

PREDICTION OF DISTRIBUTION FOR TOTAL HEIGHT AND CROWN RATIO  
USING NORMAL VERSUS OTHER DISTRIBUTIONS

Except where reference is made to the work of others, the work described in this thesis is my own or was done in collaboration with my advisory committee. This thesis does not include proprietary or classified information.

---

Tanka Prasad Acharya

Certificate of Approval:

---

Ralph S. Meldahl  
Associate Professor (Retired)  
Forestry and Wildlife Sciences

---

Greg L. Somers, Chair  
Associate Dean of Education  
Forestry and Wildlife Sciences

---

Asheber Abebe  
Assistant Professor  
Mathematics and Statistics

---

Joe F. Pittman  
Interim Dean  
Graduate School

PREDICTION OF DISTRIBUTION FOR TOTAL HEIGHT AND CROWN RATIO  
USING NORMAL VERSUS OTHER DISTRIBUTIONS

Tanka Prasad Acharya

A Thesis

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Masters of Science

Auburn, Alabama  
December 15, 2006

PREDICTION OF DISTRIBUTION FOR TOTAL HEIGHT AND CROWN RATIO  
USING NORMAL VERSUS OTHER DISTRIBUTIONS

Tanka Prasad Acharya

Permission is granted to Auburn University to make copies of this thesis at its discretion, upon request of individuals or institutions and at their expense. The author reserves all publication rights.

---

Signature of Author

---

Date of Graduation

## VITA

Tanka Prasad Acharya , son of Krishna Prasad Acharya and Devi Acharya, was born on May 27, 1958, in Sunsari, Nepal. He completed high school from Vishnu Madhyamic Vidhyalaya, Sonapur, Sunsari, Nepal in 1975. He received a Bachelor of Science degree in Zoology from Tribhuvan University in 1980. He worked as a mathematics teacher in a high school from 1980 to 1984. He received a second Bachelor of Science degree in Forestry from Tribhuvan University in 1988. He attained Yale University and received a Master of Forest Science degree in 1995. He is a faculty member of Tribhuvan University at the Institute of Forestry since 1988. He entered Auburn University in August, 2002 and received another Masters of Science degree in Probability and Statistics in August, 2004.

## THESIS ABSTRACT

### PREDICTION OF DISTRIBUTION FOR TOTAL HEIGHT AND CROWN RATIO USING NORMAL VERSUS OTHER DISTRIBUTIONS

Tanka Prasad Acharya

Master of Science, December 15, 2006  
(M.S., Auburn University, 2004)  
(M.F.S., Yale University, 1995)

Total 106 Typed Pages

Directed by Greg L. Somers

Relative to the other southern pine species, many aspects of longleaf pine (*Pinus palustris* Mill.) growth are poorly understood due to the lack of detailed studies and modeling efforts. With changes occurring in the ecosystem and even the climate, there is a need for flexible models. One such area is the crown shape and leaf area.

This thesis is part of a larger research project based at Eglin Air Force Base in Florida entitled “crown shape and leaf area distribution models for naturally regenerated longleaf pine trees at Eglin Air Force Base in Florida”. Crown shape and leaf area distribution models can also be used as an initialization model to develop biologically-based, resource driven mixed models (individual tree model between empirical growth and yield models and process models) to get more precise growth estimates of longleaf. This thesis includes three chapters. Chapter 1 explains about the larger research project at Eglin including the sections of in-depth literature review for chapters 2 and 3.

Chapter 2 explores the models which can predict the entire distribution of total height rather than just the mean total height ( $H$ ) given diameter at the breast height ( $DBH$ ). Three different models with a normal error term were compared. In the first model (Model A) total height was used as a function of  $DBH$  with a constant variance. The weighted least square approach was applied to model the distribution of  $H$  in the second model to correct the problem of increasing variance. The third model (Model C) used  $H$  as a function of  $DBH$  and variance was also used as a function of  $DBH$ . All three models have some probability of predicting negative total height which is not realistic, however, the probability of predicting negative height from the third model was very low. Finally, a gamma model was used to predict the distribution of  $H$ , this model never predicts negative height. Cross validation indicated that the gamma model was the best model to predict the distribution of total height among all the models compared in this study.

Chapter 3 explores the models which can predict the entire distribution of crown ratio rather than just the mean crown ratio ( $CR$ ) given  $H$ . Similar model forms using the assumption that the errors are normally distributed versus beta distributed errors were compared. Two different model forms were used in both the cases. The first model had mean crown ratio as a logistic function of  $H$  with a constant variance, whereas, the second model had a mean crown ratio as a logistic function of  $H$  and variance also as a function of  $H$ . The second model form with beta distributed error was found to be the best model among the models used in the study. Cross validation indicated that the beta model is capable of predicting the distribution of  $CR$  correctly. This model never suffers from the problem of predicting the  $CR$  beyond the logical range of 0 and 1.

## ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my major professor, Dr. Greg L. Somers. This thesis would have been impossible without his persistent guidance, encouragement, and patience. I am grateful to my advisory committee members, Dr. Asheber Abebe for his relentless support, encouragement, and guidance. I would also like to express my appreciation to my advisory committee member Dr. Ralph S Meldahl for his encouragement and support. Thanks to Dr. John Kush for his help in all the aspects related to the completion of this degree at Auburn University.

I would like to express my heartfelt gratitude to my mother, father and sisters for their continuous encouragement throughout my education. Finally, I would like to express my thanks to my wife Renuka, and children Richa and Rebika, for their constant encouragement from Nepal to accomplish my career objectives during difficult times.

Style of manual or journal used: Ecological Modeling

Computer software used:

SAS 9.1

Microsoft Excel

Microsoft Word



## TABLE OF CONTENTS

LIST OF TABLES .....	xi
LIST OF FIGURES .....	xii
CHAPTER 1 OVERALL INTRODUCTION .....	1
1.1 Introduction.....	1
1.2 Methods.....	10
1.3 Flow Chart .....	16
1.4 Literature Review for Height – Diameter Model.....	17
1.5 Literature Review for Crown Ratio Model .....	20
CHAPTER 2 MODEL FOR PREDICTING DISTRIBUTION OF TOTAL HEIGHT LONGLY PINE AT EGLIN AIR FORCE BASE IN FLORIDA.....	31
2.1 Introduction .....	31
2.2 Materials and Methods.....	33
2.2.1 Data and Study Area .....	33
2.2.2 Methods.....	34
2.3 Results.....	38
2.3.1 Models That Require The Assumption Of Normality Of The Random Error .....	39
2.3.2 Models That Do Not Require The Assumption Of Normality Of The Random Error.....	50

2.4 Discussion and Conclusion.....	56
2.5 Literature Cited.....	61
CHAPTER 3 PREDICTION OF DISTRIBUTION FOR CROWN RATIO OF AN	
INDIVIDUAL TREE USING NORMAL VERSUS BETA	
DISTRIBUTION MODEL FOR LONGLEAF PINE AT EGLIN	
AIR FORCE BASE.....	
	65
3.1 Introduction .....	65
3.2 Materials and Methods.....	67
3.2.1 Study Area and Sampling Of Trees .....	67
3.2.2 Methods.....	69
3.3 Results.....	72
3.3.1 Logistic Function With Normal Distribution Of	
Errors.....	72
3.3.1.1 Normal Model A.....	72
3.3.1.2 Normal Model B.....	76
3.3.2 Logistic Function With Beta Distribution Of Errors .....	79
3.3.2.1 Beta Model A.....	81
3.3.2.2 Beta Model B.....	83
3.4 Discussion and Conclusion.....	86
3.5 Literature Cited.....	89

## LIST OF TABLES

2.	1. Model Parameter Estimates .....	41
	2. Model Use to Fit the Data.....	56
3.	1. Model Parameter Estimates .....	75
	2. Model Used to Fit the Data.....	86

## LIST OF FIGURES

2.	1. Plot of Total Height Vs. Diameter at Breast Height .....36	36
	2. Plot of Standardized residual Vs. Predicted Total Height - Model A.....40	40
	3. Plot of Observed Total Height, Predicted Total Height, and Prediction Interval for Total Height against Diameter at Breast Height - Model A .....40	40
	4. Distribution of Predicted Total Height (Diameter at Breast Height = 4, 10, and 40 cms) - Model A .....41	41
	5. Plot of Standardized Residual Vs. Predicted Height - Model B Weighted Least Square .....44	44
	6. Plot of Observed Total Height, Predicted Total Height, and Prediction Interval for Total Height Vs. Diameter at Breast Height - Model B .....45	45
	7. Distribution of predicted Total Height (Diameter at Breast Height = 4, 10, and 40 cms) Model B.....46	46
	8. Variance of Total Height Vs. Average Diameter at Breast Height by Diameter at Breast Height class.....47	47
	9. Plot of Observed Total Height, Predicted Total Height, and Prediction for Interval Total Height Vs. Diameter at Breast Height - Model C .....48	48
	10. Plot of Standardized Residual Vs. Predicted Total Height - Model C.....48	48
	11. Distribution of Predicted Total Height (Diameter at Breast Height	

	= 4, 10, and 40 cms) Model C.....	49
12.	Probability of predicted Total Height less than 1.3 m Vs. Diameter at Breast Height (Model A, B, and C) .....	50
13.	Dispersion parameter $\phi$ Vs. Average Diameter at Breast Height .....	52
14.	Plot of standardized deviance residual Vs. Predicted Total Height Model D Gamma.....	54
15.	Observed Total Height, Predicted Total Height, and Prediction Interval for Total Height Vs. Diameter at Breast Height Model D Gamma .....	54
16.	Distribution of Predicted Total Height (Diameter at Breast Height = 4, 10, and 40 cms) Model D Gamma.....	55
17.	Predicted Total Height for Model A, B, C, and D Vs. Diameter at Breast Height.....	57
3.		
1.	Plot of Crown Ratio Vs. Total Height .....	68
2.	Plot of $\ln((1/\mu)-1)$ Vs. Total Height .....	73
3.	Standardized Residual Vs. Predicted Crown Ratio – Normal Model.....	74
4.	Observed Crown Ratio, Predicted Crown Ratio and Prediction Interval for Crown Ratio Vs. Total Height for Normal Model A .....	75
5.	Plot of Variance of Crown Ratio Vs. Total Height – Normal Model B .....	76
6.	Standardized Residual Vs. Predicted Crown Ratio – Normal Model B .....	77
7.	Observed Crown Ratio, Predicted Crown Ratio, and Prediction Interval for Crown Ratio Vs. Total Height - Normal Model B .....	78

8. Distribution of Crown Ratio (Total Height = 2, 5, 10, and 25 m)	
Normal Model B .....	79
9. Probability mass outside 0 and 1 for Normal Model A and B Vs. Total	
Height.....	79
10. Deviance Residual Vs. Total Height Beta Model A .....	82
11. Observed Crown Ratio, Predicted Crown Ratio, and Prediction	
Interval for Crown Ratio Vs. Total Height – Beta Model A .....	83
12. Plot of $\phi$ Vs. Total Height.....	84
13. Deviance Residual Vs. Total Height Beta Model B .....	84
14. Observed Crown Ratio, Predicted Crown Ratio, and Prediction	
Interval for Crown Ratio Vs. Total Height – Beta Model B.....	85
15. Distribution of Crown Ratio (H = 2, 5, 10, 25 m) – Beta model B .....	86
16. Predicted Crown Ratio for Normal Model A, B and	
Beta Model A, B Vs. H.....	87

## **CHAPTER 1**

### **OVERALL INTRODUCTION**

#### **1.1 INTRODUCTION**

##### **Empirical growth models**

Empirical growth models are developed by measuring trees at intervals of a year or more and relationships between change in size and certain other variables are determined (Bruce, 1990). Extensive databases from long-term permanent field plots covering different management regimes and site productivities are required to develop such models. With this approach it is rarely possible to predict the effects of change in the environmental conditions and forest management regime. These models need to be updated with additional data describing the change at the stand level or environmental conditions to be able to predict the future stand growth with better precision. Empirical growth models are criticized because they make no allowance for changes in management and environmental conditions.

Most empirical models are developed to simulate forest growth as a guide to management. The common variables included in the simple models for even-aged stands are species, site index, an index of stand density, and initial tree size usually based on age. The primary criterion for selecting the best prediction model is statistical fit to the data. These types of models need to be updated as management practices change, or if there are

changes in environmental conditions. In these situations the assumptions that the past can predict the future will be violated.

Empirical forest growth models are classified by Munro (1973) as whole-stand and single-tree models. Single-tree models are divided into distance-dependent and distance-independent models. Whole-stand models can be based on untagged re-measured plots. Stand-level models use the stand as the primary growth unit. These models work under the assumptions of similar growth pattern during the past and future (Bruce, 1990). This implies that there will be no change in the environmental conditions and management practices during the time period under consideration. They are not suitable to incorporate the effects of different environmental factors and changing management regimes. This restriction led to the development of more flexible individual tree distance-dependent and distance-independent growth and yield models.

Individual-tree models use a tree as a primary growth unit. These models also work under the assumptions of same growth pattern during the past and future. Distance-dependent models are based on mapped plots and distance-independent models are based on plots with tagged trees (Bruce, 1990). In the distance-dependent model the competition on a subject tree from its surrounding is taken into account to quantify the effect of limited growing space on the growth of the subject tree. This effect is quantified by indices based on the relative size and location of subject tree and competing trees (Daniels and Burkhart, 1975). These indices provide empirical predictions of competition, but they do not explain how different growth processes are affected by limiting resources. It is important to understand the effects of limiting resources on tree growth to predict the effect of alteration in the environmental conditions.



## **Physiologically-based Process models**

A different approach has been taken in the development of physiologically-based process models. In these models, trees are grown as the sum of metabolic processes over relatively short periods of time (hours or days). The modeled processes could include photosynthesis, mineral metabolism, respiration, carbon allocation and assimilation, absorption, and accumulation of water, minerals, and gases, translocation, and growth regulation (Dixon, 1990). These models require shorter duration of time for data collection. These models are independent of current stand conditions and management practices.

According to Lambert-Beer equation physiologically-based process models simulate photosynthetically active radiation absorption (APAR) as functions of the path length of sun rays through the crowns, leaf area encountered and an extinction coefficient that is species specific (Bartelink, 2000). APAR is then converted to net primary productivity (NPP) using some measure of the efficiency of radiation use by the species. NPP is then allocated on an hourly (Bossell, 1996) or daily (Bartelink, 2000) time step. These values are then integrated over the growing season and then allocated to individual tree components of leaf, branch, stem, and root biomass.

Developing models of tree growth as an end result of physiological processes would be is the best approach. However, this approach suffers from some problems. The number of detailed processes that need to be measured on each individual tree, commitment from a number of scientists for the development of sub-models for each process, and the information required to initialize the model severely limits the development of such models. This data intensity makes validation of the entire model very difficult. Lastly, scaling up physiological processes over both time and space represents

another form of extrapolation. Photosynthesis is a short-term event which is highly dependent on local and short-term changes in the environment, such as changes in sun angle and temperature over a day, which cannot themselves be precisely modeled. Integrating these predictions over time and space produce errors and misleading predictions (Johnsen et al., 2001). Most physiological studies are conducted on seedlings or small trees grown in controlled environment. Scaling these up to mature trees growing in a dynamic, competitive environment might lead to severely biased estimates due to the incomplete or incorrect model structures (Johnsen et al., 2001).

### **A Resource-driven Model**

Empirical growth and yield models have been traditionally used for making economic decisions for timber production. Forest inventories could be updated in order to make decisions about forest management operations such as harvesting, thinning, or other silvicultural operations. To get reliable long-term growth predictions suitable for changing environment or management conditions, one alternative would be to opt for a flexible model that combines the statistical advantages of empirical models with the structural advantages of process models. With this in mind the most important components of physiological process models is the interception of solar radiation, which in an empirical sense is related to individual tree crowns rather than the bole of a tree generally modeled in individual tree models.

Interception of solar radiation is one of the most important factors for the growth and development of a tree. It is closely related to shape of the crown and foliage mass of the tree. Foliage mass is an important factor for tree growth which determines the amount of solar radiation intercepted by a tree (Vose and Swank, 1990). Vose and Allen (1988)

have shown a direct relationship between leaf area and stem wood growth. The total amount of solar radiation intercepted by individual trees is used as one of the most important inputs in the physiological process based tree growth models (Mohren, 1987). The availability of total radiation to an individual tree in a forest stand depends primarily on relative size and position of the tree crown in relation to the neighboring trees. Only a small portion of the available radiation is intercepted by an individual tree depending on the shape of the crown and foliage distribution in the crown. Therefore, it is important to model tree crown shape and foliage distribution within a crown to predict the total amount of radiation intercepted and hence the growth of the tree.

### **Crown Shape**

Crown shape of a tree is expected to change from the seedling to mature stages. Mohren (1987) explained the change in the crown shape of a Douglas-fir (*Pseudotsuga menziesii*) trees growing in an even-aged stand. He described crown shape as a function of crown parameters, which change over time. Crown shape also changes due to the competition for resources with the neighbor trees. This change may be restricted to a few sides depending upon the size and distance of the competitor in a given azimuth. The overall change in crown shape from cone to parabola to cylinder as described by Mohren (1987) is not sufficient. The change in shape due to competition from the surrounding trees also needs to be incorporated.

Crown shape can be used to describe the competitive environment of a tree. Crown shape is a difficult attribute to model and is usually defined by crown length and convexity (Grace, 1990). Crown length and crown width have been successfully used as an indicator of individual tree growth (Mitchell, 1975). Terjung and Louie (1972) have indicated that

there is a variation in total radiation available for interception to the crown based on their shapes. They have also demonstrated the variation in availability of radiation to crowns of the same shape and surface area but with different dimensions. An upright cone with a small ratio of height to radius had 10% more radiation available than a sphere, cylinder, or reverse cone with the same surface area. Oker-Blom and Kellomaki (1982) have concluded that a narrow and tall cone-shaped crown maximizes radiation interception. Interception of light by individual trees depends upon the available light reaching the crown, sun angle, shape of the crown, distribution of leaf area within the crown, and the photosynthetic capacity of the foliage.

### **Leaf area distribution**

The quantity and distribution of foliage in a tree canopy are important factors to determine the potential of a tree to utilize solar energy and assimilate carbon through photosynthesis (Grace et al. 1987; Russell et al., 1989). Foliar distribution using leaf area or mass has been modeled using a vertical negative exponential distribution (Kinerson et al., 1987), a lognormal distribution (Schrueder and Swank, 1974), and a Weibull distribution (Schrueder and Swank, 1974, Hagihara and Hozumi, 1986, and Vose, 1988). These distributions predict leaf area as a function of height within the canopy without regard to crown volume in space. Most models assume uniform leaf area distributions throughout the crown (e.g. Mohren, 1987, Bartelink, 2000). Mitchell (1975) has modeled a more spatially explicit leaf area distribution by age with concentric shells within the crown volume. Grace et al. (1987) has assumed a random foliage arrangement within sub-canopies defined by photosynthetic capacity rather than age. The random arrangement

versus a clumped arrangement has been investigated by Norman and Jarvis (1975) and Baldocchi and Hutchison (1986).

Oker-Blom (1986) and Whitehead et al. (1990) used a Poisson approximation of the binomial process to characterize a random dispersion of leaves within sub-canopies. Radiation regime in a Sitka spruce (*Picea sitchensis*) forest canopy was studied in detail by Norman and Jarvis (1975). They measured the radiation intercepted at different levels and positions within a canopy and predicted the interception assuming random and clumped foliage distributions. Predicted radiation values close to the observed radiation values were obtained when foliage was assumed to be in clumps. Baldocchi and Hutchison (1986) also concluded that an assumption of clumped foliage gave a better result in predicting photosynthetically active radiation regime in an oak-hickory forest canopy. An assumption of clumped foliage is probably more appropriate than random foliage for the tree crown of a species like longleaf pine. Although a random or uniform distribution of leaf area within the crown volume is a reasonable simplifying assumption for most trees, the extreme clumping of foliage exhibited by longleaf pine at the tip of branches will require a more sophisticated model for the spatial distribution of leaf area.

The amount of solar radiation that can penetrate the layers of foliage and become available for interception depends upon the foliage distribution within the tree crown (Grace 1990). Foliage distribution has to be defined properly to study the interception of solar radiation by an individual tree crown. A Weibull distribution was used by Hagihara and Hozumi (1986) to model the vertical distribution of leaf area of an individual tree crown. Schreuder and Swank (1974) demonstrated that a Weibull distribution explained vertical crown biomass of white pine (*Pinus strobus* L.) and loblolly pines better than

lognormal and gamma distributions. A Weibull distribution was also used by Vose (1988) to model the vertical distribution of leaf area index in a fertilized loblolly pine stand.

Vertical distribution of the foliage does not provide information about mutual shading of foliage within a tree crown, which can affect total solar radiation intercepted. Branch growth from intercepted solar radiation cannot be predicted from the vertical distribution of foliage. Foliage distribution in three dimensions will be useful for the computation of expected foliage density at different locations in a tree crown. It can also be used to predict expected solar radiation interception at different locations in the tree crown. This information will provide more reliable estimates of solar radiation intercepted by tree crowns. It would also be useful in predicting the effect of competition and management of tree crowns by pruning, thinning, and any environmental factors affecting needle expansion and retention (Nepal, 1993).

In closed-canopy plantations of loblolly pine, branch foliage biomass was largely correlated with two factors: branch diameter and relative height (Nepal, 1993). A similar positive relationship between branch foliage biomass and branch diameter has been demonstrated for other species (Stephens, 1969). However, the relationship of branch foliage biomass to relative height is known to vary by species. Valentine et al. (1994) noted that, for loblolly pine, foliage biomass per unit of cross-sectional area of the branch fluctuated among crown strata, and fewer needles occurred on the lowest strata because these branches have scant foliage. For branches at the crown base, less foliage is sustained probably because of restricted light availability, although the branch diameter is larger (Xu and Harrington, 1998).

The horizontal distribution of branch foliage fluctuates with relative height. Foliage was shifted to the outer part on the branches lower in the crown (Maguire and Bennett, 1996). A similar finding was reported in radiata pine (*Pinus radiata* D. Don) in which the horizontal foliage distribution along a branch was significantly different at different heights within the crown (Wang et al., 1990). The horizontal distribution of foliage within branches also varied with branch size, relative dominance of the tree, and crown ratio, which can be attributed to the variation in light intensity with these branch-level variables (Xu and Harrington, 1998).

Among the major aspects of crown morphology, crown ratio (defined as the ratio of crown length to tree height) is considered one of the most important, since it strongly influences light interception and growth (Assmann, 1970). Trees with different crown ratios differ in their ability to respond to changes in environmental factors, such as release of neighboring competition (Velazquez-Martinez et al., 1992). Crown ratio also can serve as a good indicator of the growth potential of individual trees and stands, as an index of carbohydrate balance (i.e. ratio of photosynthesizing tissue to aboveground respiring tissue). Low values of crown ratio are likely to be associated with low leaf areas and possibly variation in foliage distribution within the crown (Xu and Harrington, 1998).

