. The beta distribution was then used in yield models for many species, including slash pine (Bennett and Clutter 1968, Bennett et al. 1978), loblolly pine (Lenhart and Clutter

5

1971, Lenhart 1972, Burkhart and Strub 1974), birch, and pedunculate oak (Gorgoso-Varela et al. 2008).

Bailey and Dell (1973) introduced the use of the Weibull function for modeling diameter distributions. Since then, it has been favored by many others because it can fit a variety of shapes. Another advantage is that its cumulative distribution function exists in closed form, allowing easy calculation of proportion of trees in each diameter class. The Weibull function has been widely used to model diameter distribution of loblolly pine (Smalley and Bailey 1974, Feduccia et al. 1979, Matney and Sullivan 1982, Clutter et al. 1984, Baldwin and Feduccia 1987), longleaf pine (Lohrey and Bailey 1977, Jiang and Brooks 2009), loblolly and slash pine (Brooks et al. 1992), Scots pine, Austrian pine, *Pinus halepensis* (Palahi et al. 2006), birch (Gorgoso-Varela et al. 2007), and black poplar (Andrasev et al. 2009).

The Johnson's S_B distribution (Johnson 1949a) was found by Hafley and Schreuder (1977) to fit diameter and height distributions better than five other distributions. The S_B distribution has been used by many researchers to fit diameter distribution data (Hafley et al. 1982, Smith and Hafley 1984, Newberry and Burk 1985, Rennolls and Wang 2005, Fonseca et al. 2009). Different methods for estimating the parameters of the S_B distribution have been analyzed by Zhou and McTague (1996), and Scolforo et al. (2003). Schreuder and Hafley (1977) further fitted the bivariate S_{BB} distribution (Johnson 1949b) to tree height and diameter data. The bivariate approach has been applied for modeling diameter distribution (Knoebel and Burkhart 1991, Karlsson and Norell 2005).

Diameter distributions can also be defined directly from several percentiles, as opposed to being approximated by a single statistical function. Number of diameter percentiles vary from

ten (Alder 1979) to twelve (Borders et al. 1987). Cao and Burkhart (1984) joined different functions together to form a segmented CDF using five percentile points for modeling diameter frequency data.

2.1.3. Individual-Tree Models

Individual-tree models use a tree as the basic unit, and therefore can provide detailed information about stand dynamics. They are divided into two classes, distance-dependent and distance-independent, depending on whether or not tree location coordinates are required.

Distance-independent models do not use spatial information to express competition, but they usually project tree growth as a function of current tree size and other stand variables. Stage (1973) developed PROGNOSIS for mixed species of northern Rocky Mountains. Shifley (1987) described individual-tree models which were compatible with the STEMS and TWIGS projection system. A similar approach was used by Wykoff (1990) to predict individual-tree basal area increment and calibrated for eleven conifer species having diverse ecological requirements. Other distance-independent systems include models by Hynynen (1995) and Palahi et al. (2003) for Scots pine, Mabvurira and Miina (2002) for *Eucalyptus grandis* (Hill) Maiden, Sanchez-Gonzalez et al. (2006) for cork oak, and Adame et al. (2008) for rebollo oak coppices.

Distance-dependent models, on the other hand, include a spatial competition measure. The growth of each tree is obtained as a function of stand variables, tree attributes, and a measure of competition from its neighbors. Distance-dependent approach is the most promising means of predicting yield because it mimics the actual system with greater details than other methods (Mitchell 1975). Stand simulators PTAEDA and PTAEDA2 were distance-dependent models developed by Daniels and Burkhart (1975) and Burkhart et al. (1987), respectively, for

7

loblolly pine plantations. Daniels et al. (1979) produced a similar method for modeling seeded loblolly pine stands. Distance-dependent models have been developed for Rap poplar (Faber 1991), Scots pine on a drained peatland site (Miina 1994), even-aged *Shorea robusta* stands in Nepal (Rautiainen 1999), Scots pine and Norway spruce growing in a mixed forest (Vettenranta 1999), and even-aged *Eucalyptus pilularis* in Australia (Fox et al. 2007 a, b). Distance-dependent models are suitable for intensively managed stands, but the high cost for suitable data may restrict their use to research applications.

Individual-tree growth models which provide the most detailed information for tree prediction may be inaccurate for stand-level prediction because of cumulative error from summing up individual tree predictions. To improve whole-stand predictions, individual-tree models were constrained by also taking into account attributes at the diameter-class level (Zhang et al. 1997). Cao (2006) developed a new approach in which an individual tree model was constrained by optimizing for both tree and stand levels to provide a reasonable tree- and stand-level prediction of survival and growth.

Annual prediction for tree growth and survival is a method suitable for data with irregular growth intervals. Cao (2000) introduced an iterative method in which survival probability and diameter of each tree in the plot are predicted and interim values of stand density are updated for each year to predict annual diameter growth and survival for individual trees. A similar approach was adopted by Nord-Larsen (2006). Cao and Strub (2008) evaluated four different methods to simultaneously estimate parameters of an annual tree survival and a diameter growth model.

2.1.4. Linking Models of Different Resolutions

Efforts have been made to ensure compatibility in models with different resolutions, either by aggregating individual tree or diameter distribution attributes into stand-level

8

predictions or disaggregating stand attributes to individual trees. Strub and Burkhart (1975) introduced a class-interval-free method for obtaining expected yield by summing diameter-class volumes, thus linking a diameter-distribution model with a whole-stand model. Matney and Sullivan (1982) recovered Weibull parameters from whole-stand attributes in such a way that the resulting diameter distribution when integrated yielded the same values of basal area and volume as derived from a whole-stand model. Hyink and Moser (1983) discussed parameter recovery methods to relate the existing mathematical compatibility between the diameter distribution models and stand-average models. Subsequently, Lynch and Moser (1986) obtained Weibull parameters from stand attributes such as basal area and sum of diameters for stands of mixed species.

Bailey (1980), recognizing that transformations of variables preserve the functional form of the distribution, provided a link between diameter distribution models and individual-tree growth models. He assumed that either no mortality occurred or mortality was proportionally distributed over the diameter distribution. Cao (1997) extended Bailey's (1980) work by including the case when mortality is not proportionally distributed. Qin et al. (2007) discussed related approaches to project trees and diameter-distributions through time. Matney and Schultz (2008) established a link between individual-tree models and diameter-distribution models by presenting a procedure for deriving tree diameter growth and probability of survival equations from successive diameter distributions.

Plenty of work has been done in linking individual-tree models with whole-stand models. Daniels and Burkhart (1988) presented an integrated system of forest stand models in which the components of detailed overall model are collapsed to provide compatible models at lower resolutions. Zhang et al. (1993) used relative size growth function to disaggregate stand volume

growth to a list of individual trees. Somers and Nepal (1994) presented an algorithm to adjust the results from the individual-tree models to agree with stand-level estimates. Qin and Cao (2006) developed a method to adjust coefficients of an individual-tree model to match predicted stand attributes. Implications of disaggregation in forest growth and yield modeling have been discussed by Ritchie and Hann (1997). The disaggregation approach adjusts individual-tree attributes to match stand-level predictions, which are assumed reliable. Yue et al. (2008), on the other hand, combined predictions from tree-level and stand-level models and then disaggregated the combine estimates to individual trees.

Nepal and Somers (1992) used an adjustment algorithm to create general stand-table projection model that links stand-level estimates to projections of observed stand tables. Cao and Baldwin (1999) applied the constrained least squares method to modify Nepal and Somers' (1992) algorithm. The constrained least squares method consistently provided the best results as compared to other methods (Cao 2007). Most recently, Cao (2007) incorporated tree-level and stand-level equations in a stand-table projection model and, in doing so, successfully linked individual-tree, size-class, and whole-stand models together in a working system.

2.2. Prediction of Parameters in Diameter Distribution Models

Many different probability density functions such as log-normal, exponential, gamma, beta, Johnson's S_B and Weibull have been used to describe diameter distributions. Because it can fit a variety of shapes and its cumulative distribution function exists in closed form, the Weibull distribution has become the predominant function in characterizing diameter distributions. Traditionally the Weibull parameters are directly predicted as functions of stand attributes by use of regression. Recent methods have been developed to recover the Weibull parameters from diameter moments and/or percentiles, which are predicted from regression equations.

2.2.1. Parameter-Prediction Method

In the parameter-prediction method, the values of the distribution parameters for describing diameter distributions are directly estimated from the stand attributes using regression equations. The stand attributes are age, height of the dominant and codominants, and density in terms of number of trees per unit area. Clutter and Bennett (1965) introduced this method to predict parameters of the beta distributions for old-field slash pine plantations. Smalley and Bailey (1974) applied this method to predict parameters of the Weibull function used in a model to predict yield and stand structure for loblolly pine plantations. The method was later used to predict the Weibull parameters for loblolly pine (Smalley and Bailey 1974, Feduccia et al. 1979), slash pine (Schreuder et al. 1979), mixed stand of western conifers (Little 1983), black spruce (Newton et al. 2005), *Eucalyptus grandis* (Mabvurira et al. 2002). Siipilehto (1999) used the parameter-prediction method to obtain parameters of S_B and Weibull distributions. Hyink and Moser (1983) discussed the concept of parameter prediction in a generalized framework for projecting forest yield and stand structure of even-aged and uneven-aged stands.

2.2.2. Parameter-Recovery Method

In the parameter-recovery approach, the parameters of the distribution are derived either from the diameter moments or from specific percentiles which are predicted from the stand-level attributes. Since the diameter frequency distribution characteristics such as mean diameter and diameter variance can be projected more precisely than the distribution parameters themselves, the parameter recovery method is considered superior to the parameter prediction method (Parresol 2003).

11

2.2.2.1. Parameter Recovery Based on Moments

Moment-based parameter-recovery approach uses diameter moments to estimate the parameters of the Weibull distribution. Predicted values of mean diameter, quadratic mean diameter, basal area, volume, and variance of diameter can be used to recover these parameters. Ek et al. (1975) showed how to obtain a Weibull parameter for a desired quadratic mean diameter if the remaining two parameters are known. Both Weibull scale and shape parameters were recovered from predicted stand basal area and average diameter (Cao et al. 1982, Hyink and Moser 1983, Lynch and Moser 1986), from predicted stand basal area and volume (Matney and Sullivan 1982), and from predicted arithmetic and quadratic mean diameters (Knoebel et al. 1986, Bowling et al. 1989). Burk and Newberry (1984) developed a method to recover all Weibull parameters from the first three predicted non-central moments.

The parameter-recovery approach has also been applied to predict the parameters of S_B and beta distributions. Parresol (2003) used the median and the first and second non-central moments of the distribution to recover the S_B parameters, whereas Fonseca et al. (2009) recovered all four S_B parameters from the median and the first three non-central moments. Gorgoso-Varela et al. (2008) recovered the two shape parameters of the beta distribution from the average diameter and variance.

2.2.2.2. Parameter Recovery Based on Percentiles

The percentile-based parameter-recovery approach uses regression to predict certain percentiles as functions of stand variables such as age, site index and stand density. The predicted percentiles are later used to recover the distribution parameters. The main advantage of this procedures is that diameter distribution characteristics, such as minimum diameter or

diameter percentiles which are used to recover the parameters can be predicted more precisely than the parameters itself (Knowe et al. 2005).

The parameter-recovery approach based on different diameter percentiles has been applied to various forest types. Contrary to moment-based methods that often result in a system of nonlinear equations that requires numerical techniques to solve, percentiles-based problems can be generally transformed into a linear system that can be easily solved. Lohrey and Bailey (1977) used the 24th and 93rd percentiles to recover the shape and scale parameters of the Weibull distribution.

2.2.2.3. Hybrid Methods

Hybrid methods include those methods that recover distribution parameters from both moments and percentiles. McTague and Bailey (1987) developed a technique for recovering the Weibull parameters from 10th, 63rd, and 93rd percentiles such that the resulting diameter distribution produced basal area that is consistent with the predicted basal area. In another percentile-based method, Bailey et al. (1989) computed the Weibull parameters from the predicted minimum diameter ($\widehat{D}_0$), quadratic mean diameter ($\widehat{D}_q$), 25th, 50th, and 95th percentiles. Many models have been developed using a similar approach (Brooks et al. 1992, Knowe et al. 2005, Lee and Coble 2006, Coble and Lee 2008, and Jiang and Brooks 2009). Baldwin and Feduccia (1987) developed a hybrid method for managed loblolly pine based on the minimum diameter, quadratic mean diameter, and the 93rd diameter percentile. A similar method was used by Zarnoch et al. (1991) to predict the Weibull parameters for predicting growth and yield for thinned and unthinned slash pine plantations. Bullock and Burkhart (2005) used $\widehat{D}_q$ and the 25th and 97th percentiles to characterize the juvenile diameter distributions of loblolly pine by the Weibull function.

