

Diameter Distributions of Juvenile Stands of Loblolly Pine (*Pinus taeda* L.) with Different Planting Densities

by

Bronson P. Bullock

Dissertation submitted to the Graduate Faculty of
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

FORESTRY

APPROVAL:

Harold E. Burkhardt,
Chair and Department Head

Richard G. Oderwald,
Committee Member

Philip J. Radtke
Committee Member

Oliver Schabenberger
Committee Member

Stephen P. Prisley
Committee Member

February, 2002

Blacksburg, Virginia 24061

**Diameter Distributions of Juvenile Stands of Loblolly Pine
(*Pinus taeda* L.) with Different Planting Densities**

by

Bronson P. Bullock

Harold E. Burkhardt, Chairman

ABSTRACT

Diameter distributions of juvenile loblolly pine (*Pinus taeda* L.) with different planting densities were characterized utilizing a two-parameter Weibull distribution. Trend analysis was employed to describe the effects of planting density, age, relative spacing, and rectangularity on the estimated diameter distributions for juvenile loblolly pine. A reparameterization of the two-parameter Weibull distribution was sought to reduce the dispersion of the estimated shape parameter.

Methods that quantify the amount of inter-tree spatial dependency in a particular stand were applied. Empirical semivariograms were derived for each plot over all ages to enable spatial trend recognition. Moran's I and Geary's C coefficients were estimated for ground-line diameters from ages 2 to 5, and for breast height diameters from ages 5 to 11. Though there was no discernable trend in the presence of significant spatial autocorrelation with planting density, an initial negative trend with age was present, but leveled off by age 5. A conditional autoregressive model was utilized to evaluate the amount of spatial influence stems in a stand have on one another. The occurrence of

significant spatial influences was positively associated with age through age 8, the trend then leveled off; no recognizable trend was detected with planting density.

These indices help to describe stand dynamics that are influenced by the spatial distribution of stems. Models to predict the parameters of the two-parameter Weibull distribution were developed to aid in forecasting and simulation of juvenile loblolly pine. Simulations were conducted where a spatial dependency was imposed on the diameters within a stand. The spatial structure simulation enables accurate representations of stem characteristics when simulating forest stands that include spatially-explicit information.

ACKNOWLEDGEMENTS

I would genuinely like to thank my advisor, Dr. Harold E. Burkhart, for his enduring support, patience, and extraordinary guidance throughout the past five years. The invaluable advice he gave me will remain with me throughout my professional and personal pursuits. I would also like to thank my advisory and examination committee members, Dr. Richard G. Oderwald, Dr. Stephen P. Prisley, Dr. Philip J. Radtke, Dr. Oliver Schabenberger, and Dr. Randolph H. Wynne for taking the time to aid me in my efforts. A special note of appreciation goes to Dr. Edwin J. Green for his support, inspiring example, and aid in setting me on this path so many years ago.

This study would not have been possible without the financial support of the Loblolly Pine Growth and Yield Research Cooperative maintained at Virginia Polytechnic Institute and State University. The data for this study was acquired from the Loblolly Pine Growth and Yield Research Cooperative.

I would like to thank all of the faculty, staff, and students in the Department of Forestry, the College of Natural Resources, and the Department of Statistics for furnishing me with an exceptional education. My thanks go out to all of my fellow graduate students, whose humor and mutual support made the graduate experience complete, with especial thanks to Ed Boone, Kerri Cahill, Tracy Farrell, Paul Wagner, James Walters, and Jim Westfall.

An expression of gratitude is due to my extremely close friends of almost two decades, Jonathon Barcelona, David Biondi, and Allan Johnson, with whom the foundation of my individuality was set so long ago. I would like to thank all of my

family members and especially my parents, Catherine and Louis Zacchino, and James Bullock III, whose support and encouragement over the years facilitated this achievement. A distinct acknowledgement is merited to William Kirkwood for his long-standing endeavor of intellectual stimulation to expand my horizons.

A tender note of gratitude goes to my fiancée Celia Rose for her patience and mutual sacrifice throughout these trying times. Her compassion and affection were essential to the fulfillment of my goals. Lastly, I would like to extend my thanks to Falca ‘Orla’ von Brendonhaus Bullock, my German Shepherd, whose companionship was always there.

Table of Contents

ABSTRACT	i
ACKNOWLEDGEMENTS	iii
LIST OF FIGURES	vii
LIST OF TABLES	xii
1. INTRODUCTION	1
1.1. Objectives	2
1.2. Justification	3
2. LITERATURE REVIEW	7
2.1. Diameter Distributions	7
2.2. Weibull Distribution	9
2.3. Estimation Techniques	12
2.4. Juvenile Growth	14
2.5. Spatial Effects	15
3. DATA	19
4. PRELIMINARY DISTRIBUTIONAL ANALYSIS	26
4.1. Weibull Distribution	26
4.2. Weibull Parameter Estimation	27
4.3. Weibull Distribution Validation	32
5. COMPARISONS OF ESTIMATED DISTRIBUTIONS	41
5.1. Rectangularity Effect	41
5.2. Temporal Effect	51
5.3. Planting Density Effect	60

5.4.	Relative Spacing Effect.....	64
5.5.	Weibull Reparameterization	70
5.6.	Summary of Comparisons.....	76
6.	SPATIAL DEPENDENCE.....	80
6.1.	Spatial Process Definition.....	81
6.2.	Edge Effect.....	86
6.3.	Empirical Semivariogram	87
6.4.	Moran's I Coefficient.....	96
6.5.	Geary's C Coefficient	104
6.6.	Spatial Models	111
6.7.	CAR Parameter Estimation.....	113
6.8.	Spatial Dependency Parameter Comparison.....	125
6.9.	Spatial Summary	141
7.	SPATIAL STRUCTURE SIMULATION.....	145
7.1.	Stand Simulation.....	146
7.2.	Simulation Application	148
8.	SUMMARY AND RECOMMENDATIONS.....	155
	LITERATURE CITED.....	163
	APPENDIX.....	171
	VITA.....	231

LIST OF FIGURES

Figure 1: A selected example of a spacing study replication.....	20
Figure 2: Example of shapes attainable with the two-parameter Weibull function.....	28
Figure 3: Trends for (a) the coefficient of skewness over time for diameter at breast height and (b) the average coefficient of skewness over time for diameter at breast height, comparing all planting densities (trees per acre)	37
Figure 4: Trends for (a) the coefficient of kurtosis over time for diameter at breast height and (b) the average coefficient of kurtosis over time for diameter at breast height, comparing all planting densities (trees per acre)	40
Figure 5: Trends over time for plots with a planting density of 907 trees per acre but with different planting rectangularities (1:3 and 1:1.33) for diameter at breast height (D) measurements for (a) the shape parameter estimate and (b) the scale parameter estimate	42
Figure 6: Graphical representation of the average estimated diameter at breast height (D) distributions every two years for ages 5 to 11, for plots with the same planting density of 907 trees per acre but with different planting rectangularities (1:3 and 1:1.33)	45
Figure 7: Coefficient of skewness trends over time for diameter at breast height, comparing plots with the same planting density of 907 trees per acre but with different planting rectangularities (1:3 and 1:1.33) for (a) all plots and (b) the average coefficients	48
Figure 8: Coefficient of kurtosis trends over time for diameter at breast height, comparing plots with the same planting density of 907 trees per acre but with different planting rectangularities (1:3 and 1:1.33) for (a) all plots and (b) the average coefficients	50

Figure 9: Trends of the estimated scale parameter for diameter at breast height (D) over ages 5 to 11, for plots with a planting density of (a) 302, (b) 605, (c) 1210, and (d) 2722 trees per acre	52
Figure 10: Trends of the estimated shape parameter for diameter at breast height (D) over ages 5 to 11, for plots with a planting density of (a) 302, (b) 605, (c) 1210, and (d) 2722 trees per acre.....	53
Figure 11: Residual graph for the predicted scale parameter	56
Figure 12: Residual graph for the predicted shape parameter	59
Figure 13: Trends of the estimated scale parameter for diameter at breast height (D) over all observed planting densities (trees per acre) for ages (a) 5, (b) 7, (c) 9, and (d) 11	61
Figure 14: Trends of the estimated shape parameter for diameter at breast height (D) over all observed planting densities (trees per acre) for ages (a) 5, (b) 7, (c) 9, and (d) 11	62
Figure 15: Mean relative spacing trend over ages 5 to 11 for each planting density (trees per acre)	68
Figure 16: Sample size for each relative spacing bin	69
Figure 17: Parameter estimates for all plots in each relative spacing bin, with planting density subgroups (trees per acre), for (a) the scale parameter and (b) the shape parameter.....	71
Figure 18: Residual graph for reparameterized (alpha) shape parameter estimates	74
Figure 19: Residual graph for reparameterized (delta) shape parameter estimates	75
Figure 20: Mean Weibull parameter estimates for (a) the scale parameter, and (b) the shape parameter, across all ages and planting densities (trees per acre) considered in this research for diameter at breast height	78

Figure 21: Stem map (lattice) for each plot, with tree numbers 1-49	83
Figure 22: The distance-dependent weights for tree $s_{6,3}$ (circled) on a 4-by-4 ft spacing, where the weights are calculated as the inverse of the distance between the subject tree i and neighboring tree j	85
Figure 23: Graphs of the empirical semivariograms for stand number 20, derived from the original data and the residuals from a fitted linear model, for (a) ground-line diameter (d) at age 2, and (b) diameter at breast height (D) at age 11.....	93
Figure 24: Spatial distribution of diameters for stand number 20, for (a) ground-line diameter at age 2, and (b) diameter at breast height at age 11	95
Figure 25: Moran's I coefficients significant at an alpha level of 0.10 across all ages, for the residuals from a fitted model for ground-line diameters (d) and breast height diameters (D), with a mean equal to $\frac{-1}{n-1}$	99
Figure 26: Average proportion of significant Moran's I coefficients at an alpha level of 0.10 for (a) each planting density (trees per acre) and (b) over ages 2 to 5 for ground-line diameter (d), and over ages 5 to 11 for diameter at breast height (D)	102
Figure 27: Proportion of significant Moran's I coefficients at an alpha level of 0.10 for each planting density (trees per acre) over (a) ages 2 to 5 for ground-line diameter (d), and (b) ages 5 to 11 for diameter at breast height (D).....	103
Figure 28: Geary's C coefficients significant at an alpha level of 0.10 across all ages for the residuals from a fitted model for ground-line diameters (d) and breast height diameters (D), with a mean equal to one	106
Figure 29: Average proportion of significant Geary's C coefficients at an alpha level of 0.10 for (a) each planting density (trees per acre) and (b) over ages 2 to 5 for ground-line diameter (d), and over ages 5 to 11 for diameter at breast height (D)	108

Figure 30: Proportion of significant Geary's C coefficients at an alpha level of 0.10 for each planting density (trees per acre) over (a) ages 2 to 5 for ground-line diameter (d), and (b) ages 5 to 11 for diameter at breast height (D).....	110
Figure 31: The fitted Normal and two-parameter Weibull distributions for ground-line diameter at ages (a) 2 and (b) 5.....	116
Figure 32: The fitted Normal and two-parameter Weibull distributions for breast height diameter at ages (a) 8 and (b) 11.....	117
Figure 33: Derived ρ values significant at an alpha level of 0.10 across all ages for the spatial linear model including the covariate total stem height, for ground-line diameters (d) and breast height diameters (D).....	120
Figure 34: Average proportion of significant ρ values at an alpha level of 0.10 for the spatial linear model including the covariate total stem height for (a) each planting density (trees per acre) and (b) over ages 2 to 5 for ground-line diameter (d), and over ages 5 to 11 for diameter at breast height (D).....	122
Figure 35: Proportion of significant ρ values at an alpha level of 0.10 for the spatial linear model including the covariate total stem height for each planting density (trees per acre) over (a) ages 2 to 5 for ground-line diameter (d), and (b) ages 5 to 11 for diameter at breast height (D).....	124
Figure 36: Derived spatial dependency parameter values significant at an alpha level of 0.10 and corresponding significant (alpha level = 0.10) Geary's C coefficients evaluated from the residuals of the spatial linear model, with reference lines indicating no spatial autocorrelation for the CAR model ($\rho = 0$) and Geary's C coefficient ($C = 1$).....	126
Figure 37: Linear correlations derived from spatial dependency parameters (ρ) near the lower bound of possible values, with a zero reference line, for (a) stand number 128, a 12-by-12 ft spacing, ρ value of -7.57 , and (b) stand number 65, a 4-by-4 ft spacing, ρ value of -2.51	129

Figure 38: Linear correlations derived from spatial dependency parameters (ρ) near the upper bound of possible values, for (a) stand number 128, a 12-by-12 ft spacing, ρ value of 0.67, and (b) stand number 65, a 4-by-4 ft spacing, ρ value of 0.21	130
Figure 39: Linear correlations derived from spatial dependency parameters equal to -1.0 , with a zero reference line, for (a) stand number 128, a 12-by-12 ft spacing, and (b) stand number 65, a 4-by-4 ft spacing.....	131
Figure 40: Linear correlations derived from spatial dependency parameters equal to 0.10 , with a zero reference line, for (a) stand number 128, a 12-by-12 ft spacing, and (b) stand number 65, a 4-by-4 ft spacing.....	133
Figure 41: Linear correlations derived from spatial dependency parameters (ρ) near the lower bound of possible values, with a zero reference line, for stand number 65, a 4-by-4 ft spacing, with ρ values of (a) -1.7703 and (b) -2.0232	135
Figure 42: Histograms and quantile-quantile plots of the residuals for the (a) OLS and (b) CAR models fitted to diameter at breast height (D) at age 8	137
Figure 43: Semivariograms for the OLS and CAR residuals from modeling diameter at breast height (D) at age 8	139
Figure 44: Linear correlations derived from the estimated spatial dependency parameter (-1.4532), with a zero reference line, for stand number 65, a 4-by-4 ft spacing at age 8.....	140
Figure 45: Frequency histogram for diameters (D at age 8) from the original data from stand number 145 and a random realization from the spatial simulation	152
Figure 46: Spatial distribution of diameters for (a) the original diameter at breast height data (D at age 8) from stand number 145, and (b) a random realization from the spatial simulation	153

LIST OF TABLES

Table 1: Spacing, planting density (trees per acre), and planting rectangularity for the 16 plots within each replication	21
Table 2: Number of plots established for each planting density (trees per acre).....	22
Table 3: Distribution of stems by those that exceed and those that do not exceed breast height at each age since planting	25
Table 4: Results for the Chi-square, Anderson-Darling, Cramer-von Mises, and Kolmogorov-Smirnov goodness-of-fit statistics for diameter at breast height at ages 5 to 11	35
Table 5: Analysis of variance results and average parameter estimates for the scale and shape parameters for ground-line diameters (d) and diameters at breast height (D) of plots with a planting density of 907 trees per acre but different planting rectangularities.....	44
Table 6: Analysis of variance results for the coefficients of kurtosis and skewness for plots with a planting density of 907 trees per acre but with different planting rectangularities (1:3 and 1:1.33)	49
Table 7: Analysis of variance and correlation results for the scale parameter estimates by planting density for diameter at breast height.....	54
Table 8: Analysis of variance and correlation results for the shape parameter estimates by planting density for diameter at breast height.....	57
Table 9: Analysis of variance and correlation results for the scale parameter estimates by age for diameter at breast height.....	63

Table 10: Analysis of variance and correlation results for the shape parameter estimates by age for diameter at breast height.....	65
Table 11: Empirical semivariograms for stand number 20, ground-line diameter (d) at age 2, and diameter at breast height (D) at age 11	92
Table 12: Frequency of significant Moran's I coefficients (0.10 alpha level) across all planting densities, ages, and diameter measurements considered in this research .	100
Table 13: Frequency of significant Geary's C coefficients (0.10 alpha level) across all planting densities, ages, and diameter measurements considered in this research .	107
Table 14: Results for the Anderson-Darling, Cramer-von Mises, and Kolmogorov-Smirnov goodness-of-fit statistics for fit to Normal distribution from 50 random plots for ground-line diameter (d) at ages 2 and 5, and breast height diameter (D) at ages 8 and 11	115
Table 15: Frequency of significant spatial dependency in spatial linear models with a covariate term (0.10 alpha level), across all planting densities, ages, and diameter measurements considered in this research	121
Table 16: Spatial index comparisons of the original data (breast height diameters at age 8) from stand number 145, a random realization of the spatial simulator, and the median of 100 random realizations of the spatial simulator	150

1. INTRODUCTION

The research presented herein is aimed at characterizing diameter distributions for juvenile loblolly pine (*Pinus taeda* L.) stands. Loblolly pine is an extremely important commercial tree species in the United States. Diameter distributions, which have been applied to describe the characteristics of forest stands for many decades, provide valuable information on the current condition of the stand. To adequately characterize juvenile diameter distributions, a functional form that describes the traits present in the stand diameters is required. The two-parameter Weibull distribution was employed to represent the attributes in the stand diameter data. The parametric methodology presented should aid foresters in understanding the dynamics in the diameter distributions of young loblolly pine stands.

Juvenile diameter distributions were studied to gain insight into the effects of various stand-level factors. Parameters of the derived diameter distributions were modeled as a function of stand-level characteristics to increase the utility of the derived distributions. A thorough understanding of juvenile diameter distributions should prove especially useful for operational planning of stands on short rotations that require estimates of productivity at early ages. The spatial relationship between stems in a stand changes with age. The spatial dependency between stems may have a significant impact on any forecasting and simulation techniques utilized. To account for the variation in the spatial relationships, methods that capture the amount of inter-tree spatial dependency were evaluated for the juvenile diameter data considered in this research. The methods presented herein may be extended to other tree species that have similar empirical diameter distributions to that of loblolly pine.

1.1. Objectives

The objectives of this study were to:

1. Characterize the diameter distributions of juvenile stands of loblolly pine by utilizing a two-parameter Weibull distribution
2. Compare diameter distributions with respect to planting density, age, and relative spacing
3. Test for a significant rectangularity effect on the 4-by-12 ft (1:3) and 6-by-8 ft (1:1.33) spacings
4. Derive the spatial dependence between trees to account for autocorrelation among stems in a plot using a conditional autoregressive model
5. Model the parameters of the diameter distributions of juvenile loblolly pine to aid in forecasting and simulation
6. Impose a spatial autocorrelation structure among stems when simulating spatially-explicit stands from diameter distributions

1.2. Justification

Previous work on tree diameter distributions has focused mainly on trees from crown closure to rotation age. This research elucidates trends in diameter distributions of juvenile loblolly pine. Diameter distributions can be utilized to extract size class information that enables the evaluation of stand yield and impacts from stand treatments. Tree diameter is correlated with many other descriptive characteristics that are of interest to forest managers (e.g. volume and weight). Individual stem models applied across diameter classes enable class or stand-level estimates of model output characteristics. Much research has been performed on ways to represent the diameter distributions of various tree species. The Weibull distribution has been shown to describe the diameter distributions of tree stems satisfactorily and can assume a variety of shapes. Many techniques are available to estimate the parameters of the probability density function, with maximum likelihood estimation being an advantageous fitting technique due to its asymptotic properties. Diameter distributions of juvenile loblolly pine were characterized in this study through the application of the two-parameter Weibull distribution, where the parameters were estimated by maximum likelihood techniques.

There is a need for research into even-aged juvenile stands to address how diameter distributions change over time across varying stand-level factors. One factor that has a considerable impact on future stand conditions is planting density. Planting density (trees per unit area) affects both individual stem and stand characteristics. Knowledge of how diameter distributions change with varying levels of planting density will further enable trend recognition for management decisions.

Relative spacing is a measure of the competitive status of the stand that uses the number of live stems per unit area and height of the dominant stems to ascertain stand competitiveness. As the level of competitiveness in a stand increases, the relative spacing tends toward a lower limit. Diameter distribution comparisons across varying planting densities and ages, at similar relative spacings, were conducted to aid in trend recognition as a function of the competitive status of the stand.

The diameter distributions of juvenile stands that have the same planting density, but were planted on different row-by-column spacings (rectangularity) were compared. The results of this comparison could have an impact on operational planning. Significant differences would indicate that the rectangularity of a specific planting density has an impact on the development of the diameter distribution. A lack of significant differences would suggest that similar stand characteristics result from different row-by-column spacings for juvenile stands with the same planting density, allowing for more operational flexibility in row-by-column planting spacings.

The major initial factors influencing the dynamics and growth of a juvenile stand are planting density, genetics, competing vegetation, and micro-site effects (soil, weather, topography, nutrient availability, etc.). Stem diameter is highly influenced by these factors. Spatial analysis that describes the inter-tree dependencies will enable a more precise representation of the forest stand. To achieve this, the coordinates (x,y) and diameter of all stems in a plot were employed in a distance-dependent spatial model. The spatial dependence between trees in a stem-mapped plot was derived using spatial data analysis techniques. The empirical semivariogram, Moran's I, and Geary's C coefficients are statistics that quantify the magnitude and direction (positive or negative) of spatial

autocorrelation. Further, a conditional autoregressive spatial model was employed to account for the inter-tree influence among stems. The conditional autoregressive spatial model contains a spatial dependency parameter that captures the inter-tree influences. Estimation of this spatial dependency parameter will aid in describing stand dynamics and competition processes that are influenced by the spatial interactions of stems in a stand.

The presence of spatial autocorrelation between stems in a stand indicates that the characteristics of one stem are related to the characteristics of other stems. The relationship between stems is expected to be stronger at close distances and weaker at greater distances. Knowledge of the extent to which neighboring stems influence each other would be a valuable tool to aid forest management decisions. This research tests the assumption that the characteristics of neighboring stems are uncorrelated at specific spatial scales; certain statistical procedures that are routinely used in forestry require this assumption. By accounting for the spatial autocorrelation among trees in a stand, this research aids in the simulation and forecasting of young stands over a wide range of planting densities.

Modeling the parameters of diameter distributions of juvenile loblolly pine as a function of stand characteristics is needed to assist stand forecasting and simulation. To facilitate more accurate and precise prediction in the juvenile stage, input values such as planting density, age, and height of the dominant stems were considered in modeling the diameter distribution parameters. In predicting the parameters of diameter distributions, a term that directly takes the spatial autocorrelation between stems into account may improve the precision of the prediction models.