The main goal of this research project is to develop crown shape and leaf area distribution models for naturally regenerated longleaf pine trees at Eglin Air Force Base in Florida. The models developed in this study can also be used as an initialization model to develop biologically-based, resource driven mixed models (individual tree model between empirical growth and yield models and process models) to get more precise growth estimates of longleaf. This thesis will include only the first few components of the model.

## 1.2 METHODS

### Study area

The data for this study were collected in naturally regenerated longleaf pine (*Pinus palustris* Mill.) stands at Eglin Air Force Base in Florida from September 1996 to March 2000. The total land area of Eglin Air Force Base is approximately 187,000 ha, most of which is forested. The major tree species occurring on the land base at Eglin are longleaf pine, sand pine (*Pinus clausa* (Chapm.) Vasey.), and various scrub oaks (*Quercus spp.*). The base is primarily composed of all-aged stands of longleaf pine. Frequent disturbances from firearms testing maintain the stands in an open character. This area has soil with low moisture holding capacities and limited nutrients which result in very slow growth. Trees face drought conditions during the summer even though the area receives more rainfall than any other part of Florida (Anderson et al., 1999).

### Data collection

The sampling of longleaf pine trees covered the range of tree sizes, seedlings to the largest trees observed on the base, across the range of stand densities, and the range of sites. The total height ( $H$ ) ranged from 1.4 to 27.2 m and diameter at breast height ( $DBH$ ) ranged from 1.8 to 42.3 cm. In 27 stands across the study area, 106 trees were selected at random locations from September 1996 to 2000. Stands were selected to represent all the classes of site index, basal area per ha, and tree density (number of trees per ha).

Trees growing adjacent to the subject tree were evaluated to determine whether or not they are in competition with the subject tree. A competitor is defined as a tree crown projection which is in close proximity with the projected crown of the subject tree. In general any tree whose crown projection is near that of the subject tree were recorded as a



competitor, since those trees may influence the growth of the subject tree in some way, e.g. through root competition.

Competing trees were assigned a competitor ID number (1 through however many competitors there were) and measured for location, *DBH*, height at the base of live crown (*HBLC*), *H*, and crown ratio (*CR*) in the same manner as the subject tree. Species of the competitor was also noted. Spatial location of the subject tree and its competitors was recorded. A series of pin flags were laid out on the ground along the edge of the competing tree's crown where it influences the subject tree's crown. The azimuth and distance to plot center of these flags was then recorded to provide an indication of the location of the competing tree's projected crown.

Destructive measurements of subject trees were taken on a total of 106 trees. An intensive measurement of crown architecture was done on 83 of these trees. Measurement of foliage for surface area was also done on these trees. Less intensive crown measurements were made on the remaining 23 trees and were referred to as Short Data Set (SDS) trees. SDS trees were measured for gross crown characteristics in the field. Each live branch in the crown was measured for height above ground, basal diameter (diameter 2 centimeters away from the bole), overall length, and azimuth relative to the center of the tree. The number of shoots bearing live foliage and buds on each branch was counted. Twenty percent of these shoots were measured for growth over the three most recent years. Full crown architecture trees had detailed measurements made on the subject tree's crown. Subject trees were climbed and their branches were sawed off and carefully lowered to the ground in the field after taking the standing measurements. Standing measurements

included re-measuring *DBH* or ground line diameter, *H*, *HBLC*, and distance, and azimuth relative to the plot center.

Every fifth branch was selected systematically to be a sample branch. Foliage samples were collected from a single point within each meter-radius volume starting from the base of sampled branch immediately after the branch was removed from the tree. Samples were separated by age class and packed in a sealed plastic bag in a cooler until return to the laboratory. The samples collected were transferred and retained in a freezer until the time they were measured. They were placed in a sealable bag labeled with the stand, plot, and tree number, the sample point number, age of the needle, branch number, and distance from the base of branch. The sample point was marked with an aluminum tag indicating the sample numbers collected at that point and the distance from the base of branch. Additional samples were collected with each additional meter radius from the base of the branch. Foliage samples were used to estimate leaf area, leaf biomass, and nutrient contents.

The portion of the bole containing the bases of the branches was cut into shorter lengths for ease in handling. The crown of each tree was then brought to the Auburn campus to 3-dimensionally map the foliage over the past three years using the method devised by Smith (1990, 1994). Each stem was also measured throughout its length for diameter and bark thickness. Stem discs were removed every meter up the tree for radial growth, sapwood basal area, specific gravity, and age. At the laboratory these bole sections were placed in a specially designed stand and the branches reattached. Cartesian coordinates for the spatial location of live shoot tips were obtained by referencing a grid on the floor to obtain X and Y coordinates; Z coordinates were obtained by a tape measure

with a plumb bob affixed to the free end. Coordinate sets were measured at the four points representing the termini of shoot length growth for the three most recent years. Shoot diameter during these three years was also measured.

A series of conditional probability models will be applied to predict the distributions of  $H$ ,  $HBLC$ , crown length ( $CL$ ), number of whorls in the crown, number and size (length and diameter 2 centimeter away from the bole) of the branches in each whorl, the number and location of foliar points on each branch, and lastly the number and size of needles at each foliar point. Several regression models (linear and non linear) will be screened to find the best prediction model for each step in the model with the focus on predicting the distribution of values, not just the predicted mean values, using all variables from previous steps. Monte Carlo simulations will be used to simulate the marginal distribution at each step combining the predicted distributions of all previous steps in the model.

### **Simulation Steps**

0. Diameter at the breast height ( $DBH$ ) will be measured in the field directly.
1. Conditional distribution of  $H$  given  $DBH$  will be predicted for an individual tree.
2. Conditional distribution of crown ratio ( $CR$ ) given  $H$  and  $DBH$  will be predicted for an individual tree.
3. Conditional distribution of diameter at the base of live crown ( $DHBLC$ ) given  $DBH$ ,  $H$ , and  $HBLC$  will be predicted for an individual tree.
4. Conditional distribution of bole diameter ( $BD$ ) at different heights given  $DBH$ ,  $H$ ,  $HBLC$  and  $DHBLC$  will be predicted for an individual tree.

5. Conditional distribution of number of whorls ( $NWH$ ) on the tree bole given  $DHBLC$ , crown length ( $CL$ ),  $H$ , and  $DBH$  will be predicted for an individual tree.
6. Conditional distribution of each whorl height ( $WH$ ) given  $DBH$ ,  $CL$ ,  $DHBLC$  and  $NWH$  will be predicted for an individual tree.
7. Conditional distribution of number of branches ( $NB$ ) in each whorl given  $DBH$ ,  $H$ , and relative height from the ground will be predicted for an individual tree.
8. Conditional distribution of base of the branch diameter ( $BBD$ ) given  $DBH$ ,  $H$ ,  $BD$ , relative height in the bole, and the  $NB$  in the whorl will be predicted for an individual tree.
9. Conditional distribution of maximum branch length ( $MBL$ ) given  $BBD$ ,  $BD$ , relative height within the crown, and the presence of competitors will be predicted for an individual tree.
10. Conditional distribution of number of foliar points on the tree ( $NFPT$ ) given  $DBH$ ,  $H$ ,  $DHBLC$  and  $CL$  will be predicted for an individual tree.
11. Conditional distribution of number of foliar points on each branch ( $NFBP$ ) on a branch given  $BD$  and  $NFPT$ ,  $MBL$  and the presence of competitors will be predicted for an individual tree.
12. Conditional distribution of distance of each foliar point out from the bole ( $FPD$ ) given  $BBD$ ,  $BD$ ,  $MBL$ , the presence of competitors and relative height will be predicted for an individual tree.
13. Conditional distribution of horizontal angle of branch ( $FPH$ ) given  $BBD$  and  $FPD$  will be predicted for an individual tree.

14. Conditional distribution of vertical angle of branch ( $FPI$ ) given  $FPD$ , and  $MBL$  will be predicted for an individual tree.

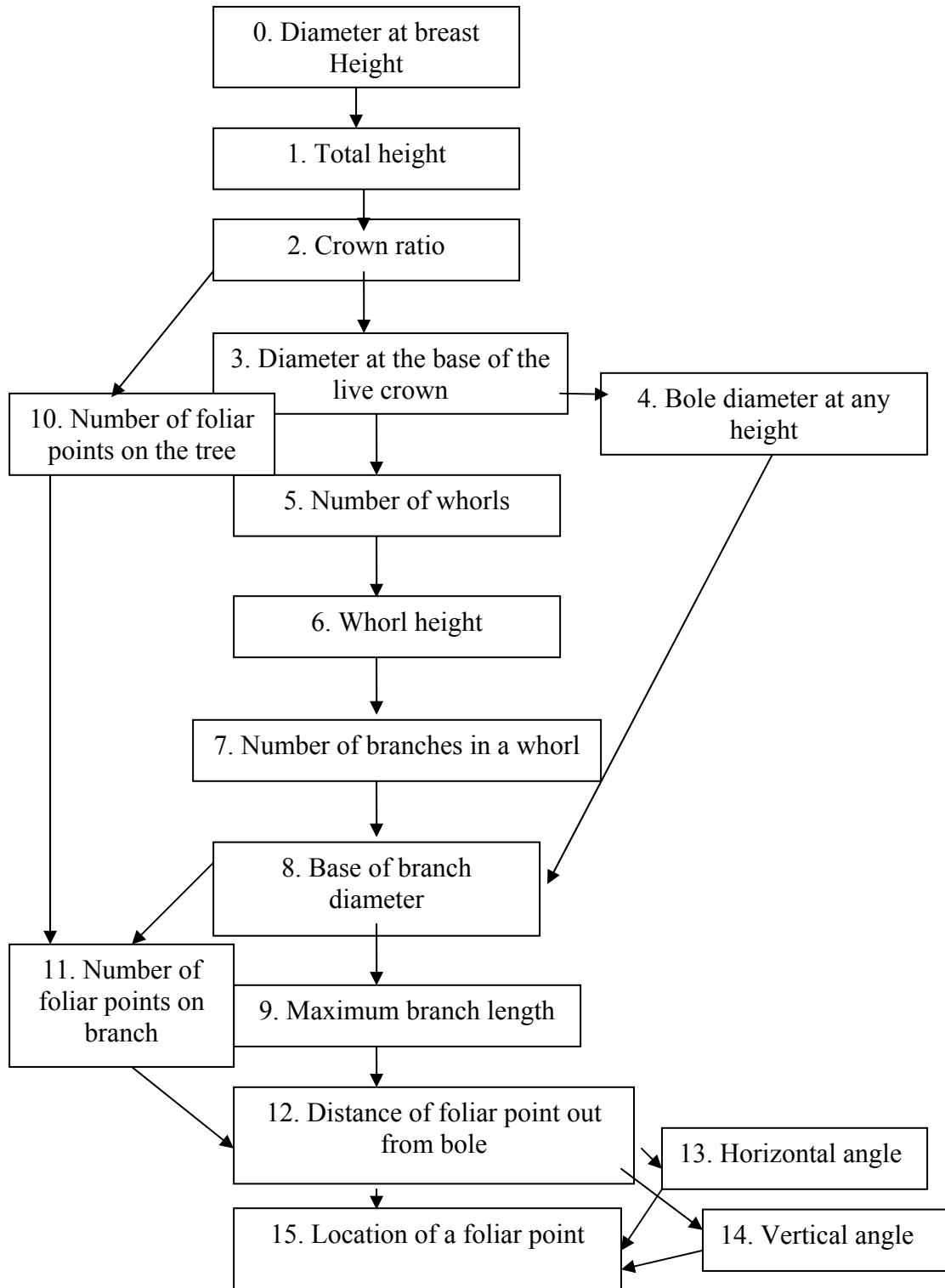
15. Conditional distribution of foliar point biomass ( $FPB$ ) given  $FPD$ , relative height above the ground and location of other foliage will be predicted for an individual tree.

This thesis is a part of a large project ‘Crown shape and leaf area distribution model for naturally regenerated longleaf pine at Eglin Air Force Base, Florida’. Out of the 15 simulation steps (Flow Chart 1.3) in the overall project this thesis includes only two steps, those are,

**Step 1** Prediction of the distribution of total height for longleaf pine at Eglin Air Force Base in Florida and

**Step 2** Prediction of the distribution of crown ratio for an individual tree using Normal versus Beta distribution model for longleaf pine at Eglin Air Force Base in Florida.

### 1.3 FLOW CHART



#### 1.4 LITERATURE REVIEW FOR HEIGHT-DIAMETER MODEL

Total tree height ( $H$ ) is often estimated from observed tree diameter at breast height ( $DBH$ ) outside bark. Measuring breast-height diameter is simpler, more accurate, and cheaper than measuring tree height. Thus models based on  $DBH$  measurements are cost effective (Peng, 2004). Tree height-diameter equations to predict height are required by forest resource managers to obtain accurate yield estimations. These estimates are used for decision making processes in forest management (Curtis, 1967). Total tree height and  $DBH$  outside bark are two essential forest inventory measures (Zhang et al., 2002). According to Botkin et al. (1972) and Avery and Burkhart (2002)  $H$  and  $DBH$  are needed to estimate timber volume, site index, and in forest growth and yield models. According to them the relationship between  $DBH$  and the total tree height is fundamental for developing growth and yield models for forest stands. Accurate height-diameter functions are essential for the prediction of stand development over time. According to Burkhart et al. (1972) and Huang et al. (1992) growth and yield models require height and diameter as basic input variables. When actual height measurements are not available, height-diameter functions can also be used to predict height growth indirectly (Larsen and Hann, 1987).

Height-diameter relationships of a given species depends on local environmental conditions and varies within a geographic region. The local environmental conditions (e. g. climate, soil and vegetation type) play a significant role in affecting the height-diameter relationships. According to a study done by Zhang et al. (2002) there are distinct variations in height-diameter relationships for trees in different ecological regions. Similar results were obtained from a study done by Peng (2004). According to him significant differences

in height-diameter relationship for black spruce were found between different ecological regions. The height-diameter relationship is generally asymptotic with height constrained to a maximum, while diameter is not (Prentice and Helmisaari, 1991). According to Niklas (1995) height may be constrained by limitations of mechanical support and increasing respirational load of stem tissue. Water stress in the upper shoots of tall trees may also limit the height growth (Ryan and Yoder, 1997).

Height-diameter relationships are generally nonlinear. Nonlinear height-diameter models have been commonly used in height predictions (Curtis et al., 1981; Arabatzis and Burkhardt, 1992). Several height-diameter models have been developed for different tree species in north America (e.g. Curtis, 1967; Larsen and Hann, 1987; Wang and Hann, 1988; Huang et al., 1992; Moore et al., 1996; Zhang, 1997; Lappi, 1997). Curtis (1967) compared a number of nonlinear height-diameter equations for second-growth Douglas-fir. Most of them provided the similar results within the range of the observed data.

A study done by Huang et al (1992) for major Alberta tree species compared twenty nonlinear height-diameter models for 16 different species. All the models were fitted using weighted nonlinear least square regression due to the presence of unequal error variance. Weighted mean square error, asymptotic t-statistics for the coefficients in the model, plots of standardized residuals against predicted height were used to check the performance of the models. Sigmoid functions, as for example, Chapman-Richards function, modified logistic function, and Weibull-type function produced the most reasonable results.

Arabatzis and Burkhardt (1992) compared the prediction performance of eight height-diameter regression models for estimating mean stand height for loblolly pine. The nonlinear model of the form  $H = ae^{b/D}$  performed the best among all the models tested.



Moore et al. (1996) have developed individual nonlinear height-diameter equations for ten major species in the inland Northwest.

In a study done by Zhang (1997) six nonlinear growth functions were fitted to tree height-diameter data for ten conifer species in the inland Northwest of the United States. The models found superior in predictions to other models are Schnute, Weibull, and Richards function. In his paper he explained that the best model is selected based on its statistical properties such as mean square error, asymptotic t-statistics for the coefficients, model  $R^2$ , and analysis of the residual, however, he suggested that the competing models should be evaluated based on their predictive ability across the range of tree diameter. He further added that model validation should be done using an independent data set.

Temesgen and Gadow (2004) have evaluated two sets of five height-diameter models. The first set of the models used total tree height as a function of *DBH*; whereas, the second set used total tree height as a function of *DBH* and other stand level variables. The analysis of residual plots and fit statistics for the models which include the stand level variables were found superior to the models with only *DBH* as an explanatory variable. However, cost of collecting data for additional variable versus the gain in precision should always be considered before selecting a model with higher number of explanatory variables for a model (Temesgen and Gadow 2004). Therefore, the principle of parsimony should always be considered before making selection and recommending a model to use.

## **1.5 LITERATURE REVIEW FOR CROWN RATIO MODEL**

The ratio of live crown length to the total tree height is commonly known as crown ratio. Crown ratio is the characteristic usually used to describe the crown size (Hynynen 1995). Measures of crown dimension are important components of forest growth and yield models, and are used in many tree and crown level growth models (Cole and Lorimer, 1994). Tree crown parameters have been used as predictor variables in diameter and height growth equations (Monseur and Sterba, 1996). Crown size is an important measure of tree vigor (Valentine et al., 1994). Crown ratio and crown length reflect the potential of a released tree to use available resources such as increased growing space and therefore are useful indicators of tree vigor (Hasenauer and Monserud, 1996), wood quality (Kershaw et al., 1990), stand density (Clutter et al., 1983), competition and survival potential (Oliver and Larson, 1996). Crown ratio has also been used as an input variable to estimate growth and mortality of individual trees (Temesgen et al., 2005).

The Crown of a tree reflects the level of competition over time (Mitchell, 1975). When density increases the crown length decreases. Reducing stand density by the application of thinning practices slows the recession of crown base (Short and Burkhart, 1992). Thus the crown length is highly influenced by the prevailing growing conditions in the past (Bella, 1971). *CR* is also considered as a measure of photosynthetic potential. Photosynthates and hormones produced in the crown control shoot growth, cambial growth, and root growth (Kozlowski, 1971). According to Monserud (1975) crown ratio is a good indicator of a tree's ability to utilize available resources for growth.

Measurement of *CR* for each tree can be time-consuming and difficult to obtain in very dense stands and for the trees where the base of live crown is obscured. Predictions of tree crown ratio have been based on allometric relations between stand and tree variables (Hasenauer and Monserud, 1996). Crown ratio can be predicted directly from tree and stand variables (Hasenauer and Monserud, 1996) or indirectly from estimates of the height to the base of live crown (Soares and Tome 2000).

According to Dyer and Burkhart (1987) “crown ratio has been predicted either directly from tree and stand variables or has been computed indirectly from estimates of total height and crown height.” They used a nonlinear model form that yields logical estimates of *CR*. The explanatory variables used in the model were stand age, tree diameter, and tree height.

Hynynen (1995) presented a nonlinear model that yields logical predictions for tree crown ratio. The model is based on data from permanent experimental plots located in even-aged Scots pine. Explanatory variables used in the model were stand dominant height, stand basal area, tree diameter, and tree height. The effect of thinning on tree crown ratio was modeled by incorporating a thinning response variable into the model.

In a study done by Hasenauer and Monserud (1996) crown ratio was predicted using a logistic function. The explanatory variables used in the model were tree size characteristics, stand density, and site factors. The size variable was found to be approximately equal in importance to stand density while the site factors explained the least amount of variation. The height/diameter ratio was the most important size variable.

Soares and Tome (2000) have used several nonlinear models for the prediction of crown ratio. The models tested were logistic, exponential, Richards and Weibull function.

The Richards function was selected for the prediction of crown ratio. As a function of age, density, diameter and dominant height.

In a study conducted by Temesgen et al (2005) crown ratio models for five species in British Columbia were fitted using the independent variables size, competition, and site. The best predictor for crown ratio was found to be the size variables. Total tree height was found to be an important predictor for all five species. The results indicated that the *CR* decreased with increase in height and with increase in competition. A logistic model was used to predict the *CR* which gives the estimates of mean predicted *CR* between 0 and 1.

## LITERATURE CITED

- Anderson, M., Somers, G. L., Smith, W. R., and Freeland, M., Ruch, D., 1999. Measuring crown dynamics of longleaf pine in sandhills of Eglin Air Force Base. Proceedings of the Second Longleaf Alliance Conference 1998. Charleston, 17-19 Nov., 46-48.
- Arabatzis, A.A., and Burkhardt, H.E., 1992. An evaluation of sampling methods and model forms for estimating height-diameter relationships in loblolly pine plantations. For. Sci. 38, 192-198.
- Assmann, E., 1970. The principles of forest yield study. Pergamon Press, New York.
- Avery, T. E., and Burkhardt, H. E., 2002. Forest Measurements. McGraw-Hill, Inc. 456 pp.
- Baldocchi, D. D., and Hutchison, B. A., 1986. On estimating canopy photosynthesis and stomatal conductance in a deciduous forest with clumped foliage. Tree phys. 2, 155-168.
- Bartelink, H. H., 2000. A growth model for mixed forest stands. For. Eco. Mang. 134, 29-43.
- Bella, I. C., 1971. A new competition model for individual trees. For. Sci. 17, 364-372.
- Bossell, H., 1996. TREEDYN3 forest simulation model. Ecol. Modeling 90:187-227.
- Botkin, D.B., Jamak, J.F., and Wallis, J.R., 1972. Some ecological consequences of a computer model of forest growth. J. Ecol. 60, 849-873.
- Bruce, D., 1990. Development of empirical forest growth models. In: Process

Modeling of Forest Growth Responses to Environmental Stress (Eds: R. Dixon, R. Meldahl, G. Ruark, W. G. Warren). Timber press, Portland, OR, U.S.A. pp 191-199.

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

Clutter, J.L., Fortson, J.C., Pienaar, L.V., Brister, G.H., and Bailey, R.L., 1983. Timber management: A quantitative approach. John Wiley & Sons, New York. 333 p.

Cole, W., and Lorimer C. G., 1994. Predicting tree growth from crown variables in managed Northern hardwood stands. For. Ecol. Manag. 67, 159-175.

Curtis, R.O., 1967. Height-diameter and height-diameter-age equations for second-growth Douglas-fir. For. Sci. 13, 365-375.

Curtis, R.O., Clendenin, G.W., DeMars, D.J., 1981. A new stand simulator for coast Douglas-fir: DFSIM user's guide. USDA For. Serv. Gen. Tech. Rep. PNW-128.

Daniels, R. F., and Burkhart, H. E., 1975. Simulation of individual tree growth and stand development in managed loblolly pine plantations. Division of Forestry and Wildlife Resources. Virginia Polytechnic Institute and State University, Blacksburg, VA 24061. Publication FWS-5-75. 69 p.

Dixon, R. K. 1990. Physiological process and tree growth. In: Process Modeling of Forest Growth Response to Environmental Stresses (Eds: R. Dixon, R. Meldahl, G. Ruark, W. G. Warren). Timber Press, Portland, OR, U.S.A. pp 21-32.

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

Grace, J. C., 1990. Modeling the interception of solar radiation energy and net photosynthesis. In: *Process Modeling of Forest Growth Responses to Environmental Stress* (Eds: R. Dixon, R. Meldahl, G. Ruark, W. G. Warren). Timber Press, Portland, OR, U.S.A. pp. 142-158.