2.2.3. CDF Regression

The cumulative-distribution-function regression (CDFR) method, introduced by Cao (2004), is similar to the parameter-prediction method in that the Weibull parameters are predicted from stand attributes, but the coefficients of these equations are obtained through minimization of the squared deviations between the observed and predicted CDF's. This new approach was found superior to all methods evaluated for prediction of diameter distributions of loblolly pine plantations by Cao (2004). The CDFR technique was also found by Newton and Amponsah (2005) and Cao and McCarty (2006) to yield the best goodness-of-fit statistics among the methods tested. Nord-Larsen and Cao (2006) applied this technique to even-aged beech with satisfactory results. Jiang and Brooks (2009), however, found that the hybrid method by Bailey et al. (1989) provided better results than the CDFR method for young longleaf pine plantations.

CHAPTER 3. MATERIALS AND METHODS

3.1. Data

Data from the Southwide Seed Source Study, which involves 15 loblolly pine (*Pinus taeda* L.) seed sources planted at 13 locations across 10 southern states (Wells and Wakeley, 1966), were used in this study. Seedlings were planted at 1.8 m × 1.8 m spacing. Each plot is of size 0.0164 ha, and trees were measured at ages of 10, 15 or 16, 20, and 25 or 27 years. Fifty randomly selected plots for each age group (200 plots in total) constituted the fit data set. To be included in the sample, a plot had to have at least 15 trees at that age. Another 50 plots were randomly selected among the remaining data to form the validation data set. The validation data consisted of four measurements for each plot, resulting in 200 plot-age combinations. The fit data were used for development of predicting equations, whereas the validation data were used to evaluate the methods. Summary statistics boxplots for the stand- and tree-level attributes for fit dataset are shown in Figure 1.
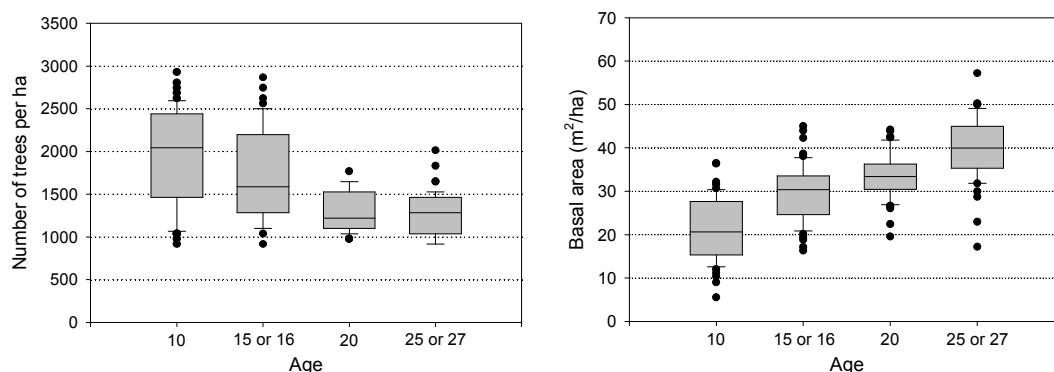
Figure 1. Box plots show the median, 25th and 75th percentiles (lower and upper extent of box), and values outside the 5th and 95th percentiles of the stand variables for four age groups in the fit data (50 plots in each age group).

(Figure 1 cont'd.)
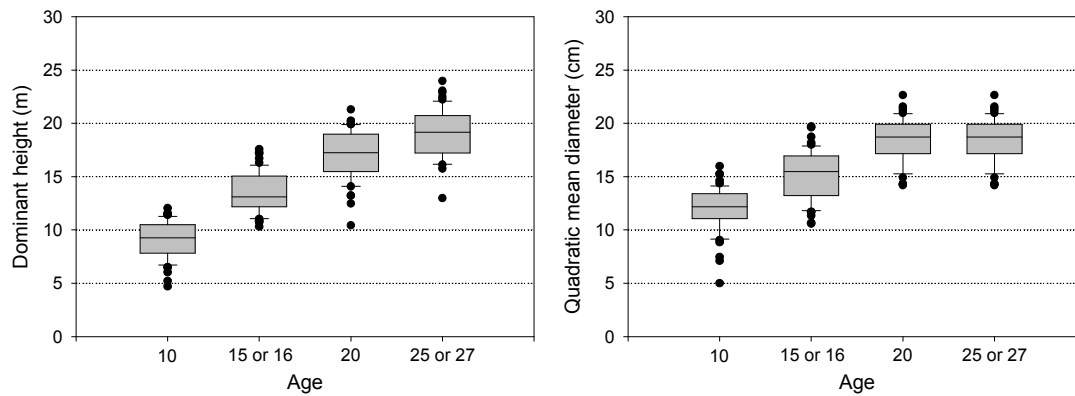
Summary statistics boxplots for validation dataset is shown in following figures.
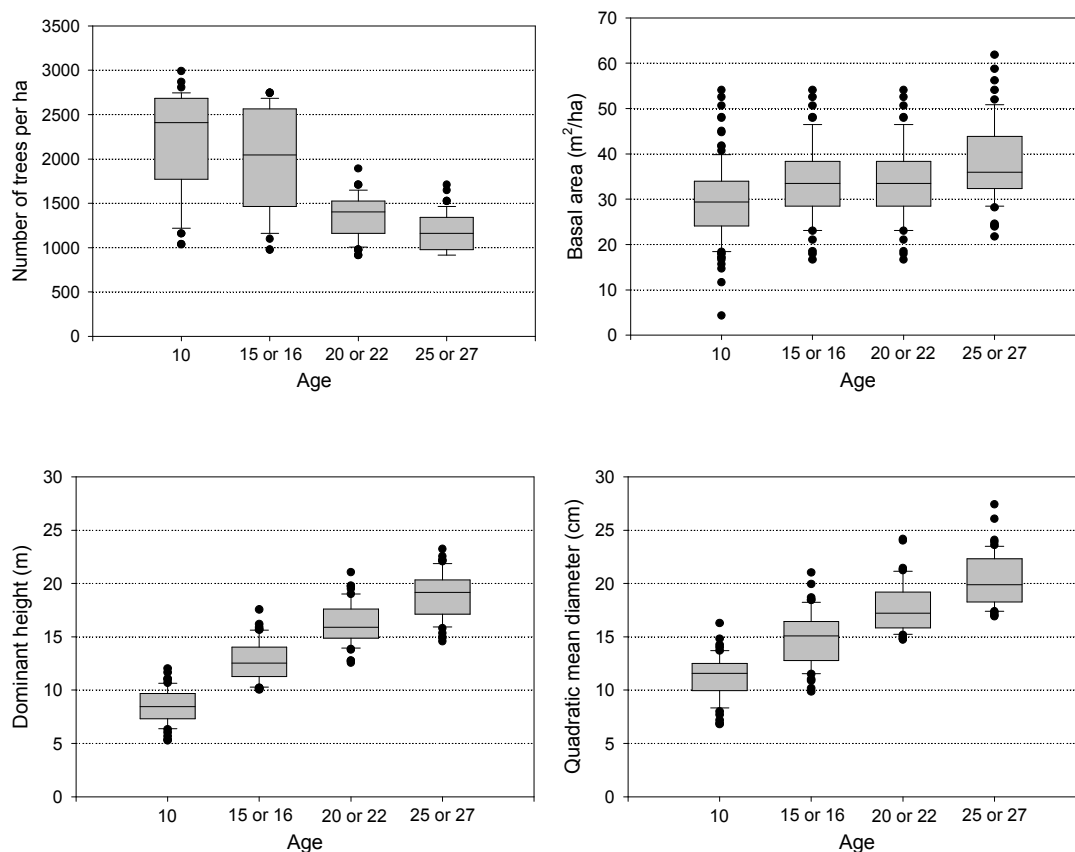
Figure 2. Box plots show the median, 25th and 75th percentiles (lower and upper extent of box), and values outside the 5th and 95th percentiles of the stand variables for four age groups in the validation data (50 plots in each age group).

16

3.2. Methods

The Weibull probability density function (PDF) used in this study to characterize diameter distribution is of the following form:

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

(1)

where *a, b*, and *c* are the location, scale, and shape parameters of the Weibull distribution, respectively, and *x* is tree diameter at breast height (dbh).

The following general form of regression equation, adopted from Cao (2004), was used to predict the Weibull parameters or diameter moments or percentiles:

$$y = exp[b_1 + b_2 RS + b_3 \ln(N) + b_4 \ln(H) + b_5/A] + \varepsilon$$

(2)

where

y = Weibull parameters, diameter percentiles, or moments (mean or variance);

$RS = \frac{\sqrt{10000/TPH}}{H}$, is the relative spacing (ratio of the average distance between trees to the average height of the dominants and codominants);

N = number of trees per hectare;

H = dominant height (average height of the dominants and codominants) in meters;

A = stand age in years;

ln (.) = natural logarithm;

b_k's = regression coefficients; and

ε = random error.

3.2.1. Ten Methods

3.2.1.1. Parameter Prediction

Method 1 – Weibull parameters

The three Weibull parameters were obtained from the fit data via maximum likelihood estimation. They were then used as dependent variables in equation (2).

3.2.1.2. Moment-based Parameter Recovery

In the moment-based parameter-recovery methods, the Weibull location parameter (a) was computed from the minimum diameter in the stand, which was predicted from equation (2):

$$a = 0.5\widehat{D}_0 , \tag{3}$$

where the caret symbol (^) above a variable name represents the predicted value of that variable.

Method 2 – $\overline{D}$ and D_{var}

Equation (2) was used to predict average diameter ($\overline{D}$) and diameter variance (D_{var}). The shape and scale parameters (b and c) were recovered from $\widehat{\overline{D}}$ and $\widehat{D}_{var}$ as follows:

$$b = \frac{\widehat{\overline{D}}-a}{G_1} , \text{ and} \tag{4}$$

$$b^2(G_2 - G_1^2) - \widehat{D}_{var} = 0. \tag{5}$$

where $G_i = \Gamma\left(1 + {}^i\!/_c\right)$, and $\Gamma(\cdot)$ is the complete gamma function.

Method 3 – D_q and D_{var}

Method 3 is similar to method 2, except that the Weibull shape and scale parameters were recovered from the predicted quadratic mean diameter ($\widehat{D}_q$) and diameter variance ($\widehat{D}_{var}$) as follows:

$$b = -a\, G_1/G_2 + [(a/G_2)^2\, (G_1^2 - G_2) + \widehat{D}_q^2/G_2]^{0.5} , \text{ and} \tag{6}$$

$$b^2(G_2 - G_1^2) - \widehat{D}_{var} = 0. \tag{7}$$

3.2.1.3. Percentile-based Parameter Recovery

Method 4 – D$_{31}$ and D$_{63}$

In this method, the minimum diameter (D_0) and the 31st and 63rd percentiles (D_{31} and D_{63}) were predicted from Equation (2). Again, the Weibull location parameter was computed from equation (4). The shape and scale parameters were calculated from

$$c = \frac{ln\left(\frac{\ln(1-0.63)}{\ln(1-0.31)}\right)}{\ln(\widehat{D}_{63}-a)-\ln(\widehat{D}_{31}-a)} \text{ , and} \tag{8}$$

$$b = \frac{\widehat{D}_{63}-a}{[-\ln(1-0.63)]^{1/c}} \text{ .} \tag{9}$$

Method 5 – D$_{50}$ and D$_{95}$

Method 5 is similar to method 4, except that the Weibull shape and scale parameters were recovered from the predicted 50th and 95th percentiles ($\widehat{D}_{50}$ and $\widehat{D}_{95}$) as follows:

$$c = \frac{ln\left(\frac{\ln(1-0.95)}{\ln(1-0.50)}\right)}{\ln(\widehat{D}_{95}-a)-\ln(\widehat{D}_{50}-a)} \text{ , and} \tag{10}$$

$$b = \frac{\widehat{D}_{50}-a}{[-\ln(1-0.50)]^{1/c}} \text{ .} \tag{11}$$

Method 6 – D$_{25}$, D$_{50}$, and D$_{95}$

This method is similar to methods 4 and 5, except that the Weibull shape and scale parameters were recovered from three predicted percentiles, $\widehat{D}_{25}$, $\widehat{D}_{50}$, and $\widehat{D}_{95}$ as follows:

$$c = \frac{ln\left(\frac{\ln(1-0.95)}{\ln(1-0.25)}\right)}{\ln(\widehat{D}_{95}-a)-\ln(\widehat{D}_{25}-a)} \text{ , and} \tag{12}$$

$$b = \frac{\widehat{D}_{50}-a}{[-\ln(1-0.50)]^{1/c}} \text{ .} \tag{13}$$

Method 7 – D$_{31}$, D$_{50}$, and D$_{63}$

Method 7 is similar to method 6, except that the Weibull shape and scale parameters were recovered from three predicted percentiles, $\widehat{D}_{31}$, $\widehat{D}_{50}$, and $\widehat{D}_{63}$ as follows:

$$c = \frac{ln\left(\frac{\ln(1-0.63)}{\ln(1-0.31)}\right)}{\ln(\widehat{D}_{63}-a)-\ln(\widehat{D}_{31}-a)} \text{, and} \tag{14}$$

$$b = \frac{\widehat{D}_{50}-a}{[-\ln(1-0.50)]^{1/c}}. \tag{15}$$

3.2.1.4. Hybrid Methods

Method 8 – $\bar{D}$ and D$_{95}$

Method 8 is similar to method 2, except that the Weibull shape and scale were recovered

from the predicted average diameter $(\widehat{\bar{D}})$ and the 95th percentile $(\widehat{D}_{95})$ as follows:

$$b = \frac{\widehat{D}_{95}-a}{[-\ln(1-0.95)]^{1/c}} \text{, and} \tag{16}$$

$$a + b\Gamma\left(1+\frac{1}{c}\right) - \widehat{\bar{D}} = 0. \tag{17}$$

Method 9 – D$_q$ and D$_{95}$

Method 9 is similar to method 8, with $\widehat{D}_q$ replacing $\widehat{\bar{D}}$. The Weibull shape and scale

parameters are the solution of

$$b = \frac{\widehat{D}_{95}-a}{[-\ln(1-0.95)]^{1/c}} \text{, and} \tag{18}$$

$$b^2 G_2 + 2abG_1 + a^2 - \widehat{D}_q^2 = 0. \tag{19}$$

Method 10 – D$_q$, D$_{25}$, D$_{50}$, and D$_{95}$

Method 10 was from Bailey et al. (1989). The difference between this method and the

rest of the methods was that the Weibull location parameter was calculated from $\widehat{D}_0$, $\widehat{D}_{50}$, and the

number of trees in the plot (n):

$$a = (n^{1/3}\widehat{D}_0 - \widehat{D}_{50})/(n^{1/3} - 1). \tag{20}$$

The scale and shape parameters are computed from:

$$c = \frac{ln\left(\frac{\ln(1-0.95)}{\ln(1-0.25)}\right)}{\ln(\widehat{D}_{95}-a)-\ln(\widehat{D}_{25}-a)} \text{, and} \tag{21}$$

$$b = -a\,G_1/G_2 + [(a/G_2)^2\,(G_1^2 - G_2) + \widehat{D}_q^2/G_2]^{0.5}. \tag{22}$$

3.2.2. Three Approaches

Equation (2) was a general equation used to predict Weibull parameters as well as diameter moments and percentiles. In this study, three approaches were investigated to obtain estimates for the regression coefficients, b_k's, in equation (2).

3.2.2.1. The SUR Approach

Because the error terms are correlated among the equations used to predict Weibull parameters, moments, or percentiles, the SUR approach was used to simultaneously estimate the regression coefficients b_k's in the system of equations with the SAS procedure Model, option SUR (SAS Institute, Inc. 1993). In method 1, the system of equations was used to predict the three Weibull parameters. In the remaining methods, the diameter moments and/or percentiles formed the set of equations. The Weibull location parameter "a" was obtained from equation (3).

3.2.2.2. The CDFR Approach

The CDF Regression approach was originally developed by Cao (2004). In this approach, the b_k's were obtained by minimizing the sum of squares of observed and predicted z, where z is CDF instead of the dependent variable:

$$\text{minimize } \sum_{i=1}^{p}\sum_{j=1}^{n_i}\left(F_{ij} - \widehat{F}_{ij}\right)^2 / n_i \tag{23}$$

where

$F_{ij} = (j–0.5)/n_i$ = observed cumulative probability of the tree j in the ith plot-age combination;

j = rank of that tree in terms of dbh (from smallest to largest) in the ith plot-age combination;

n_i = number of trees in the ith plot-age combination;

$\hat{F}_{ij} = 1 - \exp\left\{-\left[\frac{x_{ij}-a}{b}\right]^c\right\}$, value of the Weibull CDF evaluated at x_{ij};

x_{ij} = dbh of tree j in the ith plot-age combination; and

p = number of plot-age combinations.

Note that $F_{ij} = (j–0.5)/n_i$ is a Hazen's plotting position used for graphing empirical pdf's (Cunnane 1978). This method results in a CDF value less than 1 for the maximum diameter in a plot. In contrast, $F_{ij} = j/n_i$ as defined by Cao (2004) leads to a maximum CDF value of 1 for each plot, which is not theoretically correct because a predicted CDF would approach 1 but never attains the value of 1.

SAS procedure Model (SAS Institute, Inc. 1993) was used to iteratively search for the parameters of the CDF regression. Including the Weibull location parameter in the system of equations resulted in, for many plots, predicted location parameters that were much lower in value than the observed minimum diameters. Consequently, the minimum diameter was predicted separately from the system of equations.

3.2.2.3. The Modified CDFR Approach

The Modified CDFR approach is similar to the CDFR approach, except that the diameter class information is used to compute the CDF instead of individual trees as in the CDFR approach. The modified CDF regression was then fitted with SAS Procedure Model (SAS Institute, Inc. 1993). Diameter classes having 2-cm width were used and the b_k's were obtained by minimizing the function given in (23), where

$$F_{ij} = \frac{\sum_{k=1}^{j} n_{ik}}{n_i} = \text{observed cumulative probability of diameter class } j \text{ in the } i^{th} \text{ plot-age}$$

combination;

n_{ik} = number of trees in the kth diameter class in the ith plot-age combination;

n_i = total number of trees in the ith plot-age combination;

$\hat{F}_{ij} = 1 - \exp\left\{-\left[\frac{x_{ij}+1-a}{b}\right]^c\right\}$, the value of the Weibull CDF evaluated at the upper bound

of the jth diameter class; and

x_{ij} = midpoint of the jth diameter class in the ith plot-age combination.

Table 1 shows a summary of the applications of the SUR, CDFR, and Modified CDFR approaches to the ten methods included in the evaluation.

Table 1. Summary of methods for predicting Weibull parameters included in this study.

Methods	Approach a: SUR	Approach b: CDF Regression	Approach c: MCDF Regression
Parameter Prediction			
Method 1 – Weibull parameters	1a	1b	1c
Moment-based Parameter Recovery			
Method 2 – $\bar{D}$ and D_{var}	2a	2b	2c
Method 3 – D_q and D_{var}	3a	3b	3c
Percentile-based Parameter Recovery			
Method 4 – D_{31} and D_{63}	4a	4b	4c
Method 5 – D_{50} and D_{95}	5a	5b	5c
Method 6 – D_{25}, D_{50}, and D_{95}	6a	6b	6c
Method 7 – D_{31}, D_{50}, and D_{63}	7a	7b	7c
Hybrid Methods			
Method 8 – $\bar{D}$ and D_{95}	8a	8b	8c
Method 9 – Dq and D_{95}	9a	9b	9c
Method 10 – Dq, D_{25}, D_{50}, and D_{95}	10a	10b	10c

3.3. Evaluation

The following four goodness-of-fit statistics were computed for each method to evaluate the methods. The method producing the lowest values for each of the evaluation statistics is the best method. The **Anderson-Darling (AD)** statistic is calculated with

$$AD_i = -n_i - \sum_{j=1}^{n_i}(2j-1)\left[ln(u_j) + ln(1 - u_{n_i-j+1})\right]/n_i, \tag{24}$$

where

$$u_j = F(x_j) = 1 - \exp\left\{-\left[\frac{x_j-a}{b}\right]^c\right\};$$

n_i = number of trees in the ith plot-age combination; and

x_j are dbh, sorted in ascending order for each plot-age combination;

$(x_1 \leq x_2 \dots \dots \leq x_{n_i})$.

The **Kolmogorov-Smirnov (KS)** statistic is calculated with

$$KS_i = max\{max_{1\leq i\leq n_i}[(j/n_i) - u_j], max_{1\leq i\leq n_i}[(u_j - (j-1)/n_i]\}, \tag{25}$$

where n_i and u_j are previously defined in equation (24).

The **Negative Log-Likelihood (mlogL)** statistic is calculated as follows:

$$mLogL = \sum_{j=1}^{n_i}\left[ln(b) - ln(c) + (1-c)\,ln\left(\frac{x_{ij}-a}{b}\right) + \left(\frac{x_{ij}-a}{b}\right)^c\right], \tag{26}$$

where n_i is previously defined in equation (24) and x_{ij} is the dbh of the tree j in the ith plot-age combination.

The **Error Index (EI)** is computed from

$$EI_i = \sum_{k=1}^{m_i}|n_{ik} - \hat{n}_{ik}|, \tag{27}$$

where, n_{ik} and $\hat{n}_{ik}$ are the observed and predicted number of trees per hectare in diameter class k, and m_i is the number diameter classes for the ith plot-age combination.

3.3.1. Ranking of Methods

The traditional standard or ordinal ranks for m methods are 1, 2, …, m. They show the order of the method, but not how close they are to one another. In this study, the relative ranks were developed to display the relative positions of the methods. The relative rank of method i is defined as

$$R_i = 1 + \frac{(m-1)(S_i - S_{min})}{S_{max} - S_{min}},$$ (28)

where

R_i = relative rank of method i ($i = 1, 2, …, m$);

S_i = goodness-of-fit statistic produced by method i;

S_{min} = minimum value of the goodness-of-fit statistic; and

S_{max} = maximum value of the goodness-of-fit statistic.

In this ranking system, the best and the worst methods have relative ranks of 1 and m, respectively. Ranks of the remaining methods are expressed as real numbers between 1 and m. Because the magnitude, and not only the order, of the S_i's are taken into consideration, the new ranking system should provide more information than the traditional ordinal ranks. For example, relative ranks of 1, 1.2, 4.7, 4.9, 5 in the case of five methods suggest that the methods fall into two groups that were separated by a large gap.

CHAPTER 4. RESULTS AND DISCUSSION

Altogether ten methods were used to predict the parameters of Weibull function for modeling the diameter distributions. Each method was then carried out by using three different approaches (SUR, CDFR, and MCDFR) resulting in a total of thirty methods. The full model used to predict the Weibull parameters and diameter moments or percentiles was in the form of equation (2). The final models were selected in a backward elimination procedure by manually removing insignificant variables (at 5% level of significance).

4.1. The SUR Approach

The parameter estimates for predicting Weibull parameters and diameter moments and percentiles based on the SUR approach are presented in Table 2. Four goodness-of-fit statistics (AD, KS, mLogL, and EI) were computed to evaluate these methods; their means and standard deviations are shown in Table 3. The relative ranks were computed from the means of the statistics, based on the method described in section 3.3.5 (Table 4).

Table 2. Coefficients of regression equations for predicting Weibull parameters, diameter moments, and diameter percentiles from the Seemingly Unrelated Regression approach.

Dependent Variable	b_1	b_2	b_3	b_4	b_5
a	6.08327	-2.30454	-0.38324		-18.0943
b	5.58586	-2.40865	-0.44199		6.73413
c	0.81430				5.50300
D_0	11.65376	-5.78286	-0.87808	-0.58714	-6.75389
D_{25}	6.65052	-2.56361	-0.48240		-0.83737
D_{31}	6.76116	-2.75387	-0.49326		
D_{50}	7.04498	-2.77660	-0.51499		
D_{63}	7.17042	-2.76864	-0.52278		
D_{95}	7.81760	-2.63151	-0.58401		
$\bar{D}$	7.02532	-2.73406	-0.50965		-0.50489
Dq	5.58586	-2.40865	-0.44199		6.73413

Table 3. Means and standard deviations of the goodness-of-fit statistics produced by different methods based on the Seemingly Unrelated Regression approach.

Method	AD[1]		KS[2]		mLogL[3]		EI[4]	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
1	4.676	5.096	0.2642	0.1101	71.612	22.109	2340.71	882.75
2	3.475	4.101	0.2425	0.0886	71.026	21.939	2271.08	878.77
3	3.480	4.103	0.2426	0.0890	71.030	21.945	2272.17	878.52
4	9.222	8.439	0.3966	0.1227	78.116	23.368	2289.87	818.38
5	16.870	11.208	0.5305	0.1135	86.260	25.876	2449.81	861.18
6	3.743	4.904	0.2424	0.0896	71.330	23.029	2278.62	904.06
7	3.734	4.916	0.2424	0.0892	71.349	22.989	2277.90	906.29
8	4.018	5.442	0.2457	0.0926	71.634	23.709	2292.61	923.90
9	3.976	5.537	0.2432	0.0925	71.672	23.960	2284.54	917.55
10	3.867	5.064	0.2435	0.0915	71.211	23.278	2284.38	905.61

[1]   AD = Anderson-Darling statistic,
[2]   KS = Kolmogorov-Smirnov statistic,
[3]   mLogL = negative Log-Likelihood statistic, and
[4]   EI = error index.

Table 4. Relative ranks of ten methods based on the Seemingly Unrelated Regression approach.