Current spatially-explicit stand simulators initiate a stand by randomly assigning tree characteristics from a predicted distribution. The random assignment assumes that the characteristics are spatially unstructured. A significant spatial autocorrelation would negate this assumption. Accounting for this phenomenon should produce more accurate representations of stem-mapped stands as a result of the reflection of inter-tree spatial dependencies in stand simulation. Taking into account the inter-tree dependencies in the initial stand generation will affect how the stems are treated with respect to growth and competition processes as the stems are projected forward in time. Any spatial dependency between stems will ultimately have an effect on stand dynamics and response to management treatments.

2. LITERATURE REVIEW

Loblolly pine is an extremely important commercial tree species in the United States and abroad. Its natural range extends from southern New Jersey to eastern Texas, and south to central Florida (Hardin et al. 2001). Knowing the size distribution of juvenile stands will aid in making proper management decisions by allowing forest land managers to accurately depict stand structure.

2.1. Diameter Distributions

The use of diameter at breast height (D) distributions for explaining the dynamics of tree stands has been in practice for many years. Diameter distributions provide a means of characterizing a stand and enable prediction of stand yield and size class structures. The focus of the diameter distribution work presented here is for single species even-aged plantations. In general, even-aged stands tend to have unimodal diameter distributions that exhibit some degree of skewness. Previous work on diameter distributions can be separated into two classes of methodology, those that use a functional form to describe the diameter distribution, and those that use a method that does not rely on the specification of a functional form, i.e., a nonparametric procedure.

The first class describes a diameter distribution by utilizing a probability density function (pdf) as a representation of the population diameter distribution. There have been many distributional forms suggested for characterizing diameter distributions, e.g., Normal, Exponential, Beta, Johnson's S_B , Gamma, Weibull, Negative Exponential, and Lognormal. Early work by Clutter and Bennett (1965) used a transformed Beta distribution for characterizing the diameter distribution of slash pine. Burkhart (1971)

evaluated the Bennett-Clutter (Bennett and Clutter 1968) diameter distribution yield estimation technique and found that the Bennett-Clutter technique reliably predicted yields for the old-field slash pine plantations.

The methods used for estimating the parameters of the diameter distribution pdf fall into two categories, parameter prediction models and parameter recovery models. Parameter prediction models utilize a specific pdf to characterize the diameter distribution where the parameters of the pdf are estimated directly from the data. The parameter estimates are then related to stand attributes to create prediction equations for the pdf parameters. Parameter recovery models predict stand-level average attributes from the data that are then used to characterize (recover) the underlying diameter distribution (pdf) of the stand. Hyink and Moser (1983) suggest a mathematical compatibility between diameter distribution models and stand-average models, and present an overview of parameter prediction and parameter recovery methods. Brooks et al. (1992) utilized four distributional percentiles and a parameter recovery method to obtain Weibull-based diameter distributions for loblolly and slash pine.

The second class describes diameter distributions utilizing a nonparametric procedure. A nonparametric method smoothes the sample and adds flexibility in shape, to render a better representation of the population distribution that does not require specification of a density function. Droessler and Burk (1989) utilized nonparametric smoothers on sample diameter distributions. The smoothers considered were the kernel method and the frequency polygon-average shifted histogram (FP-ASH; Scott 1985) method. Compared to the three-parameter Weibull distribution, the nonparametric smoothers were found to be a better representation of the cumulative distribution function

in a 10-tree sample. Another nonparametric method was presented by Borders et al. (1987), where a percentile-based procedure was used to predict diameter distributions. The authors utilize twelve percentile points from the observed data for prediction of the diameter distribution. Haara et al. (1997) estimated basal area diameter distributions using a nonparametric k-nearest neighbor method that is based on a weighted distance. The three-parameter Weibull outperformed the k-nearest neighbor technique for stands dominated by Norway spruce, Scots pine, or broadleaves.

Diameter distributions have also been developed for shade tolerant tree species (Lorimer and Krug 1983) and uneven-aged stands (Leak 1964, Leak 1965, Moser 1976). Uneven-aged stands are more likely to have diameter distributions that are multi-modal or monotonically decreasing. These stands are best characterized by nonparametric techniques (Borders et al. 1987) for the multi-modal case, and the Negative Exponential (Moser 1976) for the monotonically decreasing case. The segmented Weibull approach presented by Cao and Burkhart (1984) is a parametric approach that will also fit multi-modal distributions.

2.2. Weibull Distribution

The first comprehensive study on diameter distributions utilizing the Weibull density function was performed by Bailey and Dell (1973). In Bailey and Dell's study the Weibull function was shown to fit the empirical diameter distributions well, and was able to assume a variety of curve shapes. Cao and Burkhart (1984) developed a segmented Weibull distribution approach for modeling diameter distributions of loblolly pine plantations. The segmented technique is especially useful for describing irregular (multimodal) diameter data. The Weibull distribution was employed by Kilkki et al.

(1989) when deriving methods for obtaining the diameter distribution of spruce using only known stand characteristics to describe the Weibull function.

A method for deriving the volume of a stand between two diameter limits that is class-interval-free was presented by Strub and Burkhart (1975). Their method made use of an assumed population density of a transformed Beta or a two-parameter Weibull distribution and an individual tree volume equation. Matney et al. (1987) used a three-parameter Weibull distribution to create a yield system for slash pine. The Weibull distribution has also been applied in the prediction of stand mortality for young loblolly pine (Somers et al. 1980).

Zutter et al. (1986) compared modified tree diameter data types that were fitted with a two-parameter Weibull distribution. The data types reviewed were the complete, left censored, and left truncated diameter data. The authors found that the fit of the Weibull distribution decreased with an increase in the level of truncation or censorship. The left truncated two-parameter Weibull distribution is recommended if the data have a left truncation point (i.e. a minimum diameter class), and will lead to a reduction in bias and an increase in precision of the fitted distribution.

Tree diameter growth models that are compatible with a stand-level diameter distributional form were derived by Bailey (1980). The growth models presented by Bailey (1980) are implied by the specific diameter distribution assumed. The seven functions examined were the Normal, Exponential, Beta, Johnson's S_B , Gamma, Weibull, and Lognormal. Lenhart (1988) developed a system for diameter distribution yield-prediction for young unthinned loblolly and slash pine. A parameter recovery method was used to obtain the parameters for the three-parameter Weibull distribution. The

estimated Weibull distribution was then used for yield prediction. Murray and von Gadow (1991) predicted changes in the diameter distribution before and after thinning. The authors imposed consistency conditions utilizing two separation parameters to ensure that the thinned and remaining diameter distribution add up exactly to the pre-thinned distribution.

Hafley and Schreuder (1977) compared various distributional forms by reviewing the fit of the Beta, Johnson's S_B , Weibull, Lognormal, Gamma, and Normal distributions. These distributions were compared with respect to their fit to diameter and height data from even-aged loblolly, shortleaf, and longleaf pine stands. Johnson's S_B distribution was found to be the most stable across a variety of data sets. Johnson's S_B distribution was further applied by Knoebel and Burkhart (1991) in a bivariate form for modeling the diameter distribution at two points in time for yellow-poplar.

Maltamo et al. (1995) compared the Beta and Weibull functions for modeling basal area distributions in stands of Scots pine and Norway spruce. The authors found no clear differences between the Beta and the two-parameter Weibull, but the two-parameter Weibull was found to represent the empirical distribution better than the three-parameter Weibull distribution. The authors reasoned that this was due to the two-parameter Weibull allowing the frequency to immediately increase from zero, without being tied to the minimum diameter. Further work with basal area distributions was performed by Maltamo (1997), where the three-parameter Weibull distributions for mixed stands were compared by species. The author found that constructing different models for each tree species yielded much more accurate results in the predicted basal area distributions.

Ek et al. (1975) studied the specification of distributional parameters for the three-parameter Weibull distribution when a specific quadratic mean diameter (diameter of the tree with mean basal area) is desired. Three yield projection methods for loblolly pine were compared by Borders and Patterson (1990). These authors evaluated the ability to project stand and stock tables by diameter class utilizing a Weibull parameter recovery model, a percentile-based model, and a basal area growth projection method. The basal area growth projection method performed best for most of the comparisons in the study.

2.3. Estimation Techniques

Burk and Newberry (1984) employed the first three non-central moments of the diameter random variable to estimate the Weibull distribution parameters using a parameter recovery technique. Rennolls et al. (1985) utilized the Weibull distribution to characterize diameter distributions of conifers. The parameters were estimated by maximum likelihood and weighted least squares regression, and a family of Weibull distributions indexed by diameter was obtained.

Shiver (1988) evaluated three methods (maximum likelihood, percentile, and moments estimation) for estimating the three-parameter Weibull distribution and the approximate number of sample trees required to describe the diameter distributions to a certain error threshold for diameters in unthinned slash pine plantations. Shiver also evaluated. It was found that maximum likelihood estimation had the best fit. Furthermore, if the estimated distribution is to have less than 10% error in any one class, the approximate number of sample trees needed is 50.

To validate the usage of a particular pdf function, various goodness-of-fit tests can be applied. Literature has differed on which distributions are preferred, due in part to

the varying methods used to estimate the parameters and the statistical tests applied for model selection. Reynolds et al. (1988) developed an error index to aid in model selection and validation. The error index is a weighted sum of absolute differences between observed and predicted diameter distributions. It is an improvement over the standard goodness-of-fit tests because it differentiates between errors in a smaller diameter class and those in a larger diameter class. Techniques for calibrating diameter distributions using additional information were performed by Kangas and Maltamo (2000) on predicted percentile and three-parameter Weibull distributions. The calibration, which utilized the number of stems and basal area in each specified diameter class, improved the accuracy of the results for the diameter distributions of Scots pine.

McTague and Bailey (1987) developed a noniterative technique for obtaining the parameters of a Weibull distribution using parameter recovery techniques. The predicted diameter distribution using the noniterative technique was consistent with the predicted basal area. Gove and Patil (1998) modeled basal area distributions using weighted distribution theory, where the density function can represent either basal area or tree frequency.

An interactive nonlinear optimization program, GINO, was applied by Gove and Fairweather (1989) to find maximum likelihood estimates (MLEs) for the two- and three-parameter Weibull distribution. Another program, WEIBUL, developed by Zutter et al. (1982) estimates the parameters for a two- or three-parameter Weibull distribution and allows for censored or truncated data. The MLEs are obtained using an iterative process that searches for parameters over a bounded solution space. Garcia (1981) presents a

simplified method that computes the parameter estimates for a three-parameter Weibull distribution using a simplified method-of-moments technique.

2.4. Juvenile Growth

Zhang et al. (1996) concluded that stand density has a significant effect on diameter and crown ratio development for juvenile loblolly pine. Sharma et al. (in press) studied the mean response to planting density of diameter at breast height, survival, height, and measures of crown size for a loblolly pine spacing study trial. The results indicated that all of the characteristics were affected by density over a 16 year period from the time of planting. The characteristic most significantly affected by density was diameter at breast height. Pienaar and Shiver (1993) found significant differences in the average diameter at breast height across varying planting densities (100-1000 trees per acre) for an old-field loblolly pine spacing study at age 8.

Harms et al. (2000) compared the growth patterns of a loblolly pine spacing trial planted in Hawaii with that of a spacing trial planted in South Carolina. Measurements of diameter at breast height were taken at ages 4, 7, 11, 20, 25, 26, and 34 years. The comparisons were made using a loblolly pine model developed by Hafley et al. (1982). In Hawaii there were significant differences in the mean stand diameter between the low planting density (3.7 meter spacing) and the high planting density (1.8 meter spacing) plots by age 4, indicating an effect on the mean of the diameter distribution from planting density. Harms et al. (2000) report that the diameter growth response to spacing is of the same magnitude when comparing the Hawaii and South Carolina locations, but that the Hawaii plots have higher overall growth.

Lee and Lenhart (1998) showed that the average diameter values for loblolly pine in East Texas in the 15 and 20 year age classes were significantly larger for lower planting densities. The planting densities were grouped into five planting density classes (≤ 400 , 500, 600, 700, and ≥ 800 trees per acre) and four age classes (5, 10, 15, and ≥ 20 years). A one-way analysis of variance (ANOVA) resulted in no significant differences ($\alpha=0.05$) in the average diameter between the five planting density groups for age classes 5 and 10 years. The average diameters were significantly larger for lower planting densities (400 and 500) for age classes 15 and 20 years than for higher planting densities (Lee and Lenhart 1998). Sharma et al. (in review) evaluated the effect of planting rectangularity on stand characteristics. The authors found that for 1-centimeter diameter classes summed over all plots with similar planting rectangularities, there were no significant differences in the empirical diameter distributions of loblolly pine for ages 5 to 16.

2.5. Spatial Effects

Neighboring stems in a forest may have strong spatial correlations. Spatial correlation is defined as the strength of the dependency and influence an observation has on its neighbors. Spatial analysis tries to account for the correlation between observations to yield more accurate models. A spatial correlation structure present in a forest stand will violate the assumption that observations made within the stand are independent of each other. Violation of this assumption may affect the results of any statistical procedure that requires an assumption of independent and identically distributed data. Magnussen (1990) modeled the spatial covariance in tree height for tree-genetics field trials. The models presented in Magnussen (1990) are adjusted to

account for the spatial covariance in the data, but the author notes that one will never be able to account for all spatial effects. A nearest neighbor approach was applied to a jack pine progeny study by Magnussen (1994) where heights and diameters were analyzed for spatial correlation. Accounting for the spatial correlation resulted in a reduction of 37 and 54% of the genetic family variance for height at ages 12 and 17, respectively. The correction did not have a large influence on the genetic family variance for diameters because the spatial correlation was weak at age 12 and almost zero at age 17. The lack of correlation among the genetic family variance for diameters was attributed to competition effects within the forest stand (Magnussen 1994).

Reed and Burkhart (1985) assessed the levels of spatial autocorrelation in tree characteristics (product, defect, species class, and basal area) utilizing Moran's I (Moran 1950) and Geary's C (Geary 1954) indices. The spatial autocorrelation of basal area did not show any strong trends with age, site index, or density for any of the data sets considered. The authors speculate that the spatial autocorrelation of tree basal area for plots with low levels of competition should be positive between neighboring stems. Plots that have intermediate levels of competition should tend to have an increasingly negative spatial autocorrelation. While plots that have extremely high levels of competition should have measures of spatial autocorrelation that are positive, possibly due to stems growing at similar rates in dense stands.

A second-order trend surface was used by Liu and Burkhart (1994) to account for competition effects on diameter and total height of loblolly pine stands. Residuals from an independent Normal random variable that was generated and fitted to the trend surface were utilized to calculate Moran's I as a measure of spatial autocorrelation in the data.

The authors found a general declining trend in the Moran indices with an increase in age. The Moran index went from positively significant to non-significant for diameter measurements.

Stoyan and Penttinen (2000) discuss the application of point and marked point processes to problems in forestry. A point process is a stochastic spatial process where events occur at random locations. A marked point process is a point process with a random variable measured at each location. Penttinen et al. (1992) use a marked point process on a spruce stand to evaluate the spatial correlation of tree diameters. The authors found no significant spatial correlation in tree diameters and noted that this may be a result of the stand having been thinned several times. Penttinen et al. (1992) further utilized marked point processes on a mixed birch-pine stand and found a strong negative spatial correlation for tree diameters. Moeur (1993) employed nearest neighbor analysis and Ripley's K (Ripley 1976) function for describing the spatial correlation between stem mapped data. The nearest neighbor analysis performed by Moeur (1993) used only the distance to the nearest tree to define the empirical cumulative distribution of distances. The empirical cumulative distribution of distances is compared with the cumulative distribution of distances that assumes spatial randomness, i.e., a Poisson process. If the empirical cumulative distribution is plotted against the cumulative distribution, the direction of departure from the cumulative distribution shows the departure from spatial randomness (i.e. clustered or uniform). Ripley's K analysis uses the distances between all pairs of trees in the data to evaluate spatial patterns. Gates and Westcott (1981) utilized zones of influence to spatially model neighboring plants and showed that

competition affects the skewness of the size distribution and correlations in plant biomass between neighboring individuals.

3. DATA

The data used for this study come from the Virginia Polytechnic Institute and State University Loblolly Pine Growth and Yield Research Cooperative. The cooperative installed a loblolly pine spacing study in 1983 at four locations in Virginia and North Carolina, with two locations each in the Coastal Plain and Piedmont regions. The objective of this study was to examine the effects of different planting densities on the stand and stem dynamics of loblolly pine in the Southeast. At each location three replicates of the study design were installed with sixteen plots per replication, for a total of 192 plots. The base design is a grid with four spacings (4, 6, 8, and 12 feet) randomly assigned on each axis (see Figure 1). Table 1 lists the spacing, planting density in trees per acre (TPA), and planting rectangularity for the sixteen plots contained in each replication. The total number of plots over the entire study established in each planting density class is given in Table 2. Each replication follows a non-systematic experimental design presented by Lin and Morse (1975) which allows the spacing to vary in two dimensions and have a constant number of trees per plot. Each plot had an initial planting of 49 trees, which were planted on a 7-by-7 grid based on the row-by-column spacing of that plot. Three rows and columns of trees act as a buffer between all plot edges to minimize the adjacency effects of surrounding plots. The loblolly pine seedlings planted were 1-0 genetically improved stock grown in the same nursery. Amateis et al. (1988) provide additional details on the layout of the study.

Individual tree measurements have been recorded annually during the dormant

Figure 1: A selected example of a spacing study replication

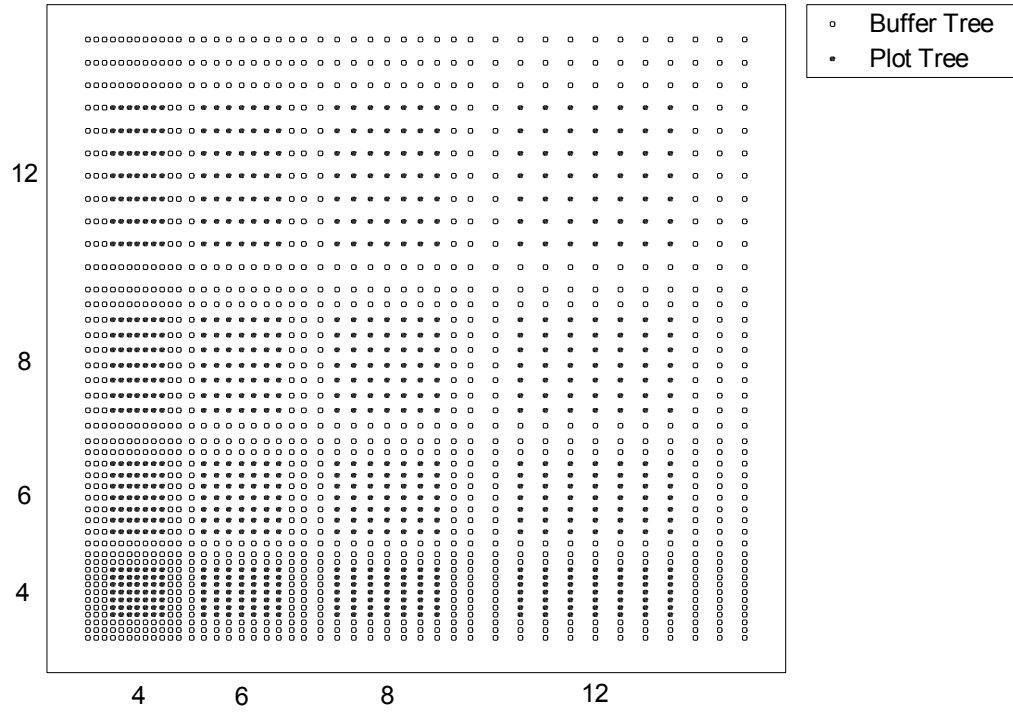


Table 1: Spacing, planting density (trees per acre), and planting rectangularity for the 16 plots within each replication

Plot Number	Spacing, ft	Planting Density	Rectangularity
1	4-by-4	2722	1:1
2	4-by-6	1815	1:1.5
3	4-by-8	1361	1:2
4	4-by-12	907	1:3
5	6-by-4	1815	1:1.5
6	6-by-6	1210	1:1
7	6-by-8	907	1:1.33
8	6-by-12	605	1:2
9	8-by-4	1361	1:2
10	8-by-6	907	1:1.33
11	8-by-8	680	1:1
12	8-by-12	453	1:1.5
13	12-by-4	907	1:3
14	12-by-6	605	1:2
15	12-by-8	453	1:1.5
16	12-by-12	302	1:1

Table 2: Number of plots established for each planting density (trees per acre)

Planting Density	No. of Plots
302	12
453	24
605	24
680	12
907	48
1210	12
1361	24
1815	24
2722	12
Total	192

season since planting. Ground-line diameter (d) was measured through age 5, and diameter at breast height (D) was measured from age 5 onward, where breast height is defined as 4.5 feet above ground level. Tree condition, total height, height to live crown, crown width (both within and between rows), and maximum crown width were recorded through age 10. After age 10, tree condition and D were measured annually, while total stem height and height to the live crown were measured every other year. Herbicide treatments were applied to control competing hardwood vegetation on all of the plots through age 4. Site index, base age 25 years, varied from 60 to 72 feet for the four locations (Sharma et al. (in press)).

Many techniques have been used to define a juvenile stand of trees. Radtke and Burkhardt (1999) utilized the basal area inflection point as the onset of competition for loblolly pine, indicating the end of the juvenile stage. The age at which stands with variable spacings reached this threshold varied from 4 to 10 years. Stand age has also been used (Lenhart 1988) to define young loblolly stands. A versatile growth and yield simulator for loblolly pine, PTAEDA2, generates a precompetitive stand at 8 years of age, which is assumed to be the end of the juvenile phase with competition beginning at 9 years since establishment (Burkhardt et al. 1987). Juvenile stands of loblolly pine were defined in this study as those stands whose age since planting is less than or equal to 11 years. As the stand density (live trees per acre) changes over time, relative spacing will be used for stand comparisons. The diameter distributions compared in this study consist of measurements at ground-line and breast height. Using breast height is problematic in young stands where all of the stems may not have surpassed 4.5 feet in total height (see Table 3). In the data used in this research 0.6% of the stems at age 5, 0.2% of the stems

at age 6, and 0.01% of the stems at age 7 had not reached breast height. By age 8 all of the stems had attained breast height.