Grace, J.C., Jarvis, P.G., and Norman, J.M., 1987. Modeling the interception of solar radiation energy in intensively managed stands. *N.Z.J. For. Sci.* 17, 193-209.

Hagihara, A., Hozumi, K., 1986. An estimate of the photosynthetic production of individual trees in a *Chamaecyparis obtusa* plantation. *Tree Phys.* 1, 9-20.

Hasenauer, H., Monserud R.A., 1996. A crown ratio model for Austrian forests. *For. Ecol. Manag.* 84, 49-60.

Huang, S., Titus, S.J., and Wiens, D.P., 1992. Comparison of nonlinear height-diameter functions for major Alberta tree species. *Can. J. For. Res.* 22, 1297-1304.

Hynynen, J., 1995. Predicting tree crown ratio for un-thinned and thinned Scots pine stands. *Can. J. For. Res.* 25, 57-62.

Johnsen, K., Samuelson, Teskey, L., McNulty, R., S., and Fox T., 2001. Process models as tools in forestry research and management. *For. Sci.* 47, 2-8.

Kershaw, J.A., Maguire, D.A., and Hann, D.W., 1990. Longevity and duration of radial growth in Douglas-fir branches. *Can. J. For. Res.* 20, 1690-1695.

Kinerson, J. A., D. A. Maguire, and D. W. Hann. 1987. Longevity and duration of radial growth in Douglas-fir branches. *Can. J. For. Res.* 20:1690-1695.

Kozlowski, T.T., 1971. *Growth and Development of Trees*. Vol. 1. Seed Germination, Ontogeny, and Shoot Growth 1971. Academic Press, New York, 443 pp.

- Lappi, J. 1997. A longitudinal analysis of height/diameter curves. *For. Sci.* 43: 555-570.
- Larsen, D.R., and Hann, D.W., 1987. Height–diameter equations for seventeen tree species in southwest Oregon. Oregon State Univ., For. Res. Lab. Pap. 4.
- Maguire, D. A., and Bennett, W. S., 1996. Patterns in vertical distribution of foliage in young costal Douglas-fir. *Can. J. For. Res.* 26, 1991-2005.
- Mitchell, K. J., 1975. Dynamics of simulated yield of Douglas-fir. *For. Sci. Mono.* 17. 39p.
- Mohren, G. M. J., 1987. Simulation of forest growth, applied to Douglas fir stands in the Netherlands. Ph. D. thesis. Agricultural University of Wageningen, The Netherlands. 184p.
- Monserud, R.A., 1975. Methodology for simulating Wisconsin Northern hardwood stand dynamics. Ph.D. Thesis, University of Wisconsin, Madison, 156 pp.
- Monseurd, R.A., and Sterba H., 1996. A basal area increment model for individual trees growing in even and uneven aged forest stands in Austria. *For. Ecol. Manag.* 80, 57-80.
- Moore, J. Zhang, A., and Stuck, L. D., 1996. Height–diameter equations for ten tree species in the Inland Northwest. *West. J. Appl. For.* 11, 132-137.
- Munro, D. D., 1973. Forest growth models – a prognosis. In: *Growth models for tree and stand simulation*. Royal Collage of Forestry, Stockholm, Sweden, 7-21 p.
- Nepal, S. K., 1993. Crown Shape Modeling For Loblolly Pine: A Frontier Approach. Ph. D. Dissertation. Auburn University, Auburn, AL 157p.



Niklas, K. J., 1995. Size-dependent allometry of tree height, diameter and trunk-taper. *Ann. Bot. (London)*, 75, 217-227.

Norman, J. M., and Jarvis, P. G., 1975. Photosynthesis in Sitka spruce (*Picea sitchensis* (Bong.) Carr.). V. Radiation penetration theory and a test case. *J. of Appl. Ecol.* 12, 839-878.

Oker-Blom, P., 1986. Photosynthetic radiation regime and canopy structure in modeled forest stands. *ACTA Forestalia Fennica* 197. Society of Forestry, Finland, 44p.

Oker-Blom, P., and Kellomaki, S., 1982. Theoretical computations on the role of crown shape in the absorption of light by forest trees. *Math. Biosci.* 59, 291-311.

Oliver, C.D., and Larson, B.C., 1996. *Forest stand dynamics*. John Wiley and Sons, Toronto.

Peng, C., 2004. Developing and evaluating tree height-diameter models at three geographic scales for black spruce in Ontario. *North. J. Appl. For.* 21, 83-92.

Prentice, I. C., and Helmisaari, H., 1991. Silvics of north European trees: compilation, comparisons and implications for forest succession modeling. *For. Ecol. Manage.* 42, 79-93.

Russell, G., Jarvis, P.G., and Monteith, J.L., 1989. Absorption of radiation by canopies and stand growth. In: *Plant canopies: their growth, form and function* (Eds: G. Russell, B. Marshall, and P.G. Jarvis). Cambridge University press, Society of Experimental Biology, Cambridge, England. pp. 21-39.

Ryan, M. G., and Yoder, B. J. 1997. Hydraulic limits to tree height and tree growth. *Bioscience*, 47, 235-242.

Schreuder, H. T., and Swank, W. T., 1974. Coniferous stands characterized with the Weibull distribution. *Can. J. For. Res.* 4, 518-523.

Short III, E.A., and Burkhart, H.. 1992. Prediction crown-height increment for thinned and unthinned loblolly pine plantations. *For. Sci.* 38, 594-610.

Smith, W.R., 1990. The static geometric modeling of three-dimensional crown competition. In: *Process Modeling of Forest Growth Responses to Environmental Stresses* (Eds: R.Dixon, R. Meldahl, G.Ruark, and W. Warren). Timber Press, Portland, OR, pp. 294-302.

Smith, W.R., 1994. An empirical evaluation of a three-dimensional crown model for predicting volume growth, *For. Ecol. Mang.* 69, 199-209.

Soares, P., and Tome, M., 2000. A tree crown ratio prediction equation for eucalyptus plantations. *Ann. For. Sci.* 58, 193-202.

Stephens, G. R., 1969. Productivity of red pine, 1. Foliage distribution in tree crown and stand canopy. *Agric. Meteorol.* 6, 275-282.

Temesgen, H., and Gadaw, K.V., 2004. Generalized height-diameter models-an application for major tree species in complex stands of interior British Columbia. *Eur. J. Forest Res.* 123, 45-51.

Temesgen, H., LeMay, V., and Mitchell, S. J., 2005. Tree crown ratio models for multi-species and multi-layered stands of southeastern British Columbia. *For. Chron.* 81,133-141.

Terjung, W. H., and Louie S. S. F., 1972. Potential solar radiation on plant shapes. *Inter. J. Biometeor.* 16, 25-43.

- Valentine, H. T., V. C. Baldwin, T. G. Gregorie, and H. E. Burkhart. 1994. Surrogates for foliar dry matter in loblolly pine. *For. Sci.* 40, 576-585.
- Velazquez-Martinez, A., Perry, D. A., and Bell, T.E., 1992. Response of above ground biomass increment, growth efficiency, and foliar nutrient to thinning, fertilization and pruning in Douglas-fir plantations in the central Oregon cascades. *Can. J. For. Res.* 22, 1278-1289.
- Vose, J. M., 1988. Pattern of leaf area distribution within crowns of nitrogen- and phosphorus- fertilized loblolly pine trees. *For. Sci.* 34, 564-573.
- Vose, J. M., and Allen, H. L., 1988. Leaf area, stem wood growth, and nutrition relationships in loblolly pine. *For. Sci.* 34, 547-563.
- Vose, J. M., and Swank, W. T. 1990. A conceptual model of forest growth emphasizing stand leaf area. In: *Process Modeling of Forest Growth Responses to Environmental Stresses* (Eds: R. Dixon, R. Meldahl, G. Ruark, and W. Warren). Timber Press, Portland, OR, pp. 274-287.
- Wang, C.H. and Hann, D.W., 1988. Height–diameter equations for sixteen tree species in the central western Willamette valley of Oregon. Oregon State Univ., *For. Res. Lab. Res. Pap.* 51.
- Wang, Y. P., Jarvis, P. G., and Benson, M. L., 1990. Two-dimensional needle area density distribution within the crowns of *Pinus radiata*. *For. Ecol. Mang.* 32, 217-237.
- Whitehead, D., Grace, J. C., and Godfrey, M. J. S., 1990. Architectural distribution of foliage in individual *Pinus radiata* D. Don crowns and the effects of clumping on radiate interception. *Tree Phys.* 7, 135-155.

Xu, M., and Harrington, T. B., 1998. Foliage biomass distribution of loblolly pine as affected by tree dominance, crown size, and stand characteristics. *Can. J. For. Res.* 28, 887-892.

Zhang, L. 1997., Cross-validation of non-linear growth functions for modeling tree height–diameter relationships. *Ann. Bot.* 79, 251-257.

Zhang, L., Peng, C., Huang, S., and Zhou, X., 2002. Development and evaluation of ecoregion-based jack pine height-diameter models for Ontario. *For. Chron.* 78, 530-538.

## CHAPTER 2

### MODEL FOR PREDICTING DISTRIBUTION OF TOTAL HEIGHT FOR LONGLEAF PINE AT EGLIN AIR FORCE BASE IN FLORIDA.

#### 2.1 INTRODUCTION

The prediction of total height from observed diameter at breast height (1.3 m above ground) has many applications in forest management and silvicultural research (Meyer, 1940). Height-diameter models have been used in the prediction of tree volume (Curtis, 1967), site index, and growth and yield of a tree and a stand (Vanclay, 1994). Total tree height and diameter at breast height are also used as important input variables in several growth and yields models (Burkhart et al., 1972). Precise estimation of tree height is one of the most important requirements in forest inventory, growth and yield estimation, and forest management (Botkin et al., 1972).

Diameter at breast height (*DBH*) of a tree is easily measured in the field; whereas measurement of total tree height (*H*) is difficult and more time-consuming. Height-diameter functions can be developed and used to estimate *H* from the measurements of *DBH* in the field. These functions can be used to predict total tree height growth, and volume when actual height measurements are not available (Larsen and Hann, 1987).

Generally the height-diameter functions have a nonlinear nature. Nonlinear height-diameter models have been used in height predictions for many tree species (Arabatzis and Burkhart, 1992). Several height-diameter models have been developed

and used for various tree species in North America (e.g. Larsen and Hann, 1987; Wang and Hann, 1988; Huang et al., 1992; Moore et al., 1996; Zhang, 1997; Lappi, 1997). The primary purpose of all of these models is the prediction of the mean tree height for a given *DBH*. Of secondary importance are estimates of the confidence interval about the predicted mean height especially near the extremes and beyond the range of the data. Selecting model forms that reduce the width of the confidence intervals while maintaining good predictions of the mean is a valid objective of model development. The width of the confidence interval depends upon two factors, the sample size and the variability of the predictions at each *DBH*. Modeling this variability along with the mean therefore has potential to reduce the confidence interval of the mean predicted height. Models that seek to predict the distribution of variables, not just the mean, have potential benefits as the models of forest growth and yield become more complex. Simulating the distribution is essential to estimate prediction errors and ultimately the confidence interval of future yields. For example, estimating the average height per tree in an inventory when tree heights are predicted from observed *DBH* normally considers only the variability in mean heights between trees. The resulting confidence interval is underestimated since it does not include the variability due to the prediction.

This variability can be easily included by using a Monte Carlo simulation of the predicted distribution of tree height for each observed tree. However, this is dependent upon the distribution, and not just the mean, being simulated well by the model. For this reason, including much more complex models involving multiple interconnected predictions, the predicted distribution of height for given *DBH* will be of equal importance as the predicted average height. The primary objectives of this study are to:

- i) Construct a model to predict the distribution of total height based on *DBH* for longleaf pine (*Pinus palustris* Mill.) trees growing in natural stands at Eglin Air Force Base in Florida.
- ii) Assess the ability of the model to predict the distribution of height based on the *DBH* measurement in the field.

## **2.2 MATERIALS AND METHODS**

### **2.2.1 DATA AND STUDY AREA**

Longleaf pine (*Pinus palustris* Mill.) is naturally distributed in most of the Atlantic and Gulf Coastal Plains from southeastern Virginia to eastern Texas and south through the northern two-thirds of peninsular Florida. This tree is also found in the Piedmont, Ridge and Valley, and mountain Provinces of Alabama and northwest Georgia (Boyer, 1991). Longleaf pine grows in warm, wet temperate climates characterized by hot summers and mild winters. More frequent fires favor longleaf pines, which are more fire adapted. (Boyer, 1991).

Data for this study were collected in naturally regenerated longleaf pine stands at Eglin Air Force Base in Florida from September 1996 to March 2000. Total land area of Eglin Air Force Base is approximately 187,000 ha, most of which is forested. The major tree species occurring on the land base at Eglin are longleaf pine, sand pine (*Pinus clausa* (Chapm.) Vasey.), and various scrub oaks (*Quercus spp.*). The base is primarily composed of all-aged stands of longleaf pine. Frequent disturbances from firearms testing maintain the stands in an open character. The major soil series include Lakeland and Troup (Overing and Watts 1989). These soil series have deep, excessively drained, rapidly permeable soils and limited nutrients which result in very slow growth. For this reason the trees face

drought conditions during the summer even though the area receives more rainfall than any other part of Florida (Anderson et al., 1999). The trees are flat topped with a site index ranging from 16 to 19 m; whereas a nearby Escambia Experimental Forest with similar age has a site index range of 18 to 26 m (Meldahl et al., 1998). The basal area per ha for Eglin in the study area was 12 m<sup>2</sup>/ha; whereas the basal area per ha for Escambia Experimental Forest was 26-34 m<sup>2</sup>/ha. (Kush et al., 1999).

Sampling of longleaf pine trees covered the range of tree sizes, from saplings to the largest trees observed on the base, across the range of stand densities, and the range of sites. Total height ( $H$ ) ranged from 1.4 to 27.2 m and diameter at breast height ( $DBH$ ) ranged from 1.8 to 42.3 cm. In 27 stands across the study area, 85 sample trees were selected at random locations. Stands were selected to represent all the classes of site index, basal area per ha, and tree density. Samples were selected from the stands which were naturally established and free of major disturbances, primarily crown-scorching fires, within the past five years. Total height and  $DBH$  were measured for each tree. Trees with major defects including forking, major breakage, and weak stems were not selected as candidate trees.

### **2.2.2 METHODS**

Least squares is a very common method of estimating parameters. This method provides optimal estimators when the assumptions of homogeneity of variance in the error term (Meng and Tsai, 1986) and linearity between response and predictor variables are satisfied. However, the plot of  $H$  against  $DBH$  depicts a nonlinear relationship. Moreover, total height and variance of total height increase as  $DBH$  increases (Figure 1) and the assumptions required for the optimality of ordinary least squares (OLS) estimators are not satisfied. Therefore, a nonlinear least square approach was used in model fitting. A non-

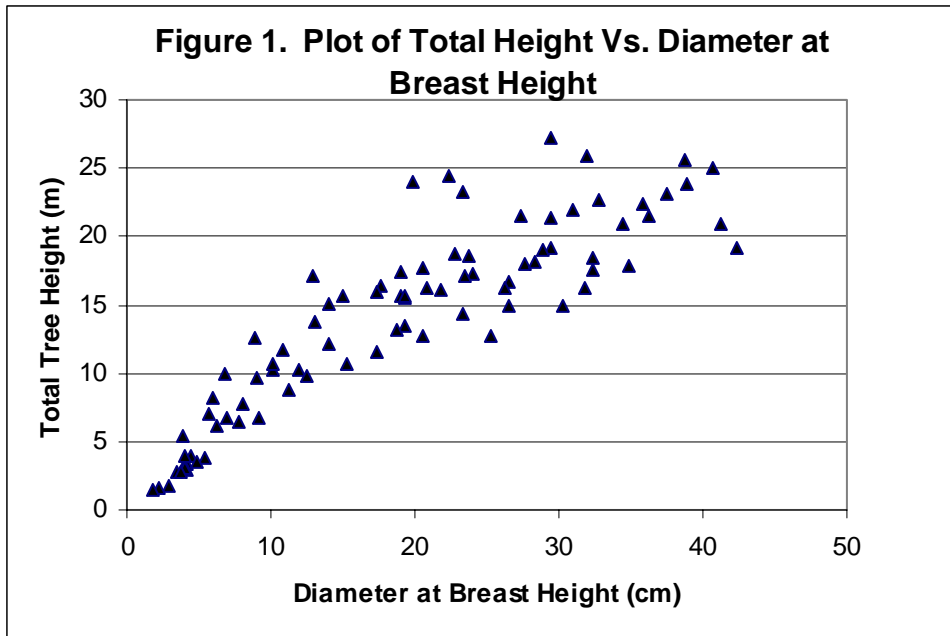


linear regression model used by Curtis et al. (1981), Larsen and Hann (1987), and Wang and Hann (1988) with different formulations was used to obtain an appropriate model form.

Model fitting was carried out with the NLMIXED procedure using a quasi-Newton algorithm (SAS Institute, Inc. 2004). NLMIXED will produce maximum likelihood estimates of any likelihood function. This procedure allows all parameters to be functions of independent variables and can be used to find estimators of parameters in complex models. The NLMIXED procedure uses the method of maximum likelihood (MLE) to obtain the estimates of the coefficients numerically. MLE selects as estimates the values of the parameters that maximize the likelihood of the observed sample. It is consistent, that is, as the sample size gets larger, the estimates converge to the true values. It is asymptotically efficient, which means that for large samples, it produces precise estimates. It is also asymptotically unbiased, that is, for large samples one expects to get the true value on average. The distributions of the estimators themselves are asymptotically normal, if the sample is large enough. These are all excellent large sample properties. For the normal distribution the least square method and MLE provide the same estimates of parameters (Casella and Berger, 2002).

The following criteria were used to decide the performance of height-diameter models:

1. Significant t-test for the parameters in the model.
2. AIC value.
3. Analysis of the standardized residual plots.
4. Kolmogorov-Smirnov test for normality of standardized residuals.



5. Simulated Type I Error rate using cross validation.

The asymptotic t-statistic for each coefficient in the model should be significant for any model to be considered appropriate for prediction of mean and prediction of distribution. It tests whether an independent variable is important to be included in the model. The model with any of the parameters insignificant at a specified level of significance will be considered a *bad* model and hence discarded. Therefore, this is the first test that any candidate model has to pass to be considered for further analysis.

Akaike (1973) developed an estimation criterion, defined as

$$AIC = -2 \log_e (l(\hat{\theta} | data)) + 2K$$

where,  $\log_e l(\hat{\theta} | data)$  is the value of the maximized

log-likelihood over the unknown parameter ( $\theta$ ), given data and the model. K is the

number of estimated parameters included in the model (Anderson et al., 2000). AIC

penalizes for the addition of parameters, and thus selects a model that fits well but has the

smallest number of parameters. The model with the lowest AIC is selected as a good candidate for the data available (Anderson et al., 2000).

In the normal distribution the standardized residual is computed as,

$$d_i = \frac{e_i}{\hat{\sigma}_i}, i = 1, 2, \dots, n, \text{ where } e_i = y_i - \hat{y}_i, y_i \text{ is observed } H, \text{ and } \hat{y}_i \text{ is the predicted } H \text{ for}$$

the  $i$ th tree. The standardized residuals have mean zero and unit variance. They are useful in analyzing the problems of model specification and outliers. Most of the standardized residuals should lie in the range of  $-3 \leq d_i \leq 3$ . Analysis of the standardized residual plots against the predicted value should show a more or less even scatter around the zero line throughout the range of the predicted values in which case equal error variance is assumed. A model that does not suffer from non-constant error variance is considered to be good candidate for further analysis. Any pattern in the standardized residual might reveal unequal error variance, or a problem of model specification (Huang et al., 1992).

The Kolmogorov-Smirnov (KS) test was used to test whether the standardized residuals follow a standard normal distribution. The KS test is based on comparison of the empirical distribution function (edf) of the data and the cumulative distribution function (cdf) of an assumed distribution. It is non-parametric and distribution free test (Hollander and Wolf, 1999). Standardized residuals for a good model will follow a standard normal distribution and will pass this test.

Cross validation checks whether the distribution obtained from an independent data set follows a specified distribution (Lilliefors, 1967). The KS test was used for this purpose. The Type I Error rates of the KS test obtained from running the simulation is compared to the nominal Type I Error rate of 5% in our study. A model will be deemed a

better predictive model than a competing model if its simulated Type I Error rate is closer to the nominal Type I Error rate than its competitor's simulated Type I Error rate.

## 2.3 RESULTS

The analysis of height-diameter relationship depicted by the scatter plot of  $H$  and  $DBH$  (Figure 1) reveals that there exists a nonlinear association. The larger trees growing in the study area are flat topped and it is reasonable to assume that the height will approach an asymptote as  $DBH$  gets larger becoming independent of  $DBH$ . Therefore, a non-linear model of the form  $H = 1.3 + e^{a_h + b_h * DBH^{c_h}}$  was used to predict the mean height in all the models; however, different formulations of variance were used to predict the distribution of total height. This model estimates the total height to be 1.3 meters as  $DBH$  approaches zero. Total height approaches an asymptote as  $DBH$  gets very large, outside the range of the observed data, which is biologically reasonable. The models were grouped into two broad categories. Category one (model A, B, and C) requires the assumption of normality of the error term and category two (model D) does not require the assumption of normality.

### 2.3.1 MODELS THAT REQUIRE THE ASSUMPTION OF NORMALITY OF THE RANDOM ERROR

#### A. Model for the prediction of total height ( $H$ ) as a nonlinear function of $DBH$

This model assumes that total tree height  $H$  is normally distributed with mean height  $\mu$  and variance in total height  $\sigma^2$ , that is,  $H \sim N(\mu, \sigma^2)$ , where

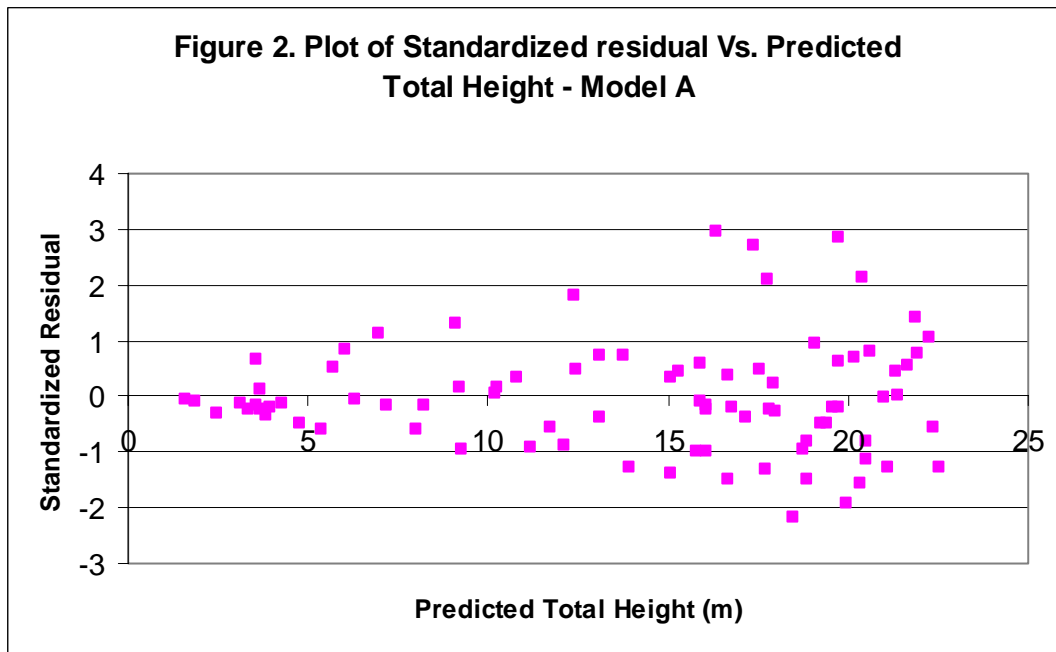
$\mu = 1.3 + e^{a_h + b_h * DBH^{c_h}}$ ,  $\sigma^2$  is an unknown constant and  $a_h$ ,  $b_h$  and  $c_h$  are unknown

coefficients to be estimated from the data. The mean  $\mu$  is expected tree height in meters,

$DBH$  is the observed diameter outside bark at breast height in centimeters, and  $e$  is the base of the natural logarithm function. The NLMIXED procedure in SAS was used to maximize the likelihood function or equivalently the log likelihood function

$$\log(\lambda(a_h, b_h, c_h, \sigma | H_1, K, H_n)) = -\frac{n}{2} \log 2\pi - n \log \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^n (H_i - \mu)^2 .$$

For the nonlinear model A all the parameters  $a_h, b_h,$  and  $c_h$  were significant (Table 1). The test of normality (Kolmogorov-Smirnov) revealed that the standardized residuals were not normal (p-value=0.0293), and the plot of standardized residual against predicted total height indicated that the error variances were not equal throughout the range of predicted height (Figure 2), that is, there exists a problem of non-constant variance.