Method	Relative Rank				Sum of Ranks	Overall Rank
	AD[1]	KS[2]	mLogL[3]	EI[4]		
1	1.8072	1.6819	1.3462	4.5062	9.3415	2.3342
2	1.0000	1.0055	1.0000	1.0000	4.0055	1.0000
3	1.0032	1.0067	1.0021	1.0549	4.0669	1.0154
4	4.8614	5.8180	5.1883	1.9462	17.8139	4.4526
5	10.0000	10.0000	10.0000	10.0000	40.0000	10.0000
6	1.1798	1.0000	1.1791	1.3797	4.7386	1.1833
7	1.1744	1.0003	1.1905	1.3434	4.7085	1.1758
8	1.3647	1.1053	1.3589	2.0841	5.9130	1.4769
9	1.3364	1.0252	1.3815	1.6778	5.4209	1.3539
10	1.2633	1.0358	1.1092	1.6697	5.0781	1.2682

[1]   AD = Anderson-Darling statistic,
[2]   KS = Kolmogorov-Smirnov statistic,
[3]   mLogL = negative Log-Likelihood statistic, and
[4]   EI = error index.

A radar chart (Figure 3) shows graphically the relative rankings of the ten methods. Each method is represented by a quadrilateral, whose area is smallest for the best method and largest for the worst method.

Figure 3. Relative ranks of ten methods based on the Seemingly Unrelated Regression approach. Method resulting in the smallest area inside the box represents the best method.

The largest area in Figure 1 belongs to method 5 (D_{50} and D_{95}), which ranked last in all four evaluation statistics. The next largest area are from method 4 (D_{31} and D_{63}) and method 1 (Weibull parameters). The areas formed by the remaining methods are almost indistinguishable from one another. This group of methods produced lower values of evaluation statistics, with method 2 ($\bar{D}$ and D_{var}) having the highest rank in three out of four statistics, and the best overall ranking. These visual results are consistent with the overall rankings presented in Table 4.

4.2. The CDFR Approach

Parameter estimates for the ten methods based on the CDFR approach are presented in Table 5. Table 6 shows the means and standard deviations of the goodness-of-fit statistics produced by the CDFR approach. The relative ranks of the ten methods are presented in Table 7.

Table 5. Coefficients of regression equations for predicting Weibull parameters, diameter moments, and diameter percentiles from the CDF Regression approach.

Method	Dependent Variable	b_1	b_2	b_3	b_4	b_5
1	b	4.95758	-1.88134	-0.38762	0.21365	3.23251
	c	6.67966	-5.34300	-0.40226	-0.70784	4.37674
2	$\overline{D}$	5.98814	-2.16514	-0.44552	0.15114	0.53527
	D_{var}	5.96375		-0.63579	0.55288	
3	Dq	5.38139	-1.69082	-0.39486	0.22224	0.37597
	D_{var}	-1.54062	4.92116		1.38324	-4.30579
4	D_{31}	8.02868	-4.31979	-0.60891	-0.15306	0.77814
	D_{63}	7.01308	-3.26449	-0.52529		0.84566
5	D_{50}	6.98359	-3.54718	-0.53793		0.83566
	D_{95}	6.03076	-2.30443	-0.44884	0.14701	0.75331
6	D_{25}	-23.6158	18.74934	1.75346	3.60340	
	D_{50}	7.06889	-2.95733	-0.51957		0.61498
	D_{95}	-33.704	26.36538	2.56939	4.78983	
7	D_{31}	9.85415	-11.4056	-0.09254	-1.63660	
	D_{50}	6.95752	-2.80171	-0.50311		
	D_{63}	9.63102	-11.5747		-1.71161	
8	$\overline{D}$	3.68454		-0.25071	0.47662	
	D_{95}	5.69550	-1.98785	-0.42379	0.18930	0.52450
9	Dq	6.85882	-1.80606	-0.45488		-1.19910
	D_{95}	6.31597	-2.1821	-0.45753	0.08693	
10	D_0	8.82900	-4.91421	-0.78470		
	D_{50}	7.01792	-2.75970	-0.51166		
	Dq	5.62421	-1.85712	-0.41224	0.18915	0.41041
	D_{25}	7.56094	-3.75144	-0.6378		-5.64058
	D_{95}	6.03824		-0.44965		-11.1105

Table 6. Means and standard deviations of the goodness-of-fit statistics produced by different methods based on the CDF Regression approach.

Method	AD[1] Mean	AD[1] Standard Deviation	KS[2] Mean	KS[2] Standard Deviation	mLogL[3] Mean	mLogL[3] Standard Deviation	EI[4] Mean	EI[4] Standard Deviation
1	2.841	2.881	0.2348	0.0803	71.349	21.468	2233.72	826.44
2	2.850	2.859	0.2350	0.0799	71.421	21.590	2237.63	834.55
3	2.811	2.762	0.2344	0.0798	71.459	21.543	2235.35	827.78
4	2.852	2.950	0.2347	0.0807	71.333	21.498	2234.42	830.69
5	2.849	2.911	0.2347	0.0803	71.356	21.524	2235.35	832.50
6	2.872	2.814	0.2365	0.0793	71.596	21.660	2242.70	835.87
7	2.869	2.720	0.2374	0.0806	71.706	21.849	2246.10	839.18
8	2.834	2.852	0.2345	0.0802	71.396	21.533	2235.48	829.80
9	2.846	2.799	0.2357	0.0799	71.517	21.525	2236.59	832.04
10	2.775	2.770	0.2343	0.0788	71.418	21.746	2239.22	837.69

[1] AD = Anderson-Darling statistic,
[2] KS = Kolmogorov-Smirnov statistic,
[3] mLogL = negative Log-Likelihood statistic, and
[4] EI = error index.

Table 7. Relative ranks of ten methods based on the CDF Regression approach.

Method	Relative Rank AD[1]	Relative Rank KS[2]	Relative Rank mLogL[3]	Relative Rank EI[4]	Sum of Ranks	Overall Rank
1	7.1218	2.5913	1.3799	1.0000	12.0930	1.6249
2	7.9527	3.0275	3.1214	3.8425	17.9441	3.3998
3	4.3597	1.3709	4.0345	2.1850	11.9500	1.5815
4	8.1072	2.2959	1.0000	1.5089	12.9120	1.8733
5	7.8638	2.3274	1.5583	2.1850	13.9345	2.1835
6	10.0000	7.3540	7.3412	7.5283	32.2234	7.7313
7	9.7026	10.0000	10.0000	10.0000	39.7026	10.0000
8	6.4529	1.6347	2.5052	2.2795	12.8723	1.8613
9	7.5735	5.1944	5.4221	3.0864	21.2765	4.4106
10	1.0000	1.0000	3.0346	4.9984	10.0330	1.0000

[1] AD = Anderson-Darling statistic,
[2] KS = Kolmogorov-Smirnov statistic,
[3] mLogL = negative Log-Likelihood statistic, and
[4] EI = error index.

The radar chart below (Figure 4) shows graphically the relative rankings of the ten methods.

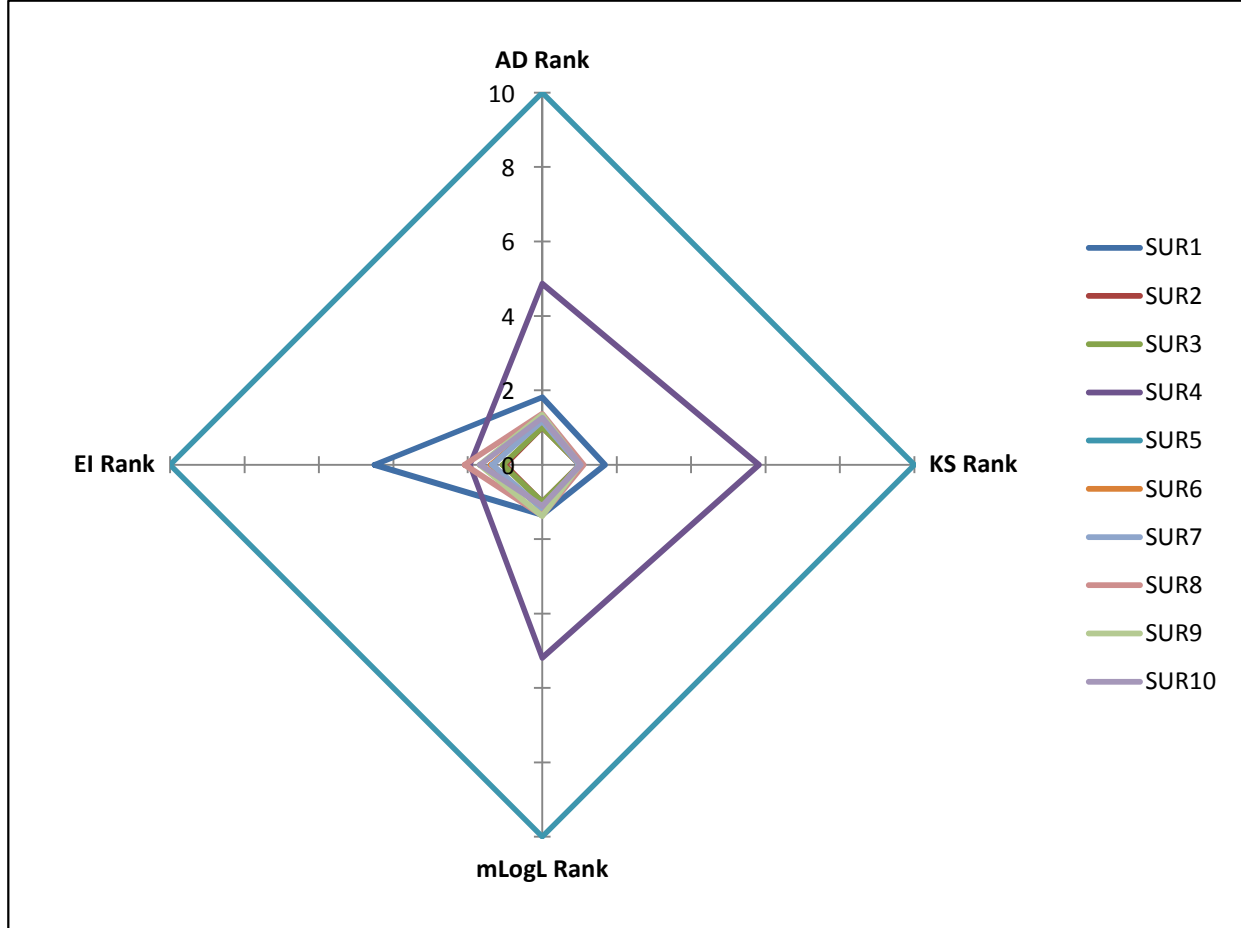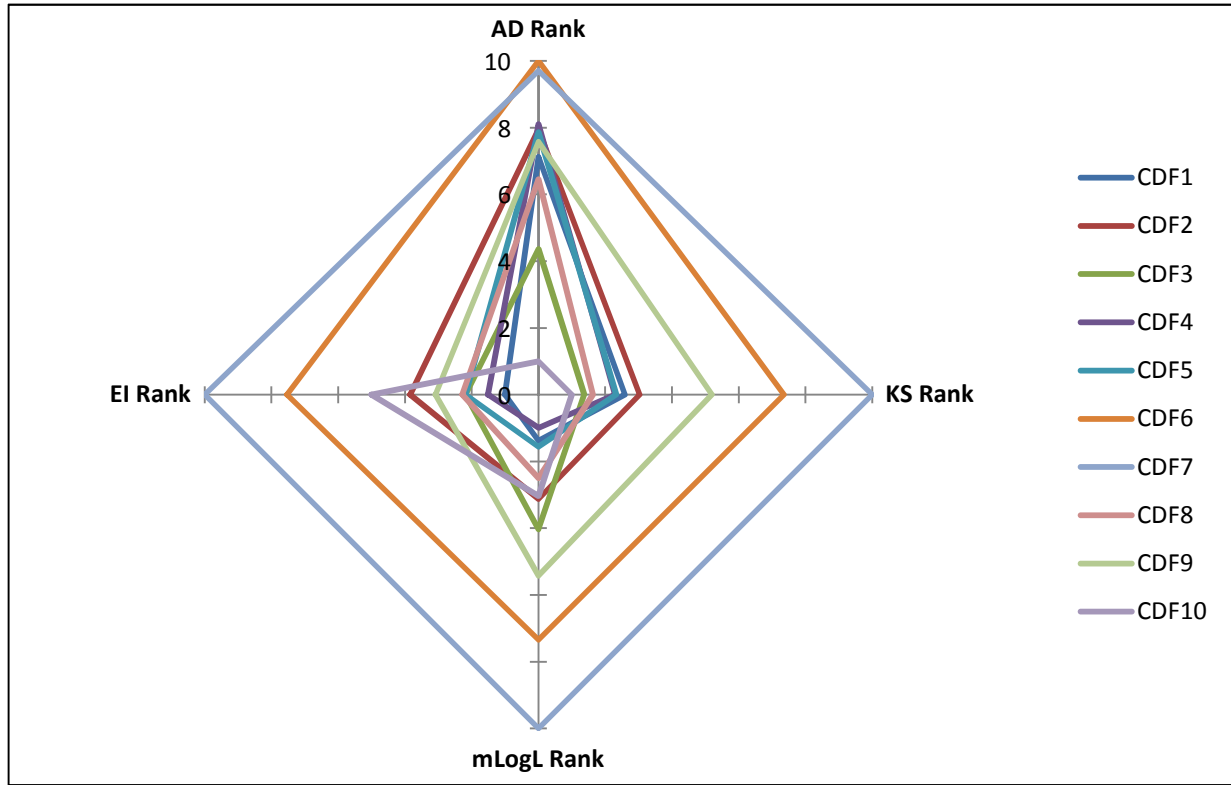Figure 4. Relative ranks of ten methods based on the CDF Regression approach. Method resulting in the smallest area inside the box represents the best method.