Table 3: Distribution of stems by those that exceed and those that do not exceed breast height at each age since planting

Live Stems	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11
> 4.5 ft H	1719	7585	8954	9133	9149	9144	9104	9053	9000	8922
≤ 4.5 ft H	7572	1654	254	51	20	1	0	0	0	0
Total	9291	9239	9208	9184	9169	9145	9104	9053	9000	8922

Where H = Total stem height

4. PRELIMINARY DISTRIBUTIONAL ANALYSIS

The characterization of diameter data by a specific distributional form requires that the distributional form selected be flexible enough to accurately represent possible resulting data values. The two-parameter Weibull distribution has been shown to adequately characterize tree diameter distributions (Bailey and Dell 1973). To predict the diameter distribution of juvenile loblolly pine utilizing the two-parameter Weibull distribution, an overall model containing easily obtained stand-level characteristics as regressors is desired. Stand descriptors such as planting density, relative spacing, age, site index, and descriptive statistics of diameters and heights are necessary as predictors. A thorough investigation into the influence that these descriptors have on the prediction of the diameter distribution parameters of a juvenile stand was performed.

4.1. Weibull Distribution

The Weibull distribution was applied to tree diameter distributions by Bailey and Dell (1973) because of the ease with which the parameters can be estimated and the resultant distribution integrated, and the distribution's flexibility to assume a variety of shapes. The probability density function (pdf) for the two-parameter Weibull of a random variable x , is:

$$f(x) = \frac{c}{b} \left(\frac{x}{b} \right)^{c-1} \exp \left[- \left(\frac{x}{b} \right)^c \right], \quad (1)$$
$$x \geq 0, b > 0, c > 0$$

where b is the scale parameter, c is the shape parameter, and $\exp[\cdot]$ is the exponential function (the base of the natural logarithm). The random variable x is stem diameter

(ground-line or breast height), and is assumed to be a continuous random variable. The scale parameter is related to the range of the diameters in the distribution, and the shape parameter determines the skewness of the distribution. The two-parameter Weibull distribution is utilized rather than the three-parameter Weibull distribution to allow the lower end of the distribution to shift with the data by setting the location parameter to zero (Maltamo et al. 1995). A visual representation of the variety of shapes that the two-parameter Weibull can represent is given in Figure 2. The shape of the distribution will be right skewed for $1.0 < c < 3.6$, approximately Normal for $c = 3.6$, left skewed for $c > 3.6$, and monotonically decreasing for $c \leq 1.0$. Note that when $c = 1.0$ the Exponential distribution is obtained.

The cumulative distribution function (cdf) for the two-parameter Weibull is:

$$F(x) = 1 - \exp\left[-\left(\frac{x}{b}\right)^c\right], \quad (2)$$

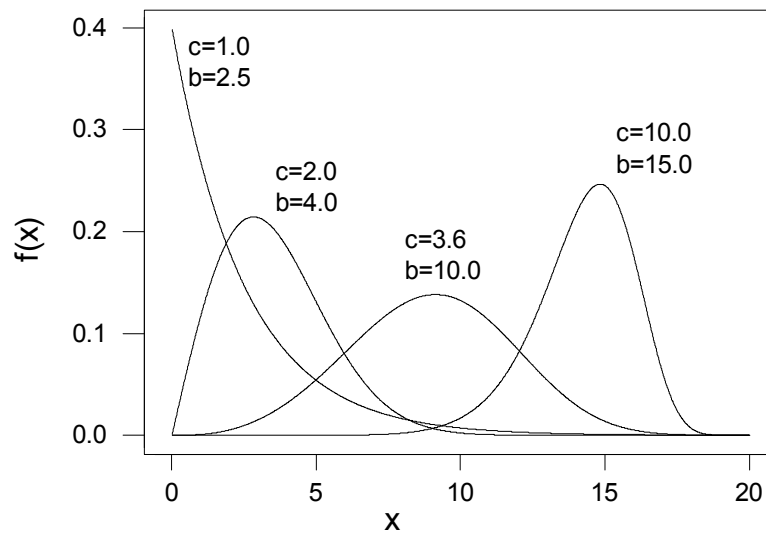
$$x \geq 0, b > 0, c > 0$$

Although there has not been a biological basis presented for the Weibull distribution, the flexibility with which it can assume a variety of unimodal shapes make it particularly useful for describing diameter distributions. Since most multi-modal diameter distributions occur in stands that are uneven-aged (Moser 1976), stands that have been thinned (Murray and von Gadow 1991), or stands that are not single-species (Lorimer and Krug 1983), the Weibull distribution was deemed adequate for this study.

4.2. Weibull Parameter Estimation

The parameters for the Weibull distribution were estimated using maximum likelihood. Some of the statistical properties that make maximum likelihood estimation

Figure 2: Example of shapes attainable with the two-parameter Weibull function



desirable are that they are consistent (asymptotically unbiased) and asymptotically Normal (Bickel and Doksum 1977). Tree diameters (ground-line or breast height) of stems in the same stand may be correlated with other stems that are near the subject tree (Reed and Burkhart 1985, Liu and Burkhart 1994). The term ‘near’ has been left vague because there are many ways to evaluate the spatial effects of inter-tree competition and correlation (see, for example, Daniels 1976, Biging and Dobbertin 1992, Moeur 1993, Magnussen 1994).

In maximum likelihood estimation, the likelihood of a sample of n observations in the data vector $\underline{x} = (x_1 \dots x_n)'$, given the parameter vector $\underline{\theta} = (b \ c)'$ for the two-parameter Weibull, is termed $L(\underline{\theta}; \underline{x})$, where

$$L(\underline{\theta}; \underline{x}) = \prod_{i=1}^n \left(\frac{c}{b} \left(\frac{x}{b} \right)^{c-1} \exp \left[- \left(\frac{x}{b} \right)^c \right] \right) \quad (3)$$

To obtain the maximum likelihood estimate of $\underline{\theta}$, $L(\underline{\theta}; \underline{x})$ is maximized given the observed data. Taking the logarithm of equation (3) yields the log-likelihood function of the diameters (ground-line or breast height) $\log L(\underline{\theta}; \underline{x})$.

$$\log L(\underline{\theta}; \underline{x}) = n \log c - nc \log b + (c-1) \sum_{i=1}^n \log x_i - \sum_{i=1}^n \left(\frac{x_i}{b} \right)^c \quad (4)$$

The estimating equations that follow, (6) and (8), result from taking the partial derivatives of the log-likelihood function with respect to the scale and shape parameters, b and c respectively, and equating them to zero. The subscript $_0$ denotes initial values for the iterative procedure and the subscript $_1$ denotes the subsequently derived values.

$$\frac{\partial \log L}{\partial b} = \frac{-nc}{b} + \left(\frac{c \sum_{i=1}^n x_i^c}{b^{c+1}} \right) \quad (5)$$

$$\hat{b}_1 = \left(\frac{\sum_{i=1}^n x_i^{\hat{c}_0}}{n} \right)^{\frac{1}{\hat{c}_0}} \quad (6)$$

$$\frac{\partial \log L}{\partial c} = \frac{n}{c} + \sum_{i=1}^n \log x_i - n \log b - \sum_{i=1}^n \left(\frac{x_i}{b} \right)^c \cdot \log \left(\frac{x_i}{b} \right) \quad (7)$$

$$\hat{c}_1 = \left(\frac{\sum_{i=1}^n x_i^{\hat{c}_0} \cdot \log x_i}{\sum_{i=1}^n x_i^{\hat{c}_0}} - \frac{\sum_{i=1}^n \log x_i}{n} \right)^{-1} \quad (8)$$

The variance-covariance matrix for the two-parameter Weibull is presented in Cohen (1965) for complete, singly censored, and multiple censored samples. The variance-covariance matrix presented below is analogous to the complete sample case presented by Cohen (1965). The variance-covariance matrix for $\hat{\theta}$ utilizes the maximum likelihood estimates of the Weibull parameters.

$$\begin{bmatrix} -\frac{\partial^2 \log L}{\partial c^2} \Big|_{\hat{b}, \hat{c}} & -\frac{\partial^2 \log L}{\partial c \partial b} \Big|_{\hat{b}, \hat{c}} \\ -\frac{\partial^2 \log L}{\partial c \partial b} \Big|_{\hat{b}, \hat{c}} & -\frac{\partial^2 \log L}{\partial b^2} \Big|_{\hat{b}, \hat{c}} \end{bmatrix}^{-1} \cong \begin{bmatrix} Var(\hat{c}) & Cov(\hat{c}, \hat{b}) \\ Cov(\hat{c}, \hat{b}) & Var(\hat{b}) \end{bmatrix} \quad (9)$$

Differentiating (5) and (7) the following equations are obtained:

$$-\frac{\partial^2 \log L}{\partial c^2} \Big|_{\hat{b}, \hat{c}} = \frac{n}{c^2} + \frac{1}{b^c} \cdot \sum_{i=1}^n x_i^c \cdot \log x_i - \frac{\log b}{b^c} \cdot \sum_{i=1}^n x_i^c \quad (10)$$

$$-\frac{\partial^2 \log L}{\partial b^2} \bigg|_{\hat{b}, \hat{c}} = \frac{-nc}{b^2} + \frac{c^2}{b^{c+2}} \cdot \sum_{i=1}^n x_i^c + \frac{c}{b^{c+2}} \cdot \sum_{i=1}^n x_i^c \quad (11)$$

$$-\frac{\partial^2 \log L}{\partial c \partial b} \bigg|_{\hat{b}, \hat{c}} = \frac{n}{b} - \frac{c}{b^{c+1}} \cdot \sum_{i=1}^n x_i^c \cdot \log x_i - \frac{1}{b^{c+1}} \cdot \sum_{i=1}^n x_i^c + \frac{c \log b}{b^{c+1}} \cdot \sum_{i=1}^n x_i^c \quad (12)$$

The correlation coefficient between the shape and scale parameters can be easily derived as follows:

$$\rho_{\hat{c}, \hat{b}} = \frac{Cov(\hat{c}, \hat{b})}{\sqrt{Var(\hat{c}) \cdot Var(\hat{b})}} \quad (13)$$

A macro¹ was written to iteratively solve for the maximum likelihood estimates utilizing a trust-region optimization procedure for each plot at each age. The scale and shape parameters are initialized utilizing the percentiles (Dubey 1967, Bailey and Dell 1973, and Zarnoch and Dell 1985) of the observed data in equations (14) and (15). The iterative technique updates the starting values for each subsequent iteration utilizing the previous resulting values.

$$\hat{b}_0 = P_{50} \quad (14)$$

$$\hat{c}_0 = \frac{\log\left(\frac{\log(1-.17)}{\log(1-.97)}\right)}{\log\left(\frac{P_{17}}{P_{97}}\right)}, \quad (15)$$

where P_i = the i^{th} percentile

The Weibull distribution parameter estimates were derived for each of the 192 plots for d at ages 2 to 5, and for D at ages 5 to 11. The parameter estimates enable comparisons of many types to be performed.

To overlay the estimated density function on a histogram of the data with width h, the following function is utilized:

¹ SAS Institute Inc., Cary, NC, USA

$$f(x) = \frac{c \cdot h \cdot 100\%}{b} \left(\frac{x}{b}\right)^{c-1} \exp\left[-\left(\frac{x}{b}\right)^c\right], \quad (16)$$

$$x \geq 0, b > 0, c > 0$$

The histogram interval (width) is determined in SAS using the method of Terrell and Scott (1985). The total sample size is used to derive the number of midpoints and hence the interval width. The minimal interval width is predetermined due to the ground-line and breast height diameters being measured to the nearest tenth-inch diameter class. Example graphs of the D histograms with the overlaid Weibull distributions are presented in Appendix I for a plot from each planting density over an age range of 5-11 years. These graphs provide insight into the fit and flexibility of the Weibull distribution.

4.3. Weibull Distribution Validation

To validate the application of the two-parameter Weibull distribution, goodness-of-fit statistics (Chi-square, Anderson-Darling, Cramer-von Mises, and Kolmogorov-Smirnov) were computed to compare the fit for each individual plot at each age and graphical comparisons of the data and estimated distributions were produced. A macro was written in SAS to calculate these goodness-of-fit statistics and simultaneously overlay each data histogram with the estimated Weibull distribution. The goodness-of-fit statistics test the null hypothesis that the empirical data come from the specified distribution.

The Chi-square goodness-of-fit test statistic is defined as follows:

$$\chi_{obs}^2 = \sum_{i=1}^m \frac{(O_i - E_i)^2}{E_i} \sim \chi_{(m-p-1)}^2, \quad (17)$$

where O_i = the observed percentage of data in the i^{th} histogram interval

E_i = the expected percentage of data in the i^{th} histogram interval

m = the number of histogram intervals

p = the number of estimated parameters

The degrees of freedom (df) for the test are $df_{\text{Chi-sq}} = m - p - 1$. The Chi-square goodness-of-fit test is dependent on the histogram interval chosen, and may lead to different test statistics and p-values based on the specific histogram interval utilized. Hence, differing conclusions on the null hypothesis that the data follow a specific cdf may be drawn for the same data set.

The Anderson-Darling, Cramer-von Mises, and Kolmogorov-Smirnov test statistics are based on the empirical distribution function (edf) of each plot. The edf is a step function with a height of $\frac{1}{n}$ for the ordered data set $x_{(1)}, \dots, x_{(n)}$, and is denoted $F_n(x)$.

The edf assumes that the observations are independent and can be defined as:

$$F_n(x) = \begin{cases} 0, & x < x_{(1)} \\ \frac{i}{n}, & x_{(i)} \leq x < x_{(i+1)} \quad i = 1, \dots, n-1 \\ 1, & x \geq x_{(n)} \end{cases} \quad (18)$$

The edf is an estimate of the cdf, where the edf describes the proportion of observations less than or equal to x , and the cdf describes the probability that a specific observation is less than or equal to x . The Anderson-Darling, Cramer-von Mises, and Kolmogorov-Smirnov tests measure the differences between the edf, $F_n(x)$, and the cdf, $F(x)$, for each plot. They are superior to the Chi-square test because they are independent of the histogram interval.

The Anderson-Darling test statistic is defined as follows:

$$AD = n \int_{-\infty}^{+\infty} \frac{(F_n(x) - F(x))^2}{F(x)[1 - F(x)]} dF(x) \quad (19)$$

The Cramer-von Mises test statistic is defined as follows:

$$CM = n \int_{-\infty}^{+\infty} (F_n(x) - F(x))^2 dF(x) \quad (20)$$

The Kolmogorov-Smirnov test statistic is given by:

$$KS = \sup_x |F_n(x) - F(x)| \quad (21)$$

where $\sup_x$ = the supremum evaluated over x

The p-values for each plot across ages 5 to 11 were computed for the Chi-square, Anderson-Darling, Cramer-von Mises, and Kolmogorov-Smirnov goodness-of-fit statistics. The proportion of the 192 plots at each age that rejected the null hypothesis that the data come from the estimated cdf is reported in Table 4, enabling comparisons of the estimated two-parameter Weibull distributions to the observed diameter data. Out of the 1344 fitted diameter distributions (α -level of 0.20), 36.8% were rejected by the Chi-square test, 0.8% were rejected by the Anderson-Darling test, 0.7% were rejected by the Cramer-von Mises test, and the Kolmogorov-Smirnov test rejected 3.1%. The Chi-square test is dependent on the choice of the histogram interval, which was determined by the method presented by Terrell and Scott (1985), and may give a false sense of lack-of-fit for the distributions. Given this inconsistency in the Chi-square goodness-of-fit statistic, only the Anderson-Darling, Cramer-von Mises, and Kolmogorov-Smirnov test statistics were used to evaluate the distributional fit.

In general, with increasing age, a decreasing trend in the number of plots that reject the null hypothesis is present. The negative trend present indicates that the Weibull

Table 4: Results for the Chi-square, Anderson-Darling, Cramer-von Mises, and Kolmogorov-Smirnov goodness-of-fit statistics for diameter at breast height at ages 5 to 11

	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Overall
Chi-square								
NS	114	117	109	129	118	134	128	849
S	78	75	83	63	74	58	64	495
P								0.368
Cramer-von Mises								
NS	190	189	189	192	192	192	191	1335
S	2	3	3	0	0	0	1	9
P								0.007
Anderson-Darling								
NS	189	188	189	192	192	192	191	1333
S	3	4	3	0	0	0	1	11
P								0.008
Kolmogorov-Smirnov								
NS	175	182	184	192	190	189	191	1303
S	17	10	8	0	2	3	1	41
P								0.031

Where NS = The number of plots that are non-significant at the 0.20 alpha level

S = The number of plots that are significant at the 0.20 alpha level

P = Proportion of rejections of the null hypothesis

distribution does an increasingly better job of characterizing the diameter distribution of the plots. The goodness-of-fit statistics test the null hypothesis that the data come from a specific distribution with specified parameters. Utilizing an alpha level of 0.20 all three test statistics indicate that there is not strong enough evidence to reject the null hypothesis in all but a few cases. Though other distributional forms are possible, the two-parameter Weibull distribution has been shown in this research to be useful in characterizing diameter distributions of juvenile loblolly pine.

The two-parameter Weibull distribution has the location parameter set equal to zero, which enables the pdf to immediately increase from zero. To address the concern that the two-parameter Weibull distribution may artificially make all of the estimated distributions left skewed because of the exclusion of the location parameter, the skewness of each individual plot was calculated for D at each age. The coefficient of skewness is defined as:

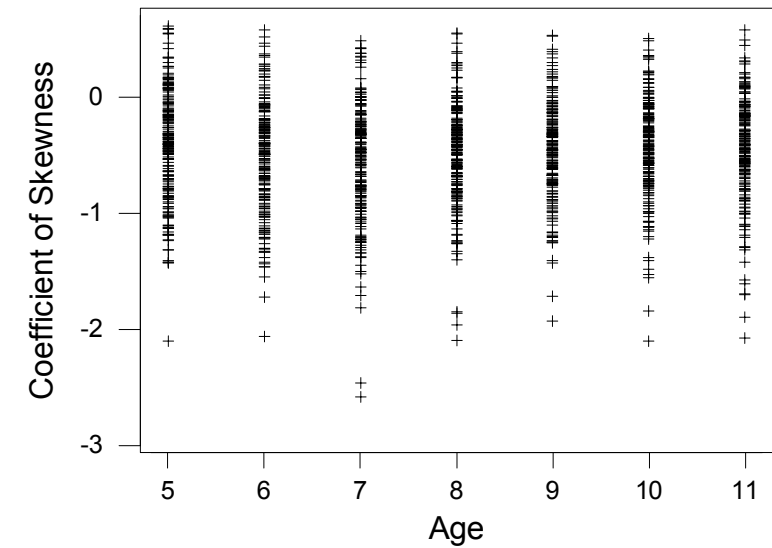
$$c_s = \left(\frac{n}{(n-1)(n-2)} \right) \cdot \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^3, \quad (22)$$

where s = the standard deviation of the sample

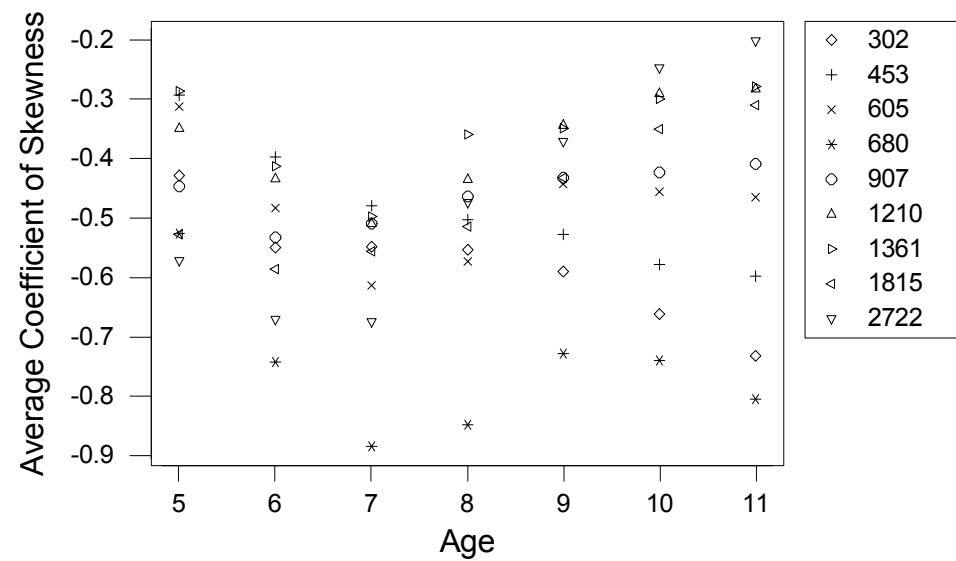
The deviation of each observation from the overall mean is cubed, hence maintaining the direction of the deviation. These deviations are summed and a large positive value indicates that the distribution is right skewed, and a large negative value indicates left skewness. The results show that for the 1344 coefficients of skewness estimated for D (192 plots, ages 5-11), only 177 coefficients (13.17%) were greater than zero (Figure 3 (a)), indicating that the observed diameter distributions are mostly left

Figure 3: Trends for (a) the coefficient of skewness over time for diameter at breast height and (b) the average coefficient of skewness over time for diameter at breast height, comparing all planting densities (trees per acre)

(a)



(b)



skewed. The skewness of the empirical diameter distributions should become right skewed as the trees mature, though the juvenile trees studied herein exhibit mostly left skewness. Figure 3 (b) displays the average coefficient of skewness for each of the planting densities for ages 5 to 11. A trend between the average coefficient of skewness and planting density can be discerned at age 8. The more densely planted stands tend to be less skewed to the left as age increases, while those stands with lower planting densities tend to have a higher degree of skewness to the left. The early onset of competition in the more densely planted stands may account for this phenomenon in the coefficients of skewness. As competition sets in, one would expect diameter distributions to move from a left skewed distribution to a more symmetric distribution. The coefficients of skewness indicate that most of the empirical diameter distributions are left skewed, which supports the estimates of the shape parameter from the two-parameter Weibull distribution being larger than 3.6 for most plots.