The fitted curve for prediction of total height does go through the center of the observed total height (Figure 3). This is indicative of a good predictive ability of Model A for the prediction of mean total height. This model works under the assumption of equal

variance and the prediction intervals for the whole range of predicted height are the same (Figure 3) which is not realistic. The prediction interval is too wide for smaller *DBH* and too narrow for larger *DBH*.

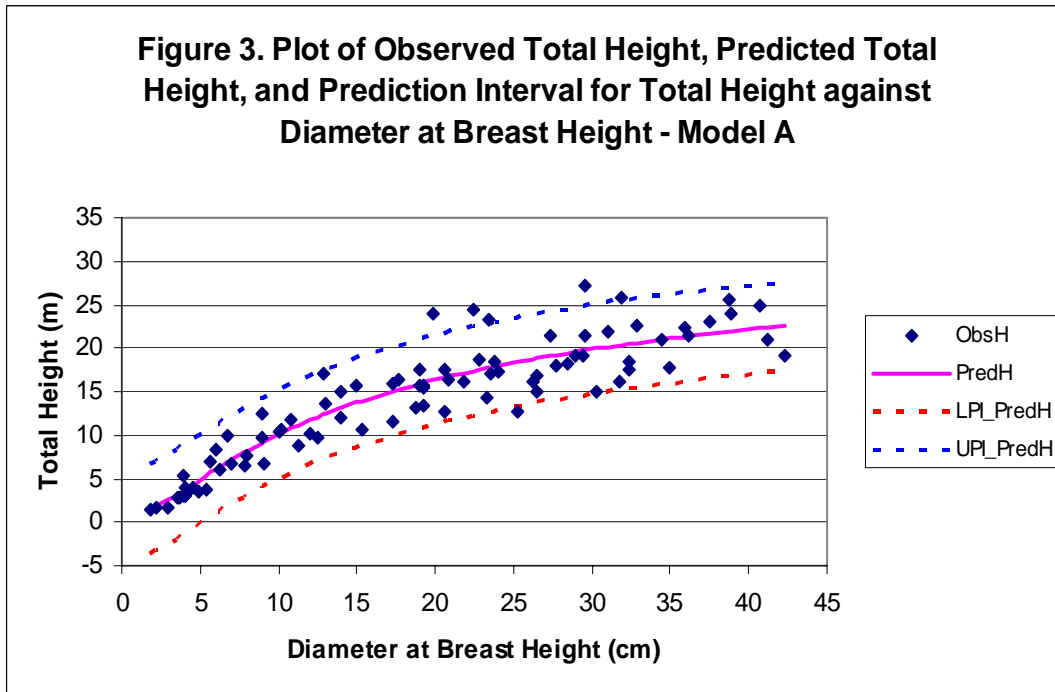
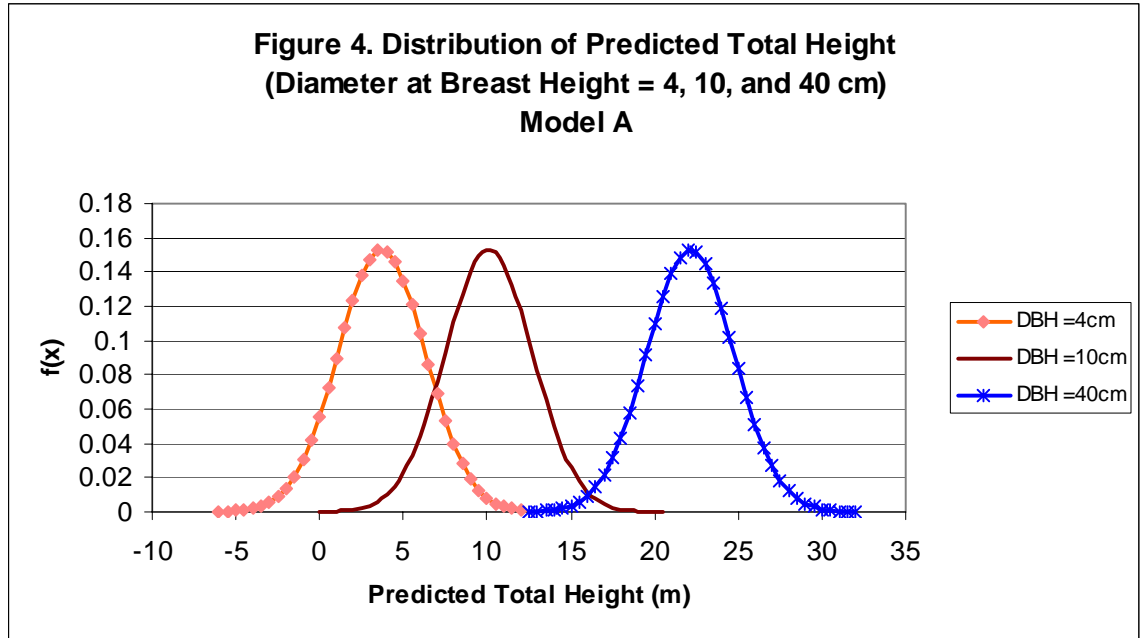


Figure 4 reveals that for all three examples *DBH* (4, 10, and 40 cm) the variance is the same, however, there is a shift in the mean heights. This is true for all the trees with different *DBH* within the range of observed data set (Figure 3).



**Table 1. Model Parameter Estimates**

Model	Coeff	Estimate	P-val	LCL	UCL	AIC	K-S (P-val)
A	$\hat{a}_h$	3.5232	<.0001	3.0352	4.0112	412	0.0293
	$\hat{b}_h$	-7.3457	0.0002	-11.1079	-3.5836		
	$\hat{c}_h$	-0.7360	0.0002	-1.1095	-0.3626		
	$\hat{\sigma}^2$	6.7871	<.0001	4.7171	8.8571		
B	$\hat{a}_h$	3.3395	<.0001	3.0354	3.6437	391.8	<0.010
	$\hat{b}_h$	-8.7253	<.0001	-11.2247	-6.2258		
	$\hat{c}_h$	-0.8846	<.0001	-1.1367	-0.6325		
	$\hat{\delta}$	0.05496	<.0001	0.03820	0.07172		
C	$\hat{a}_h$	3.2688	<.0001	3.1017	3.4359	373.7	0.0611
	$\hat{b}_h$	-9.8979	<.0001	-10.3702	-9.4257		

	$\hat{c}_h$	-0.9780	<.0001	-1.1055	-0.8505		
	$\hat{b}_v$	-69.2005	0.0449	-136.80	-1.5980		
	$\hat{c}_v$	-2.4469	<.0001	-3.4770	-1.4168		
	$\hat{\delta}$	8.8303	<.0001	5.3966	12.2639		
D	$\hat{a}_h$	3.2585	<.0001	3.0974	3.4197	380.2	>0.150
	$\hat{b}_h$	-9.8464	<.0001	-10.9006	-8.7921		
	$\hat{c}_h$	-.9797	<.0001	-1.1.22	-0.8572		
	$\hat{\phi}$	0.05256	<.0001	0.03667	0.06845		

**B. Model for the prediction of total height ( $H$ ) as a *weighted* nonlinear function of  $DBH$**

Observing an increasing variation of response variable as the values of explanatory variable increase is common in forestry (Huang et al., 1992). That the variation in total height  $H$  increases as the  $DBH$  increases is obvious from the scatter plot of  $H$  against  $DBH$  (Figure 1). Using a weighting factor in the regression analysis has been the standard approach to correct this problem in forestry (Meng and Tsai 1986). Generally the weight selected is inversely proportional to the error variance (Huang et al., 1992). However, weighted regression changes the estimates of parameters and the standard errors of the estimates (Ratkowsky, 1990). Weighted statistics are not as straightforward as un-weighted statistics and it is difficult to make clear interpretations (Carroll and Ruppert, 1988).

Theoretically, the weights are required to be known exactly in weighted regression (Huang et al., 1992). This is not the case in many practical applications and estimated



weights must be used instead. It is usually difficult to evaluate the effect of having to use estimated weights. While a slight variation in the weights usually do not affect a regression analysis significantly, if the weights are estimated from a small sample, the results of an analysis could be severely affected (Carroll and Ruppert, 1998). Weighted least squares (WLS) estimators of model parameters are also sensitive to the effects of outliers.

For the nonlinear model B, total height  $H$  is taken to be normally distributed with mean height  $\mu$  and variance  $\sigma^2 = \delta \times DBH^{2k}$ , that is,  $H \sim N(\mu, \sigma^2)$ , where

$\mu = 1.3 + e^{a_h + b_h \times DBH^{c_h}}$ . This is equivalent to the nonlinear regression

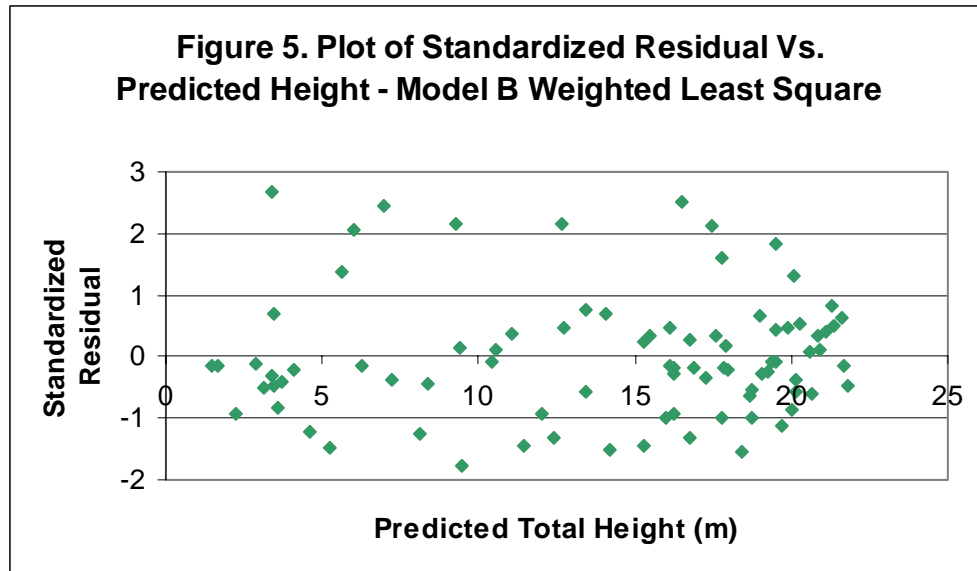
model  $H = \mu + \sqrt{\omega(DBH)} \times \varepsilon$ , where  $\varepsilon \sim N(0, \sigma^2)$  and  $\omega(DBH) = DBH^{2k}$  is a weight

function. To perform weighted nonlinear regression both sides of the nonlinear regression

model are divided by  $\sqrt{\omega(DBH)}$  which gives  $H^* = \mu^* + \varepsilon$ , where  $H^* = H / \sqrt{\omega(DBH)}$

and  $\mu^* = \mu / \sqrt{\omega(DBH)}$ .

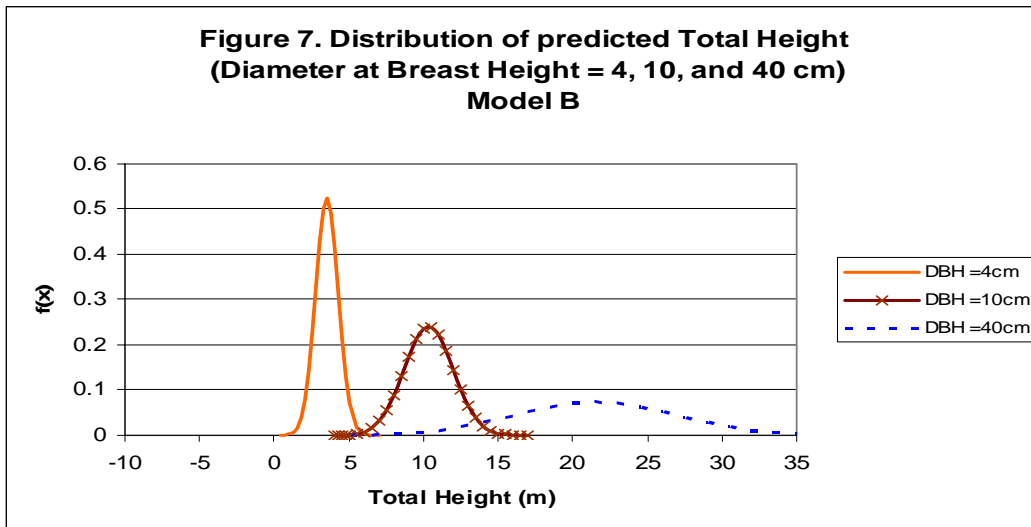
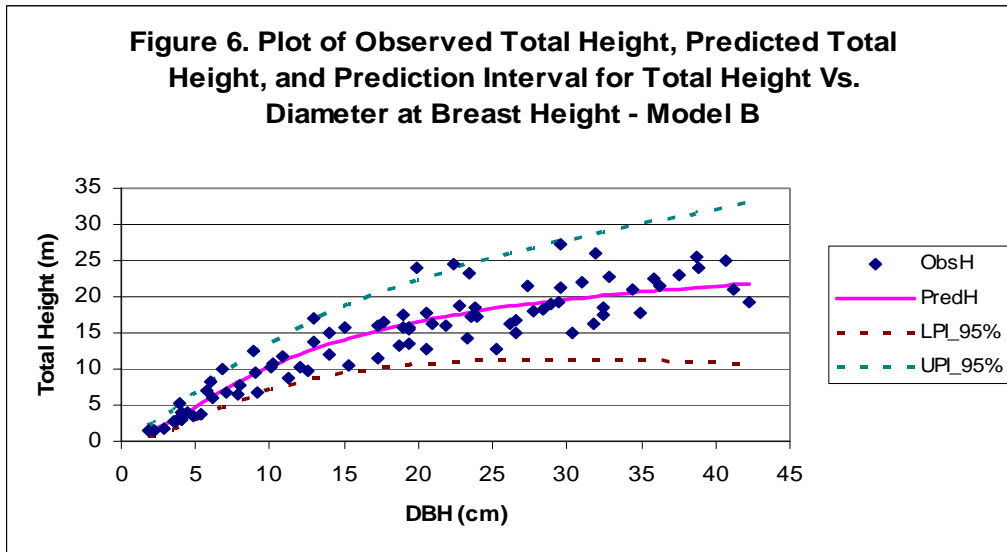
The weighted nonlinear regression approach (Model B) was applied to correct the problem of non-constant variance that was prevalent in Model A. Huang et al. (1992) used a weighting factor of the form  $w_i = DBH_i^{2k}$  as it required only a single unknown parameter and goes to 0 as DBH goes to 0. The nonlinear Model B was fitted with WLS using ten alternative values for  $k$  ( $k = 0.5, 0.6, 0.7, 0.75, 0.8, 0.85, 0.9, 1, 1.5, 2.0$ ). The highest stabilization of error variance was achieved with  $k = 0.85$  (Figure 5). This weight is also similar to the weight applied by Larsen and Hann (1987), and Farr et al. (1989).



The model parameter estimates for the mean were obtained by maximizing the normal log likelihood where the mean and variance of height are simultaneously specified as  $1.3 + e^{a_h + b_h \times DBH^{c_h}}$  and  $\delta \times DBH^{2 \times 0.85}$ , respectively. The NLMIXED procedure allows for specification for all parameters of a distribution not just the mean. This way, difficulties with transforming the response variable to obtain comparable AIC values were avoided. All the parameters for Model B were significant (Table 1). After the application of WLS, the plot of standardized residuals against predicted total height revealed that the error variances were approximately constant throughout the range of predicted height (Figure 5); however, Kolmogorov-Smirnov test of normality (Table 1) indicated that the errors did not follow the standard normal distribution (p-value < 0.010). Therefore, the WLS approach was appropriate to correct the problem of non-constant error variance; however, non-normality of the errors still remains a problem. The fitted line for prediction does go through the center of the observed data suggesting that the predictions for the mean height achieved from Model B with weight  $DBH^{2 \times 0.85}$  are acceptable. The prediction interval for total

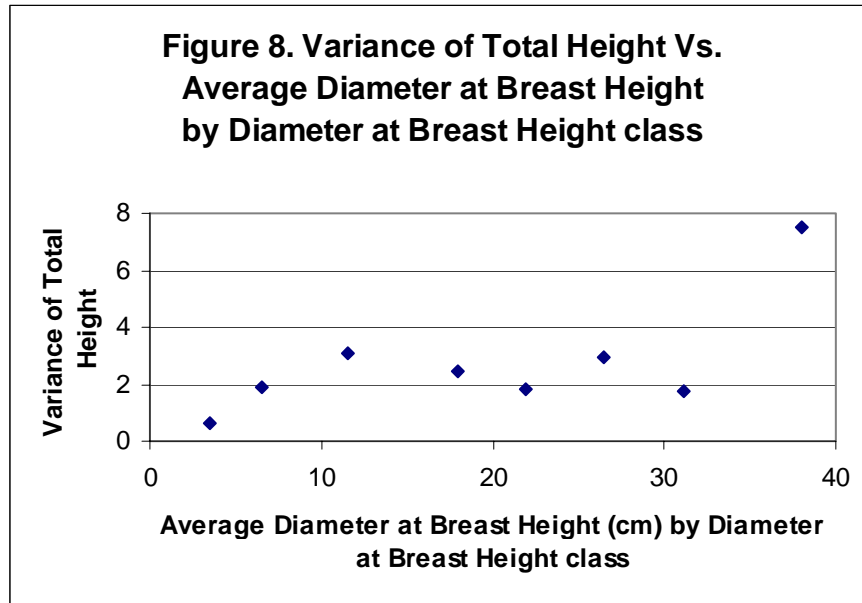
height increases as *DBH* increases across the range of *DBH* (Figure 6) which agrees with the behavior of the observed data in our study. The AIC for the model is lower than model A indicating a slight improvement in the model.

Figure 7 reveals that as the *DBH* increases both mean and variance of total height also increase. This is obvious from the shift in mean with different spread of the normal distribution plots obtained for *DBH* 4, 10, and 40 cm in the figure.



**C. Model for the prediction of total height as a nonlinear function of *DBH* and also variance of total height as a nonlinear function of *DBH***

Since NLMIXED provides MLE of all parameters simultaneously, the variance of *H* can be any function of *DBH*, not just the limited power function used in weighted regression. A model for predicted mean *H* was developed based on the analysis of the scatter plots of *H* against *DBH* (Figure 1). A plot of variance versus *DBH* was made by dividing the data into eight classes. A nonlinear function is evident (Figure 8), but the variance appears to approach an asymptote instead of going to infinity as assumed in WLS where  $\sigma^2 = \delta \times DBH^{2k}$ . Therefore, an exponential function similar to that used for  $\mu$  was used for variance.

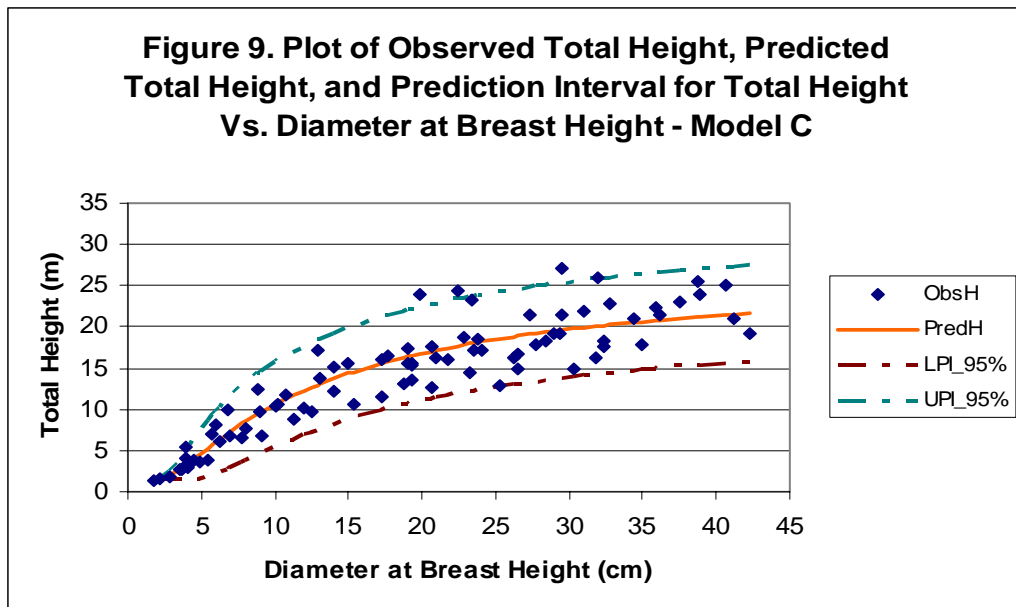


Total height *H* is taken to be normally distributed with mean height  $\mu$  and variance  $\sigma^2$ , that is,  $H \sim N(\mu, \sigma^2)$ , where the mean height of *H* is given by

$$\mu = 1.3 + e^{a_h + b_h \times DBH^{c_h}} \text{ and the variance of } H \text{ is given by } \sigma^2 = \delta \times e^{b_v \times DBH^{c_v}} . \text{ This model}$$

predicts a total height of 1.3 m and variance of height equal to zero when  $DBH$  approaches zero (Figure 9). Both the mean and variance of total height approach asymptotes as the  $DBH$  gets very large implying both the mean and variance become independent of  $DBH$ . In the model above  $a_h$ ,  $b_h$ ,  $c_h$ ,  $b_v$ ,  $c_v$  and  $\delta$  are unknown parameters to be estimated.