Both the overall ranks of the methods (Table 7) and the radar chart (Figure 4) show that the ten methods can be grouped into three general groups. The good methods include methods 1, 3, 4, 8, and 10; the difference in values of goodness-of-fit statistics was small for the methods in this group, with method 10 (Dq, $D_{25,}$ D_{50}, and D_{95}) being ranked highest overall. The intermediate group consists of methods 2 ($\bar{D}$ and D_{var}), 5 (D_{50}, and D_{95}), and 9 (D_q and D_{95}). The last group, which produced higher values for goodness-of-fit statistics, consisted of method 7 (D_{31}, D_{50}, and D_{63}) and method 6 (D_{25}, D_{50}, and D_{95}), with method 7 being ranked last in three statistics and next to last in another.

4.3. The Modified CDFR Approach

The Modified CDFR approach is similar to the CDF regression approach, except that information from diameter classes rather than from individual tree is used to compute the CDF. The coefficients of different regression equations used in the Modified CDFR approach are presented in Table 8. Means and standard deviations for four goodness-of-fit are shown in Table 9. Table 10 shows the relative ranks of ten methods for the Modified CDFR approach.

Table 8. Coefficients of regression equations for predicting Weibull parameters, diameter moments, and diameter percentiles from the Modified CDFR approach.

Method	Dependent Variable	b_1	b_2	b_3	b_4	b_5
1	b	5.15256	-1.96590	-0.40184	0.19179	3.01959
	c	2.50016	-1.56311		-0.35497	
2	$\bar{D}$	5.95666	-2.07355	-0.43396	0.14015	
	D_{var}	4.94930		-0.64333	0.87043	
3	Dq	5.84880	-1.95729	-0.42525	0.15834	
	D_{var}	4.77006		-0.62542	0.88947	
4	D_{31}	6.46465	-2.84150	-0.46979		-0.55942
	D_{63}	6.18670	-2.45359	-0.45093	0.08071	
5	D_{50}	6.42425	-2.70680	-0.46115		-0.92123
	D_{95}	6.05735	-2.16099	-0.43847	0.12113	
6	D_{25}	7.26725		-0.42708	-0.50969	-16.67660
	D_{50}	5.86451	-2.01292	-0.42603	0.14917	
	D_{95}	10.07989		-0.51252	-1.12632	-22.41020
7	D_{31}	3.13363	2.48766	-0.10052		-17.63450
	D_{50}	6.94084	-2.72662	-0.50077		
	D_{63}	2.40152	3.64880			-19.97640
8	$\bar{D}$	5.31485	-1.66236	-0.38853	0.22775	
	D_{95}	3.46464		-0.24781	0.53863	
9	Dq	7.01802	-2.68187	-0.50919		
	D_{95}	7.41697	-2.51196	-0.52833		
10	D_0	8.82900	-4.91421	-0.78470		
	D_{50}	7.01792	-2.75970	-0.51166		
	Dq	5.89517	-2.00289	-0.41768	0.14861	
	D_{25}	2.74741	3.51696			-18.26170
	D_{95}	4.54473	4.43843		-0.45469	-24.58270

Table 9. Means and standard deviations of the goodness-of-fit statistics produced by different methods based on the Modified CDFR approach.

Method	AD[1]		KS[2]		mLogL[3]		EI[4]	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
1	3.123	3.867	0.2332	0.0840	70.883	22.118	2232.21	843.95
2	3.069	3.786	0.2327	0.0839	70.758	21.883	2226.69	837.75
3	3.063	3.767	0.2327	0.0838	70.756	21.865	2226.29	836.86
4	3.107	3.887	0.2328	0.0837	70.856	22.131	2232.07	844.60
5	3.143	3.920	0.2337	0.0843	70.900	22.140	2233.83	846.36
6	3.060	3.559	0.2328	0.0824	70.827	21.716	2227.15	832.18
7	3.040	3.563	0.2336	0.0822	70.785	21.674	2225.33	825.54
8	3.079	3.827	0.2324	0.0845	70.778	21.910	2226.78	835.96
9	3.098	3.828	0.2340	0.0839	70.817	21.939	2228.55	840.53
10	2.960	3.618	0.2318	0.0816	70.683	21.812	2224.90	831.70

[1]   AD = Anderson-Darling statistic,
[2]   KS = Kolmogorov-Smirnov statistic,
[3]   mLogL = negative Log-Likelihood statistic, and
[4]   EI = error index.

Table 10. Relative ranks of ten methods based on the Modified CDFR approach.

Method	Relative Rank				Sum of Ranks	Overall Rank
	AD[1]	KS[2]	mLogL[3]	EI[4]		
1	9.0175	6.4587	9.2982	8.3673	33.1417	8.5731
2	6.3690	4.6915	4.0810	2.8040	17.9454	4.6240
3	6.0464	4.4430	4.0366	2.4009	16.9269	4.3594
4	8.2405	4.9461	8.1855	8.2262	29.5982	7.6523
5	10.0000	8.6323	10.0000	10.0000	38.6323	10.0000
6	5.9038	5.0367	6.9500	3.2676	21.1581	5.4589
7	4.9487	8.1288	5.2152	1.4334	19.7260	5.0868
8	6.8535	3.5829	4.9378	2.8947	18.2689	4.7081
9	7.7934	10.0000	6.5581	4.6786	29.0301	7.5047
10	1.0000	1.0000	1.0000	1.0000	4.0000	1.0000

[1]   AD = Anderson-Darling statistic,
[2]   KS = Kolmogorov-Smirnov statistic,
[3]   mLogL = negative Log-Likelihood statistic, and
[4]   EI = error index.

The radar chart below (Figure 5) shows graphically the relative rankings of the ten methods in the Modified CDFR approach.
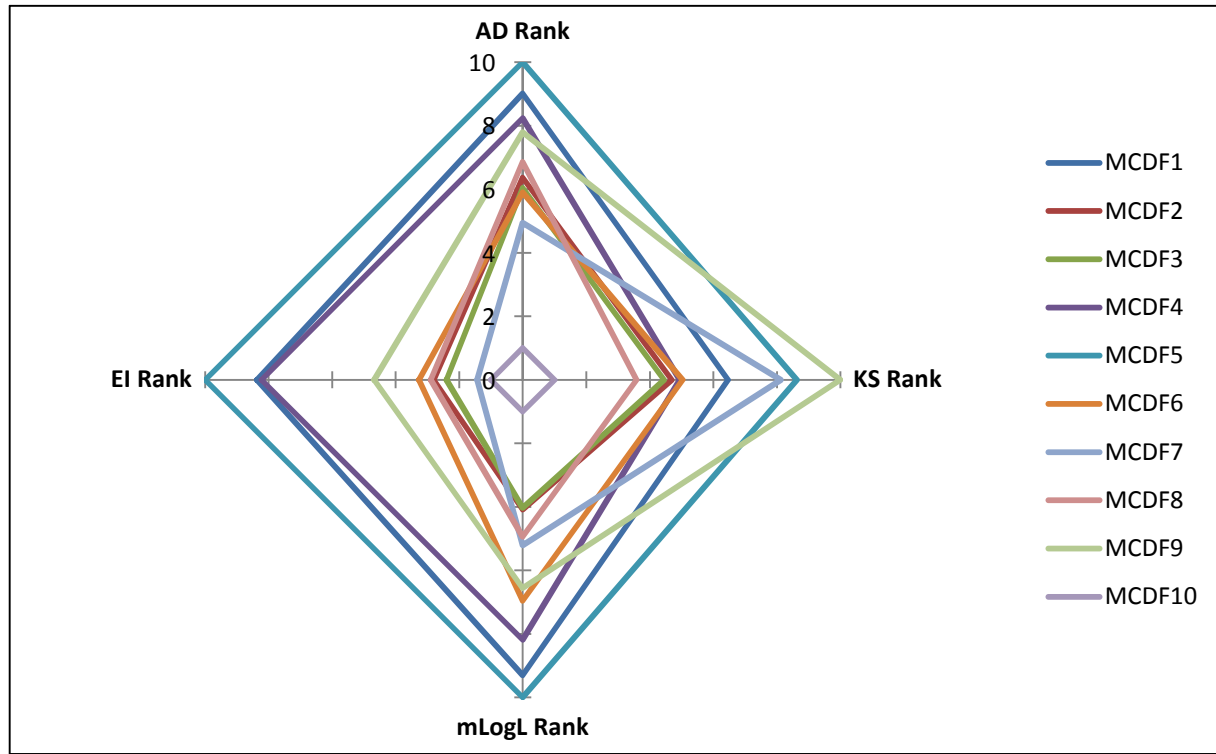
Figure 5. Relative ranks of ten methods based on the Modified CDFR approach. Method resulting in the smallest area inside the box represents the best method.

The radar chart (Figure 5) also reveals three groups, though not as distinct as those in the previous two approaches. The worst group comprises methods 1, 4, 5, and 9, with method 5 being ranked last overall. Method 10 (D_q, D_{25}, D_{50}, and D_{95}) is the sole occupant of the best group, ranking first in all statistics. The intermediate group includes the remaining methods.

4.4. Comparison of Three Approaches

Relative ranks for all thirty methods based on goodness-of-fit statistics were also computed to find the optimum method (Table 11). Method 10 from the Modified CDFR approach ranked as the best method among all thirty methods evaluated in this study, whereas method 5 based on the SUR approach was the poorest performer.

Table 11. Overall comparison of the thirty methods based on four goodness-of-fit statistics.

Method		Relative Rank				Sum of Ranks	Overall Rank
		AD[1]	KS[2]	mLogL[3]	EI[4]		
SUR[5]	1	4.9124	4.1445	2.7294	15.9326	27.7189	6.8536
	2	2.4407	2.0418	1.6386	6.9545	13.0756	3.1807
	3	2.4505	2.0457	1.6452	7.0950	13.2364	3.2210
	4	14.2648	17.0010	14.8371	9.3773	55.4801	13.8168
	5	30.0000	30.0000	30.0000	30.0000	120.0000	30.0000
	6	2.9914	2.0248	2.2031	7.9267	15.1460	3.7000
	7	2.9747	2.0257	2.2388	7.8338	15.0730	3.6817
	8	3.5574	2.3520	2.7695	9.7306	18.4095	4.5186
	9	3.4708	2.1031	2.8409	8.6900	17.1048	4.1913
	10	3.2469	2.1362	1.9829	8.6694	16.0353	3.9231
CDFR[6]	1	1.1366	1.2925	2.2393	2.1373	6.8057	1.6081
	2	1.1552	1.3074	2.3737	2.6414	7.4777	1.7766
	3	1.0750	1.2508	2.4441	2.3474	7.1173	1.6862
	4	1.1586	1.2824	2.2100	2.2275	6.8785	1.6263
	5	1.1532	1.2835	2.2530	2.3474	7.0372	1.6661
	6	1.2008	1.4553	2.6993	3.2951	8.6506	2.0708
	7	1.1942	1.5458	2.9045	3.7335	9.3780	2.2533
	8	1.1217	1.2598	2.3261	2.3642	7.0718	1.6748
	9	1.1467	1.3815	2.5512	2.5073	7.5867	1.8040
	10	1.0000	1.2381	2.3670	2.8464	7.4515	1.7700
Modified CDFR	1	1.7165	1.1299	1.3717	1.9426	6.1607	1.4463
	2	1.6059	1.0878	1.1380	1.2308	5.0625	1.1708
	3	1.5924	1.0819	1.1360	1.1792	4.9896	1.1525
	4	1.6841	1.0939	1.3219	1.9245	6.0244	1.4121
	5	1.7576	1.1816	1.4032	2.1514	6.4938	1.5298
	6	1.5864	1.0960	1.2666	1.2901	5.2391	1.2151
	7	1.5465	1.1696	1.1888	1.0554	4.9604	1.1452
	8	1.6261	1.0615	1.1764	1.2424	5.1064	1.1818
	9	1.6654	1.2141	1.2490	1.4706	5.5992	1.3054
	10	1.3815	1.0000	1.0000	1.0000	4.3815	1.0000

[1] AD = Anderson-Darling statistic,
[2] KS = Kolmogorov-Smirnov statistic,
[3] mLogL = negative Log-Likelihood statistic,
[4] EI = error index,
[5] SUR = Seemingly Unrelated Regression, and
[6] CDFR = Cumulative Distribution Function Regression.