The kurtosis of each plot was also derived for D at each age. Kurtosis is a measure of the heaviness of the tails and the peakedness of the sample when compared to the Normal distribution. A negative kurtosis value indicates that the sample is platykurtic (flat topped, thinner tails), while a positive value for the kurtosis indicates that the sample is leptokurtic (peaked, fatter tails). The coefficient of kurtosis is defined as follows:

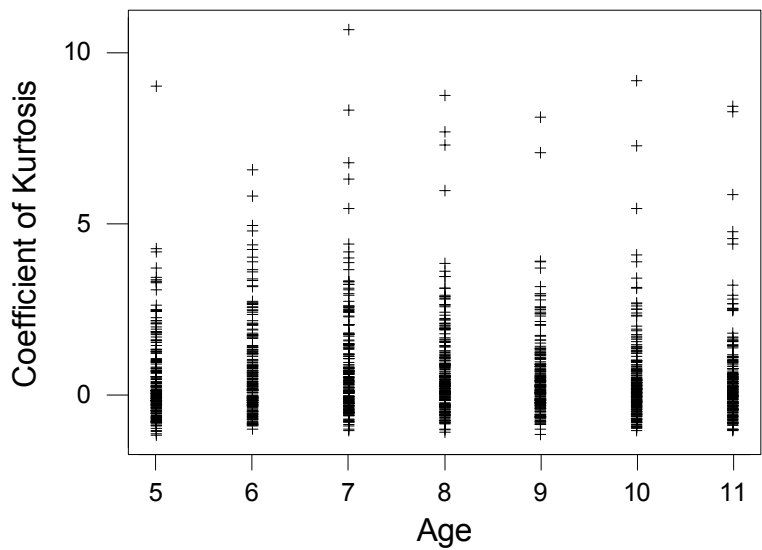
$$c_k = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \cdot \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^4 - \frac{3(n-1)}{(n-2)(n-3)} . \quad (23)$$

The deviations from the overall mean are raised to the fourth power, hence the direction of the deviation cannot be ascertained from the kurtosis value. For the 1344

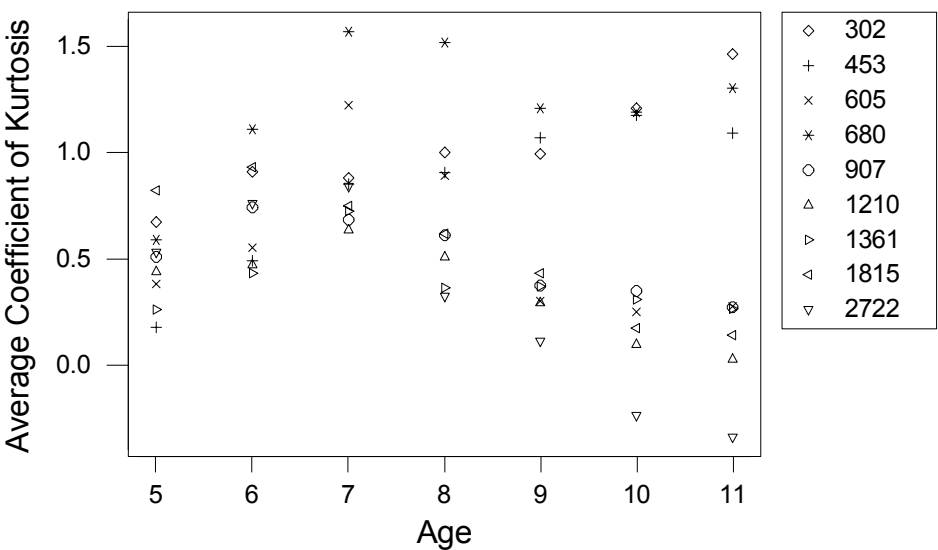
coefficients of kurtosis estimated for D, 829 coefficients (61.68%) were greater than zero (Figure 4 (a)). The large number of kurtosis coefficients greater than zero indicates that a majority of the observed diameter distributions are more peaked with fatter tails when compared to the Normal distribution. The average coefficient of kurtosis is presented in Figure 4 (b) for each of the planting densities for ages 5 to 11. One would expect data that follows a Weibull distribution (with a shape parameter ≈ 3.6) to be more peaked and have a fatter tail when compared to the Normal distribution. The coefficients of kurtosis support the use of the Weibull distribution because a majority of the empirical diameter distributions have a c_k greater than zero. A similar trend to the one noted in Figure 3 (b) is present in Figure 4 (b) between the average coefficient of kurtosis and the planting density. The trends present indicate that at higher planting densities the distributional peakedness and heaviness of the tails is diminishing. Beginning at age 8, the difference between the more densely planted stands and those planted at lower densities was noted, and the divergence between the trends increases with age. The disparity between the planting densities for the coefficient of kurtosis is attributed to the earlier onset of competition in the more densely planted stands.

Figure 4: Trends for (a) the coefficient of kurtosis over time for diameter at breast height and (b) the average coefficient of kurtosis over time for diameter at breast height, comparing all planting densities (trees per acre)

(a)



(b)



5. COMPARISONS OF ESTIMATED DISTRIBUTIONS

To elucidate the trends in the diameter distributions of juvenile loblolly pine, comparisons were made on the estimated parameters of the two-parameter Weibull distribution. These comparisons were performed on constant factors (i.e. planting density) and non-constant factors (i.e. age and relative spacing) for each plot.

5.1. Rectangularity Effect

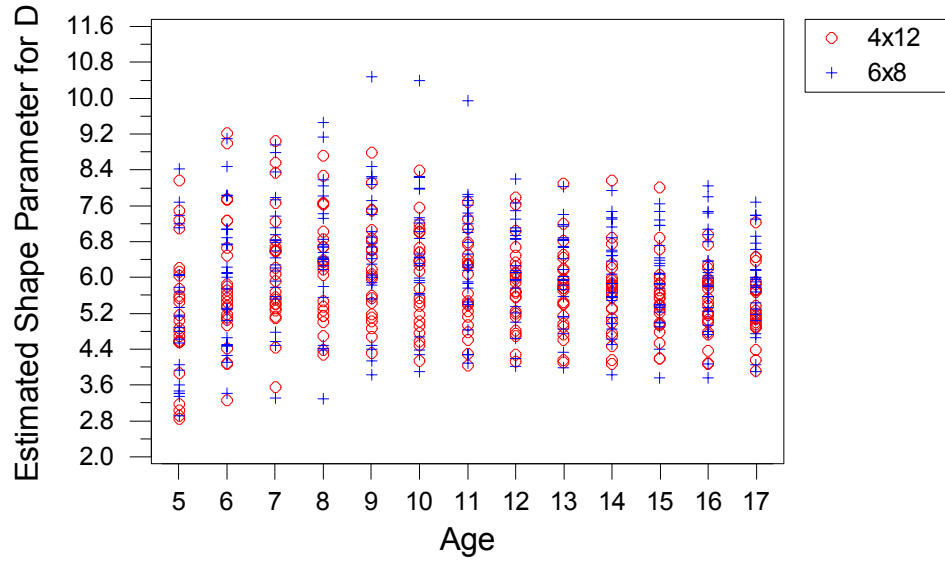
Plot rectangularity is defined as the ratio of the row-by-column spacing. The density of a plot has been previously defined in terms of the planting density. It is hypothesized that there may be a significant rectangularity effect between plots with identical planting densities but different row-by-column spacings. To test this hypothesis, all plots with a planting density of 907 TPA were compared; these plots include the 4-by-12 ft (1:3) and the 6-by-8 ft (1:1.33) spacings. All tests herein assume that row orientation is negligible, i.e. aspect and slope differences are non-significant.

Sharma et al. (in press) showed that for loblolly pine plots with the same planting density, rectangularity did not significantly influence the mean value of D , while the planting density of a plot was shown to have a significant effect on the mean value of D . Sharma et al. (in review) did not find any significant differences in the empirical diameter distributions of 1-centimeter diameter classes with different planting rectangularities. The diameter classes Sharma et al. (in review) used were derived for each age from the summation of stands planted at the same density but having different rectangularities.

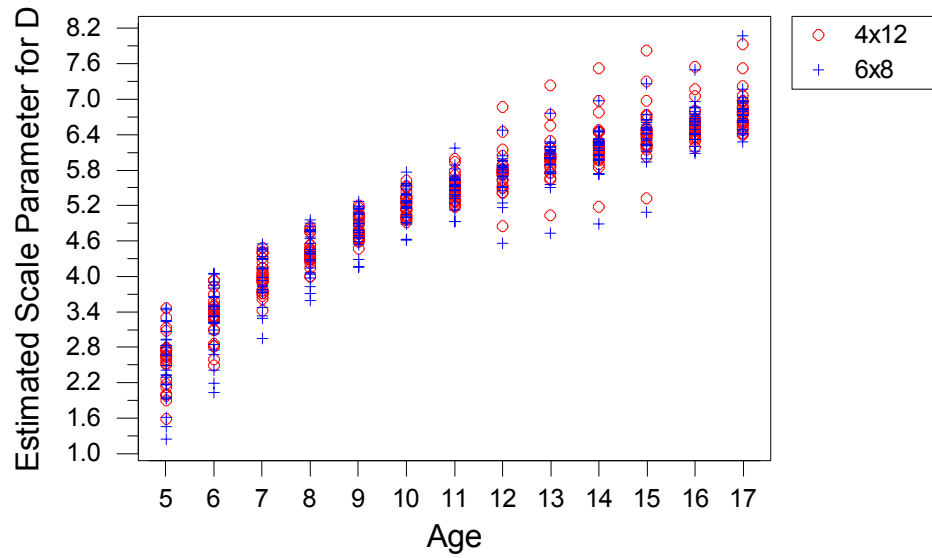
Graphical representations of the estimated parameters for the two-parameter Weibull for plots with a planting density of 907 TPA are presented in Figure 5 (a) for the

Figure 5: Trends over time for plots with a planting density of 907 trees per acre but with different planting rectangularities (1:3 and 1:1.33) for diameter at breast height (D) measurements for (a) the shape parameter estimate and (b) the scale parameter estimate

(a)



(b)



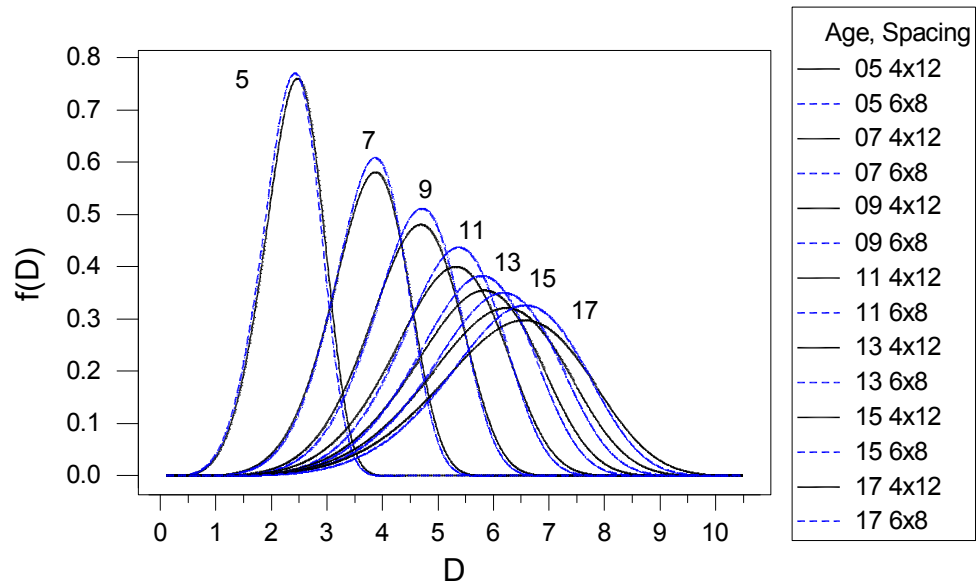
shape parameter, and in Figure 5 (b) for the scale parameter. These graphs illustrate the lack of differences between plots with a planting rectangularity of 1:3 and 1:1.33 and planting density of 907 TPA. An ANOVA test for mean differences was performed on plots with a rectangularity of 1:3 and 1:1.33 on a basis of the estimated Weibull parameters at each age (α -level of 0.05). The results are presented in Table 5, with 24 observations (plots) in each spacing and 46 error degrees of freedom for each test of mean differences. Table 5 also presents the average estimated Weibull parameters for the plots with a rectangularity of 1:3 and 1:1.33 for ages 5 to 17. The average parameter estimates were inserted to create Figure 6, a representation of the average distributions every two years (ages 5-17) for D. Figure 6 shows that the estimated average diameter distributions for the 6-by-8 ft plots are consistently more peaked than the 4-by-12 ft plots.

No significant differences were found between the average scale parameter estimates for plots with a rectangularity of 1:3 and 1:1.33. The decreasing trend in the p-values for the shape parameter estimates becomes significant at age 16 (p-value = 0.043), and then returns to non-significance at age 17. The rejection of the null hypothesis that there are no differences in the means of the scale parameters at age 16 may be due to a type I error occurrence. The type I error rate is the probability of rejecting the null hypothesis given that the null hypothesis is true. In the tests presented above, the type I error rate was controlled for the comparisonwise error, increasing the chance of at least one type I error occurring when multiple tests are performed. There were seventeen individual ANOVA tests for mean differences performed for each parameter estimate. Assuming the comparisons are independent, the approximate chance of making at least

Table 5: Analysis of variance results and average parameter estimates for the scale and shape parameters for ground-line diameters (d) and diameters at breast height (D) of plots with a planting density of 907 trees per acre but different planting rectangularities

Diameter	Age	Scale	Shape	4x12	4x12	6x8	6x8
		Parameter p-value	Parameter p-value	Shape	Scale	Shape	Scale
d	2	0.592	0.702	4.030	1.105	4.140	1.059
d	3	0.757	0.739	5.174	2.067	5.319	2.029
d	4	0.837	0.769	5.677	2.868	5.813	2.838
d	5	0.905	0.663	6.072	3.605	6.271	3.622
D	5	0.755	0.937	5.202	2.571	5.167	2.522
D	6	0.760	0.898	5.923	3.303	5.978	3.260
D	7	0.881	0.485	6.193	3.982	6.465	3.967
D	8	0.960	0.428	6.227	4.406	6.537	4.401
D	9	0.942	0.292	6.204	4.824	6.620	4.829
D	10	0.825	0.172	6.017	5.211	6.536	5.227
D	11	0.796	0.096	5.872	5.495	6.453	5.514
D	12	0.587	0.126	5.791	5.768	6.252	5.710
D	13	0.537	0.179	5.711	6.033	6.097	5.963
D	14	0.433	0.172	5.628	6.242	6.022	6.151
D	15	0.487	0.113	5.524	6.457	5.954	6.370
D	16	0.878	0.043	5.436	6.607	6.011	6.593
D	17	0.812	0.054	5.397	6.808	5.909	6.784

Figure 6: Graphical representation of the average estimated diameter at breast height (D) distributions every two years for ages 5 to 11, for plots with the same planting density of 907 trees per acre but with different planting rectangularities (1:3 and 1:1.33)



one type I error for each set of parameter tests is:

$$1 - (1 - \alpha_c)^c \quad (24)$$

$$1 - (1 - 0.05)^{17} = 0.582 ,$$

where α_c = the comparisonwise error rate

c = the number of comparisons

Though one cannot say with certainty, the rejection of the null hypothesis for the shape parameter at age 16 could be a type I error occurrence. To control for the experimentwise error, the Bonferroni inequality can be utilized. The Bonferroni adjustment involves dividing the stated alpha level for the experiment by the number of comparisons to be made, in this case:

$$\alpha_c = \frac{\alpha_e}{c} = \frac{0.05}{17} = 0.00294 , \quad (25)$$

where α_e = the experimentwise error rate

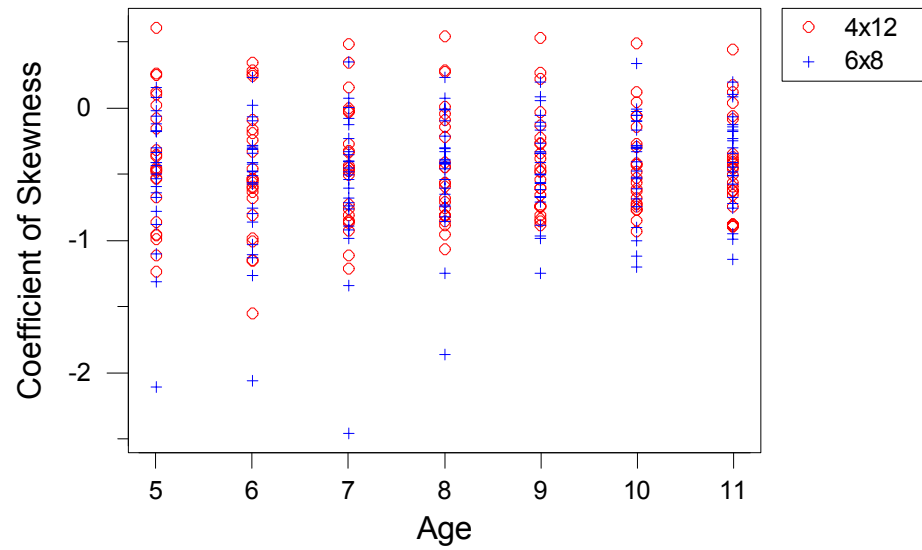
Specifying the alpha level that controls for the experimentwise error rate for all seventeen comparisons, none of the tests for mean differences are significant. Since this study is aimed at quantifying the diameter distributions of juvenile loblolly pine, and there were no significant differences between the mean Weibull parameter estimates for both the shape and scale parameters through age 15 when controlling for either the comparisonwise or experimentwise error rates, the 4-by-12 ft and the 6-by-8 ft plots were combined into one overall group for further analysis. The non-significance of the ANOVA tests indicates that any difference between the population means cannot be discerned from the samples of plots with a planting density of 907 TPA and rectangularities of 1:3 and 1:1.33.

To further aid in discerning any differences between plots with different planting rectangularities but having the same planting density of 907 TPA, the coefficients of skewness and kurtosis were evaluated. The coefficients of skewness (see equation (22)) estimated for D show that of the 336 c_s estimates (48 plots, ages 5-11), only 43 coefficients (12.80%) were greater than zero for plots with a planting rectangularity of 1:3 and 1:1.33 (Figure 7 (a)). These results indicate that the empirical diameter distributions are mostly left skewed for the plots with a planting density of 907 TPA. Figure 7 (b) displays the average coefficient of skewness for the 4-by-12 ft and 6-by-8 ft plots, over time. An ANOVA test for mean differences between these average skewness parameters is presented in Table 6. The results show that there are no significant differences in the mean skewness coefficients for the 4-by-12 ft and 6-by-8 ft plots for ages 5 to 11.

For the 336 coefficients of kurtosis (see equation (23)) estimated for D, 214 c_k estimates (63.69%) were greater than zero for the plots with a planting rectangularity of 1:3 and 1:1.33 (Figure 8 (a)), signifying that most of the observed diameter distributions have thicker tails and are more peaked than the Normal distribution. The average coefficient of kurtosis is presented in Figure 8 (b) for the 4-by-12 ft and 6-by-8 ft plots. A test for mean differences between these plots (Table 6) shows no significant differences for ages 5 to 11. The lack of significant differences between plots with a rectangularity of 1:3 and 1:1.33 for the coefficients of skewness and kurtosis further justifies combining the 4-by-12 ft and 6-by-8 ft plots. Based on this result and the lack of significant differences in the scale and shape parameter estimates, plots planted at the

Figure 7: Coefficient of skewness trends over time for diameter at breast height, comparing plots with the same planting density of 907 trees per acre but with different planting rectangularities (1:3 and 1:1.33) for (a) all plots and (b) the average coefficients

(a)



(b)

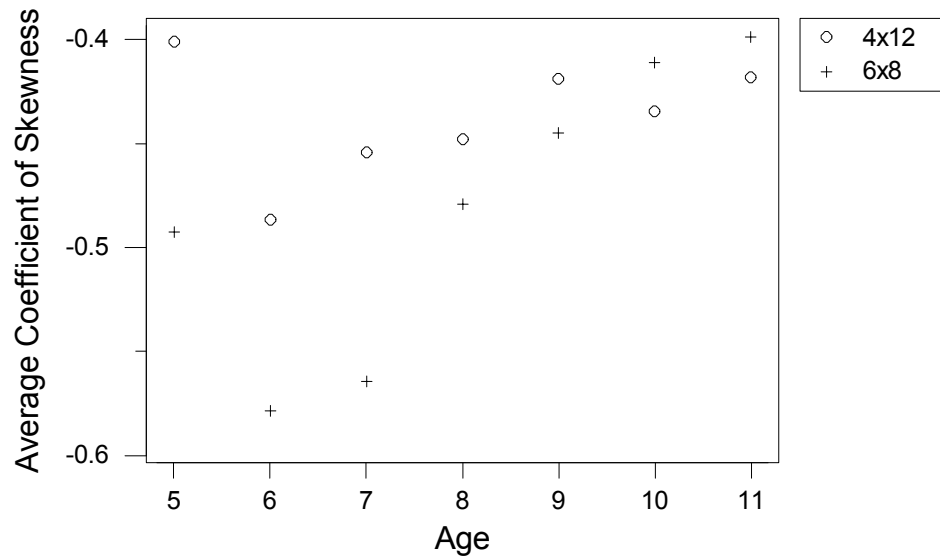


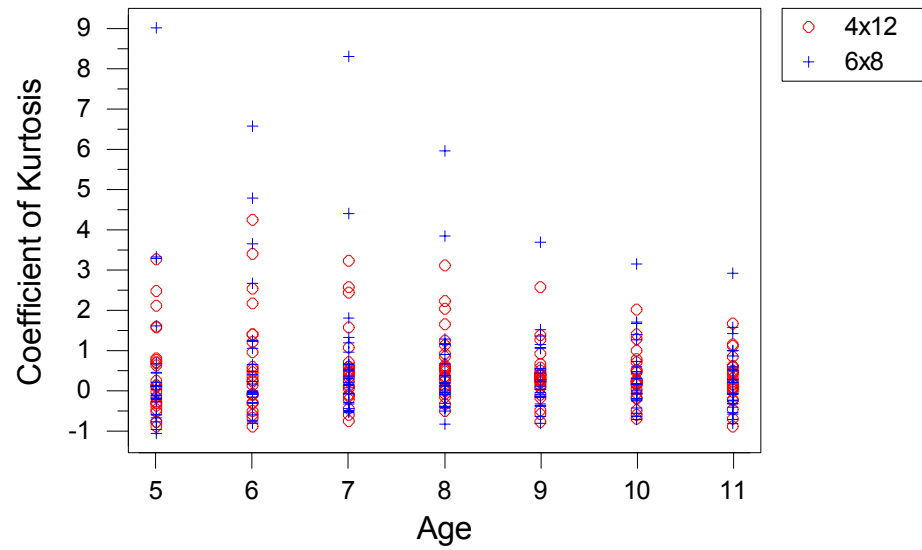
Table 6: Analysis of variance results for the coefficients of kurtosis and skewness for plots with a planting density of 907 trees per acre but with different planting rectangularities (1:3 and 1:1.33)