All the parameters for model C were found to be highly significant ( $p$ -value $<.0001$ ) and the AIC for model C was 373.7 which is less than either model A or B (Table 1). A plot of the standardized residual (Figure 10) indicated that the problem of non-constant variance has been corrected. The test of normality (Kolmogorov-Smirnov) for standardized residual was only marginally significant under the null hypothesis of a standard normal ( $p$ -value=0.0611) (Table 1). Similar to model A and B this model was able to predict the mean total height very well. Prediction intervals of total heights increased as the  $DBH$  of a tree increased and approached a constant as  $DBH$  increased (Figure 10). The variance of total height also increases from a value of zero to a constant spread in the prediction as  $DBH$  increases.



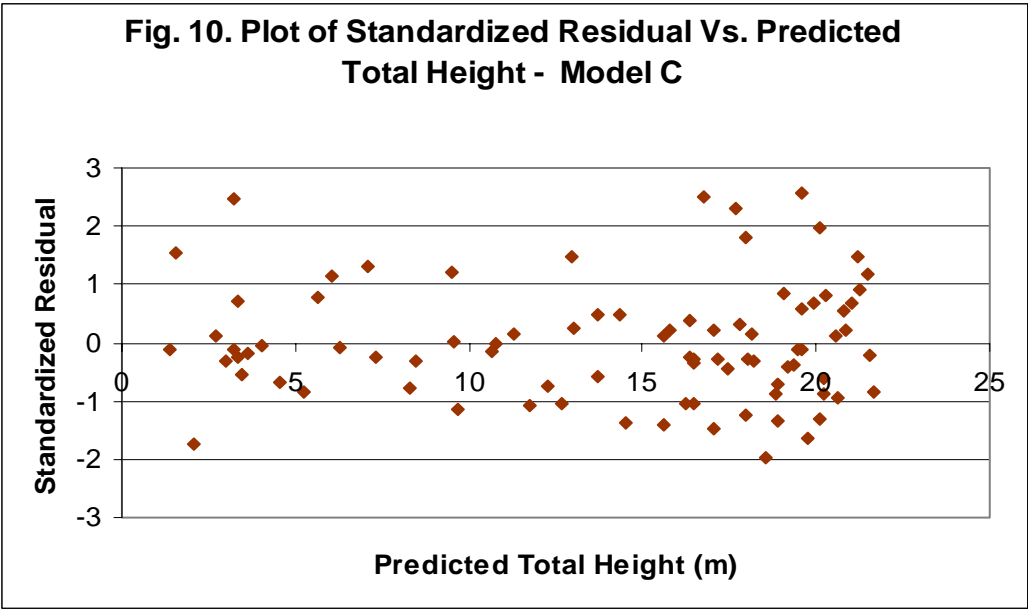
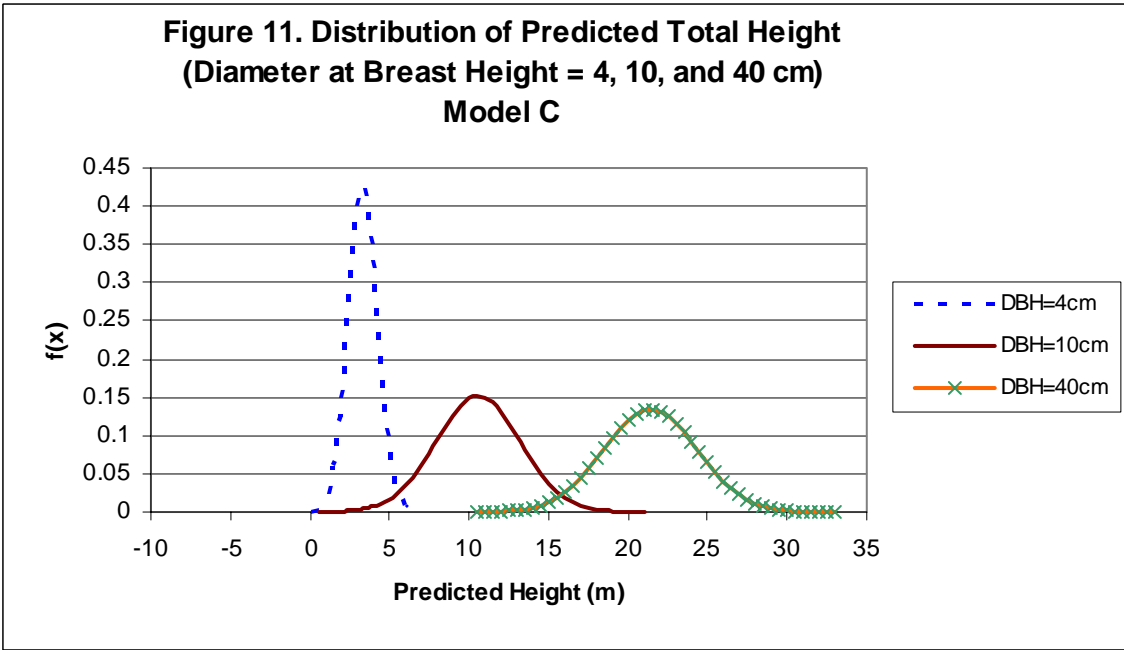


Figure 11 depicts that as *DBH* increases the mean height and variance of height also increases, however, the variance tends to stabilize at higher *DBH*.



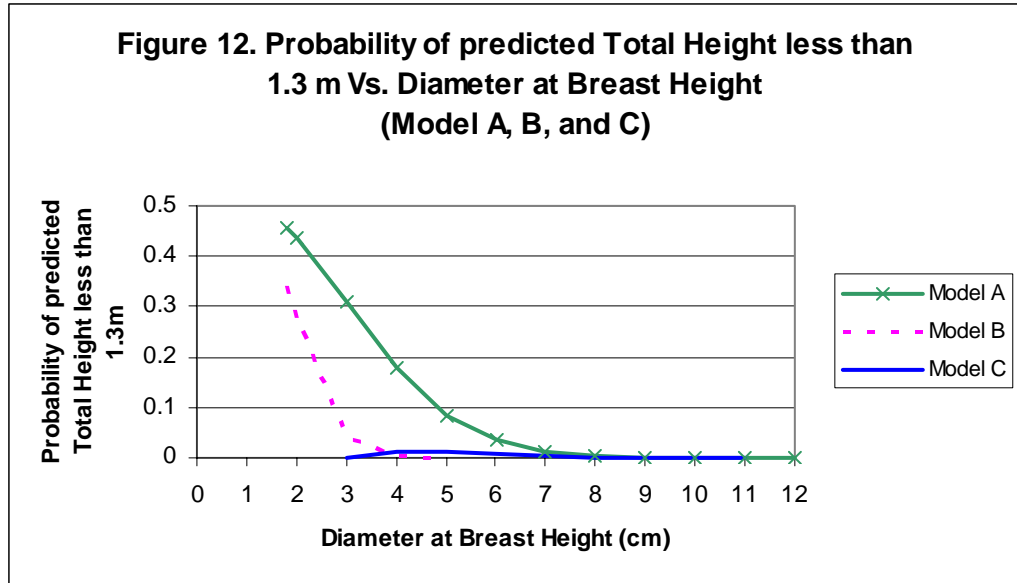
A cross-validation simulation using the fitted Model C was run 1000 times in SAS to test the ability of Model C to predict the distribution of total height based on *DBH*

measurements in the field. During the simulation 70 training data points were selected randomly out of 85 using the *Survey Select* procedure of SAS and the model was fitted to estimate the parameters. The estimated parameters were then used to predict  $H$  for the independent (test) data set with 15 data points. Then the standardized residuals were computed and tested against the standard normal (K-S) at the 0.05 level. The result obtained from the simulation indicated that p-values  $< 0.05$  were obtained approximately 28% of the time. For stable models this observed rejection rate should approximate the rejection rate used in each simulation, 5%. An observed rejection rate of 28% indicates model C does not consistently produce standardized residuals that follow a standard normal distribution.

### **2.3.2 MODELS THAT DO NOT REQUIRE THE ASSUMPTION OF NORMALITY OF THE RANDOM ERROR**

#### **D. Gamma Model for the prediction of total height as a function of $DBH$**

Figure 12 gives the plot of the probability that total height is less than 1.3 m estimated by the three models for  $DBH$  values in the range of 2 to 12 cm. It reveals that the probability of getting a predicted total height of less than 1.3 m is the highest for model A followed by models B and C, respectively. For model A the  $\Pr(H < 1.3 \text{ m})$  is the highest (45%) when a tree has a  $DBH$  of approximately 2 cm and the probability declines sharply and approaches 0 as  $DBH$  approaches 8 cm. Similarly, for model B the estimated  $\Pr(H < 1.3 \text{ m})$  is the highest (35%) when a tree has a  $DBH$  of approximately 2 cm and the probability declines sharply and approaches 0 as  $DBH$  approaches 4.5 cm. Interestingly for model C



the estimated  $\Pr(H < 1.3 \text{ m})$  is almost 0 when a tree has the *DBH* less than or equal to 3 cm, the probability is highest (1.5 %) when a tree has a *DBH* of approximately 4 cm and the probability declines and approaches 0 as *DBH* approaches 7 cm. Since predicted height less than 1.3 m for trees with positive *DBH* is not possible, this property of these models argues against a normal distribution for the error. A more applicable assumption would be a distribution with a minimum value such as the gamma distribution.

A random variable  $X$  is said to follow the three parameter gamma distribution if its probability density function (pdf) is

$$f(x | \gamma, \alpha, \beta) = \frac{(x - \gamma)^{\alpha-1} \exp\left(-\frac{x - \gamma}{\beta}\right)}{\beta^{\alpha} \Gamma(\alpha)} \quad x \geq \gamma; \alpha, \beta > 0,$$

where  $\alpha$  is a shape parameter,  $\beta$  is a scale parameter, and  $\gamma$  is a location parameter.

Since the model must predict a total height of 1.3 m when *DBH* approaches zero, the location parameter  $\gamma$  was set to 1.3. The shape parameter  $\alpha$  influences the peakedness of



the distribution, while the scale parameter  $\beta$  influences the spread of the distribution. The mean of the  $G(1.3, \alpha, \beta)$  distribution is  $\alpha\beta + 1.3$  and the variance is  $\alpha\beta^2$ .

A convenient reparameterization of the  $G(1.3, \alpha, \beta)$  density to allow direct modeling of the mean would result from defining  $\alpha = 1/\phi$  and  $\beta = (\mu - 1.3)\phi$ . This results in the pdf

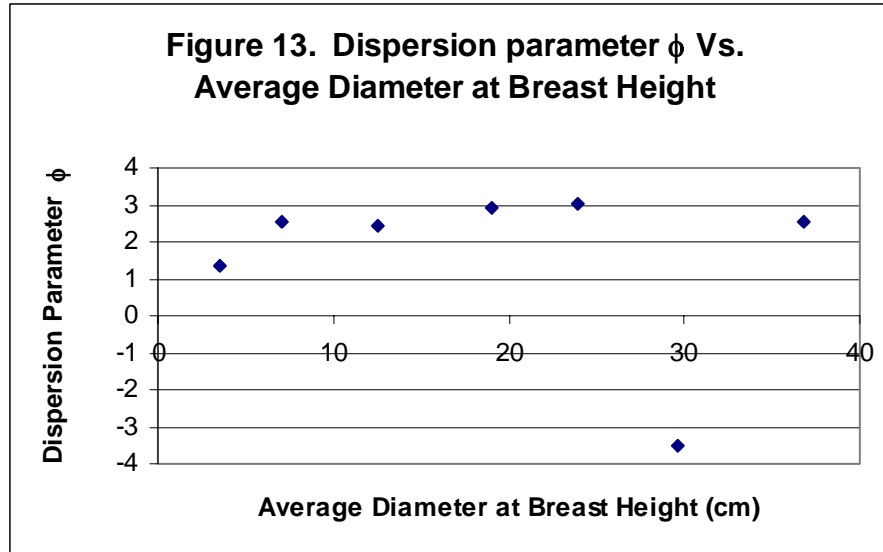
$$f(H | 1.3, \mu, \phi) = \frac{(H - 1.3)^{(\frac{1}{\phi})-1} \exp^{-\frac{(H-1.3)}{\phi(\mu-1.3)}}}{(\phi(\mu - 1.3))^{\frac{1}{\phi}} \Gamma\left(\frac{1}{\phi}\right)}.$$

Maximizing the gamma log likelihood function  $\lambda(1.3, \mu, \phi | H)$  with  $\mu = 1.3 + e^{(a_h + b_h \times DBH^{c_h})}$  and  $\phi$  unspecified, the log likelihood function for individual observation that is distributed according to the gamma pdf above is given by

$$\lambda(1.3, \mu, \phi | H) = \frac{\log(H - 1.3) - \log \phi - \log(\mu - 1.3)}{\phi} - \frac{(H - 1.3)}{(\mu - 1.3) \times \phi} - \log(H - 1.3) - \log \Gamma\left(\frac{1}{\phi}\right)$$

where  $\phi$  is the dispersion parameter,  $\mu$  is the mean height,  $H$  is the total observed height with a range  $[1.3, \infty)$ .

The data set was divided into 7 classes based on  $DBH$ . The mean and  $\phi$  of the Gamma distribution for each  $DBH$  class was computed using NLMIXED. Figure 13 indicates that  $\phi$  is not a function of  $DBH$ . The same functional form for  $\mu$  as the previous model was used. NLMIXED in SAS was used to fit the Gamma model (Model D). All the parameters were highly significant (p-value < .0001) (Table 1).



For the gamma distribution, a standardized deviance residual was used which is

computed as  $d_{i,r} / \sqrt{\hat{\phi} \times (1 - h_{ii})}$  where,

$$d_{i,r} = [\text{sign}((y_i - 1.3) - (\hat{\mu}_i - 1.3)) \sqrt{2 \times ((y_i - 1.3) - (\hat{\mu}_i - 1.3)) / (\hat{\mu}_i - 1.3)} - \log((y_i - 1.3) / (\hat{\mu}_i - 1.3))] / \sqrt{\hat{\phi} \times (1 - h_{ii})}$$

,  $d_{i,r}$  is deviance residual,  $h_{ii}$  is the diagonal element of the hat matrix,  $y_i$  is the observed

height, and  $\hat{\mu}_i$  is the predicted height (McCullagh and Nelder, 1989). While the Pearson

residual is appropriate and equivalent to the deviance residual for normally distributed

errors, it is not appropriate when the errors are not symmetric, such as gamma distribution.

The reason is that the contribution of each observation to the deviance is a chi-square

distributed quantity whose square root (with the right sign) is a normally distributed

random variable (McCullagh and Nelder, 1989).

Analysis of standardized deviance residual plot revealed that the error variance is

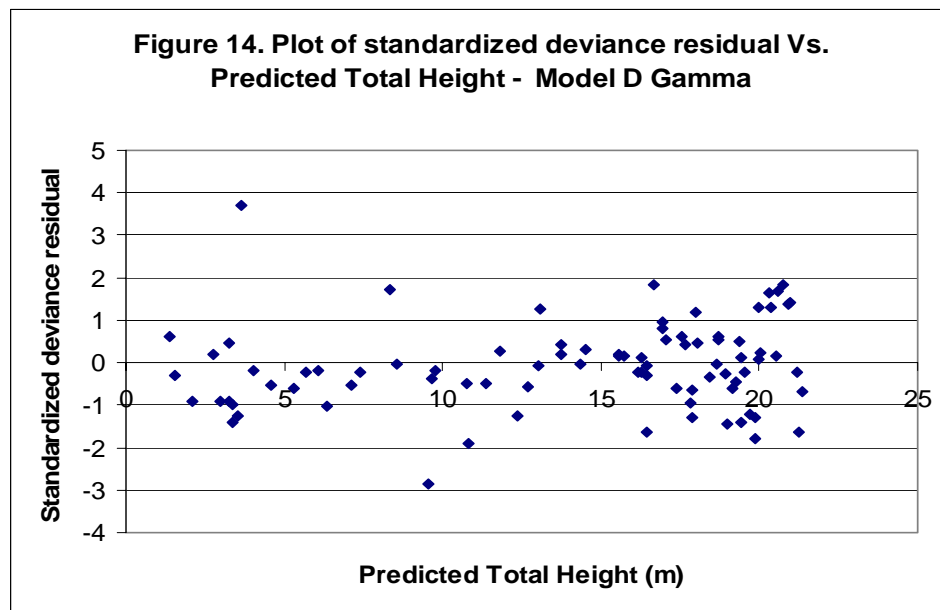
approximately constant throughout the range of the predicted total height (Figure 14) and

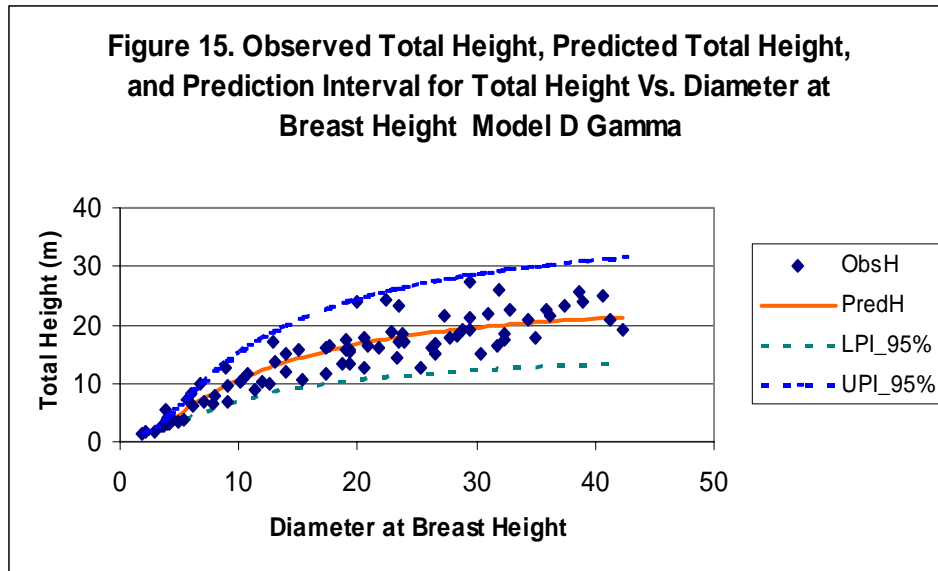
Kolmogorov-Smirnov test for normality (Table 1) revealed that the standardized deviance

residuals were normal (p-value > 0.15). Figure 15 indicated that predicted height increased

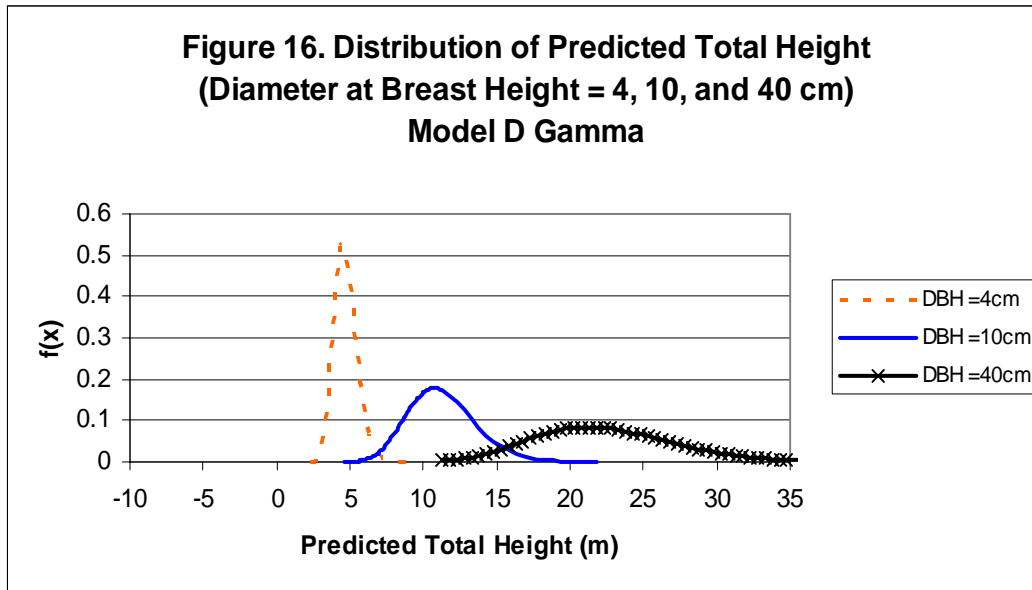
as the *DBH* increased; however, the rate of increase in total height decreased as the *DBH* attained a value of 30 cm or greater and approached an asymptote. The fitted line for prediction passes through the center of the observed data indicating that the predictions achieved from the Gamma model are performing as expected for the prediction of mean. AIC for the model was 380.2 which is larger than model C.

It becomes obvious from Figure 15 that the width of the prediction interval increases as *DBH* increases and tends to be constant as *DBH* gets larger which agrees with the observed data and is very similar to model C. As *DBH* goes to infinity  $\mu$  goes to the asymptotic value of  $1.3 + e^{a_h}$ . The variance of the gamma distribution is  $(\mu - 1.3)^2 \phi$ . The variance also approaches an asymptotic value of  $e^{2a_h} \phi$ . As *DBH* goes to zero,  $\mu$  goes to 1.3 m and variance goes to zero. These are the same properties as model C.





The distribution of the predicted total height obtained from the gamma distribution (Figure 15) depicted that the range of simulated total height was reasonable and the majority of the values were centered close to the observed value of total height in the data set for the corresponding *DBH*. The distribution of total height using estimated parameters from the gamma model with *DBH* of 4, 10, and 40 cm were obtained (Figure 16). As discussed earlier, it is theoretically impossible to get values of mean total height that are less than 1.3 m. Furthermore, since this model is restricted to values greater than 1.3 m, no simulated heights below 1.3 m are possible.



A simulation using the fitted Gamma model was also run 1000 times in SAS to test the predictive ability of this model in a cross validation. Once again 70 training data points were randomly selected out of 85 using *Survey select procedure* in SAS and the model was fitted to estimate the parameters. The estimated parameters were then used to predict  $H$  from the independent data set of the remaining 15 data points. Then the standardized deviance residuals were computed and tested against the standard normal distribution using the Kolmogorov-Smirnov test. The result obtained from the simulation indicated that p-values  $< 0.05$  was obtained approximately 9% of the total number of runs in the simulation. This is close to the nominal level of 5%. It should also be noted that model D performed slightly better than model C with only 4 parameters as compared to model C with 6 parameters.