4.4.1. The SUR vs. the CDFR and Modified CDFR Approaches

The radar plot based on the relative ranks of thirty methods (Figure 6) shows three

distinct groups. The Modified CDFR approach was best, followed by the CDFR approach.
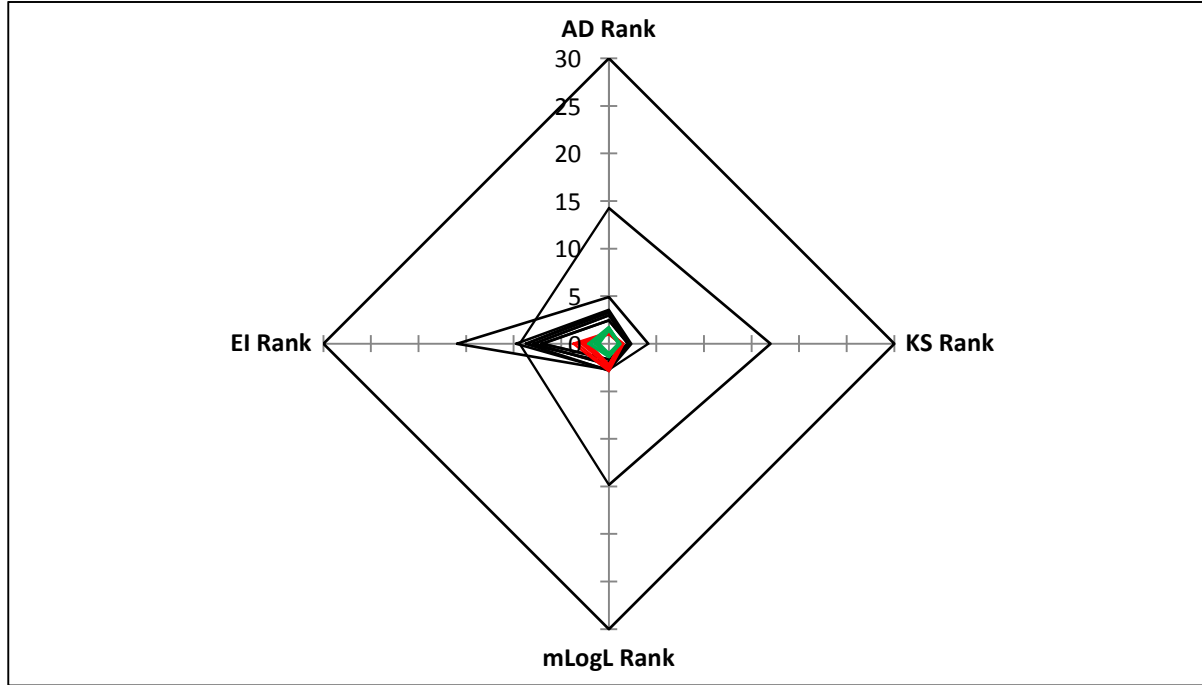

Figure 6. Overall comparison of the SUR (black), the CDFR (red), and the MCDFR (green)
approaches for all 30 methods. Method resulting in the smallest area inside the box
represents the best method.

The SUR approach was a distant third. The overall ranks ranged from 1.000 to 1.530 for

the Modified CDFR approach, 1.608 to 2.253 for the CDFR approach, and 3.181 to 30.000 for

the SUR approach (Table 11). The SUR approach consistently yielded higher AD, KS, and EI

statistics than did the other two approaches. There was, however, some overlap in the mLogL

statistic between the SUR and the CDFR approaches. On the other hand, the Modified CDFR

produced lower values for all four evaluation statistics than did the SUR approach.

These results indicate that the CDFR and the Modified CDFR approaches were clearly

superior to the SUR approach. However, the use of more extensive dataset and locally important

reference data and considering other explanatory variables could further validate the application of the methods for larger areas. The SUR approach did not perform as well as the other two approaches because it is based on point estimates: its objective is to minimize the squared difference between the observed and predicted Weibull parameters, diameter moments, and/or percentiles. The CDFR and Modified CDFR approaches, on the other hand, are based on the entire distribution: aiming at minimizing the squared difference between the observed and predicted CDF.

4.4.2. The CDFR vs. the Modified CDFR Approach

Figure 7 is Figure 6 redrawn after removing three worst methods, all of which were from the SUR approach, to better show the relative performance of methods based on the CDFR and Modified CDFR approaches.
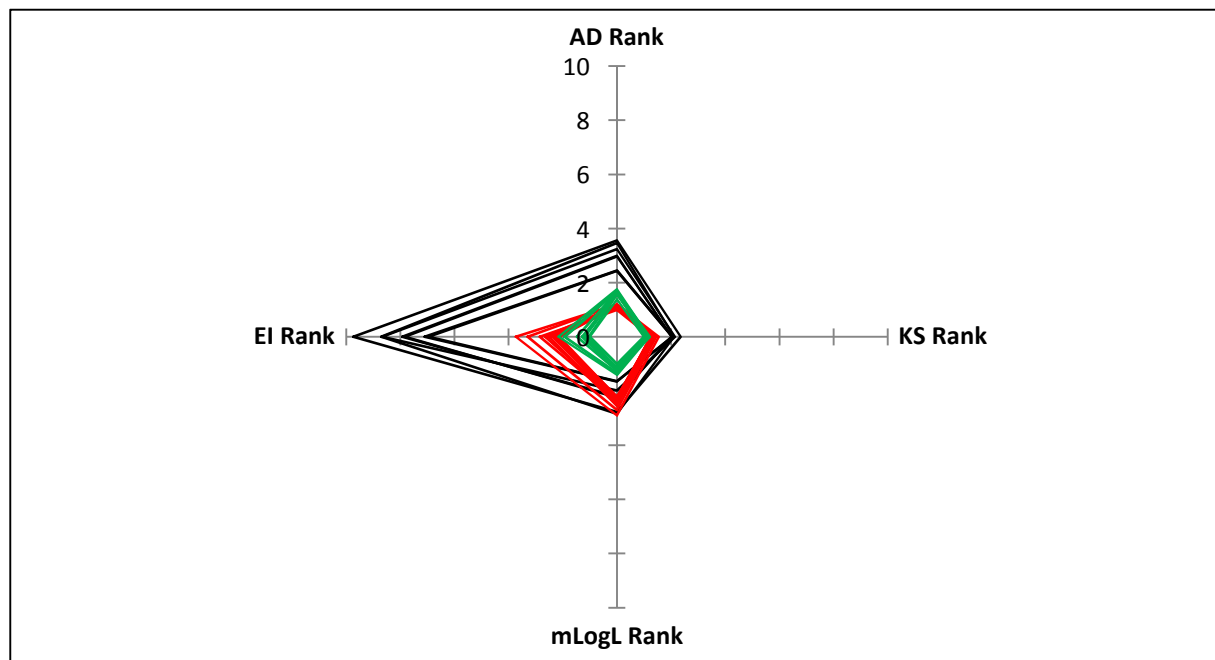

Figure 7. Comparison of the SUR (black), the CDFR (red), and the MCDFR (green) approaches for the best 27 methods. Method resulting in the smallest area inside the box represents the best method.

The Modified CDFR approach performed better in terms of the KS, EI, and mLogL statistics, but consistently produced higher AD values as compared to the CDFR. The reason for increased AD values might be because the CDF in the Modified CDFR approach is calculated based on the histogram of the diameter classes, and the AD statistic gives more weight to the tails than does the KS statistic (Cirillo and Hüsler 2009). It makes sense that the Modified CDFR approach would produce low EI values because both its CDF and the error index are based on the histogram of the diameter distribution.

4.4.3. The Best and Worst Methods

All ten methods from the Modified CDFR approach performed well. Method 10 (D_q, D_{25}, D_{50}, and D_{95}) was clearly the best overall method. This method also ranked best in the CDFR approach and performed relatively well in the SUR approach (rank of 1.268 out of 10).

The worst method belonged to method 5 (D_{50} and D_{95}), which ranked last in both SUR and Modified CDFR approaches. Method SUR 5 also scored lowest in all four evaluation statistics among thirty methods.

CHAPTER 5. SUMMARY AND CONCLUSIONS

Diameter distribution is an effective method for describing stand properties because important variables such as volume, value, conversion cost, and product specifications are dependent on tree diameter. Because of its flexibility and relative simplicity, the Weibull function has been widely used in describing diameter and different methods have been used in predicting Weibull parameters. A successful diameter-distribution model requires good prediction of its parameters. However, there is no strong rationale to prove one method better than another method. In this study, new methods were developed by (1) extending the CDF regression technique to various parameter recovery methods, and (2) modifying the CDFR approach such that the CDF is based on information from diameter classes rather than from individual diameters. All methods were ranked based on their goodness-of-fit statistics to determine the optimum method and optimum approach for predicting Weibull parameters.

The Modified CDFR approach consistently provided better results than did the SUR and the CDFR approaches. This approach was superior to the CDFR approach for all evaluation statistics, except for the AD-statistic. The CDFR approach also performed better than the SUR approach for all methods. The results are consistent with the findings by Cao (2004). The poor performance of the SUR approach may be because its objective is to optimize the Weibull parameters, diameter percentiles, or moments instead of optimizing the distribution itself as in the CDFR and the Modified CDFR approaches.

When the SUR approach is used, method 2 ($\overline{D}$ and D_{var}) produced the lowest values for three goodness-of-fit statistics (AD, mlogL, and EI) and ranked the best among the ten methods evaluated. Method 6 (D_{25}, D_{50}, and D_{95}) for recovering Weibull parameters, produced the best

result in terms of KS statistics and similar results as other methods for other evaluation statistics was ranked the second to method 2. Method 5 (D_{50} and D_{95}) was the poorest performer (produced highest values for goodness-of-fit statistics) in the SUR approach.

Method 10 (D_q, D_{25}, D_{50}, and D_{95}) was the best method in the CDFR approach. However, this method was unable to produce the best results for all goodness-of-fit statistics. It was ranked 1st for AD and KS statistics 3rd for mLogL, and 5th for the error index. Method 7 (D_{31}, D_{50}, and D_{63}) was the poorest performer in this approach.

Method 10 (Bailey et al. 1989) provided the best result for every evaluation statistics based on the MCDFR approach and ranked the overall best among the methods evaluated. Method 5 (D_{50}, and D_{95}) was the poorest performer in the Modified CDFR approach. For any given method, its overall ranking from the Modified CDFR approach was better than that from the SUR and CDFR approaches.

If a choice is to be made among these three approaches, we recommend the Modified CDFR approach over the SUR and CDFR approaches, based on the findings of this study. On the other hand, given a particular approach, method 2 is recommended for SUR, and method 10 is recommended for the CDFR, and MCDFR approaches.

It should be mentioned that the inclusion of multiple diameter percentiles in methods 6, 7, and 10 might, for some data sets, result in illogical crossing over of predicted values of these percentiles (i.e. $\widehat{D}_{25} > \widehat{D}_{50}$ or $\widehat{D}_{50} > \widehat{D}_{95}$). In that case, constraints would have to be placed to ensure logical prediction. We also found that use of quadratic mean diameter for recovering Weibull parameters to be significant because every method that included Dq performed better than other methods for our data set. Results from this study might not be representative for all

data sets. There are sufficient reasons to believe that the SUR approach is not as effective as the other two approaches. However, the performance difference between the CDFR and the Modified CDFR might vary with data sets. Future research using other loblolly pine data sets as well as data from other species would further verify the findings of this study.

LITERATURE CITED

Adame, P., J. Hynynen, I. Canellas, and M. D. Rio. 2008. Individual-tree diameter growth model for rebollo oak (*Quercus pyrenaica* Willd.) coppices. For. Ecol. Manag. 255:1011-1022.

Alder, D. 1979. A distance-independent tree model for exotic conifer plantations in East Africa. For. Sci. 25: 59-71.

Andrasev, S., M. Bobinac, and S. Orlovic. 2009. Diameter structure models of Black Poplar selected clones in the section Aigeiros (Duby) obtained by the Weibull distribution. Sumarski List 133: 589-603.

Bailey, R.L. 1980. Individual tree growth derived from diameter distribution models. For. Sci. 26: 626-632.

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

Bailey, R.L., T.M. Burgan, and E. J. Jokela. 1989. Fertilized mid-rotation-aged slash pine plantations—Stand structure and yield prediction models. South. J. Appl. For. 13:76-80.