		Kurtosis	Skewness
Age 5	F	0.07	0.43
	p-value	0.796	0.517
Age 6	F	0.07	0.42
	p-value	0.797	0.520
Age 7	F	0.22	0.58
	p-value	0.640	0.449
Age 8	F	0.09	0.06
	p-value	0.769	0.805
Age 9	F	< 0.01	0.06
	p-value	0.979	0.809
Age 10	F	0.01	0.05
	p-value	0.920	0.826
Age 11	F	0.01	0.04
	p-value	0.943	0.849

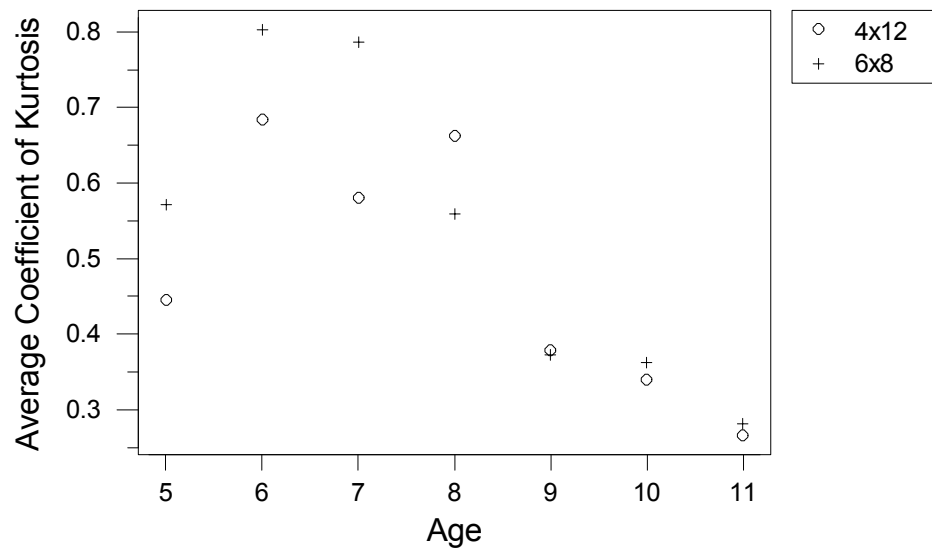
With an error degrees of freedom of 46 for each test, $df_{\text{err}} = 46$

Figure 8: Coefficient of kurtosis trends over time for diameter at breast height, comparing plots with the same planting density of 907 trees per acre but with different planting rectangularities (1:3 and 1:1.33) for (a) all plots and (b) the average coefficients

(a)



(b)



same density (907 TPA) but having different planting rectangularities were combined for all further analysis.

5.2. Temporal Effect

The estimated Weibull parameters were compared at the same planting density over time to elicit any temporal trends and allow for inferences on the juvenile diameter distributions. For diameter at breast height, the comparisons were made over a 7-year period, from ages 5 to 11. The scale and shape parameters were compared separately and the results on an initial TPA basis are presented. Graphical illustrations of the trends in the estimated scale (Figure 9 (a-d)) and shape (Figure 10 (a-d)) parameters are presented for four planting densities from those considered in this research. These graphs illustrate a positive trend between age and the scale and shape parameter estimates. Plots with lower planting densities had visibly steeper trends for both the shape and scale parameter estimates. A curvilinear association between the estimated parameters and age was noted in some cases, implying that linearization may be necessary in modeling the parameters as a function of age.

An ANOVA test for mean differences was used to compare the Weibull parameter estimates to see if the scale or shape parameters for a specific planting density differed significantly over time. Table 7 reveals significant differences in the estimated mean scale parameter across ages 5-11 for all planting densities (all p-values < 0.001), indicating that age has a significant effect on the estimated mean scale parameter of the Weibull distribution for D. One expects this to be the case due to the scale parameter

Figure 9: Trends of the estimated scale parameter for diameter at breast height (D) over ages 5 to 11, for plots with a planting density of (a) 302, (b) 605, (c) 1210, and (d) 2722 trees per acre

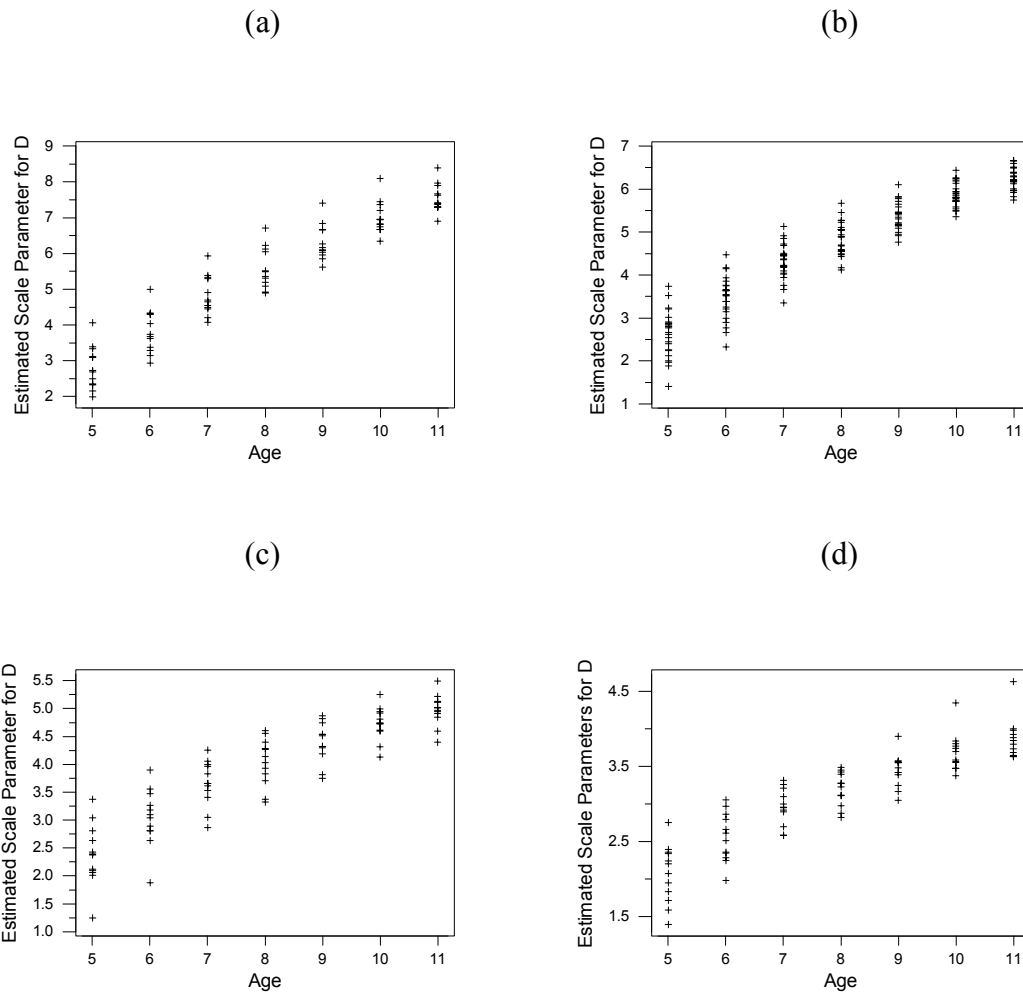


Figure 10: Trends of the estimated shape parameter for diameter at breast height (D) over ages 5 to 11, for plots with a planting density of (a) 302, (b) 605, (c) 1210, and (d) 2722 trees per acre

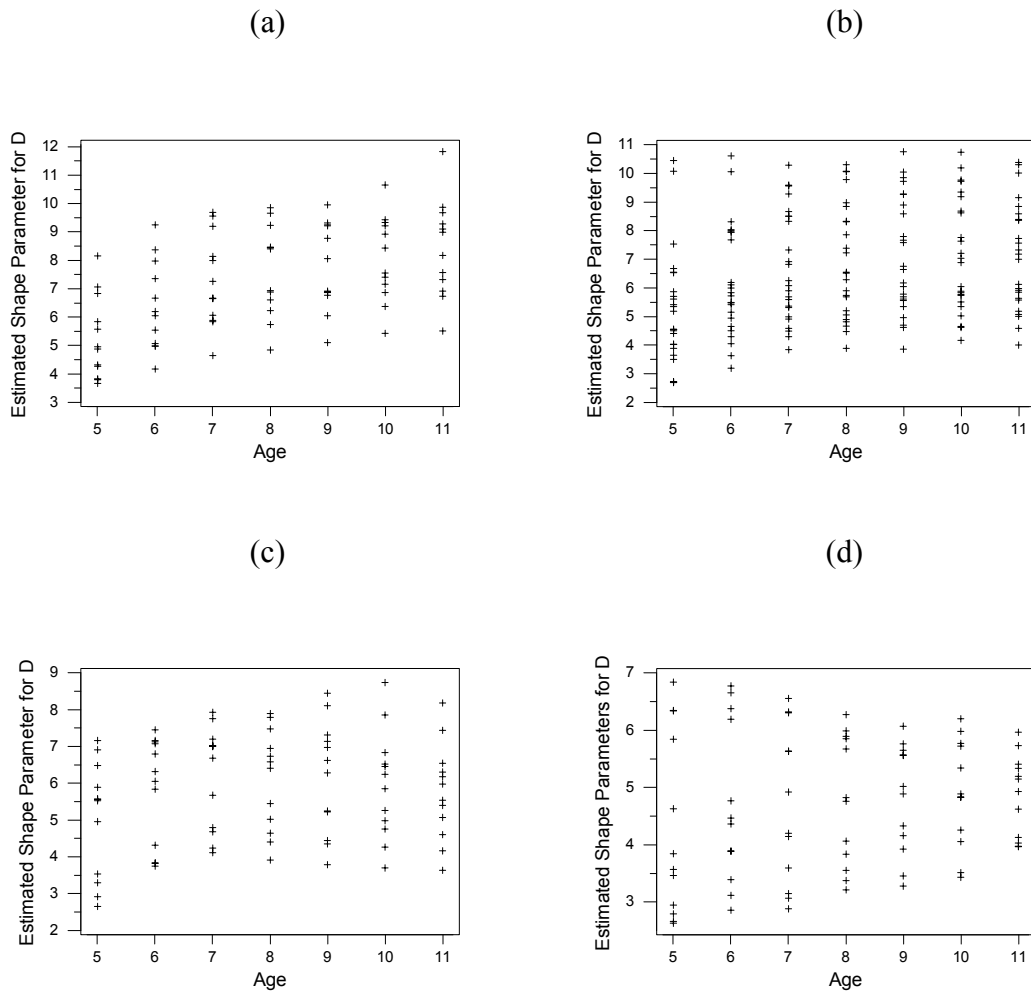


Table 7: Analysis of variance and correlation results for the scale parameter estimates by planting density for diameter at breast height

	302 TPA		453 TPA		605 TPA		680 TPA		907 TPA	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age 5	2.816	0.602	2.716	0.643	2.633	0.542	2.662	0.631	2.546	0.526
Age 6	3.819	0.599	3.636	0.634	3.471	0.505	3.504	0.607	3.281	0.471
Age 7	4.833	0.550	4.543	0.499	4.306	0.407	4.284	0.434	3.975	0.346
Age 8	5.575	0.576	5.154	0.477	4.847	0.396	4.781	0.417	4.403	0.312
Age 9	6.305	0.499	5.751	0.401	5.372	0.333	5.293	0.316	4.827	0.264
Age 10	7.007	0.464	6.326	0.329	5.866	0.277	5.760	0.271	5.219	0.245
Age 11	7.543	0.396	6.754	0.278	6.226	0.248	6.077	0.233	5.504	0.247
df _{err}	77		161		161		77		329	
F	124.16		217.15		252.45		93.91		418.82	
p-value	< 0.001		< 0.001		< 0.001		< 0.001		< 0.001	
Pearson	0.947		0.935		0.941		0.927		0.928	
p-value	< 0.001		< 0.001		< 0.001		< 0.001		< 0.001	

	1210 TPA		1361 TPA		1815 TPA		2722 TPA	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age 5	2.385	0.550	2.340	0.498	2.266	0.451	2.071	0.388
Age 6	3.047	0.514	2.972	0.454	2.847	0.403	2.558	0.327
Age 7	3.658	0.407	3.564	0.324	3.339	0.330	2.957	0.244
Age 8	4.037	0.421	3.904	0.303	3.627	0.320	3.204	0.225
Age 9	4.393	0.357	4.241	0.269	3.914	0.284	3.445	0.222
Age 10	4.731	0.302	4.554	0.256	4.181	0.265	3.688	0.255
Age 11	4.969	0.282	4.785	0.249	4.385	0.259	3.867	0.276
df _{err}	77		161		161		77	
F	60.00		151.70		120.17		61.09	
p-value	< 0.001		< 0.001		< 0.001		< 0.001	
Pearson	0.894		0.907		0.888		0.896	
p-value	< 0.001		< 0.001		< 0.001		< 0.001	

Where TPA = Trees per acre

SD = Standard deviation

df_{err} = Error degrees of freedom

being related to the range of the diameters in the distribution and this range should change over time. Therefore, modeling the scale parameter as a function of age should prove useful. The Pearson correlation between the scale parameter and age, using one observation from each plot, is also given in Table 7 and shows significant positive correlation across all planting densities. The following linear regression shows the strength of the relationship between the estimated scale parameter, age, planting density, and average height of the dominant stems.

$$\text{Log}(\hat{b}) = 0.1425 - 0.02536\text{Log}(A) - 0.2114\text{Log}(TPA) + 0.8705\text{Log}(H) \quad (26)$$

$$R^2 = 0.940, \text{MSE} = 0.006,$$

where $A = \text{Age}$

$H = \text{Average height of the dominant stems, in ft}$

The high R-square value indicates that age, planting density, and average height of the dominant stems are good predictors of the scale parameter, and will aid in compiling a prediction model. A graph of the residuals derived from equation (26) is presented in Figure 11. A model fitted to the untransformed scale parameter, b , was originally sought, but a trend was present in the residuals. The ladder of transformations was used to systematically transform the parameter until the trend in the residuals was no longer present. The logarithm transformation stabilized the variance of the residuals and removed the trend, as can be seen in Figure 11.

The shape parameter results given in Table 8 for the ANOVA tests yield varying results based on the planting density. For the lower initial densities, 302-907 TPA, there is a significant difference in the mean shape parameter estimates for ages 5-11. For the higher initial densities, 1210-2722 TPA, the differences in the mean shape parameter

Figure 11: Residual graph for the predicted scale parameter

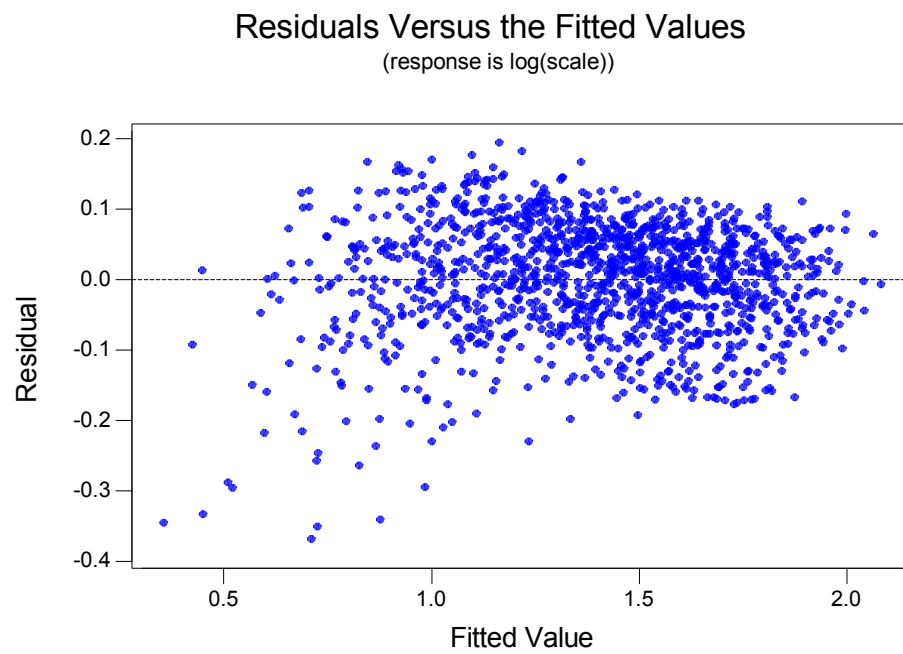


Table 8: Analysis of variance and correlation results for the shape parameter estimates by planting density for diameter at breast height

	302 TPA		453 TPA		605 TPA		680 TPA		907 TPA	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age 5	5.267	1.457	5.182	1.835	5.260	2.018	5.184	1.463	5.184	1.487
Age 6	6.386	1.568	6.243	1.784	6.222	1.950	6.052	1.428	5.951	1.474
Age 7	7.303	1.627	6.919	1.659	6.714	1.935	6.384	1.376	6.329	1.333
Age 8	7.608	1.623	7.186	1.646	6.954	2.003	6.536	1.278	6.382	1.335
Age 9	7.763	1.534	7.396	1.428	7.137	2.011	6.748	1.019	6.412	1.353
Age 10	8.069	1.512	7.511	1.274	7.139	1.979	6.687	1.093	6.276	1.310
Age 11	8.416	1.714	7.611	1.411	7.075	1.869	6.664	1.143	6.162	1.205
df _{err}	77		161		161		77		329	
F	5.70		7.27		2.96		2.30		4.87	
p-value	< 0.001		< 0.001		0.009		0.043		<0.001	
Pearson	0.522		0.420		0.271		0.329		0.187	
p-value	< 0.001		< 0.001		< 0.001		0.002		0.001	

	1210 TPA		1361 TPA		1815 TPA		2722 TPA	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age 5	5.041	1.570	4.923	1.381	4.848	1.232	4.326	1.602
Age 6	5.795	1.462	5.373	1.302	5.356	1.142	4.728	1.423
Age 7	6.176	1.393	5.606	1.040	5.546	1.096	4.701	1.368
Age 8	6.105	1.373	5.616	0.960	5.448	1.097	4.775	1.136
Age 9	6.163	1.529	5.451	0.887	5.367	1.060	4.806	0.953
Age 10	5.954	1.463	5.316	0.924	5.306	0.947	4.903	0.938
Age 11	5.755	1.309	5.261	0.858	5.256	0.844	4.869	0.711
df _{err}	77		161		161		77	
F	0.91		1.18		1.04		0.31	
p-value	0.494		0.318		0.401		0.931	
Pearson	0.122		0.050		0.063		0.128	
p-value	0.268		0.522		0.414		0.245	

Where TPA = Trees per acre

SD = Standard deviation

df_{err} = Error degrees of freedom

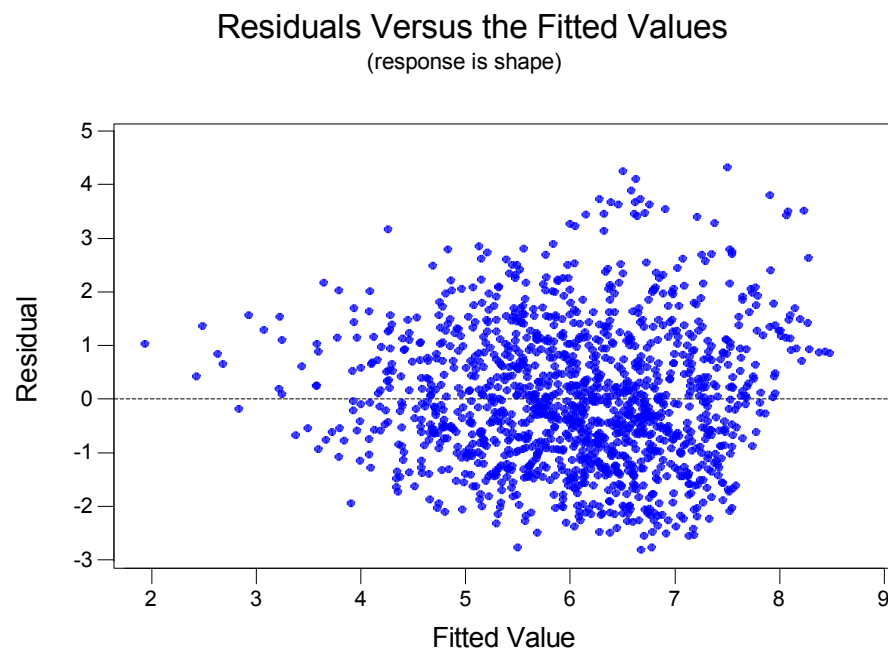
estimates are non-significant (all p-values > 0.31). Hence, for the ages considered in this study, age has a significant effect on the mean estimated shape parameter for lower planting densities, but not for those plots with higher planting densities. The variation in the estimated shape parameters is large, making a trend in the means hard to discern. The correlation between the shape parameter estimates and age is also given in Table 8. The Pearson correlation is derived from the individual plot values, and the results show a significant positive correlation (all p-values ≤ 0.002) between the shape parameter estimates and age for plots with lower planting densities (302-907 TPA). Those plots with higher planting densities do not exhibit evidence (all p-values > 0.24) of a significant correlation between age and the shape parameter estimates. Age may not aid much in modeling the shape parameter at higher planting densities, so inclusion of other factors was reviewed. The following linear regression does not show a strong trend between the estimated shape parameter, age, planting density, and average height of the dominant stems.

$$\hat{c} = 6.055 - 4.922\text{Log}(A) - 1.003\text{Log}(TPA) + 5.359\text{Log}(H) \quad (27)$$

$$R^2 = 0.388, MSE = 1.64$$

The low R-square value indicates that the shape parameter may require other model forms and components to increase the predictive ability. A graph of the residuals is presented in Figure 12. The residuals do not have a discernable trend, and the variance is relatively constant. Though relatively constant, the variance of the residuals is large and further efforts to reduce the dispersion of the estimated shape parameter were conducted, as discussed in section 5.5.