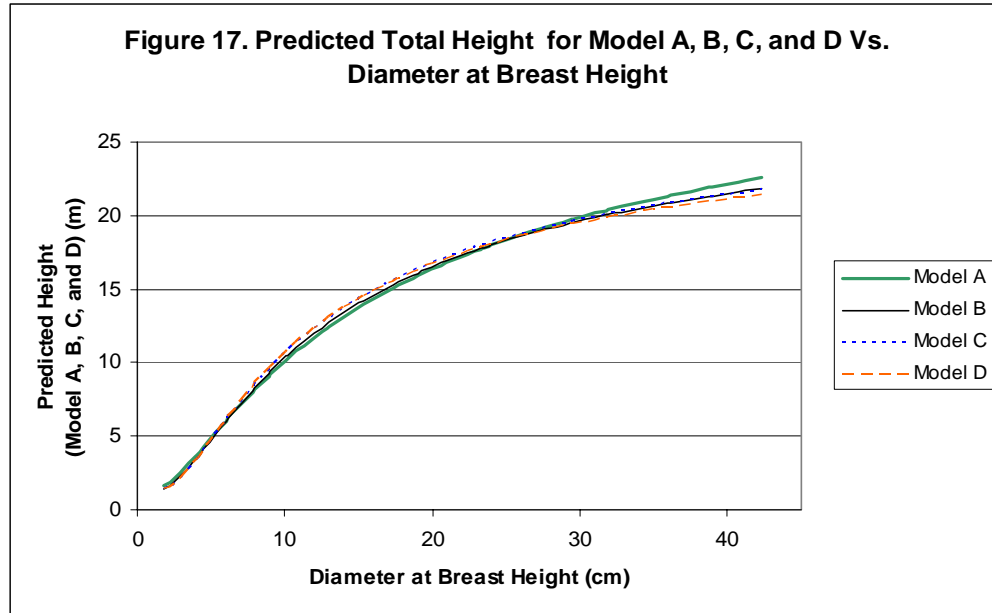
## 2.4 DISCUSSION AND CONCLUSION

Models with the same mean and different formulations of variance were used to compare the performance of the models for the prediction of the distribution of total height (table 2).

**Table 2. Models used to fit the data.**

Model	Distribution	Location (Mean)	Variance
A	Normal	$\mu = 1.3 + e^{a_h + b_h \times DBH^{c_h}}$	$\sigma^2$
B	Normal	$\mu = 1.3 + e^{a_h + b_h \times DBH^{c_h}}$	$\delta \times DBH^{2k}$
C	Normal	$\mu = 1.3 + e^{a_h + b_h \times DBH^{c_h}}$	$\delta \times e^{(b_v \times DBH^{c_v})}$
D	Gamma	$\mu = 1.3 + e^{a_h + b_h \times DBH^{c_h}}$	$(\mu - 1.3)^2 \phi$

It became evident that all the models were performing similarly for the purpose of predicting the mean height (Figure 17); however, there was a considerable difference in their ability to predict the distribution of total height.



Analysis of the standardized residuals for Model A revealed that they did not follow a standard normal distribution (Table 1), and also the error variance was not constant throughout the range of predicted  $H$  (Figure 2). This model had the highest AIC. The prediction interval also included values below 1.3 m of total height for smaller values of  $DBH$  (Figure 3). For example, a tree with  $DBH$  of 4 cm had a distribution of height with values below 1.3 m (Figure 4). Therefore, this was not an appropriate model for the prediction of total height.

Weighted regression model with the best weight was fitted in SAS to correct the problem of non-constant variance. Analysis of residuals after the application of weighted regression revealed that the problem of unequal error variance was corrected (Figure 5); however, the standardized residuals still did not follow a standard normal distribution (Table 1). The AIC was less than the AIC for Model A. The prediction interval for the total height increased as the  $DBH$  increased throughout the range of the  $DBH$ , however, the width of the prediction interval got very large near the larger end of the  $DBH$  which did not

agree with the observed data. It became evident from Figure 6 and 7 that the prediction interval included the values of total height less than 1.3 m for the smaller *DBH* e.g. *DBH* of 4 cm (Figure 7) though it is theoretically not possible to get a tree with total height of less than 1.3 m even when *DBH* approaches zero. Therefore, this was also not an appropriate model for the prediction of distribution of total height.

Next a non-linear regression model C was tried, where, predicted *H* was as a non-linear function of *DBH* and variance was also a non-linear function of *DBH*. Assumptions of constant variance and normality of error were checked using residual plot and Kolmogorov-Smirnov tests, respectively. Analysis of standardized residuals depicted that error variance was stabilized (Figure 9); however, the small p-value of the Kolmogorov-Smirnov test of normality is suggestive of non-normality (Table 1). Model C indicated that the prediction interval for the total height increased as the *DBH* increased up to a point and then stayed constant throughout the range of the *DBH* which agreed with the observed data. The AIC was the lowest among all the models; however, it became evident from Figure 10 and 11 that the prediction interval included the values of total height less than 1.3 m for the smaller *DBH* e.g. *DBH* of 4 cm (Figure 11). Type I Error rate obtained from running a simulation 1000 times was 28%. This error rate was very high compared with the nominal Type I Error rate of 5% in the study. Model C was a better model than A and B, however, this was also not an adequate model for the prediction of distribution of total height.

Models with the assumption of normality of errors are reasonable for the prediction of the mean height; however, they are not suitable for estimation of the distribution of total height since the range of a normally distributed random variable spans from negative infinity to positive infinity. So, there is a possibility of obtaining negative predictions,



which are not realistic. The normal distribution is symmetric and bell shaped distribution and hence cannot explain any kind of skew ness (positive or negative). In order to overcome these problems a gamma distribution was assessed since it does have a minimum value and has the potential for significant skewness.

The gamma model (Model D) was found to be the best model among all the nonlinear models used to fit the data in this study. This model needed the estimation of only four parameters to predict the distribution of  $H$ ; whereas, the competing model C needed six parameters to be estimated. This model had all the parameters highly significant (Table 1). The AIC was 380.2 which is close to the AIC for model C. Analysis of the plot of standardized residuals (Figure 14) revealed that the error variance was constant across the whole range of predicted values of total height and the residuals were normally distributed (Table 1). The prediction interval was found to be reasonable and the prediction interval did not include a value of total height less than 1.3 m even when  $DBH$  approached zero. The error rate obtained from the simulation of the model for cross validation using Kolmogorav-Smirnov test was reasonable (approximately 9%). This was close to the nominal Type I Error rate of 5% in the study which reveals that the gamma model agrees with the data to a greater extent than the other models. Therefore, this model was adequate to fit the data well and was the best model for the prediction of distribution of total height.

The MLE technique in NLMIXED procedure makes it possible to model mean  $\mu$  and variance  $\sigma^2$  of total height simultaneously in any distribution. This procedure empowers the modeler to design models to predict the distribution of total height when the mean and variance are both nonlinear functions of  $DBH$ . WLS was a first attempt to model the variance as a function of independent variable in the study. It is restricted to a single

parameter that must be estimated outside the fitting procedure; however, NLMIXED allows the specification of the functional form of the variance along with the functional form of mean and obtain the MLE estimates of all the coefficients simultaneously. While the normal assumption for the error term is reasonable for the prediction of the mean response for a given predictor and the confidence interval on the prediction, it is no longer necessary to be restricted to a normal assumption on the error. A normal distribution assumption can have serious problems when modeling the entire distribution for simulation purposes e.g. it can produce predictions lower than theoretically possible and the distribution is restricted to be symmetric about the mean response. NLMIXED can handle any distribution as long as the log likelihood function can be specified e.g. gamma distribution which has a minimum value and can model skewed error distributions.

## 2.5 LITERATURE CITED

Anderson, D. R., Burnham, K. P., and Thompson, W. L., 2000. Null hypothesis testing: Problems, prevalence, and an alternative. *J. Wildl. Manage.* 64, 912-923.

Anderson, M., Somers, G. L., Smith, W. R., Freeland, M., and Ruch, D., 1999. Measuring crown dynamics of longleaf pine in sandhills of Eglin Air Force Base. *Proceedings of the Second Longleaf Alliance Conference 1998. Charleston, 17-19 Nov.,* 46-48.

Akaike, H., 1973. Information theory as an extension of the maximum likelihood principle. In: *Second International Symposium on Information Theory* (Eds: B. N. Petrov and F. Csaki). Akademiai Kiado, Budapest. pp 267-281

Arabatzis, A.A., and Burkhart, H.E., 1992. An evaluation of sampling methods and model forms for estimating height-diameter relationships in loblolly pine plantations. *For. Sci.* 38, 192-198.

Botkin, D.B., Jamal, J.F., and Wallis, J.R., 1972. Some ecological consequences of a computer model of forest growth. *J. Ecol.* 60, 849-873.

Boyer, W.D., 1991. *Pinus palustris* Mill. (Longleaf Pine). In: *Silvics of North America 1.* (Eds: Bruns, R. M. and Honkala, B. H. Tech. Coords. ). Conifer. Agric. Handb. 654. Washington, DC: U.S. Department of Agriculture. pp 405-412.

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

Carroll, R.J., and Ruppert, D., 1988. Transformation and weighting in regression. Chapman and Hall, London.

- Casella, G., and Berger, R. L. 2002. *Statistical Inference*. Duxbury, Thompson Learning Inc. 660pp.
- Curtis, R.O., 1967. Height-diameter and height-diameter-age equations for second-growth Douglas-fir. *For. Sci.* 13, 365-375.
- Curtis, R.O., Clendenin, G.W., and DeMars, D.J., 1981. A new stand simulator for coast Douglas-fir: DFSIM user's guide. USDA For. Serv. Gen. Tech. Rep. PNW-128.
- Farr, W.A., DeMars, D.J., and Dealy, J.E. 1989. Height and crown width related to diameter for open-grown western hemlock and Sitka spruce. *Can. J. For. Res.* 19, 1203-1207.
- Hollander, M., and Wolf, D. A., 1999. *Nonparametric Statistical Methods*. Second Edition. John Wiley and Sons, Inc. 787 pp.
- Huang, S., Titus, S.J., and Wiens, D.P., 1992. Comparison of nonlinear height-diameter functions for major Alberta tree species. *Can. J. For. Res.* 22, 1297-1304.
- Kush, J. S., Boyer, W. D., Meldahl, R. S., and Ward, G. A., 1999. Precommercial thinning intensity in longleaf pine: effect on product volume and value. In: *Proceedings of the second Longleaf alliance conference* (Eds: Kush, J. S.). Longleaf Alliance Report No. 4. 196 pp.
- Lappi, J. 1997. A longitudinal analysis of height/diameter curves. *For. Sci.* 43: 555-570.
- Larsen, D.R., and Hann, D.W., 1987. Height-diameter equations for seventeen tree species in southwest Oregon. Oregon State Univ., For. Res. Lab. Pap. 49, 16 pp.
- Lilliefors, H. W., 1967. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *American Statistical Association Journal*, 399-402.

- McCullagh, P., and Nelder, J. A. 1989. *Generalized Linear Models*, 2<sup>nd</sup>. London: Chapman & Hall. 511 pp.
- Meldahl, R. S., Kush, J. S., Rayamajhi, J. N., and Farrar, R. M. 1998. Productivity of natural stands of Longleaf pine in relation to competition and climatic factors. In: *The productivity and sustainability of southern forest ecosystems in changing environment* (Eds: Mickler, S. F., Fox, S.). Springer-Verlag New York, Inc. 892 pp.
- Meng, C.H., and Tsai, W.Y., 1986. Selection of weight for a weighted regression of tree volume. *Can. J. For. Res.* 16, 671-673.
- Meyer, H.A. 1940. A mathematical expression for height curves. *J. For.* 38, 415-420.
- Moore, J., Zhang, A., and Stuck, L. D., 1996. Height–diameter equations for ten tree species in the Inland Northwest. *West. J. Appl. For.* 11, 132-137.
- Overing, J. D., and Watts, F. C., 1989. *Soil survey of Walton County, Florida*. USDA, Soil Conservation Service, 234 pp.
- Ratkowsky, D.A., 1990. *Handbook of nonlinear regression*. Marcel Dekker, Inc., New York.
- SAS Institute, Inc., 2004. *SAS/STAT User's guide: Statistics, version 9.1.*, Cary, N.C. 5136 pp.
- Vanclay, J.K., 1994. *Modeling forest growth and yield: Applications to mixed tropical forests*. CAB International, United Kingdom, 329 pp.
- Wang, C.H. and Hann, D.W., 1988. Height–diameter equations for sixteen tree species in the central western Willamette valley of Oregon. *Oregon State Univ., For. Res. Lab. Res. Pap.* 51, 7 pp.

Zhang, L., 1997. Cross-validation of non-linear growth functions for modeling tree height–diameter relationships. *Ann. Bot.* 79, 251-257.

## CHAPTER 3

### PREDICTION OF DISTRIBUTION FOR CROWN RATIO OF AN INDIVIDUAL TREE USING NORMAL VERSUS BETA DISTRIBUTION MODEL FOR LONGLEAF PINE IN EGLIN AIR FORCE BASE.

#### 3.1 INTRODUCTION

The ratio of live crown length to the total tree height is commonly known as crown ratio. Crown ratio (*CR*) is the characteristic usually used to describe the crown size (Hynynen, 1995) which is an important element of forest growth and yield models (Cole and Lorimer, 1994). According to Short and Burkhart (1992) and Valentine et al. (1994) it is frequently used to predict individual tree growth. In many diameter and height growth equations tree crown parameters are used as explanatory variables (Monseurd and Sterba, 1996) and crown ratio has been used as a predictor of tree vigor (Hasenauer and Monserud, 1996), stand density, competition and survival potential of trees within a forest (Oliver and Larson, 1996).

Tree crown ratio can be predicted directly from tree variables such as total height (*H*) and diameter at the breast height (*DBH*) (Hasenauer and Monserud 1996). It can also be predicted indirectly from estimates of the height to the base of live crown (*HBLC*) (Dyer and Burkhart 1987). Hasenauer and Monserud (1996) and Ritchie and Hann (1987) have used the logistic functions to predict the mean tree crown ratio which is always between 0

and 1. The logistic function with normal distribution of errors works under the assumption that the errors are normally, identically, and independently distributed with mean zero and a constant variance, i.e.  $\varepsilon \sim NIID(0, \sigma^2)$ . Theoretically, under the assumption of normality the error term can assume a value from negative infinity to positive infinity and the prediction interval for  $CR$  can be well beyond 0 and 1 which is not reasonable. Normal distribution approach works fine for the prediction of mean  $CR$ , however, it might not work for the prediction interval of  $CR$ , in all other circumstances.

An alternative to assuming a normal distribution of errors is to assume the errors follow a beta distribution. This distribution assures that the prediction intervals are always between 0 and 1 (Cassella and Berger, 2002). Therefore, this study is an attempt to compare the distribution of predicted  $CR$  using Normal versus Beta distributed errors with the same functional form for the mean  $CR$  (i.e. logistic). The fitted model obtained from this study can be used for simulation purposes when the entire distribution of predicted  $CR$  is needed.

The primary objectives of this study are to:

- i) Compare the prediction of distribution for  $CR$  using Normal versus Beta distributed errors for longleaf pine (*Pinus palustris* Mill.)
- ii) Assess the ability of these models to predict the distribution of crown ratio.



## **3.2 MATERIALS AND METHODS**

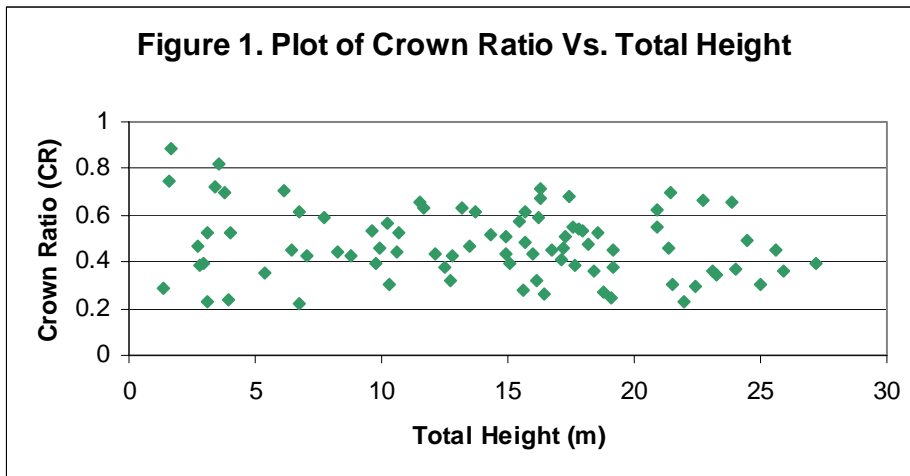
### **3.2.1 STUDY AREA AND SAMPLING OF TREES**

Longleaf pine is naturally distributed in most of the Atlantic and Gulf Coastal Plains from southeastern Virginia to eastern Texas and south through the northern two-thirds of peninsular Florida. This tree is also found in the Piedmont, Ridge and Valley, and mountain Provinces of Alabama and northwest Georgia (Boyer, 1991).

Longleaf pine grows in warm, wet temperate climates characterized by hot summers and mild winters. More frequent fires favor longleaf pines, which are more fire adapted (Boyer, 1991).

Data for this study were collected in naturally regenerated longleaf pine (*Pinus palustris* Mill.) stands at Eglin Air Force Base in Florida from September 1996 to March 2000. Total land area of Eglin Air Force Base is approximately 187,000 ha, most of which is forested. The major tree species occurring on the land base at Eglin are longleaf pine, sand pine (*Pinus clausa* (Chapm.) Vasey.), and various scrub oaks (*Quercus spp.*). The base is primarily composed of all-aged stands of longleaf pine. Frequent disturbances from fire arms testing maintain the stands in an open character. The major soil series include Lakeland, and Troup (Overing and Watts 1989). These soil series have deep, excessively drained, rapidly permeable soils and limited nutrients which result in very slow growth. This is the reason the trees face drought conditions during the summer even though the area receives more rainfall than any other part of Florida (Anderson et al., 1999). The larger trees are flat topped with a site index ranging from 16 to 19 m; whereas stands at the nearby Escambia Experimental Forest with similar age has a site index range of 18 to 26 m (Meldahl et al., 1998). The basal area per ha for Eglin in our study area was 12 m<sup>2</sup>/ha;

whereas the basal area per ha for Escambia Experimental Forest was 26-34 m<sup>2</sup>/ha. (Kush et al., 1999). Therefore, the trees are open grown with little crown competition resulting in a similar crown ratio throughout the entire range of the total height of the tree (Fig. 1) with comparatively higher variance of *CR* in the shorter trees (i.e. < 5 m ).



The sampling of longleaf pine trees covered the range of tree sizes, from saplings to the largest trees observed on the base, across the range of stand densities and sites. The trees were selected to cover the entire range of observed *CR* in the study area. The total height (*H*) ranged from 1.4 to 27.2 m and diameter at breast height (*DBH*) ranged from 1.8 to 42.3 cm. In 27 stands across the study area, 85 sample trees were selected at random locations. Stands were selected to represent all classes of site index, basal area per hectare, and tree density. Samples were selected from the stands which were naturally established free of major disturbances, primarily crown-scorching fires, within the past five years. Total height, *DBH*, and height to the base of live crown (*HBLC*) were measured for each tree. Trees with major defects including forking, major breakage, and weak stem were not selected as a sample. The observed *CR* was computed as  $CR = (H - HBLC) / H$ .

### 3.2.2 METHODS

Model fitting was carried out with the NLMIXED procedure using a quasi-Newton algorithm (SAS Institute, Inc. 2004). NLMIXED will produce maximum likelihood estimates of any likelihood function. This procedure allows all parameters to be functions of independent variables and can be used to find estimators of parameters in complex models. NLMIXED procedure uses the method of maximum likelihood (MLE) to obtain the estimates of the coefficients numerically. MLE selects as estimates the values of the parameters that maximize the likelihood of the observed sample. It is consistent, that is, as the sample size gets larger, the estimates converge to the true values. It is asymptotically efficient, which means that for large samples, it produces precise estimates. It is also asymptotically unbiased, that is, for large samples one expects to get the true value on average. The distributions of the estimators themselves are asymptotically normal, if the sample is large enough. These are all excellent large sample properties. For the normal distribution the least square method and MLE provide the same estimates of parameters (Casella and Berger, 2002).

The following criteria were used to compare the performance of height-diameter models:

1. Significant t-test for the parameters in the model.
2. AIC value.
3. Analysis of the standardized residual plots.
4. Kolmogorov-Smirnov test for normality of standardized residuals.
5. Simulated Type I Error rate using cross validation.

The asymptotic t-statistic for each coefficient in the model should be significant for any model to be considered appropriate for prediction of the parameters of the distribution. It tests the importance of an independent variable and whether it should be included in the model. The model with any of the parameters insignificant at a specified level of significance will be considered a *bad* model and hence discarded. Therefore, this is the first test that any candidate model has to pass to be considered for further analysis.

Akaike (1973) developed an estimation criterion, which he defines as

$AIC = -2 \log_e(l(\hat{\theta} | data)) + 2K$  where,  $\log_e l(\hat{\theta} | data)$  is the value of the maximized log-likelihood over the unknown parameter ( $\theta$ ).  $K$  is the number of estimated parameters included in the model (Anderson et al. 2000). AIC penalizes for the addition of parameters, and thus selects a model that fits well but has the smallest number of parameters. The model with the lowest AIC is selected as a good candidate for the data available (Anderson et al., 2000).

In the normal distribution the standardized residual is computed as,

$d_i = \frac{e_i}{\hat{\sigma}}, i = 1, 2, \dots, n$ , where  $e_i = y_i - \hat{y}_i$ ,  $y_i$  is observed CR, and  $\hat{y}_i$  is the predicted CR for the  $i$ th tree. The standardized residuals have mean zero and unit variance. They are useful in analyzing the problems of model specification and outliers. Most of the standardized residuals should lie in the range of  $-3 \leq d_i \leq 3$ .

Analysis of the standardized residual plots against the predicted value should show a more or less even scatter around the zero line throughout the range of the predicted values in which case equal error variance is a valid assumption. A model that does not suffer from non-constant error variance is considered to be good candidate for further analysis. Any

pattern in the standardized residual might reveal unequal error variance, or problem of model specification (Huang et al., 1992).

The Kolmogorov-Smirnov (KS) test was used to test whether the standardized residuals follow a standard normal distribution. The KS test is based on comparison of the empirical distribution function (ecdf) of the data and the cumulative distribution function (cdf) of an assumed distribution. It is non-parametric and distribution free test (Hollander and Wolf, 1999). Standardized residuals for a good model will follow a standard normal distribution and will pass this test.

A cross-validation simulation for the competing models was run 1000 times (Singh et al., 2006) in SAS to test the ability of the model to predict the distribution of *CR* based on *H* measurements in the field. During the simulation 70 training data points were selected randomly out of 85 using the *Survey Select* procedure of SAS and the models were fitted to estimate the parameters. The estimated parameters were then used to get the predictions for the distribution of *CR* from independent (test) data set with 15 data points. Then the residuals were computed and analysis was performed during each run of the simulation and p-values were checked for Kolmogorov-Smirnov test on the standardized residuals. Cross validation checks whether the distribution obtained from an independent data set follows a specified distribution (Lilliefors, 1967). The Type I Error rates of the KS test obtained from running the simulation is compared to the nominal Type I Error rate of 5% (Liu et al., 2006). A model was deemed a better predictive model than a competing model if its simulated Type I Error rate was closer to the nominal Type I Error rate than its competitor's simulated Type I Error rate.