Baldwin, V.C. Jr., and D.P. Feduccia. 1987. Loblolly pine growth and yield prediction for managed West Gulf plantations. USDA For. Serv. Res. Pap. SO-236, 27p.

Beck, D.E., and L. Della-Bianca. 1972. Growth and yield of thinned yellow-poplar. USDA For. Serv. Res. Pap. SE-101, 20p.

Bennett, F.A., and J.L. Clutter. 1968. Multiple-product yield estimates for unthinned slash pine plantations - pulpwood, saw timber, gum. USDA For. Serv. Res. Pap. SE-35, 21 p.

Bennett, F.A., F.T. Lloyd, B.F. Swindel, and E.W. Whitehorne. 1978. Yields of veneer and associated products from unthinned, old-field plantations of slash pine in the north Florida and South Georgia flatwoods. USDA For. Serv. Res. Pap. SE-176. 80 p.

Borders, B.E., and R.L. Bailey. 1986. A compatible system of growth and yield equations for slash pine fitted with restricted three-stage least squares. For. Sci. 32: 185-201.

Borders, B.E., and W.D. Patterson. 1990. Projecting stand tables: A comparison of the Weibull diameter distribution method, a percentile-based projection method, and a basal area growth projection method. For. Sci. 36: 413-424.

Borders, B.E., R.A. Souter, R.L. Bailey, and K.D. Ware. 1987. Percentile-based distributions characterize forest stand tables. For. Sci. 33: 570-576.

Bowling, E.H., H.E. Burkhart, T.E. Burk, and D.E. Beck. 1989. A stand-level multispecies growth model for Appalachian hardwoods. Can. J. For. Res. 19: 405-412.

Brooks, J.R., B.E. Borders, and R.L. Bailey. 1992. Predicting diameter distributions for site-prepared loblolly and slash pine plantations. South. J. Appl. For. 16: 130-133.

Buckman, R. E. 1962. Growth and yield of red pine in Minnesota. U.S. Department of Agriculture, Technical Bulletin 1272. Washington, DC. 50 p.

Bullock, B. and H. Burkhart. 2005. Juvenile diameter distributions of loblolly pine characterized by the two-parameter Weibull function. New Forests 29: 233-244.

Burk, T.E., and J.D. Newberry. 1984. A simple algorithm for moment-based recovery of Weibull distribution parameters. For. Sci. 30: 329-332.

Burkhart, H.E., and M.R Strub. 1974. A model for simulation of planted loblolly pine stands. In: J. Fries (Editor), Growth Models for Tree and Stand Simulation. Royal College of Forestry, Stockholm, pp. 128-135.

Burkhart, H.E., and P.T. Sprinz. 1984. Compatible cubic volume and basal area projection equations for thinned old-field loblolly pine plantations. For. Sci. 30: 86-93.

Burkhart, H.E., K.D. Farrar, and R.L. Amateis. 1987. Simulation of individual tree growth and stand development in loblolly pine plantations on cutover, site-prepared areas. Sch. For. Wildl. Res; Va. Polytech Int. & State Univ., Pub. No. FWS-1-87-43 p.

Burkhart, H.E., R.C. Parker, and R.G. Oderwald. 1972. Yields for natural stands of loblolly pine. Div. For. Wildl. Resour. Va. Polytech. Inst. State Univ. Publ. FWS-2-72, 63 p.

Cao, Q.V. 1997. A method to distribute mortality in diameter distribution models. For. Sci. 43: 435-442.

Cao, Q.V. 2000. Prediction of annual diameter growth and survival for individual trees from periodic measurements. For. Sci. 46: 127-131.

Cao, Q.V. 2004. Predicting parameters of a Weibull function for modeling diameter distribution. For. Sci. 50: 682-685.

Cao, Q.V. 2006. Predictions of individual-tree and whole-stand attributes for loblolly pine plantations. For. Ecol. Manag. 236: 342-347.

Cao, Q.V. 2007. Incorporating whole-stand and individual-tree models in a stand table projection system. For. Sci. 53: 45-49.

Cao, Q.V., and H. E. Burkhart. 1984. A segmented distribution approach for modeling diameter frequency data. For. Sci. 30:129-137.

Cao, Q.V., and M. Strub. 2008. Evaluation of four methods to estimate parameters of an annual tree survival and diameter growth model. For. Sci. 54: 617-624.

Cao, Q.V., and S. M. McCarty. 2006. New methods for estimating parameters of Weibull functions to characterize future diameter distributions in forest stands. P. 338-340 in Proc. of the South. Silvic. Res. Conf. USDA For. Serv. Gen. Tech. Rep. SRS-92.

Cao, Q.V., and V.C. Baldwin, Jr. 1999. A new algorithm for stand table projection models. For. Sci. 45: 506-511.

Cao, Q.V., H.E. Burkhart, and R.C. Lemin. 1982. Diameter distributions and yields of thinned loblolly pine plantations. Research Publication FWS-1-82. Blacksburg, VA: Virginia polytechnic Institute and State University, School of Forestry and Wildlife. 62 p.

Cirillo, P., and J. Hüsler. 2009. On the upper tail of Italian firms' size distribution. Physica A: Statistical Mechanics and its Applications. 388: 1546-1554.

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

Clutter, J.L. and F. A. Bennett. 1965. Diameter distributions in old-field slash pine plantations. Ga. For. Res. Counc. Rep. 13, 9 p.

Clutter, J.L., and E.P. Jones, Jr. 1980. Prediction of growth after thinning of old-field slash pine plantations. USDA For. Serv. Res. Pap. SE-217, 14 p.

Clutter, J.L., W.R. Harms, G.H. Brister, and J.W. Rhenney. 1984. Stand structure and yields of site-prepared loblolly pine plantations in the lower coastal plain of the Carolinas, Georgia, and North Florida. USDA For. Serv. Gen. Tech. Rep. SE-27, 173 p.

Coble, D. W., and Y. J. Lee. 2008. A new diameter distribution model for unmanaged slash pine plantations in East Texas. South. J. Appl. For. 32: 89-94.

Coble, D.W. 2009. A new whole-stand model for unmanaged loblolly and slash pine plantations in East Texas. South. J. Appl. For. 33: 69-76.

Corral-Rivas, J., S. S. Orois, H. Kotze and K. van Gadow. 2009. Testing the suitability of the Nepal-Somers stand table projection method for *Eucalyptus grandis* plantations in South Africa. Southern Forests. 71: 207-214.

Cunnane, C., 1978. Unbiased plotting-positions – A review: Journal of Hydrology, v. 37: 205-222.

Daniels, R.F., and H.E. Burkhart. 1975. Simulation of individual tree growth and stand development in managed loblolly pine plantations. Div. For. Wildl. Resour. Va. Polytech. Inst. State Univ. Publ. FWS-5-75, 69 p.

Daniels, R.F., and H.E. Burkhart. 1988. An integrated system of forest stand models. For. Ecol. Manag. 23: 159-177.

Daniels, R.F., H.E. Burkhart, G.D. Spittle, and Somers, G.L. 1979. Methods for modeling individual tree growth and stand development in seeded loblolly pine stands. Sch. For. Wild. Resour. Va. Polytech. Inst. State Univ. Publ. Fws-1-79, 50 p.

Ek, A.R. 1974. Nonlinear models for stand table projection in northern hardwood stands. Can. J. For. Res. 4:23-27.

Ek, A.R., J.N. Issos, and R.L. Bailey. 1975. Solving for Weibull diameter distribution parameters to obtain specified mean diameters. For. Sci. 21: 290-292.

Faber, P.J. 1991. A distance-dependent model of tree growth. For. Ecol. Manage. 41: 111 -123.

Feduccia, D.P., T.R. Dell, W.F. Mann, Jr., and B.H. Polmer. 1979. Yields of unthinned loblolly pine plantations on cutover sites in the West Gulf region. USDA For. Serv. Res. Pap. So-148, 88 p.

Fonseca, T.F., C.P. Marques, and B.R. Parresol. 2009. Describing maritime pine diameter distributions with Johnson's SB distribution using a new all-parameter recovery approach. For. Sci. 55: 367-373.

Fox, J. C., H.Q. Bi, and P.K. Ades. 2007(a). Spatial dependence and individual-tree growth models I. Characterising spatial dependence. For. Ecol. Manag.245: 10-19.

Fox, J. C., H.Q. Bi, and P.K. Ades. 2007(b). Spatial dependence and individual-tree growth models II. Modelling spatial dependence. For. Ecol. Manag.245: 20-30.

Gorgoso-Varela, J.J., A. Rojo-Alboreca, E. Afif-Khouri, and M. Barrio-Anta. 2008. Modeling diameter distributions of birch (*Betula alba* L.) and pedunculate oak (*Quercus robur* L.) stands in northwest Spain with the beta distribution. Invest. Agrar. Sist. Recur. For. 17: 271-281.

Gorgoso-Varela, J.J., J.G. Álvarez-González, A. Rojo, and J.A. Grandas-Arias. 2007. Modelling diameter distributions of *Betula alba* L. stands in northwest Spain with the two-parameter Weibull function. Invest. Agrar. Sist. Recur. For. 16: 113-123.

Hafley, W. L., W. D. Smith, and M. A. Buford. 1982. A new yield prediction model for unthinned loblolly pine plantations. South. For. Res. Center, School of For. Res., N.C. State Univ. Tech. Rep.1. 65 p.

Hafley, W.L., and H.T. Schreuder. 1977. Statistical distributions for fitting diameter and height data in even-aged stands. Can. J. For. Res. 7: 481-487.

Hyink, D.M., and J.W. Moser, Jr. 1983. A generalized framework for projecting forest yield and stand structure using diameter distributions. For. Sci. 29: 85-95.

Hynynen, J. 1995. Predicting the growth response to thinning for Scots pine stands using individual-tree growth models. Silva Fennica 29: 225-246.

Jiang, L.C., and J.R. Brooks. 2009. Predicting diameter distributions for young longleaf pine plantations in Southwest Georgia. South. J. Appl. For. 33: 25-28.

Johnson N.L. (1949a). Systems of frequency curves generated by methods of translation. Biometrika 36:149-176.

Johnson N.L. (1949b). Bivariate distributions based on simple translation systems. Biometrika 36:297-304.

Karlsson, K., and L. Norell. 2005. Predicting the future diameter of stems in Norway spruce stands subjected to different thinning regimes. Can. J. For. Res. 35: 1331-1341.

Knoebel, B. R., and H. E. Burkhart. 1991. A bivariate distribution approach to modeling forest diameter distribution at two points in time. Biometrics 47: 241-253.

Knoebel, B.R., H.E. Burkhart, and D.E. Beck. 1986. A growth and yield of thinned yellow-poplar. For. Sci. Monograph 27, 41 p.

Knowe, S.A., S.R. Radosevich, and R.G. Shula. 2005. Basal area and diameter distribution prediction equations for young Douglas-Fir plantations with hardwood competition: Coast ranges. West. J. Appl. For. 20: 77-93.

Lee, Y.J., and D.W. Coble. 2006. A new diameter distribution model for unmanaged loblolly pine plantations in East Texas. South. J. Appl. For. 30: 13-20.

Lenhart, J.D. 1972. Cubic volume yields for unthinned old-field loblolly pine plantations in the Interior West Gulf Coastal Plain. Stephen F. Austin State Univ. Tex. For. Pap. 14, 46 p.

Lenhart, J.D. 1996. Total and partial stand-level yield prediction for loblolly and slash pine plantations in east Texas. South. J. Appl. For. 20: 36-41.

Lenhart, J.D., and J.L. Clutter. 1971. Cubic-foot yield tables for old-field loblolly pine plantations in the Georgia Piedmont. Ga. For. Res. Counc. Rep. 22 - Set. 3, 12 p.

Little S.N. 1983. Weibull diameter distribution for mixed stands of western conifers. Can. J. For. Res. 13:85-88.

Lohrey, R.E., and R.L. Bailey. 1977. Yield tables and stand structure for unthinned long leaf pine plantations in Louisiana and Texas. USDA For. Ser. Res. Pap. SO-133. 55p.

Lynch, T.B., and Moser, J.W., Jr. 1986. A growth model for mixed species stands. For. Sci. 32: 697-706.

Mabvurira, D., and J. Miina. 2002. Individual-tree growth and mortality models for *Eucalyptus grandis* (Hill) Maiden plantations in Zimbabwe. For. Ecol. Manag. 161: 231-245.

Mabvurira, D., M. Maltamo, and A. Kangas. 2002. Predicting and calibrating diameter distributions of Eucalyptus grandis (Hill) Maiden plantations in Zimbabwe. New Forests 23: 207-223.

MacKinney, A. L., F. Schumacher, and L.E. Chaiken. 1937. Construction of yield tables for nonnormal loblolly pine stands. J. Agric. Res. 54:531-545.