Figure 12: Residual graph for the predicted shape parameter



5.3. Planting Density Effect

Comparisons of the estimated Weibull parameters were made at each age across the range of planting densities considered in this investigation to further aid in illuminating trends in the scale and shape parameter estimates. For the juvenile loblolly pine data considered in this research, ANOVA tests for mean differences were computed between planting densities at a specific age. To aid in trend recognition, graphs of the data at specific ages by planting density are presented for the scale (Figure 13 (a-d)) and shape (Figure 14 (a-d)) parameter estimates. Visual inspection of the graphs reveals a general increasing trend for both the scale and shape parameter estimates with a decrease in planting density. For the scale parameter estimates, the steepness of the trend increases sharply with age and becomes nonlinear. The trend is less pronounced for the shape parameter estimates, but still shows an increase with age.

The results of the ANOVA tests on the scale parameter estimates (Table 9) indicate that there are significant mean differences based on the planting density at each age (all p-values < 0.003). Hence, planting density has a significant effect on the mean scale parameter value for ages 5 to 11. The Pearson correlation coefficients between the estimated scale parameter and the planting density are presented in Table 9 for each age. There is an increasingly negative correlation with an increase age (all p-values < 0.001), indicating a larger disparity between the more dense stands and the less dense stands with an increase in age. Biologically, this makes sense because one would expect trees in a very dense stand to have a smaller spread of diameters than those trees in a very sparse

Figure 13: Trends of the estimated scale parameter for diameter at breast height (D) over all observed planting densities (trees per acre) for ages (a) 5, (b) 7, (c) 9, and (d) 11

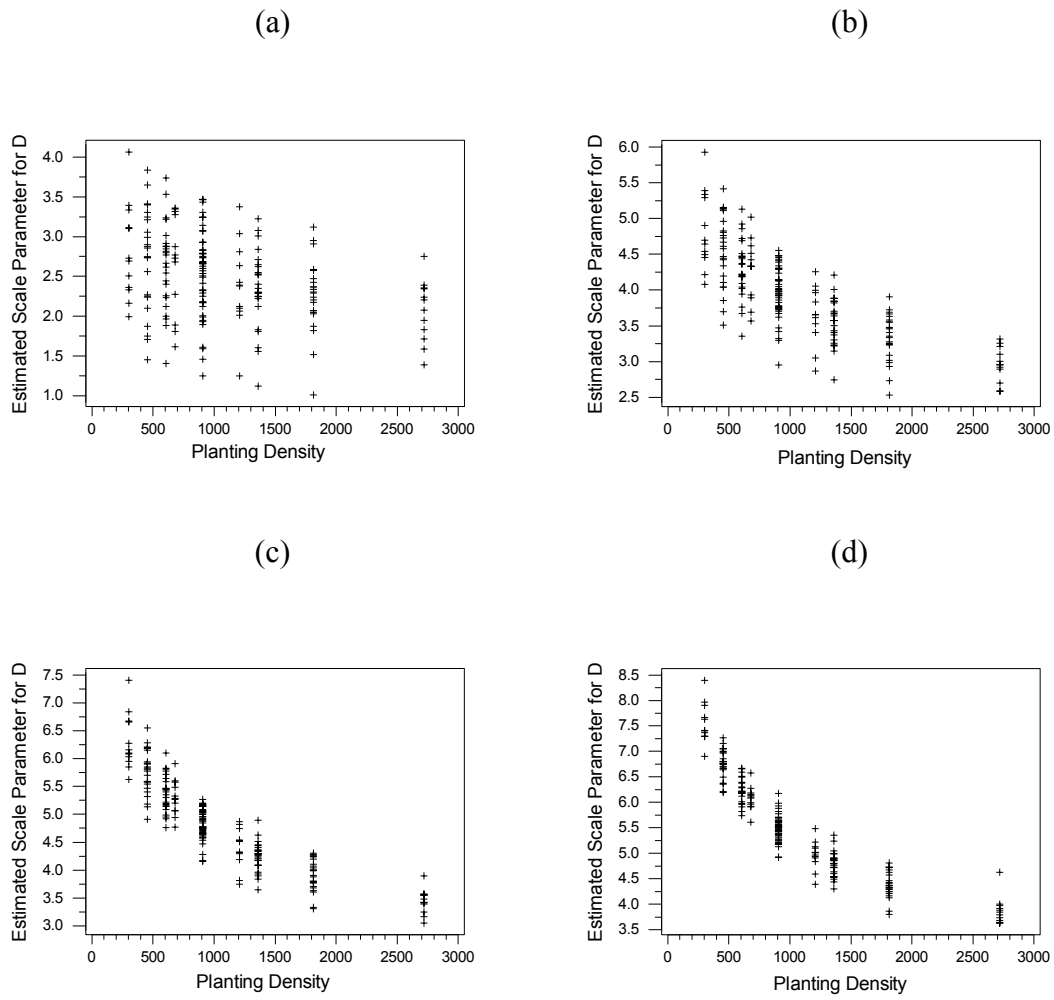


Figure 14: Trends of the estimated shape parameter for diameter at breast height (D) over all observed planting densities (trees per acre) for ages (a) 5, (b) 7, (c) 9, and (d) 11

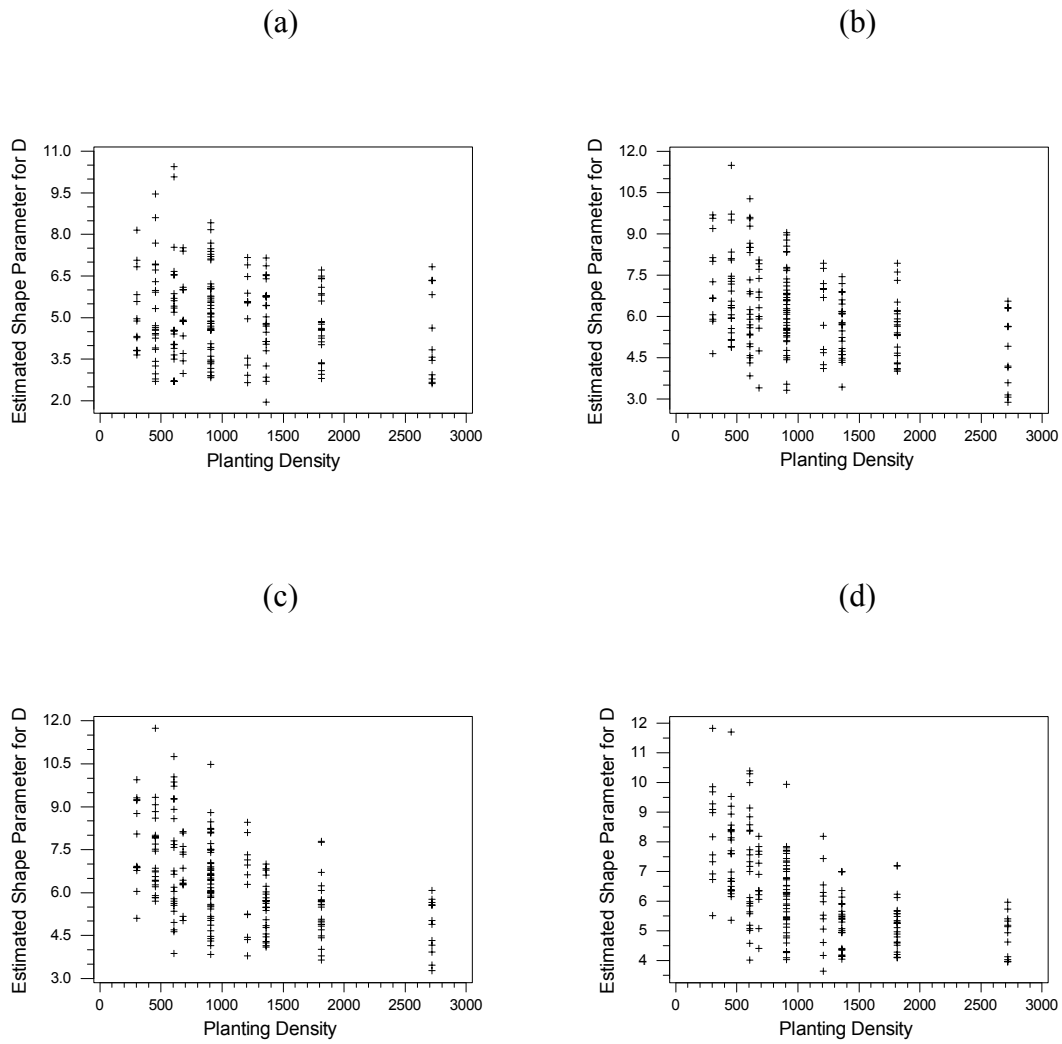


Table 9: Analysis of variance and correlation results for the scale parameter estimates by age for diameter at breast height

TPA	Age 5		Age 6		Age 7		Age 8		Age 9		Age 10		Age 11	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
302	2.816	0.602	3.819	0.599	4.833	0.550	5.575	0.576	6.305	0.499	7.007	0.464	7.543	0.396
453	2.716	0.643	3.636	0.634	4.543	0.499	5.154	0.477	5.751	0.401	6.326	0.329	6.754	0.278
605	2.633	0.542	3.471	0.505	4.305	0.407	4.847	0.395	5.372	0.333	5.866	0.277	6.226	0.248
680	2.662	0.631	3.504	0.607	4.284	0.434	4.781	0.417	5.293	0.316	5.760	0.271	6.077	0.233
907	2.546	0.527	3.281	0.471	3.974	0.346	4.403	0.312	4.827	0.264	5.219	0.245	5.504	0.247
1210	2.385	0.550	3.047	0.514	3.658	0.407	4.037	0.421	4.393	0.357	4.731	0.302	4.969	0.282
1361	2.340	0.498	2.972	0.454	3.564	0.324	3.904	0.303	4.241	0.269	4.554	0.256	4.785	0.249
1815	2.265	0.451	2.847	0.403	3.339	0.330	3.626	0.320	3.914	0.284	4.181	0.265	4.385	0.259
2722	2.070	0.388	2.558	0.327	2.957	0.244	3.204	0.225	3.445	0.222	3.688	0.255	3.867	0.276
df _{err}	183		183		183		183		183		183		183	
F	3.22		10.64		38.97		66.84		133.56		225.45		317.71	
p-value	0.002		< 0.001		< 0.001		< 0.001		< 0.001		< 0.001		< 0.001	
Pearson	-0.343		-0.543		-0.761		-0.818		-0.869		-0.888		-0.893	
p-value	< 0.001		< 0.001		< 0.001		< 0.001		< 0.001		< 0.001		< 0.001	

Where TPA = Trees per acre

SD = Standard deviation

df_{err} = Error degrees of freedom

stand. As the stems age and move into more intense competition with each other, the differences between the dispersion of the diameters in low- and high-density stands should increase.

A summary of ANOVA tests performed on the shape parameter estimates are also presented in Table 10. No significant differences in the mean shape parameters at age 5 were found across the varying planting densities. The shape parameter estimate is related to the skewness of the empirical distribution, implying that at age 5, there may be no difference between the skewness of the diameter distributions for varying planting densities. Further ANOVA tests reject the null hypothesis that there are no differences in the mean shape parameter by planting density for ages 6 to 11 (p-value < 0.05 at age 6, all other p-values < 0.001). Correlation analysis showed a significant negative correlation between the shape parameter and initial TPA for ages 5 to 11 (Table 10). The Pearson correlation becomes increasingly negative with age, indicating that as the planting density increases, the shape parameter estimate decreases. Hence, stands that are planted at lower densities have larger skewness parameters, resulting in diameter distributions that are increasingly skewed to the left.

5.4. Relative Spacing Effect

Relative spacing is a unitless statistic that is defined as the average distance between trees divided by the average height of the dominant canopy (Avery and Burkhart 2002). Relative spacing is utilized to enable comparisons to be made that take the competitive status of an individual plot into account, as measured by height and density. The use of a non-chronological measure was employed as a potentially more meaningful

Table 10: Analysis of variance and correlation results for the shape parameter estimates by age for diameter at breast height

TPA	Age 5		Age 6		Age 7		Age 8		Age 9		Age 10		Age 11	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
302	5.267	1.457	6.386	1.568	7.303	1.627	7.608	1.623	7.763	1.534	8.069	1.512	8.416	1.714
453	5.182	1.835	6.243	1.784	6.919	1.659	7.186	1.646	7.396	1.428	7.511	1.274	7.611	1.411
605	5.260	2.018	6.222	1.950	6.714	1.935	6.954	2.003	7.137	2.011	7.139	1.979	7.075	1.869
680	5.184	1.463	6.052	1.428	6.384	1.376	6.536	1.278	6.748	1.019	6.687	1.093	6.664	1.143
907	5.184	1.487	5.951	1.474	6.329	1.333	6.382	1.335	6.412	1.353	6.276	1.310	6.162	1.205
1210	5.041	1.570	5.795	1.462	6.176	1.393	6.105	1.373	6.163	1.529	5.954	1.463	5.755	1.309
1361	4.923	1.381	5.373	1.302	5.606	1.040	5.616	0.960	5.451	0.887	5.316	0.924	5.261	0.858
1815	4.848	1.232	5.356	1.142	5.546	1.096	5.448	1.097	5.367	1.060	5.306	0.947	5.256	0.844
2722	4.326	1.602	4.728	1.423	4.701	1.368	4.775	1.136	4.806	0.953	4.903	0.938	4.869	0.711
df _{err}	183		183		183		183		183		183		183	
F	0.54		2.02		4.86		6.67		9.25		11.67		14.71	
p-value	0.826		0.046		< 0.001		< 0.001		< 0.001		< 0.001		< 0.001	
Pearson	-0.145		-0.278		-0.405		-0.455		-0.504		-0.523		-0.547	
p-value	0.044		< 0.001		< 0.001		< 0.001		< 0.001		< 0.001		< 0.001	

Where TPA = Trees per acre

SD = Standard deviation

df_{err} = Error degrees of freedom

variable in explaining the variation in the Weibull parameter estimates. Plots with a very high planting density are going to reach competitive status much sooner than plots with a low planting density. To facilitate comparisons, the Weibull parameter estimates were compared across the levels of relative spacing present over all plots for ages 5 to 11. The equation to derive relative spacing is given by:

$$RS_{jk} = \frac{\sqrt{\frac{43560}{N_{jk}}}}{\bar{H}_{jk}} , \quad (28)$$

where RS_{jk} = relative spacing at age j , for plot k , $j = 5, \dots, 11$, $k = 1, \dots, 192$

N_{jk} = live trees per acre at age j , and plot k

$\bar{H}_{jk}$ = average height of the dominant stems in ft at age j , and plot k

The height of the dominant stems was determined as the average height of all undamaged stems with D greater than that of the quadratic mean diameter. The quadratic mean diameter is defined as the diameter of the tree with mean basal area, where basal area is defined as the cross sectional area at D .

$$\bar{D}_{qjk} = \sqrt{\frac{\bar{BA}_{jk}}{0.005454}} , \quad (29)$$

where $\bar{D}_{qjk}$ = quadratic mean diameter, in., at age j , for plot k

$\bar{BA}_{jk}$ = mean basal area, in sq ft, at age j , for plot k

0.005454 = conversion factor to obtain diameter in in. from area in sq ft

$$\bar{BA}_{jk} = \frac{0.005454 \sum_{i=1}^n D_{ijk}^2}{n} \quad (30)$$

$$\bar{H}_{jk} = \frac{\left(\sum_{l=1}^m ht_l \right)}{m} , \quad (31)$$

where ht_l = total stem height, in feet, with $D_{ijk} \geq \bar{D}_{qjk}$, $l = 1 \dots m$

To compare plots with ‘similar’ relative spacing, a range of this statistic that would be considered similar was necessary. For this study a criterion of $RS_{jk} \pm 0.01$ was set to identify plots at equivalent stages in stand competition. Plots within the specified range of relative spacing were not considered to be different. A graph of the mean relative spacing for each planting density for ages 5 to 11 is presented in Figure 15, illustrating the general decreasing nature of relative spacing with an increase in stand age. The decrease in relative spacing is due to the relationship between height growth of the canopy and mortality occurrence within the stand. The decrease in relative spacing will eventually reach a species-specific lower bound. The lower bound has been estimated for loblolly pine to occur at a relative spacing of 0.15 (Lemin and Burkhart 1983). Due to mortality as the stand continues to age, the relative spacing will start to rise slightly above the lower bound.

Plot comparisons were made with respect to all plots that have a relative spacing that meets the criterion for similarity. The criterion enabled plots with different planting densities to be compared at different ages on a basis of relative spacing. Bins were used to group the plots into similar categories. The relative spacing bins were created using the above stated criterion, and ranged from a minimum of 0.11 to a maximum of 1.09. A graph of the sample size within each bin is given in Figure 16, revealing that most of the plot relative spacing values fall between 0.2 and 0.6. The number of observations decreases sharply as the bins go towards the minimum and maximum relative spacing values. The trend in sample size may inhibit inferences made from bins that contain very few plot observations.

Figure 15: Mean relative spacing trend over ages 5 to 11 for each planting density (trees per acre)

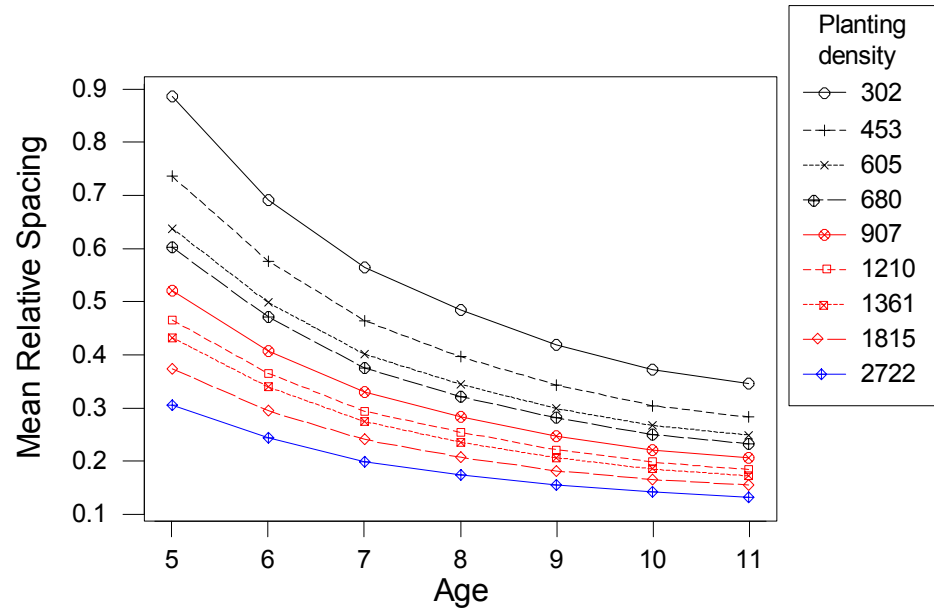
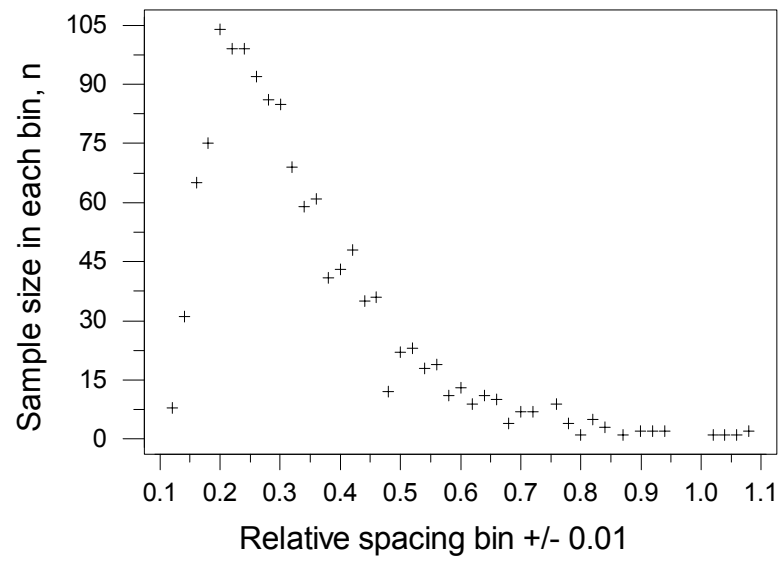


Figure 16: Sample size for each relative spacing bin



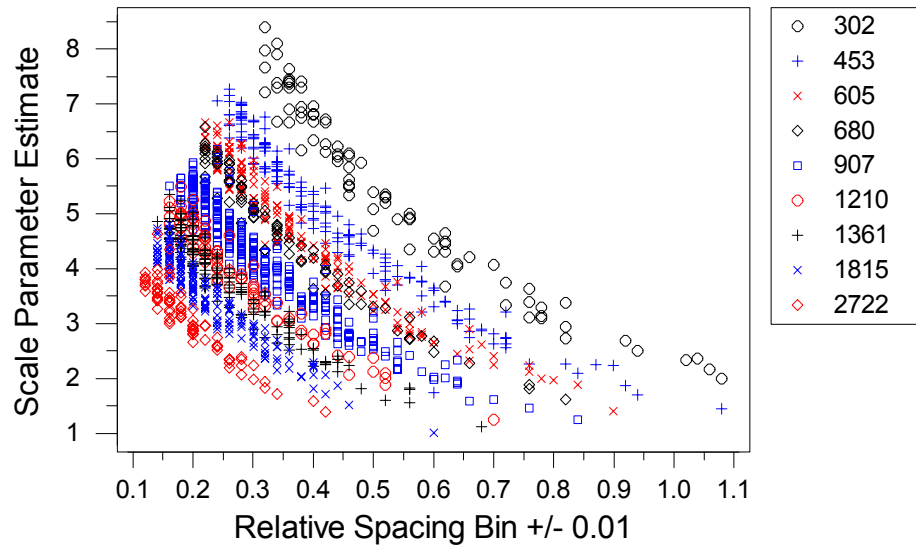
The scale and shape parameter estimates from the two-parameter Weibull distribution are presented for all plots by bin group in Figure 17 (a) and (b), respectively. These graphs do not show any discernable trends between relative spacing and the Weibull parameter estimates beyond that which is explained by planting density and age. The scale parameter estimates show a trend with planting density as the relative spacing bins increase in value, but the variance of the estimates within each bin is still large. The shape parameter estimates have a large dispersion over the range of the relative spacing bins. Planting density was also evaluated across the relative spacing bins for both the scale and shape parameter estimates (Figure 17 (a-b)). Though trends with planting density were present for the scale parameter, no strong trends were apparent for the shape parameter estimates. The variation in the parameter estimates is too great for relative spacing to be adequately used as a predictor variable of Weibull parameters. Given the inability of relative spacing to capture any trends in the parameter estimates, it was not considered further.