### 3.3 RESULTS

Prediction of the mean crown ratio should always be between 0 and 1 for a model to work properly (Hynynen, 1995). The form of Logistic function commonly used to model crown ratio is  $CR = \frac{1}{1 + e^{-x}}$  (Hasenauer and Monserud, 1996; Temesgen et al., 2005), where  $CR$  is the crown ratio and  $x$  is a linear combination of explanatory variables and coefficients. Logistic function was first introduced in 1845 by Verhulst (Hasenauer and Monserud, 1996) for population growth and has been used to model individual tree mortality (Neter and Maynes, 1970) and crown ratio (Hatch, 1980, Ritchie and Hann, 1987, and Hasenauer and Monserud, 1996).

#### 3.3.1 LOGISTIC FUNCTION WITH NORMAL DISTRIBUTION OF ERRORS

##### 3.3.1.1 Normal Model A

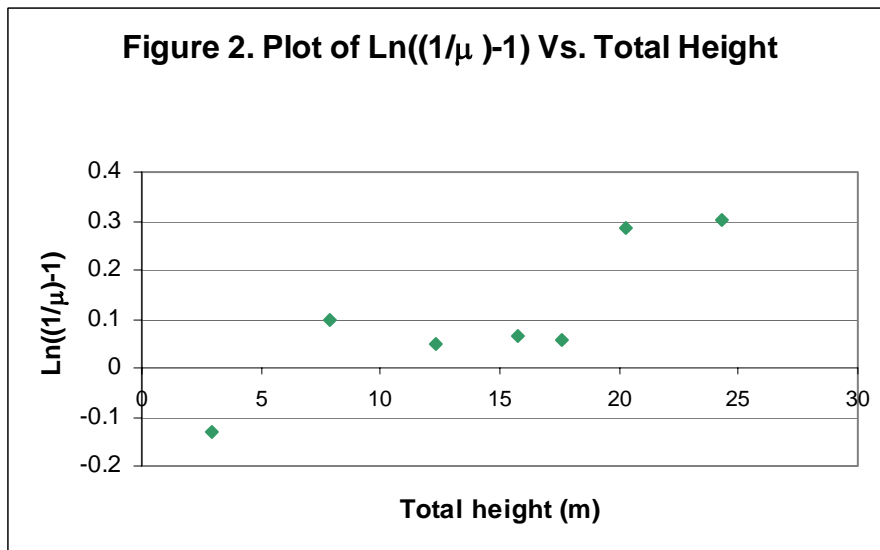
In order to investigate the appropriate functional relationship of mean  $CR$  to  $H$  the data was divided into 7 classes based on total tree height. Mean  $CR$  for each height class was computed and  $\ln((1/\mu)-1)$  against  $H$  was plotted to examine the functional relationship of mean  $CR$  to the total height. Figure 2 indicates that there is an approximate linear relationship between  $\ln((1/\mu)-1)$  and  $H$ .

The normal model assumes that tree  $CR$  is normally distributed with mean crown ratio  $\mu$  and variance in crown ratio  $\sigma^2$ , that is,  $CR \sim N(\mu, \sigma^2)$ , where  $\mu = \frac{1}{1 + e^{-(a+bH)}}$ ,  $\sigma^2$  is an unknown constant and  $a$  and  $b$  are unknown coefficients to be estimated from the data. The parameters of this model were estimated using the method of maximum likelihood (NLMIXED, SAS) or equivalently the log likelihood function

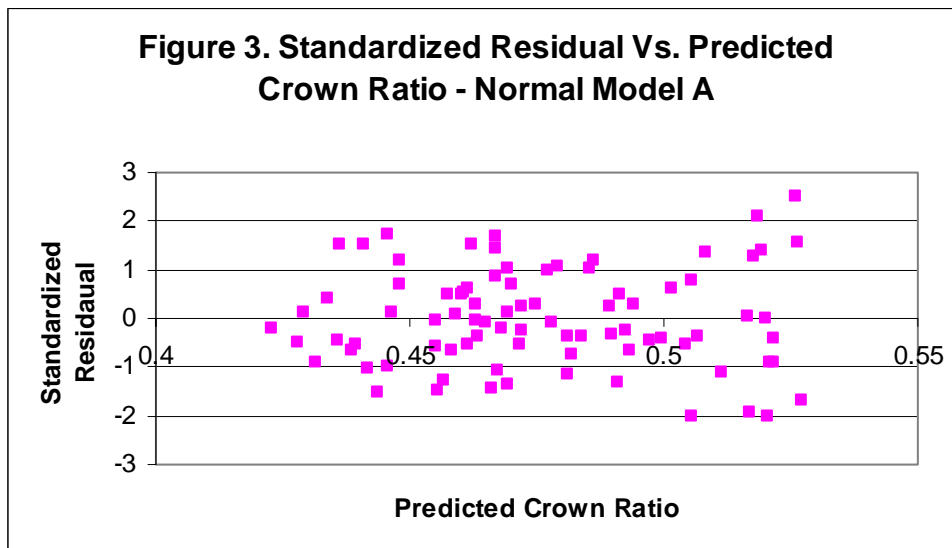
$$\lambda(\mu, \sigma^2 | CR_1, K, CR_n) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (CR_i - \mu)^2$$

A logistic model of the form  $\mu = \frac{1}{1 + e^{-(a+bH)}}$  was fitted initially, however, the coefficient  $a$  was found to be insignificant (p-val=0.35) and  $b$  was marginally significant (p-val = 0.07) at 5% level of significance. So, the coefficient  $a$  was set to 0 which forces the function through the origin (Figure 2). Thus the model was refitted as  $\mu = \frac{1}{1 + e^{-(bH)}}$  (Normal Model A).

Since tree density for Eglin in our study area is very low (12 m<sup>2</sup>/ha) and the site index limited to very low sites, inclusion of stand characteristics was not successful. Simple models using just  $DBH$  and  $H$  were compared. The AIC for the model with  $DBH$  was -80.9 and for  $H$  was -83.2. Interactions and other combinations of  $H$  and  $DBH$  were also attempted, however, they did not improve the model AIC. Therefore, the model with the explanatory variable  $H$  was selected as a potential candidate for further analysis.

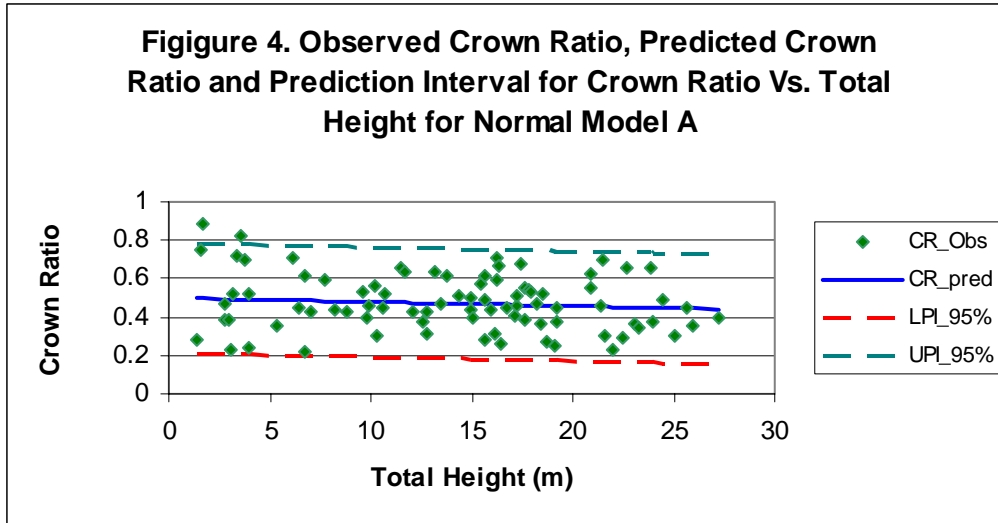


For the Normal model A, the parameter  $\sigma^2$  was highly significant (p-value < .0001) and the parameter  $b$  was also significant (p-value = 0.0345) at 5% level of significance (Table 1). The test of normality (Kolmogorov-Smirnov) revealed that the standardized residuals followed a normal distribution (p-value >0.15) (Table 1), however, the plot of standardized residual against predicted crown ratio (Figure 3) revealed a discernible trend which indicates a potential problem with model specification. The value of AIC was -84.4 (Table 1).



The fitted line for prediction of crown ratio goes through the center of the observed crown ratio (Figure 4). This is indicative of good predictive ability of Normal model A for the mean  $CR$ . This model assumes equal variance and therefore the prediction intervals for the entire range of predicted crown ratio are parallel (Figure 4) which is questionable for the smaller trees.



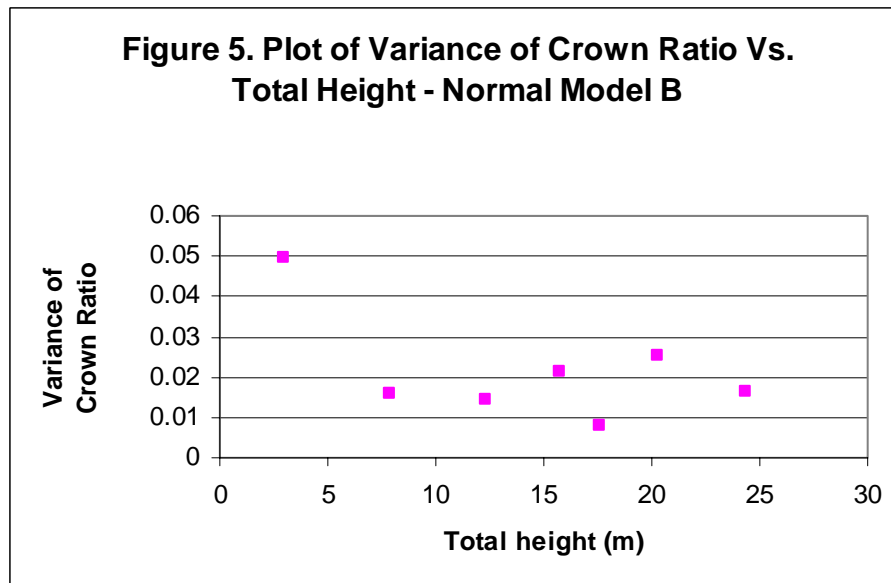


**Table 1. Model Parameter Estimates**

Model	Coeff	Estimate	P-val	LCL	UCL	AIC	K-S (P-val)
Normal A	$\hat{b}$	-0.00866	0.0345	-0.01667	-0.00065	-84.4	>0.1500
	$\hat{\sigma}^2$	0.02070	<.0001	0.01439	0.02701		
Normal B	$\hat{b}$	-0.00888	0.0130	-0.01584	-0.00192	-90.6	>0.1500
	$\hat{c}$	-4.3269	<.0001	-4.7465	-3.9072		
	$\hat{d}$	2.9355	0.0171	0.5351	5.3358		
Beta A	$\hat{b}$	-0.00836	0.0401	-0.01632	-0.00039	-85.8	>0.1500
	$\hat{\sigma}^2$	11.0598	<.0001	7.8279	14.2916		
Beta B	$\hat{b}$	-0.00882	0.0124	-0.01569	-0.00195	-94.7	>0.1500
	$\hat{c}$	2.9025	<.0001	2.5025	3.3026		
	$\hat{d}$	-3.0962	0.0053	-5.2486	-0.9438		

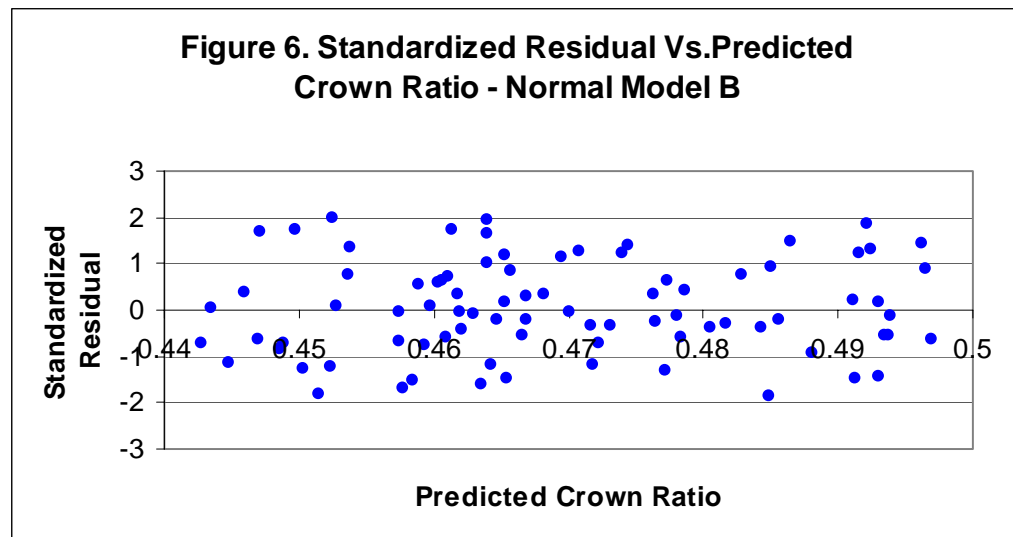
### 3.3.1.2 Normal Model B

In order to investigate the appropriate functional relationship between variance of  $CR$  and  $H$  the data was divided into 7 classes based on total tree height. Variance of  $CR$  for each height class was computed. Variance of  $CR$  against  $H$  (Figure 5) was plotted to examine the functional relationship between the variance of  $CR$  to the total height. Based on the plot the functional relationships chosen to fit the variance was  $\sigma^2 = e^{c+d/H}$ , however, the same functional form for the mean as model A, that is,  $\mu = \frac{1}{1 + e^{-(b \times H)}}$  was used (Normal Model B). The functional form  $\sigma^2 = e^{c+d/H}$  produces the variance which approaches a lower asymptote as  $H$  approaches its higher values.

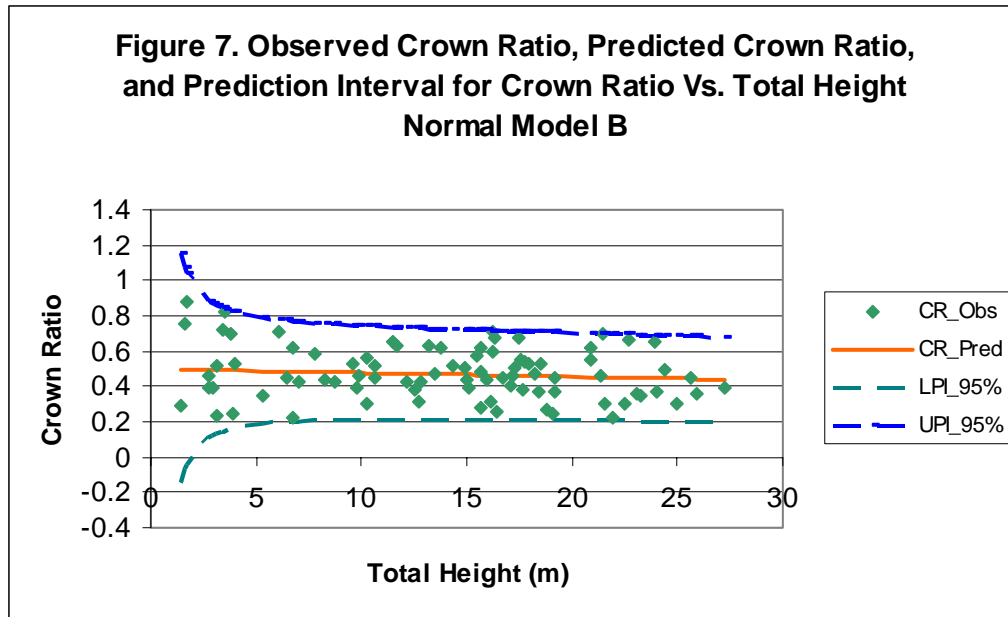


The normal model B assumes that tree  $CR$  is normally distributed with mean crown ratio  $\mu$  and variance in crown ratio  $\sigma^2$ , that is,  $CR \sim N(\mu, \sigma^2)$ , where  $b$ ,  $c$  and  $d$  are unknown coefficients to be estimated from the data. The same parameter estimation method and log likelihood function were used as the normal model A.

All the parameters  $b$ ,  $c$  and  $d$  for the Normal model B were found to be significant (Table 1) at 5% level of significance. The test of normality (Kolmogorov-Smirnov) revealed that the standardized residuals followed a normal distribution (p-value >0.15) (Table 1), and the plot of standardized residual against predicted crown ratio indicated that there was no detectable trend in the standardized residuals (Figure 6), therefore no further problems with model specification were indicated. The value of AIC was -90.6 (Table 1) which is a slight improvement over model A. Therefore, this model was a good candidate for the prediction of the distribution of crown ratio.

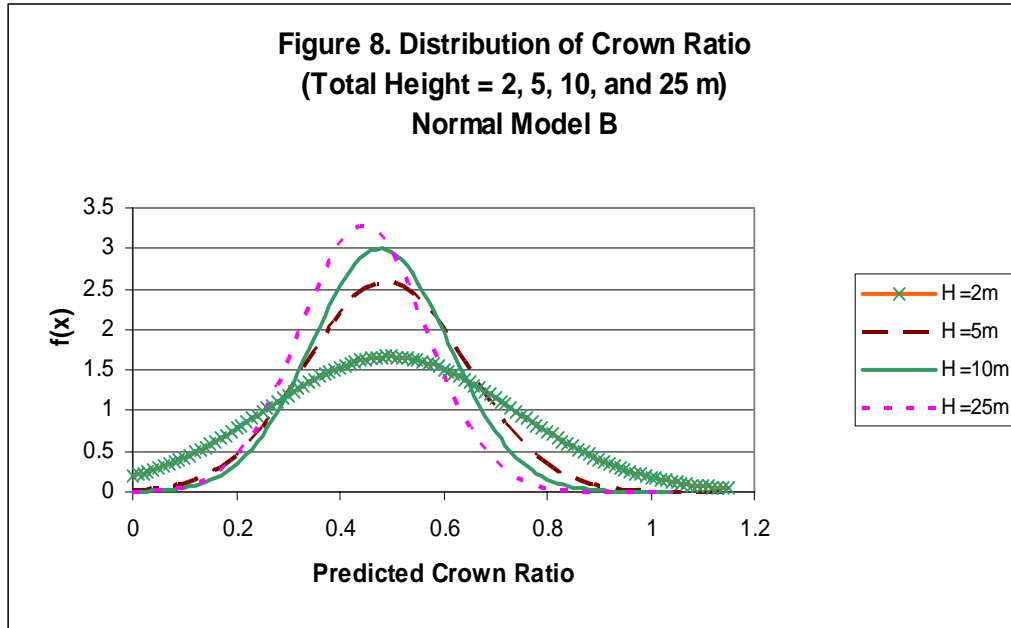


The fitted line for prediction of crown ratio goes through the center of the observed crown ratio (Figure 7). This is an indicative of a good predictive ability of Logistic model B for the mean  $CR$ . The prediction interval for  $CR$  is wide for the trees with smaller heights (1.3 m to 5 m) compared to the trees with larger heights.



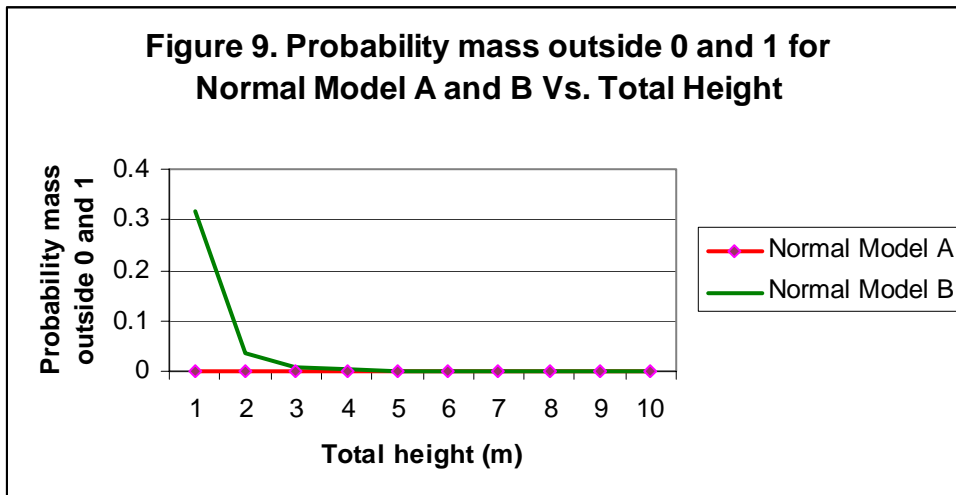
A normal distribution with mean of predicted crown ratio and standard deviation was used to generate the distribution of  $CR$  for the total heights 2, 5, 10 and 25 m (Figure 8). This figure indicates that there is not much change in mean  $CR$ , however, there is a decrease in variance of  $CR$  as the height increase. Crown ratio values greater than 1 and less than 0 are possible in the distribution when  $H = 2$  m (Figure 8).

The nominal value of the probability of Type I Error was set at 0.05. The result obtained from the cross validation indicated that p-values  $< 0.05$  were obtained approximately 2% of the total number of runs in the simulation. This suggests that the distribution of standardized residual has a lighter tail than the standard normal distribution. From statistical theory we know that if the model is correctly specified then the standardized residual follow an approximate standard normal distribution (Pierce and Schaffer 1986). Since the distribution of error deviates from standard normal the optimality of this model for the prediction of the distribution of crown ratio is questionable.



### 3.3.2 LOGISTIC FUNCTION WITH BETA DISTRIBUTION OF ERRORS

The total probability of getting  $CR$  outside the logical range  $[0,1]$  is very low and is approximately zero for Normal model A for the entire range of  $H$  (Figure 9), however, the total probability of getting  $CR$  less than 0 and greater than 1 from Normal model B increases as total height approaches 1.3 m. The total probability is approximately 25% when  $H = 1.5$  m and the probability goes down to 4% when  $H = 2$  m. Since predicted



crown ratio outside the range [0, 1] is not possible for any height, there is the potential for a better model if a different assumption of the error distribution is made. Therefore, a beta distribution was used to model  $CR$  and its prediction interval. This distribution is restricted to be between 0 and 1 for the entire distribution instead of just the mean (Cassella and Berger, 2002). Beta distributions are usually used to model proportions and since  $CR$  is the also a proportion of total height that supports the crown of a tree, the beta distribution should be a good alternative (Casella and Berger, 2002).

A random variable  $CR$  is said to follow the two parameter beta distribution if its probability density function (pdf) is  $f(CR | \alpha, \beta) = \frac{1}{B(\alpha, \beta)} CR^{\alpha-1} (1 - CR)^{\beta-1}$ ,

$0 < CR < 1, \alpha > 0, \beta > 0$ , where  $\alpha$  and  $\beta$  are shape parameters and  $B(\alpha, \beta)$

specifies the beta function. The mean and variance of the beta ( $\alpha, \beta$ ) are  $\mu = \frac{\alpha}{\alpha + \beta}$  and

$$\sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}.$$

Beta distribution will attain many different shapes as the value of parameters  $\alpha$  and  $\beta$  change. When  $\alpha > 1$  and  $\beta = 1$  the pdf will be strictly increasing; the pdf will be strictly decreasing when  $\alpha = 1$  and  $\beta > 1$ ; it will assume U-shaped when  $\alpha < 1$  and  $\beta < 1$ , or unimodal when  $\alpha > 1$  and  $\beta > 1$  (Casella and Berger, 2002).