MacKinney, A.L., and L.E. Chaiken. 1939. Volume, yield and growth of loblolly pine in the mid-atlantic region. U.S. For. Serv., Appalachian For. Exp. Sta. Tech. Note 33. 30 p.

Matney, T.G. and E.B. Schultz. 2008. Deriving tree diameter growth and probability of survival equations from successive diameter distributions. For. Sci. 54: 31-35.

Matney, T.G., and A.D. Sullivan. 1982. Compatible stand and stock tables for thinned and unthinned loblolly pine stands. For. Sci. 28: 161-171.

McTague, J. P. and R. L. Bailey. 1987. Compatible basal area and diameter distribution models for thinned loblolly-pine plantations in Santa-Catarina, Brazil. For. Sci. 33: 43-51.

Miina, J. 1994. Spatial growth model for Scots pine on drained peatland. Silva Fennica 28: 15-27.

Mitchell, K.J. 1975. Dynamics and simulated yield of Douglas-fir. For. Sci. Monograph 17, 38 p.

Murphy, P.A. 1983. A nonlinear timber yield equation system for loblolly pine. For. Sci. 29: 582-591.

Nepal, S.K., and G.L. Somers. 1992. A generalized approach to stand table projection. For. Sci. 38: 120-133.

Newberry, J.D., and T.E. Burk. 1985. S_B distribution-based models for individual tree merchantable volume-total volume ratios. For. Sci. 31: 389-398.

Newton, P.F., and I.G. Amponsah. 2005. Evaluation of Weibull-based parameter prediction equation systems for black spruce and jack pine stand types within the context of developing structural stand density management diagrams. Can. J. For. Res. 35: 2996-3010.

Newton, P.F., Y. Lei, and S.Y. Zhang. 2005. Stand-level diameter distribution yield model for black spruce plantations. For. Ecol. Manag. 209: 181-192.

Nord-Larsen, T. 2006. Modeling individual-tree growth from data with highly irregular measurement intervals. For. Sci. 52: 198-208.

Nord-Larsen, T., and Q. V. Cao. 2006. A diameter distribution model for even-aged beech in Denmark. For. Ecol. Manag. 231: 218-225.

Ochi, N., and Q.V. Cao. 2003. A comparison of compatible annual growth models. For. Sci. 49: 285-290.

Palahi, M., T. Pukkala, and A. Trasobares. 2006. Modelling the diameter distribution of *Pinus sylvestris*, *Pinus nigra* and *Pinus halepensis* forest stands in Catalonia using the truncated Weibull function. Forestry. 79**:** 553-562.

Palahi, M., T. Pukkala, J. Miina, and G. Montero. 2003. Individual-tree growth and mortality models for Scots pine (*Pinus sylvestris* L.) in north-east Spain. Ann. For. Sci. 60: 1-10.

Parresol, B.R. 2003. Recovering parameters of Johnson's S_B distribution. Res. Pap. SRS-31. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 9 p.

Pienaar, L.V., and J.W. Rheney. 1993. Yield prediction for mechanically site-prepared slash pine plantations in the southeastern coastal plain. South. J. Appl. For. 17: 163-173.

Pienaar, L.V., and W.M. Harrison. 1988. A stand table projection approach to yield prediction in unthinned even-aged stands. For. Sci. 34: 804-808.

Pienaar, L.V., and W.M. Harrison. 1989. Simultaneous growth and yield prediction equations for *Pinus elliottii* plantations in Zululand. So. Afr. For. J. 149: 48-53.

Qin, J., and Q.V. Cao. 2006. Using disaggregation to link individual-tree and whole-stand growth models. Can. J. For. Res. 36: 953-96.

Qin, J., Q.V. Cao, and D.C. Blouin. 2007. Projection of a diameter distribution through time. Can. J. For. Res. 37: 188-194.

Rautiainen, O. 1999. Spatial yield model for *Shorea robusta* in Nepal. For. Ecol. Manag. 119: 151-162.

Rennolls, K., and M. Wang. 2005. A new parameterization of Johnson's S_B distribution with application to fitting forest tree diameter data. Can. J. For. Res. 35: 575-579.

Ritchie, M.W., and D.W. Hann. 1997. Implications of disaggregation in forest growth and yield modeling. For. Sci. 43: 223-233.

Sanchez-Gonzalez, M., M.D. Rio, I. Canellas, and G. Montero. 2006. Distance independent tree diameter growth model for cork oak stands. For. Ecol. Manag. 225: 262-270.

SAS Institute Inc. 1993. SAS/ETS® user's guide. Version 6. 2nd ed. SAS Institute Inc., Cary, N.C.

Schreuder, H. T., W. L. Hafley, and F. A. Bennett. 1979. Yield prediction for unthinned natural slash pine stands. For. Sci. 25: 25-30.

Schreuder, H.T., and W. L. Hafley. 1977. A useful bivariate distribution for describing stand structure of tree heights and diameters. Biometrics. 33: 471-478.

Scolforo, J.R.S., F.C.V. Tabai, R.L.G. de Macedo, F.W. Acerbi, Jr., and A.L. de Assis. 2003. SB distribution's accuracy to represent the diameter distribution of pinus taeda, through five fitting methods. For. Ecol. Manag. 157: 489-496.

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

Siipilehto, J. 1999. Improving the accuracy of predicted basal area diameter distribution in advanced stands by determining stem number. Silva Fenn. 33: 281-301.

Smalley, G.W., and R.L. Bailey. 1974. Yield tables and stand structure for loblolly pine plantations in Tennessee, Alabama, and Georgia highlands. USDA For. Serv. Res. Pap. So-96, 81p.

Smith, W.D., and W.L. Hafley. 1984. Multiproduct yield tables for single-thinned loblolly pine plantations. South. For. Res. Ctr. 77 p.

Somers, G.L., and S.K. Nepal. 1994. Linking individual-tree and stand-level growth models. For. Ecol. Manag. 69: 233-243.

Stage, A.R. 1973. Prognosis model for stand development. USDA For. Serv. Res. Pap. INT-137, 32 p.

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

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

Van Deusen, P.C. 1988. Simultaneous estimation with a squared error loss function. Can. J. For. Res. 18: 1093-1096.

Vanclay, J.K., 1991. Modelling the growth and yield of tropical moist forests. Queens. For. Ser. Brisbane, 151 p.

Vettenranta, J. 1999. Distance-dependent models for predicting the development of mixed coniferous forests in Finland. Silva Fennica 33: 51-72.

Wells, O.O, and P.C. Wakeley. 1966. Geographic variation in survival, growth, and fusiform-rust infection of planted loblolly pine. For. Sci. Monogr. 11. 40 p.

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

Yue, C., U. Kohnle, and S. Hein. 2008. Combining tree- and stand-level growth models: a new approach to growth prediction. For. Sci. 54: 553-566.

Zarnoch, S.J., D.P. Feduccia, V.C. Baldwin, Jr., and T.R. Dell. 1991. Growth and yield predictions for thinned and unthinned slash pine plantations on cutover sites in the West Gulf region. USDA For. Serv. Res. Pap. SO-264. 32p.

Zhang, L., J.A. Moore, and J.D. Newberry. 1993. Disaggregating stand volume growth to individual trees. For. Sci. 39: 295-308.

Zhang, S., R.L. Amateis, and H.E. Burkhart. 1997. Constraining individual tree diameter increment and survival models for loblolly pine plantations. For. Sci. 43: 414-423.

Zhou, B.L. and J.P. McTague. 1996. Comparison and evaluation of five methods of estimation of the Johnson system parameters. Can. J. For. Res. 26: 928-935.

APPENDIX: EXAMPLE SAS PROGRAM

```
data one;
input plot age n hd tph d;

            *age = stand age,
            *n = number of trees per plot;
            *hd = dominant height of the stand;
            *tph = number of trees per hectare;
            *d = tree diameter at breast height;

  hdlog = log(hd);
  tphlog = log(tph);
  rs = (sqrt(10000/tph))/hd;
                                          *rs = relative spacing;

  a = 0.5*(exp(7.36711 - 2.73286*rs - 0.57797*tphlog - 7.41606/age));

*Computation of CDF in the Modified CDFR approach;

  if first.plot
  then do;
    cdf = nd/n;                            *nd = number of trees in a
                                           *diameter class;
      nclass = 1;
      d = dclass + 1;
      output;
      end;
  else do;
      do while (dclass > dclass1);       *Condition for empty diameter class;
        d = dclass1 + 1;
        nclass = nclass + 1;
        output;
        dclass1 = dclass1 + 2;
        end;
    cdf = cdf + nd/n;
    nclass = nclass + 1;
    d = dclass + 1;
    output;
    end;
dclass1 = dclass + 2;
retain cdf nclass dclass1;

*SAS Procedure Model for fitting CDF - Method 2;

proc model itprint;

  parms b11=5.224 b12=-1.548 b13=-0.384 b14=0.251
  b21=-5.911 b23=0.01 b24=2.545;

  pdq = exp(b11 + b12*rs + b13*tphlog + b14*hdlog);
  pdvar = exp(b21 + b23*tphlog + b24*hdlog);

  c = 1.1;                                *Starting value for c;
  c1 = c;
  g1 = gamma(1 + 1/c);
  g2 = gamma(1 + 2/c);
  b = ((-a*g1)/g2) + (((((a/g2)**2)*(g2*g2-g1) + (pdq*pdq/g2))**0.5);
  fOld = b*b * (g2 - g1**2) - pdvar;
  c = c1 + 1;
```

```
   g1 = gamma(1 + 1/c);
   g2 = gamma(1 + 2/c);
   b = ((-a*g1)/g2) + ((((a/g2)**2)*(g2*g2-g1) + (pdq*pdq/g2))**0.5);
   fNew = b*b * (g2 - g1**2) - pdvar;
   do j = 1 to 20 while (fOld*fNew > 0);
        c1 = c;
        c = c1 + 1;
        g1 = gamma(1 + 1/c);
      g2 = gamma(1 + 2/c);
        b = ((-a*g1)/g2) + ((((a/g2)**2)*(g2*g2-g1) + (pdq*pdq/g2))**0.5);
        fNew = b*b * (g2 - g1**2) - pdvar;
        end;                                   *At this step, fOld and fNew;
                                               *are of opposite signs;
                                               *solution is in the interval (c1, c);
   inc = c - c1;
   do while (abs(fNew) > 1e-8);                *Secant method to solve fNew = 0;
                                               *1e-8 = 0.00000001 is the tolerance;
     inc = -fNew * inc / (fNew - fOld);
        c = c + inc;
        fOld = fNew;
        g1 = gamma(1 + 1/c);
      g2 = gamma(1 + 2/c);
        b = ((-a*g1)/g2) + ((((a/g2)**2)*(g2*g2-g1) + (pdq*pdq/g2))**0.5);
        fNew =b*b * (g2 - g1**2) - pdvar;
        end;

   if d > a
     then fhat = 1 - exp(-(((d-a)/b)**c));
        else fhat = 0;
   cdf = fhat;
   _weight_ = 1/nclass;
   fit cdf;
   title 'Modified CDFR Approach - Method 2';


*SAS Procedure Model for fitting CDF – Method 10;

proc model itprint;
  parms b11=6.484 b12=-2.356 b15=0.1
        b31=7.026 b32=-2.762 b34=0.192 b35=0.464
        b41=7.807 b42=-2.615 b43=-0.583 b44=0.251;

  d0 = exp(8.828995 - 4.91421*rs - 0.7847*tphlog);
  d50 = exp(7.017923 - 2.7597*rs - 0.51166*tphlog);
  d25 = exp(b11 + b12*rs + b15/age);
  d95 = exp(b31 + b32*rs + b34*hdlog + b35/age);
  dq = exp(b41 + b42*rs + b43*tphlog + b44*hdlog);

  a = (((n**(1/3))*d0) - d50)/((n**(1/3) - 1));
  c = (log(log(0.05)/log(0.75)) / (log(d95-a) - log(d25-a)));
  g1 = gamma(1 + 1/c);
  g2 = gamma(1 + 2/c);
  b = ((-a*g1)/g2) + ((((a/g2)**2)*(g2*g2-g1) + (dq*dq/g2))**0.5);

  if d > a
    then cdf = 1 - exp(-(((d-a)/b)**c));
    else cdf = 0;
  fit cdf;
  _weight_ = 1/nclass;

  title 'Modified CDFR Approach - Method 10';
run;
quit;
```

VITA

Krishna Prasad Poudel was born on July 20, 1983, in Lekhnath-15 Nepal. He attended the Institute of Forestry, Pokhara campus, Pokhara, Nepal, in 2003 and received a Bachelor of Science degree in forestry in 2007. In August 2009, he enrolled in the Graduate School of the Louisiana State University and Agricultural and Mechanical College in Baton Rouge, Louisiana. He is currently a candidate for the degree of Master of Science in forestry, which will be awarded in August 2011.