5.5. Weibull Reparameterization

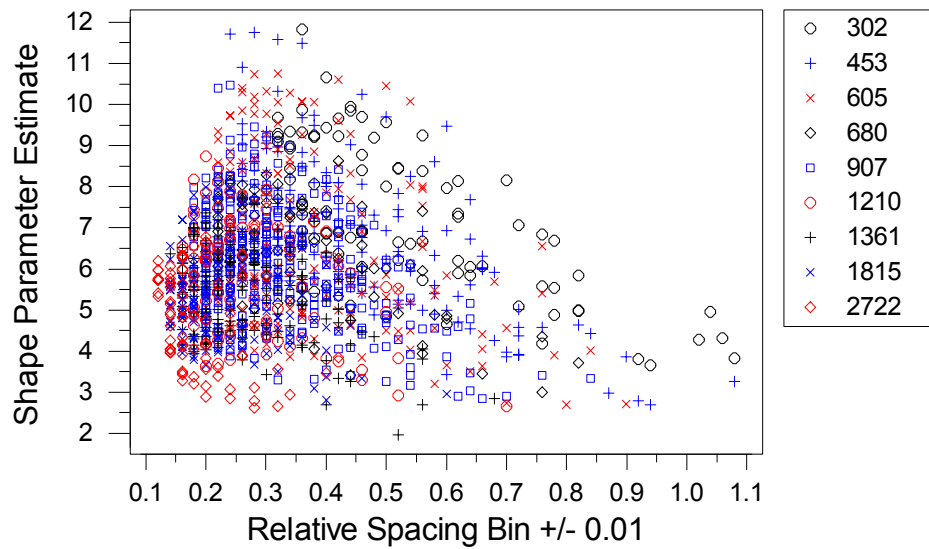
To aid in reducing the bias and variance of the estimated shape parameter, a reparameterization of the Weibull distribution was conducted. The reparameterized Weibull should enable better estimates of juvenile diameter distributions to be predicted from the modeled parameters. The dispersion of the shape parameter estimates can be viewed in Figure 10 and Figure 14. The shape parameter exhibits a large amount of dispersion with respect to age and planting density, and no clear trend is present for modeling. The substantial dispersion in the shape parameter estimates may be due to the

Figure 17: Parameter estimates for all plots in each relative spacing bin, with planting density subgroups (trees per acre), for (a) the scale parameter and (b) the shape parameter

(a)



(b)



amount of parameter-effects nonlinearity within the pdf function selected (Ratkowsky 1983). Parameter-effects nonlinearity is a function of how the parameters are represented in the density function, and may be reduced by reparameterizing the model. The reparameterization considered for the shape parameter is as follows:

$$\alpha = \log(c)$$

with the following pdf resulting:

$$f(x^*) = \frac{e^\alpha}{b} \left(\frac{x^*}{b} \right)^{e^\alpha - 1} \exp \left[- \left(\frac{x^*}{b} \right)^{e^\alpha} \right], \quad (32)$$

where $e^\alpha = e^{\log(c)} = c$

Although the parameters are represented differently, the reparameterization will not change the actual fit to the data. The log-likelihood function was altered to reflect this change in the macro written to compute the MLE of the Weibull function. The modified log-likelihood function should aid in reducing the bias and variance of the parameter estimates. Ratkowsky (1983) states that the behavior of a model is determined in conjunction with a data set, and may vary widely from model to model. Because of the similarities in the data sets (repeated annual measurements of diameters) and wide dispersion of the estimated shape parameters, a review of the parameter-effects nonlinearity was conducted.

Maximum likelihood parameter estimates were calculated for the reparameterized Weibull distribution given by equation (32). The scale parameter estimates were exactly the same as under the original, unmodified, distribution. The shape parameter estimates, $\hat{\alpha}$, are comparable to the shape parameter estimates given originally by c if they are transformed back using $e^{\hat{\alpha}}$. A model fit to the alpha values is given by:

$$\hat{\alpha} = 1.593 - 0.8693 \text{Log}(A) - 0.1580 \text{Log}(TPA) + 0.9594 \text{Log}(H) \quad (33)$$

$$R^2 = 0.407, MSE = 0.045$$

The results of this fit are an improvement over the fit given in equation (27), but the gain from the reparameterization seems to be minimal, with an R-square increase of 0.019. The residuals for this model are presented in Figure 18 and show a similar spread as those given in Figure 12 for the unaltered shape parameter estimates. Another reparameterization that may aid in reducing the parameter-effects nonlinearity was considered by moving down the ladder of transformations. The next transformation considered for the shape parameter is given by:

$$\delta = -c^{-0.5}$$

with the following pdf:

$$f(x^{**}) = \frac{\delta^{-2}}{b} \left(\frac{x^{**}}{b} \right)^{\delta^{-2}-1} \exp \left[- \left(\frac{x^{**}}{b} \right)^{\delta^{-2}} \right], \quad (34)$$

$$\text{where } \delta^{-2} = (-c^{-0.5})^{-2} = c$$

The resulting maximum likelihood parameter estimates for the scale parameter are the same as those derived without any reparameterization. Modeling of the new shape parameter, $\hat{\delta}$, was conducted with the following fitted model resulting:

$$\hat{\delta} = 0.4743 + 0.1877 \text{Log}(A) + 0.03209 \text{Log}(TPA) - 0.2089 \text{Log}(H) \quad (35)$$

$$R^2 = 0.407, MSE = 0.0026$$

Figure 19 presents a graph of the residuals from the fitted model. The reparameterization of the shape parameter using the form given by δ does not improve

Figure 18: Residual graph for reparameterized (alpha) shape parameter estimates

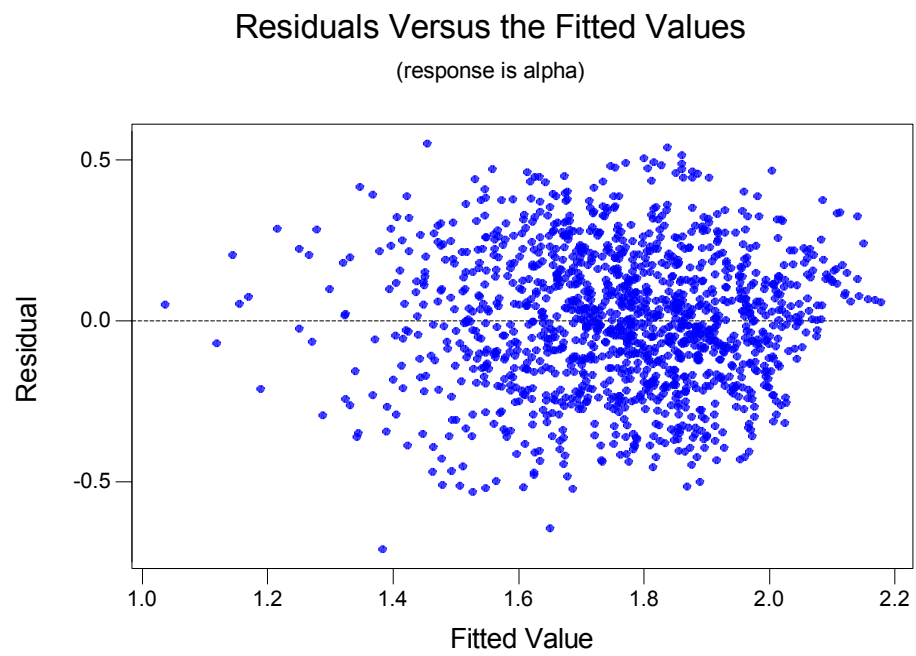
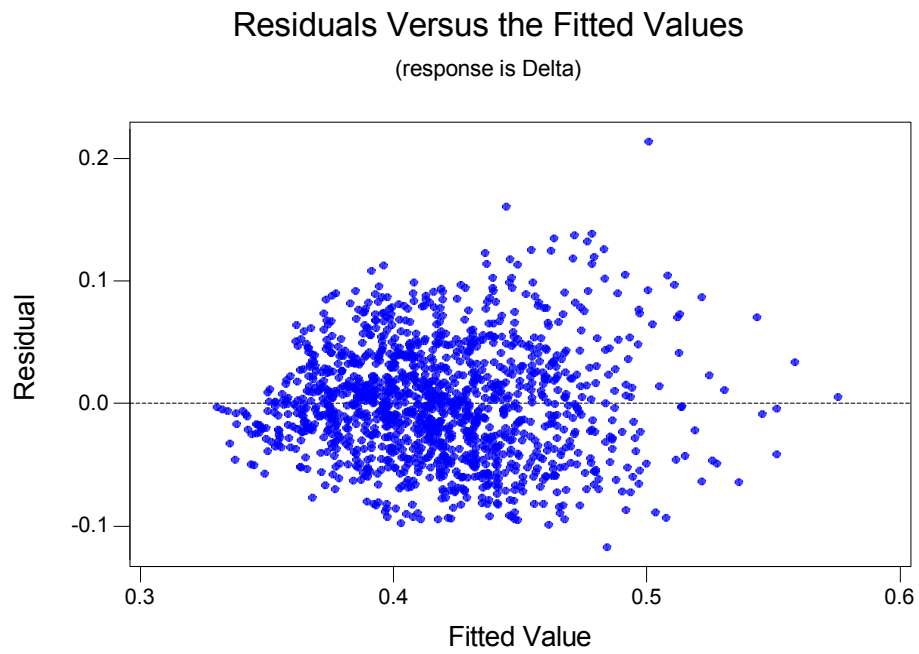


Figure 19: Residual graph for reparameterized (delta) shape parameter estimates



the fit of the estimated shape parameter when compared to the earlier reparameterization given by α . Overall, the considered reparameterizations of the shape parameter in the Weibull distribution did not reduce the variation in the parameter estimates, implying that the dispersion of the shape parameter estimates is not largely influenced by any parameter-effects nonlinearity in the model. The wide dispersion in the shape parameter estimates for a specific planting density and age seems to be natural variation inherent in the system.

5.6. Summary of Comparisons

The previous sections have reviewed in depth the trends discernable in the Weibull parameter estimates. A rectangularity effect was evaluated on all plots that had a planting density of 907 TPA and plot rectangularities of 1:3 and 1:1.33. No significant differences were found in either the scale or shape parameter estimates between plots with different rectangularities over the juvenile age range considered in this study. Further, these plots did not have significantly different coefficients of skewness or kurtosis as determined from the empirical diameter data. Based on the lack of significant differences in these statistics, plots with the same planting density of 907 TPA and differing plot rectangularities (1:3 and 1:1.33) were combined. The results in this research support those determined by Sharma et al. (in review), where the diameter distributions of loblolly pine stems planted at different rectangularities were not significantly different. The lack of significant differences implies that different plot rectangularities do not result in different diameter distributions.

The mean scale parameter estimates within each planting density have significant temporal differences over ages 5 to 11. Significant positive Pearson's correlations were

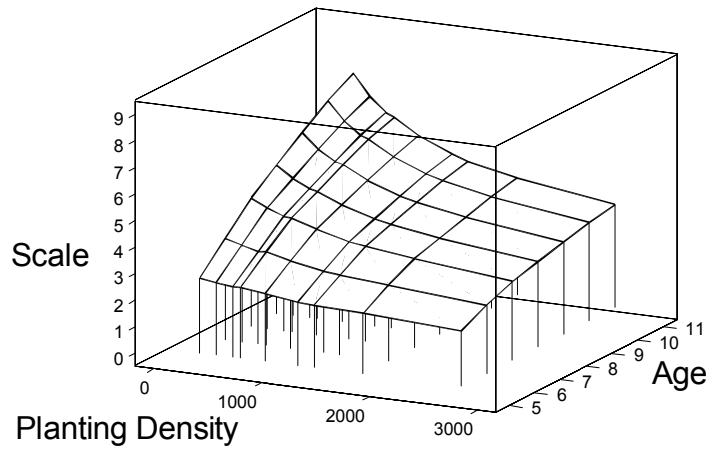
found between the scale parameter and age, across all planting densities, showing that as age increases there is an increase in the estimated scale parameter. At each age a significant planting density effect was present for the mean scale parameter estimate. The planting density effect had significantly negative Pearson's correlations between the scale parameter and planting density at each age, which signifies that an increase in planting density is associated with a decrease in the scale parameter. These effects can be viewed in Figure 20 (a), a three-dimensional representation of the mean values (Table 7) for the Weibull scale parameter estimates.

Plot relative spacing did not aid in elucidating trends in the estimated scale parameters (Figure 17 (a)). The prediction equation (26) for the scale parameter does an excellent job ($R^2 = 0.940$) of fitting the parameter estimates as a function of age, planting density, and mean height of the dominant stems. The prediction equation presented is recommended for use in estimating the scale parameter of the two-parameter Weibull distribution to obtain the diameter distribution of juvenile loblolly pine.

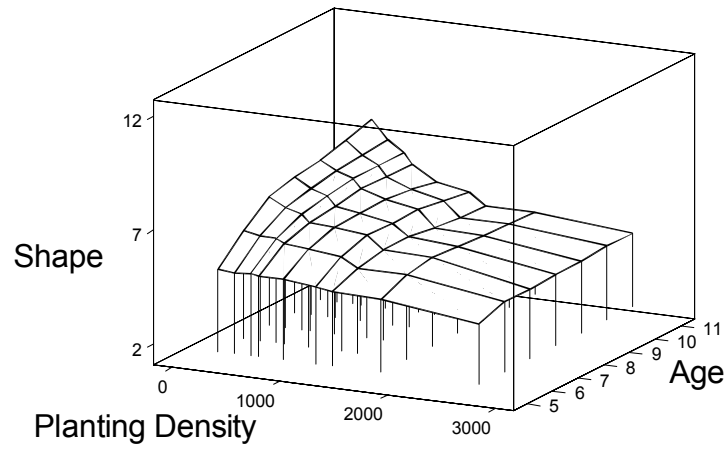
The shape parameter estimates for plots with lower planting densities (302-907 TPA) had significant differences in the means for ages 5 to 11. These plots also had a significantly positive Pearson's correlation coefficient, indicating that an increase in age is associated with an increase in the shape parameter estimate. Those plots with higher planting densities (1210-2722 TPA) did not have significantly different mean shape parameter estimates over ages 5 to 11, nor were the Pearson's correlation coefficients significant ($p\text{-values} > 0.24$). There were no significant differences in the mean estimated shape parameters across planting densities at age 5. Ages 6 through 11 show a

Figure 20: Mean Weibull parameter estimates for (a) the scale parameter, and (b) the shape parameter, across all ages and planting densities (trees per acre) considered in this research for diameter at breast height

(a)



(b)



significant planting density effect for the mean shape parameter estimate. The Pearson's correlation was significant and negative over ages 6 to 11, signifying an increase in planting density is associated with a decrease in the estimated shape parameter. Figure 20 (b) is a three-dimensional representation of the mean values (Table 8) for the shape parameter estimates and illustrates these trends.

Relative spacing was computed for each plot at each age to aid in discerning trends in the shape parameter estimates. The dispersion of the estimated shape parameters was large and inhibited trend recognition (Figure 17 (b)). A reparameterization of the Weibull distribution was conducted to reduce the influence of any parameter-effects nonlinearity on the shape parameter estimates. Due to a large amount of natural variation in the shape parameter estimates across varying planting densities and ages, the reparameterization did not greatly reduce the dispersion of the shape parameter estimates or improve the fit of the original prediction equation (27). Equation (33) is recommended ($R^2 = 0.407$) to estimate the shape parameter of juvenile loblolly pine diameter distributions and can be easily transformed back to the original parameter form of c .

6. SPATIAL DEPENDENCE

Quantifying spatial dependency should aid in describing stand dynamics and competition processes that are influenced by the spatial distribution of the stems. The presence of a high inter-tree competition-based spatial dependency in a stand may, over time, dominate any spatial dependencies present due to microsite differences. Microsite differences are those site qualities that vary across a relatively small range, including factors such as soil type, soil moisture, available sunlight, nutrient availability, and topography. If trees are assumed to not have any influence on each other at very young ages, then the microsite effects (occurring at scales that can be differentiated within the stand) will account for the initial spatial effects within the stand, and one would expect this to be a positive correlation. As stems grow and start influencing each other, the competition effects may counteract any microsite effects. As competition increases, there may be a change in the spatial relationship between the stems towards an increasingly negative correlation due to heightened competition for resources.

To aid in the characterization of diameter distributions, a methodology to account for the inter-tree spatial correlation effect was sought. The assumption that the diameter observations are independent of each other implies that there are no interactions between trees manifesting themselves in the diameter observations. To test the independence assumption, a means of quantifying the spatial location of each stem is needed. For plot k , the spatial locations of specific trees are denoted $\{s_1, \dots, s_n\}$ and the diameter data collected at these locations are denoted $Z = \{Z(s_1), \dots, Z(s_n)\}'$ where $i=1, \dots, n$ is the associated tree number. The plots studied herein were randomized within each

replication (see Amateis et al. 1988). Randomization controls for unwanted bias and neutralizes the effect of spatial correlation at the randomized scale (Cressie 1993). The neutralization of spatial correlation is only applicable to the randomized scale, not to spatial scales larger or smaller than that of the randomization scale. Measures of the inter-tree spatial correlation were employed to aid in elucidating any spatial trends in the data. These measures include the empirical semivariogram, Moran's I coefficient, and Geary's C coefficient. Further, a conditional autoregressive (CAR) model was utilized to evaluate the amount of spatial dependence within each plot, and to track this relationship over time.

6.1. Spatial Process Definition

The collection of spatial sites, T , is termed the *lattice* and is contained in two-dimensional real space $\{s_i \in T \subset \mathfrak{R}^2\}$. The stochastic spatial process is $\{Z(s) : s \in T\}$, with the actual realization of the spatial process termed $\mathbf{Z}(s)$ for the vector of all observations. The collection of spatial sites, T , is a discrete set of fixed points for lattices. The discrete set of points does not allow for any possibility of points occurring between any two lattice points within the domain T . The use of lattices for analyzing the data considered within this study is justified by the interest in the spatial dependency between stems in a plot. Further, these stems were systematically planted and surrounding competition that was not a part of the study design was removed or suppressed. Other methods of spatial data analysis are available (i.e. geostatistical and point pattern analysis) but deemed inappropriate for this study. Geostatistical data have fixed locations throughout a continuous domain, such that samples could be placed between existing locations. Point pattern data are continuous and random in the domain,

where the focus is on the actual location of the points and their spatial distribution. The lattices for the 10 different row-by-column spacings (see Table 1) are rectangular and only have meaning when the stem map is available for each plot. Figure 21 shows the general lattice design for each plot with the initial tree locations numbered 1 to 49, and can be visualized as a lattice representing any of the rectangular spacings. Thinking of each stem in a plot as part of a lattice will enable the utilization of a conditional autoregressive spatial model. From Figure 21 the lattice for each plot is termed as follows:

$$T = \{s_{u,v} : u = 1, \dots, 7; v = 1, \dots, 7\}. \quad (36)$$

The spatial neighborhood is the criterion that defines the spatial connectivity of trees. The criterion used herein is a distance-weighted measure for all stems within the plot such that for a specific subject tree, all stems are considered as neighbors with an influence that is inversely related to the distance from the subject tree. The use of an inverse distance-weighted measure has been successfully applied to many tree competition indices (see for example, Daniels 1976, Daniels et al. 1986, and Biging and Dobbertin 1992). The neighborhood weight matrix, denoted A , is a symmetric matrix with dimensions $A_{49 \times 49}$ for a plot with no mortality. The A matrix contains all of the neighborhood information for every stem with its respective neighbor locations and weights. The dimensions of the neighborhood weight matrix will change over time as mortality occurs, but will always remain square.