In order to use the logistic equation for the mean, as was used previously, it is necessary to reparameterize the beta distribution in terms of  $\mu$  and  $\phi$  where  $\mu = \frac{\alpha}{\alpha + \beta}$

and  $\phi = \alpha + \beta$ . In terms of  $\mu$  and  $\phi$  the beta distribution would be

$$f(CR | \mu, \phi) = \frac{1}{B[\mu\phi, (1-\mu)\phi]} CR^{\mu\phi-1} \times (1-CR)^{(1-\mu)\phi-1}$$

### 3.3.2.1 Beta Model A

The Beta distribution was used to fit the data where the mean  $CR$   $\mu = \frac{1}{1 + e^{-(b \times H)}}$  and dispersion parameter  $\phi$  was a constant. The estimates of  $\mu$  and  $\phi$  were obtained simultaneously using the method of maximum likelihood from NLMIXED procedure in SAS. In the model  $b$  is an unknown coefficient to be estimated from the data and  $H$  is total tree height.

The log likelihood function for individual observations that are distributed according to beta distribution is given by

$$\lambda(\mu_i, \phi | CR) = \log \Gamma(\phi) - \log \Gamma(\mu_i \phi) - \log \Gamma((1 - \mu_i) \phi) + (\mu_i \phi - 1) \times \log(CR) + \{(1 - \mu_i) \phi - 1\} \times \log(1 - CR)$$

Where  $\mu_i$  is predicted mean  $CR$  with a range  $[0,1]$ , and  $\phi$  is the dispersion parameter. It is important to note that variance  $\sigma^2$  is related to both  $\mu$  and  $\phi$  since

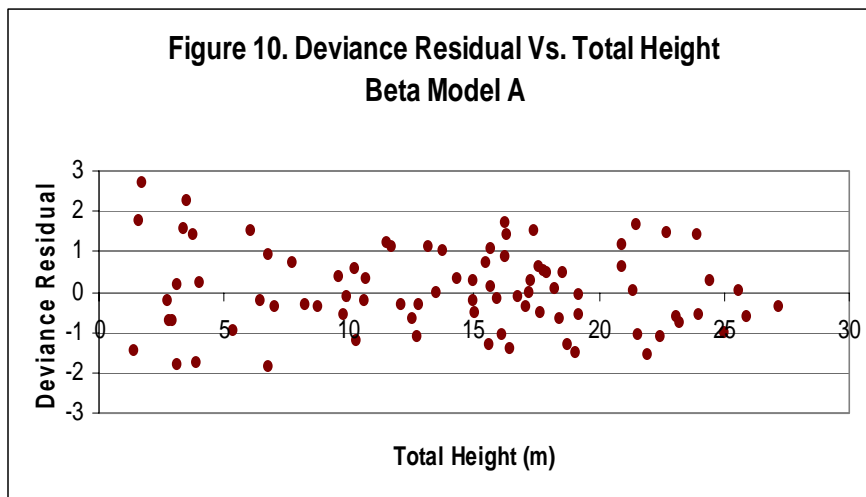
$$\sigma^2 = \frac{\mu(1-\mu)}{1+\phi} \text{ (Ferrari and Cribari-Neto, 2004). Therefore } \phi \text{ and } \sigma^2 \text{ are inversely related}$$

to each other and  $\sigma^2$  approaches to 0 as  $\mu$  approaches either 0 or 1.

Standardized residuals used for the normal models A and B are not appropriate for asymmetric distributions such as the Beta because they tend to depart from normal distribution (Pierce and Schaffer 1986). Deviance residuals are closer to standard normal when the distribution of error term is not normal (Pierce and Schaffer, 1986) and the

contribution of each observation to the deviance is a chi-square distributed quantity whose square root (with the right sign) is a normally distributed random variable (McCullagh and Nelder, 1989). For the beta distribution, a deviance residual was used which is computed as  $r^d_i = \text{sign}(y_i - \hat{\mu}_i) \{2(l_i(\tilde{\mu}_i, \hat{\phi}) - l_i(\hat{\mu}_i, \hat{\phi}))\}^{1/2}$  where,  $l_i(\tilde{\mu}_i, \hat{\phi})$  is the likelihood estimate obtained from the saturated model and  $l_i(\hat{\mu}_i, \hat{\phi})$  is the likelihood estimate obtained from the fitted model,  $y_i$  is the observed CR,  $\hat{\mu}_i$  is the predicted CR obtained from the fitted model, and  $\tilde{\mu}_i$  is equal to  $y_i$  in the saturated model (Ferrari and Cribari-Neto, 2004).

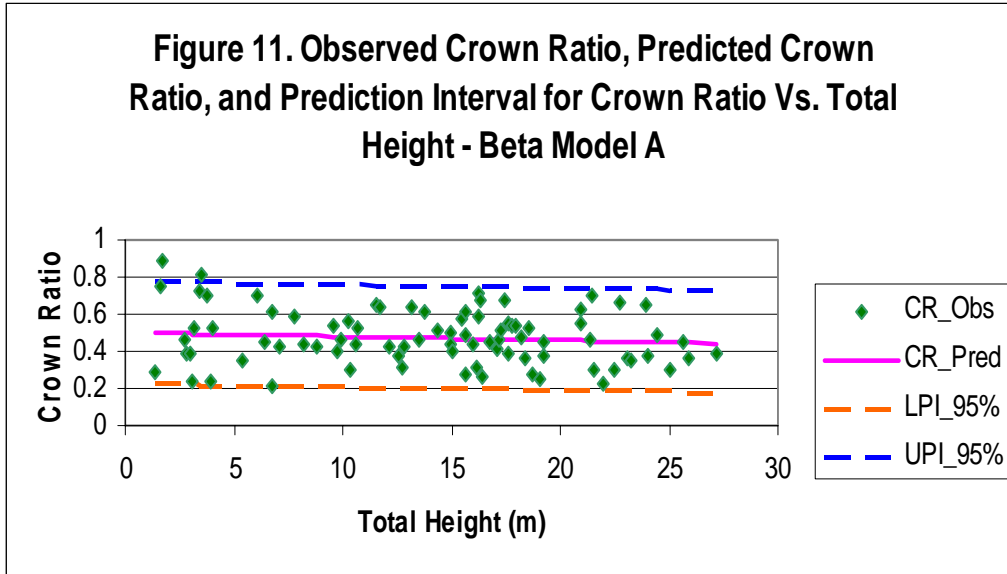
The parameters  $b$  and  $\sigma^2$  were significant (Table 1) for the Beta model A. The test of normality (Kolmogorov-Smirnov) depicted that the deviance residuals followed a normal distribution (p-value >0.15) (Table 1), however, the plot of deviance residual against predicted  $H$  indicated a discernible trend (Fig. 10). The value of AIC was -85.8 (Table 1) which is a slight improvement over normal model A.



The fitted line for prediction of crown ratio goes through the center of the observed crown ratio (Figure 11) and looks very similar to the normal models. This is indicative of



the good predictive ability of Beta model A for the mean  $CR$ . The prediction intervals for the entire range of total height are parallel (Figure 11) which is questionable for the smaller trees.



### 3.3.2.2 Beta Model B

In order to investigate the appropriate functional relationship between  $\phi$  and  $H$  the data was divided into 7 classes based on  $H$ . Mean  $H$  was computed and  $\phi$  was predicted for each class using NLMIXED procedure in SAS. Based on the plot of  $\phi$  versus total height (Figure 12) the functional relationships chosen to fit the data was  $\phi = e^{c+d/H}$ . It is important to note that since the relationship between  $\sigma^2$  and  $\phi$  is an inverse one, this is the same functional form as used for the variance in normal model B with a change in sign for the parameters.



It becomes obvious from Figure 14 that the width of prediction interval increases as the value of total height decreases. The prediction interval gets prominently larger as the total height gets < 5 m.

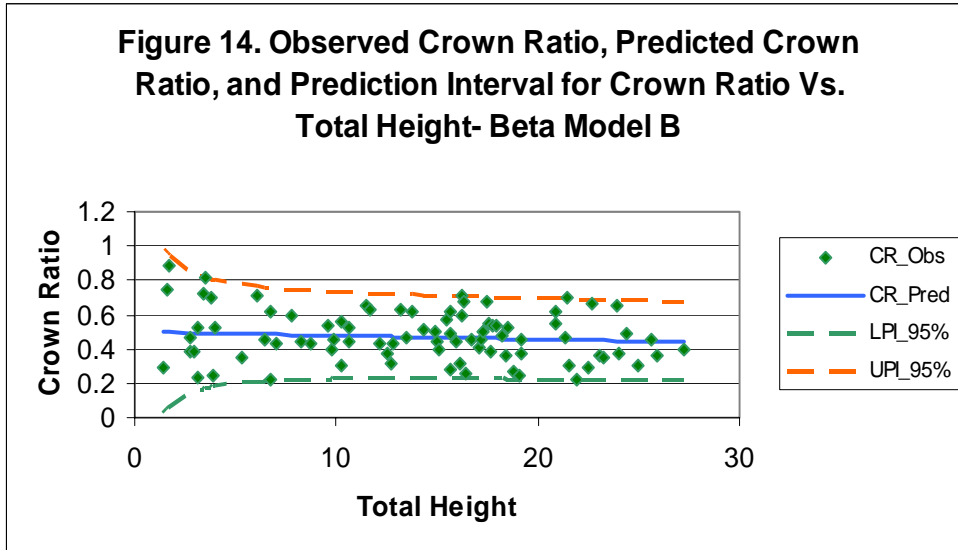
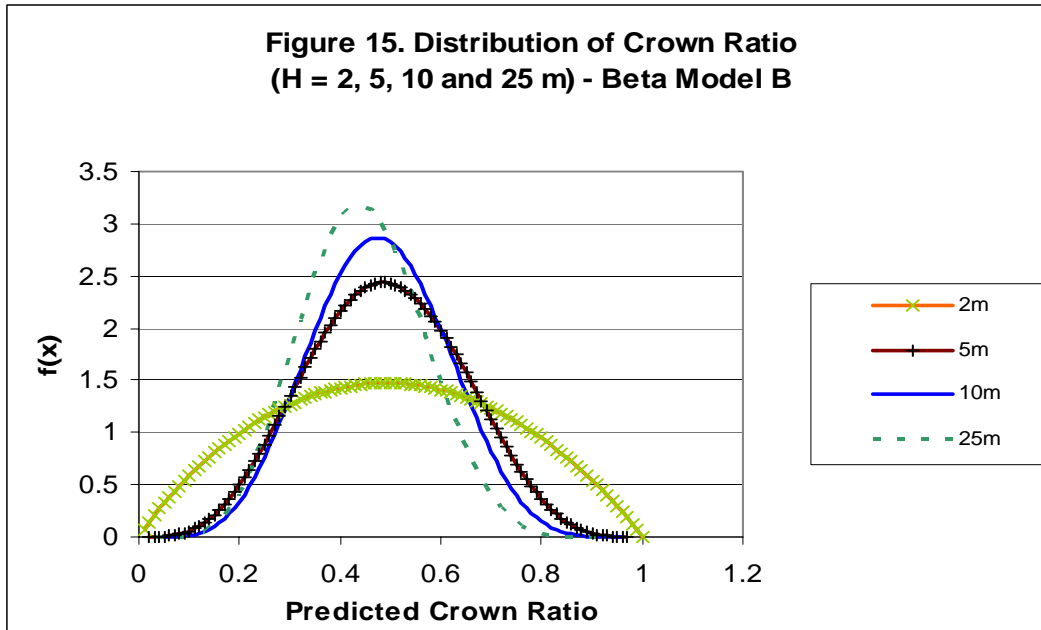


Figure 15 provides examples of the distribution for tree heights of 2, 5, 10, and 25 m. The shape of the distribution also changes from flat to more peaked as the total height increases from 2 m to 25 m (Fig. 15). Distribution of crown ratio never exceeds the range of [0, 1].

A cross validation using the fitted Beta model B was run identical to the Normal model B. The result obtained from the simulation indicated that p-values < 0.05 was obtained approximately 4 % of the total number of runs. This is very close to the nominal level of 5%. This means that deviance residuals follow approximately the standard normal distribution suggesting that model B is the most optimal model among the models considered in this paper for the prediction of distribution of crown ratio.



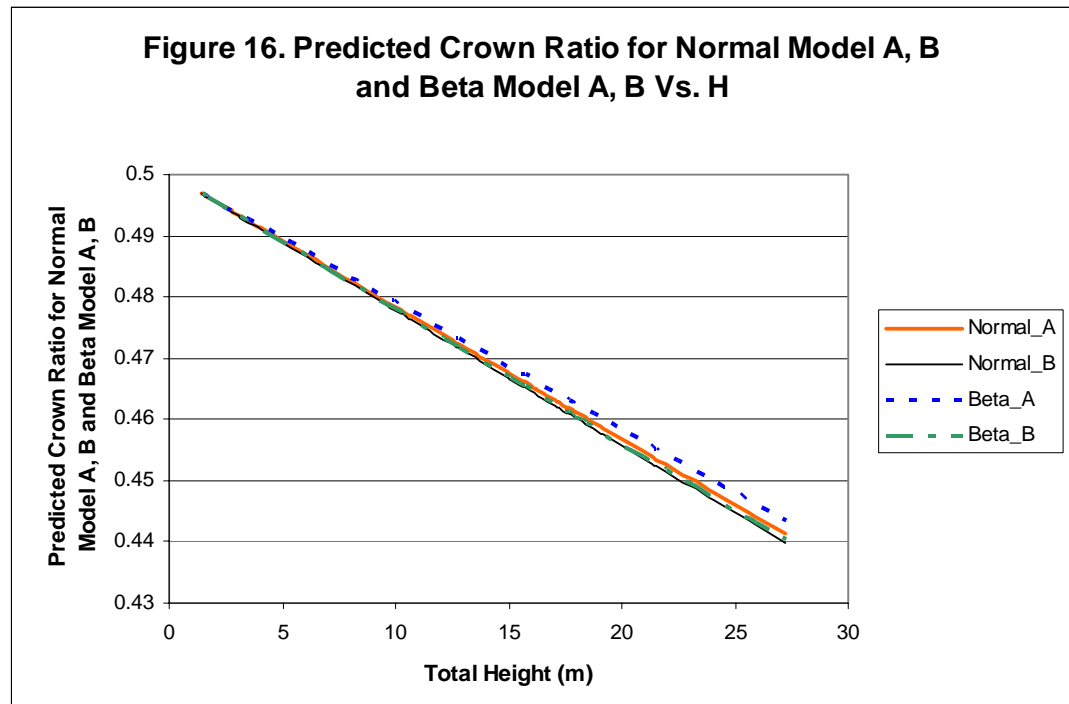
### 3.4 DISCUSSION AND CONCLUSION

The models with same mean and different formulations of variance were used to compare the performance of the models for the prediction of distribution of crown ratio (table 2).

**Table.2. Models used to fit the data.**

Model	Distribution	Location (Mean)	Variance
Normal A	Logistic	$\mu = \frac{1}{1 + e^{-(b \times H)}}$	$\sigma^2$
Normal B	Logistic	$\mu = \frac{1}{1 + e^{-(b \times H)}}$	$\sigma^2 = e^{c+d/H}$
Beta A	Beta	$\mu = \frac{1}{1 + e^{-(b \times H)}}$	$\sigma^2 = \frac{\mu(1-\mu)}{1+\phi}$
Beta B	Beta	$\mu = \frac{1}{1 + e^{-(b \times H)}}$	$\sigma^2 = \frac{\mu(1-\mu)}{1+e^{c+d/H}}$

It became evident that all the models predicted similarly for the purpose of predicting the mean (Figure 16); however, there was some difference in their ability to predict the distribution of crown ratio.



Analysis of the residuals for all four models revealed that they followed a normal distribution (Table 1), however the residuals for Normal model A and Beta model A displayed a detectable trend which indicates that there might be a problem of model specification. The prediction interval for the entire range of predicted  $CR$  was parallel for Normal model A and Beta model A, however, Normal model B and Beta model B indicated that the prediction interval for the crown ratio increased as the  $H$  decreased. Increase in prediction interval became significant when  $H$  attained a value  $< 5m$ . Normal model A had the highest AIC followed by Beta model A, Normal model B, and Beta model B had the lowest. These values of AIC indicate that Normal model A had the lowest

performance, whereas, Beta model B had the highest performance among all four models. Type I Error rate obtained from running a simulation 1000 times was 2% for Normal model B, and 4% for Beta model B. These error rates are close compared with the nominal Type I Error rate of 5% in the study. This result indicates that Beta model B is a little bit better than Normal model B for predicting the distribution of crown ratio.

The prediction interval for *CR* obtained from Beta model never crossed beyond the logical range [0, 1], whereas, from the Normal models the prediction interval crossed beyond the range, for example, the prediction interval for a tree with  $H = 1.5\text{m}$  reveals that there is approximately 25% probability of getting a predicted crown ratio value which is outside the logical range [0,1] from Normal model B. By definition the Beta model cannot extend beyond this range and is therefore a more logical assumption to make.

Models with the assumption of normality of errors are good for the prediction of the mean *CR*; however, they are not suitable for estimation of distribution of crown ratio since the range of a normally distributed random variable spans from negative infinity to positive infinity. So, there is a possibility of obtaining predictions which are not realistic. The normal distribution is symmetric and bell shaped and hence cannot explain any kind of skewness (positive or negative). In order to overcome these problems a Beta distribution was assessed since the prediction can never go beyond the logical range of crown ratio, that is, [0,1]. It can also incorporate the skewness in the distribution if any.

### 3.5 LITERATURE CITED

Anderson, D. R., Burnham, K. P., and Thompson, W. L., 2000. Null hypothesis testing: Problems, prevalence, and an alternative. *J. Wildl. Manage.* 64, 912-923.

Anderson, M., Somers, G. L., Smith, W. R., Freeland, M., and Ruch, D., 1999. Measuring crown dynamics of longleaf pine in sandhills of Eglin Air Force Base. Proceedings of the Second Longleaf Alliance Conference 1998. Charleston, 17-19 Nov., 46-48.

Akaike, H., 1973. Information theory as an extension of the maximum likelihood principle. In: Second International Symposium on Information Theory (Eds: B. N. Petrov and F. Csaki). Akademiai Kiado, Budapest. Pp 267-281.

Boyer, W.D., 1991. *Pinus palustris* Mill. (Longleaf Pine). In: Silvics of North America 1. Conifer (Eds: Burns, R. M. and Honkala, B. H. Tech. Coords.). Agric. Handb. 654. Washington, DC: U.S. Department of Agriculture, 405-412.

Casella, G., and Berger, R. L. 2002. Statistical Inference. Duxbury, Thompson learning Inc. 660pp.

Cole, W., and Lorimer C. G., 1994. Predicting tree growth from crown variables in managed Northern hardwood stands. *For. Ecol. Manag.* 67, 159-175.

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

Ferrari, S. L. P., and Cribari-Neto, F., 2004. Beta Regression for Modeling Rates and Proportions. *Journal of applied statistics.* 31, 799-815.

Hasenauer, H., and Monserud R.A., 1996. A crown ratio model for Austrian forests. *Forest Ecology and Management* 84, 49-60.

- Hatch, C.R. 1980., Modeling tree crown size using inventory data. *Mitt. Forstl. Bundes-versuchsanst. Wien*, 130, 93-97.
- Hollander, M., and Wolf, D. A., 1999. *Nonparametric Statistical Methods*. Second Edition. John Wiley and Sons, Inc. 787 pp.
- Huang, S., Titus, S.J., and Wiens, D.P., 1992. Comparison of nonlinear height-diameter functions for major Alberta tree species. *Can. J. For. Res.* 22, 1297-1304.
- Hynynen, J., 1995. Predicting tree crown ratio for un-thinned and thinned Scots pine stands. *Can. J. For. Res.* 25, 57-62.
- Kush, J. S., Boyer, W. D., Meldahl, R. S., and Ward, G. A., 1999. Precommercial thinning intensity in longleaf pine: effect on product volume and value. In: *Proceedings of the second Longleaf alliance conference* (Eds: Kush, J. S. ). Longleaf Alliance Report No. 4. 196 pp.
- Lilliefors, H. W., 1967. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. . *J. Am. Stat. Assoc.* 399-402.
- Liu, A., Schisterman E. F., and Wu, C., 2006. Multistage evaluation of measurement error in a reliability study. *Biometrics* 1-7.
- McCullagh, P., and Nelder, J. A. 1989. *Generalized Linear Models*, 2<sup>nd</sup>. London: Chapman & Hall. 511 pp.
- Meldahl, R. S., Kush, J. S., Rayamajhi, J. N., and Farrar, R. M. 1998. Productivity of natural stands of Longleaf pine in relation to competition and climatic factors. In: *The productivity and sustainability of southern forest ecosystems in changing environment* (Eds: Mickler, S. F., Fox, S.,). Springer-Verlag New York, Inc. 892 pp.



Monseurd, R.A., and Sterba H., 1996. A basal area increment model for individual trees growing in even and uneven aged forest stands in Austria. *For. Ecol. Manag.* 80, 57-80.

Neter, H., and Maynes, S.E., 1970. On the appropriateness of the correlation coefficient with a 0, 1 dependent variable. *J. Am. Stat. Assoc.* 65: 501-509.

Oliver, C.D., and Larson, B.C., 1996. *Forest stand dynamics*. John Wiley and Sons, Toronto. 424 pp.

Overing, J. D., and Watts, F. C., 1989. *Soil survey of Walton County, Florida*. USDA, Soil Conservation Service, 234 pp.

Pierce, D. A., and Schaffer, D. W. 1986. Residuals in generalized linear models. *J. Am. Stat. Assoc.* 81: 977-986.

Ritchie M., Hann, D., 1987. Equations for predicting height to crown base for fourteen tree species in Southwest oregon, Forest Research Laboratory, Oregon State Univ., Corvallis, Res. paper No. 50, 14 p.

SAS Institute, Inc., 2004. *SAS/STAT User's guide: Statistics, version 9.1.*, Cary, N.C. 5136 pp.

Short III, E.A., and Burkhart, H.. 1992. Prediction crown-height increment for thinned and unthinned loblolly pine plantations. *For. Sci.* 38: 594-610.

Singh, P., Abebe, A., and Mishra, S. N., 2006. Simultaneous testing for the successive differences of exponential location parameters. *Communications in Statistics-Simulation and Computation.*, 35, 547-561.

Temesgen, H., LeMay, V., and Mitchell, S. J., 2005. Tree crown ratio models for multi-species and multi-layered stands of southeastern British Columbia. *The Forestry Chronicle* 81,133-141.

Valentine, H.T., Ludlow, A.R., and Furnival, G.M., 1994. Modeling crown rise in even-aged stands of Sitka spruce or loblolly pine. *For. Ecol. Manag.* 69: 189-197.