A stem that dies will still have an effect on surrounding trees due to the release of

Figure 21: Stem map (lattice) for each plot, with tree numbers 1-49

$S_{u,v}$	1	2	3	4	5	6	7
1	1	2	3	4	5	6	7
2	14	13	12	11	10	9	8
3	15	16	17	18	19	20	21
4	28	27	26	25	24	23	22
5	29	30	31	32	33	34	35
6	42	41	40	39	38	37	36
7	43	44	45	46	47	48	49

Where $s_{u,v}$ = The tree located in row u , column v , $u,v = 1,\dots,7$

s_i = Individual site location, $i = 1,\dots,49$

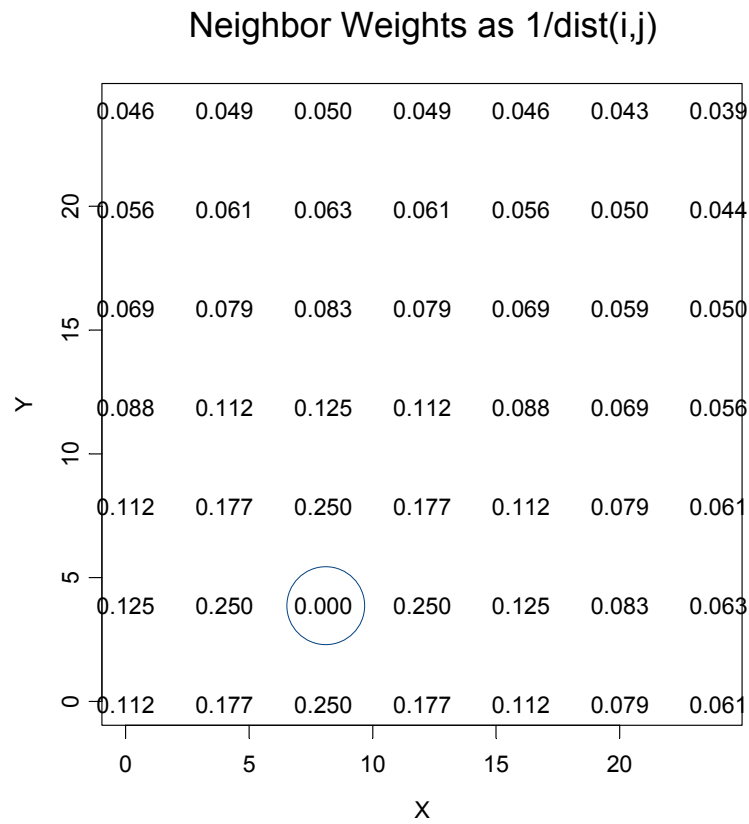
sunlight, nutrients, etc. The effect of a dead stem will be present for a certain amount of time and will decrease as time progresses and the surrounding stems fully utilize the additional resources. For this study all mortality was represented for one remeasurement period beyond the last live remeasurement period. Since the exact time of stem mortality and the values of stem characteristics at the time mortality occurs were not known, the last measured stem characteristics were used for the next remeasurement period; after the mortality was represented for one remeasurement period the spatial location was dropped from the lattice for that plot. Specific entries in the weight matrix are defined as follows for a tree at location $s_{u,v}$:

$$A_{i,j} = \begin{cases} \frac{1}{dist_{ij}}, & \text{neighbor} \\ 0, & \text{non-neighbor} \\ 0, & i = j \end{cases} \quad (37)$$

where $dist_{ij}$ = the Euclidean distance between the subject tree i and neighbor tree j

For instance, a specific tree, $s_{6,3}$, on a 4-by-4 ft spacing would have the distance-dependent weight matrix given by Figure 22. The rapid decrease in weights can be viewed as the distance to neighboring trees increases. Note that, in this example, it is assumed that no mortality is present. If a stem has been dead for more than one year, the mortality would be represented by an empty space. Although the distance-dependent weights would not change once established, the matrix has to be recalculated for each plot at each age to account for mortality. The symmetry of the weight matrix is apparent, in that the weight between subject tree i and neighbor j ($A_{i,j}$) is equal to the weight between subject tree j and neighbor i ($A_{j,i}$).

Figure 22: The distance-dependent weights for tree $s_{6,3}$ (circled) on a 4-by-4 ft spacing, where the weights are calculated as the inverse of the distance between the subject tree i and neighboring tree j



6.2. Edge Effect

Each replication within the spacing study follows a non-systematic design where the spacing varies in two directions and the number of initial stems per plot is held constant. The boundaries selected for the plots within the spacing study resulted from the randomized experimental design and were not based on any biological basis that would indicate a natural boundary. Hence, trees near the boundaries will not have all of their neighbors accounted for, resulting in an *edge effect*. Cressie (1993) notes several methods that address edge effects. A torus, which wraps the two-dimensional space into a donut-shaped surface, could be assumed and would remove all boundaries within a given plot. Application of a torus would be inappropriate for the diameter data considered here due to the spatial correlation between stems within a stand decreasing as the distance between the stems increases. Another method to address any edge effects would be to reduce the data set so that only locations with all of their neighbors present would be utilized; this could be used if, for example, a ‘queens’ neighborhood definition were employed (all surrounding stems, including those on a diagonal, are neighbors). Use of this rule would reduce the data set by $\left(\frac{2(u+v-2)}{uv} 100\% \right) = 49\%$ for the 7-by-7 lattice considered herein. A third method to address this concern would be to assume that any unobserved boundary value is equal to the mean of the process. While this would minimize the effect of the unobserved edge observations, it would not aid in the explanation of any spatial trends.

To minimize any edge effects, the previously defined weight matrix using all the stems on the plot as neighbors for each subject tree was used. Though this neighborhood

definition does not account for any possible influence from stems outside of the plot, the distance-weighting scheme should aid in describing any spatial trends in the data. Hence, no allowance for edge effect was made in this investigation.

6.3. Empirical Semivariogram

Data that have known coordinate locations enable queries into the relationship between the data and the distance of separation between pairs of observations. Matheron (1963) defines the *semivariogram* as one-half of the average squared difference between points separated by a distance of h . The variable h is defined as the Euclidean distance $\|h\|$ if it is calculated as a distance with magnitude only. If the variable h includes direction, then h will be a vector with both distance and direction. For the data considered in this investigation, the covariance between any two points should only be a function of the distance between them and not depend on direction, implying that the spatial process is isotropic (Kaluzny et al. 1998) and that all pairs of observations separated by a distance of $\|h\|$ can be compared. The isotropy assumption implies that the row orientation within a plot is negligible (i.e. aspect and slope differences are non-significant). The empirical semivariogram is defined as follows:

$$\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{N(h)} (Z(s_i) - Z(s_j))^2 \quad (38)$$

where $N(h)$ = the set of all pairwise locations that are separated by the Euclidean distance $\|h\|$

$|N(h)|$ = the number of distinct pairs in the set $N(h)$

$Z(s_i), Z(s_j)$ = data values at spatial locations s_i and s_j

The empirical semivariogram has three distinct parameters that are commonly used to define its properties. The *sill* is the $\lim_{h \rightarrow \infty} \gamma(h)$, which represents the upper asymptote of the semivariogram and is equal to the variance of the random field under second-order stationarity. The *range* is the distance at which the data are no longer autocorrelated and the sill is achieved. The range may not exist if the semivariogram does not reach an upper asymptote. The practical range is the point at which the semivariogram achieves 95% of the sill, as the lag distances in the data may not extend out to the range. The *nugget effect* is a representation of both the micro-scale variation that is not picked up by the shortest lag distance and any measurement error that is inherent in the system. The nugget effect is defined when $\|h\| \rightarrow 0$, that is, when the distance between two spatial sites goes to zero and there is still a positive covariance between the observations. For further discussion and examples, see Haining (1990), Cressie (1993), and Schabenberger and Pierce (2002).

Empirical semivariograms were calculated for each plot from ages 2 to 5 for ground-line diameter and from ages 5 to 11 for breast height diameters. Due to the distribution of spatial locations within each plot, the pairwise distances were grouped into lag classes (distance bins) that are created from a lag increment for the distance h and a tolerance level for the lag increment. The following two practical rules were given by Journel and Huijbregts (1978) to aid in choosing the lag increment and the number of lags: 1) the maximum distance considered in computing the semivariogram should be less than one-half the maximum distance between two spatial locations within the plot, and 2) the minimum number of pairs in any lag class should be greater than 30. These

rules were heeded in the creation of the empirical semivariograms, although in a few cases the number of pairs in a specific lag class was slightly less than 30.

The empirical semivariograms were computed using two different representations of the data. First, the empirical semivariograms were computed using the raw diameter data without removing any trends. The data would then contain both the large-scale variation of any mean trends in the data and the small-scale variation of measurement error and spatial autocorrelation with other stems, resulting in the $\hat{\gamma}(h)$ estimating the covariance of the differences in the observed diameters for each lag class. Second, the semivariograms were computed utilizing the residuals (observed minus predicted values) from a fitted model. A simple linear model was fit to the diameter data as follows:

$$Z = \beta_0 + \beta_1 X + \varepsilon \quad (39)$$

where Z = stem diameter

X = the covariate, total stem height

β_0, β_1 = coefficients to be estimated

ε = the error term

The residuals of this simple linear model were then used to calculate the empirical semivariogram as a function of the small-scale variation. Though no distributional assumption is required to calculate the semivariograms, the distribution of ε is assumed to be Gaussian with mean zero. To evaluate this assumption qq plots (quantile-quantile) of the residuals versus the standard Normal distribution were calculated for each set of diameter data fit. Residual plots and histograms of the residuals were also examined and did not show strong deviations from normality. Because there was no change in the number of observations or spatial locations, the lag classes and number of paired

observations within each class were the same for both methods used to calculate the empirical semivariogram.

Detrending of the mean and covariate resulted in a general reduction of $\hat{\gamma}(h)$ from the empirical semivariograms estimated using the raw data. The detrended data should yield a more accurate representation of any small-scale spatial trends due to the model accounting for large-scale variation in the data. Graphical comparisons of the empirical semivariograms are presented in Appendix II for an example stand from each planting density over ages 2 to 5 for ground-line diameters and ages 5 to 11 for breast height diameters. The example stands appearing in Appendix II are the same example stands utilized in Appendix I. The distances employed in graphing the empirical semivariograms are the averages of all pairwise distances (in ft) within a specific lag class.

The empirical semivariograms presented in Appendix II cover the entire range of planting densities and ages. There are no strong trends visible in the $\hat{\gamma}(h)$ estimates of covariance between points with a distance of separation equal to h . There are some cases where there seems to be either a positive or negative trend when comparing only the first and second lag points. When these cases are viewed with respect to the entire semivariogram, the trend is not readily distinguishable from a white noise process. The inability to divulge spatial autocorrelations is troublesome because if microsite effects occur over scales that are discernable across a plot, then a positive autocorrelation is expected to occur at very young ages; the spatial relationship is then expected to tend towards a negative autocorrelation as the stand ages and competition sets in. The lack of visible trends in the semivariograms may indicate that there is not a large amount of

spatial autocorrelation among neighboring juvenile loblolly pine stems for the planting densities considered in this study.

As an example of the empirical semivariogram analysis, stand number 20 was chosen from the sample plots given in Appendices I and II. The planting density of this plot is 907 TPA with a planting rectangularity of 4-by-12 ft (1:3); results are given for d at age 2 and D at age 11 (Table 11). The first manner in which the empirical semivariogram was calculated uses the raw diameter data and is denoted $\hat{\gamma}(h)_{raw}$, the second calculation utilizes the residuals from the linear model fitted to the data by equation (39) and is denoted by $\hat{\gamma}(h)_{resid}$. For this example there was no mortality present by age 11, so the empirical semivariograms employed the same distance bins and have the same number of observations in each bin.

The gamma values in Table 11 clearly show that by accounting for the mean trend in the data and that of the covariate total stem height, the covariance between observations over all lag distances is reduced. The reduction in the covariance justifies the use of the linear model to aid in describing any small-scale spatial trends inherent in the data. At age 11, the derived gamma values were larger than those derived at age 2 over all distance bins, indicating, as expected, that the covariance within diameters at a specific distance increase with age. Figure 23 (a-b) provides graphical representations of the derived empirical semivariogram values at each bin.

The graphs presented in Figure 23 (a-b) display the semivariance of the data by the distance of separation, enabling characterization of a spatial trend as a function of the distance between two stems to be identified. The direction to the neighboring tree from

Table 11: Empirical semivariograms for stand number 20, ground-line diameter (d) at age 2, and diameter at breast height (D) at age 11

Distance bin, ft	d at age 2		D at age 11		$ N(h) $
	$\hat{\gamma}(h)_{raw}$	$\hat{\gamma}(h)_{resid}$	$\hat{\gamma}(h)_{raw}$	$\hat{\gamma}(h)_{resid}$	
4.000	0.03464	0.02544	0.6538	0.4048	42
8.000	0.03229	0.02385	0.5431	0.3265	35
12.000	0.03407	0.01707	0.6371	0.3801	70
12.649	0.03319	0.02505	0.4624	0.3325	72
14.831	0.03580	0.02381	0.4867	0.3011	81
16.971	0.03229	0.01916	0.4566	0.2345	48
20.000	0.04180	0.02331	0.7823	0.4319	50
23.324	0.04167	0.02380	0.8050	0.6219	24
24.558	0.04243	0.02393	0.5854	0.3226	152
26.833	0.04683	0.03026	0.5247	0.3139	52
28.844	0.03483	0.02247	0.5462	0.5010	30
31.241*	0.04600	0.03084	0.6285	0.4414	20
33.941*	0.02550	0.02492	1.1635	0.6699	10
36.394	0.05129	0.02822	0.5668	0.3236	116
37.947	0.05281	0.03483	0.4978	0.3630	32

* Indicates bins where $|N(h)|$ is below the minimum number of pairs and were consequently removed from the empirical semivariograms

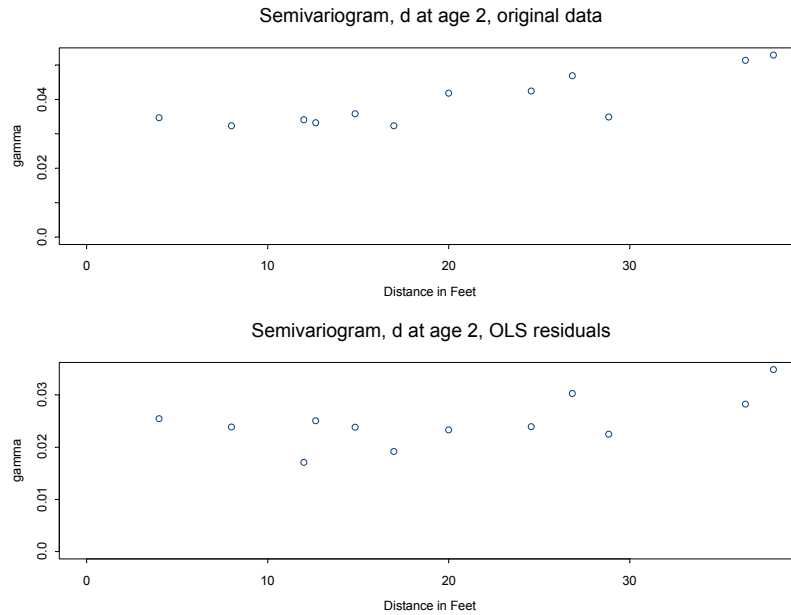
Where $\hat{\gamma}(h)_{raw}$ = The estimated gamma value derived from the raw diameter data

$\hat{\gamma}(h)_{resid}$ = The estimated gamma value derived from the residuals of the fitted model

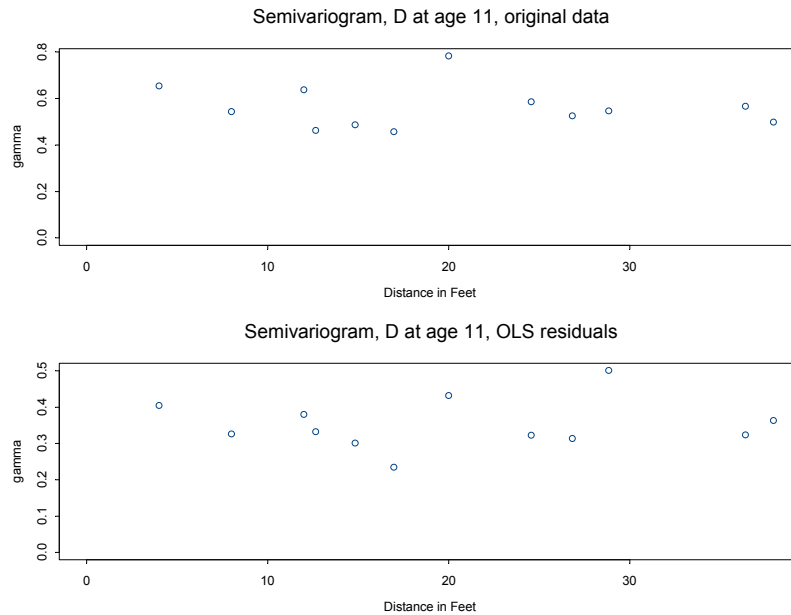
$|N(h)|$ = The number of distinct pairs in each distance bin

Figure 23: Graphs of the empirical semivariograms for stand number 20, derived from the original data and the residuals from a fitted linear model, for (a) ground-line diameter (d) at age 2, and (b) diameter at breast height (D) at age 11

(a)



(b)



the subject tree is not taken into account in this analysis. From the near horizontal trend in the semivariogram plots, it is apparent that none of the four graphs indicated strong spatial autocorrelations; though this may be due to an inability to visually pick up spatial trends in the data.

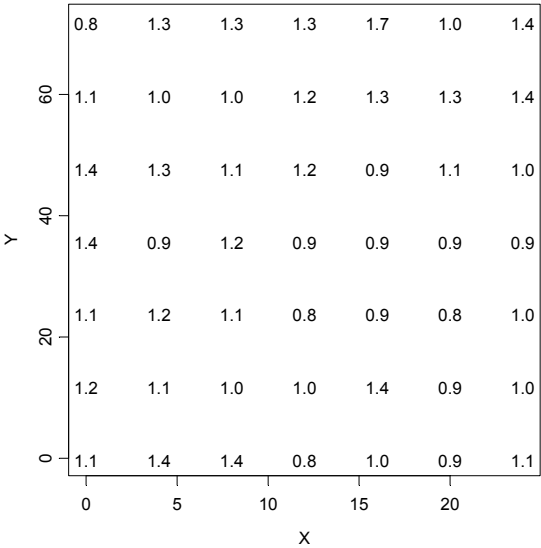
Graphs of the spatial distribution of the actual diameter measurements taken were also calculated and evaluated (Figure 24 (a-b)). The plots of the actual diameter measurements give a visual representation of possible spatial relationships between stems. Stems that have similar values to their neighbors would indicate a positive spatial correlation, whereas a negative spatial correlation would be indicated if stems had differing values when compared to their neighbors. Residual analysis was conducted for the fitted linear models. Scatter plots, histograms, and qq plots (quantile-quantile) of the residuals were generated and evaluated, but not included in this example.

The process demonstrated in the previous example was repeated for all 192 plots eleven times (ground-line diameter, ages 2 to 5; diameter at breast height, ages 5 to 11). Based on visual inspection, the graphical comparisons conducted did not show any strong spatial trends in the data. Examination of the empirical semivariograms when grouped by stand age, diameter measurement (ground-line or breast height), or planting density did not reveal any strong spatial trends in the plots considered in this study. Further analyses with other indices of spatial dependence were conducted to elucidate spatial trends in the data that were not detected with the empirical semivariograms.

Figure 24: Spatial distribution of diameters for stand number 20, for (a) ground-line diameter at age 2, and (b) diameter at breast height at age 11

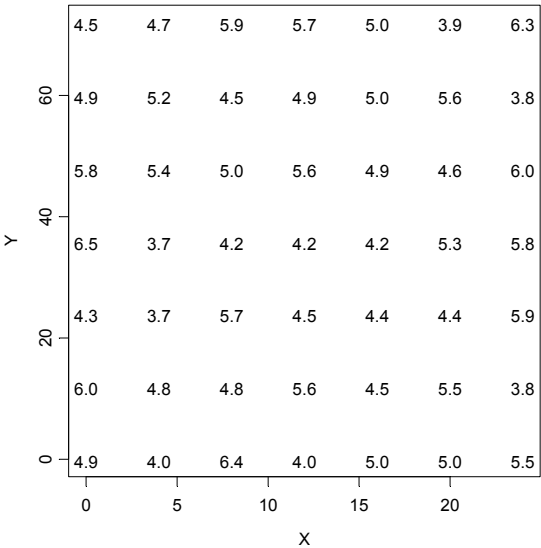
(a)

Spatial distribution of ground-line diameters at age 2



(b)

Spatial distribution of breast height diameters at age 11



6.4. Moran's I Coefficient

Testing for significant spatial autocorrelation in lattices with continuous data can be accomplished using Moran's I coefficient (Moran 1950), given by:

$$I = \frac{n}{\sum_i \sum_j A_{i,j}} \frac{\sum_i \sum_j A_{i,j} u_i u_j}{\sum_i u_i^2} \quad (40)$$

where $A_{i,j}$ = individual weight between subject stem i and neighboring stem j , obtained from the neighborhood matrix A

$u_i = Z(s_i) - \bar{Z}$, the i^{th} data value minus the plot mean value

$u_j = Z(s_j) - \bar{Z}$

n = the number of stems

The residuals from equation (39) were used as the data values ($Z(s_i)$) in calculating Moran's I coefficient. Residual plots and histograms were evaluated to establish the symmetry of the residuals around zero. Though not required for calculation of Moran's I coefficient, a normality assumption was inspected with qq plots and histograms of the residuals; no extreme violations were apparent. The neighborhood matrix A , whose elements are the observational weights, is used to reflect the influence of the surrounding stems on the subject stem. Moran's I coefficient is asymptotically normally distributed under the null hypothesis with a mean equal to $\frac{-1}{n-1}$ and variance given by σ_I^2 . Negative autocorrelation is indicated if Moran's I coefficient is less than the asymptotic mean, and positive autocorrelation is indicated if the resulting value is greater than the asymptotic mean. The direction (positive or negative) of the

autocorrelation is induced by the product of the $u_i u_j$ terms, which are estimates of the covariance between observations at two spatial locations.

There are two approaches to calculating the moments of Moran's I coefficient, randomization and normality, with both approaches having the same first moment. Schabenberger and Pierce (2002) give an excellent summary of these approaches and further details can be found in Cliff and Ord (1981) and Haining (1990). The normality approach uses the asymptotic normal distribution of Moran's I to calculate the p-values for testing the null hypothesis of no spatial autocorrelation. The test statistic is given by

$$z_{obs} = \frac{I - E[I]}{\sigma_I}, \quad (41)$$

where $E[I]$ = the first moment of Moran's I coefficient under the null hypothesis, and the null hypothesis is rejected if $|z_{obs}| > z_{\alpha/2}$. Haining (1990) warns that the asymptotic normality assumption will not hold if testing occurs with a sparse weight matrix.

The randomization approach considers the data to be fixed and a random draw from the $n!$ different possible permutations of the data to the lattice sites. The p-value can then be estimated by either ranking the observed Moran's I coefficient with the $n!-1$ other permutations, or by comparing the observed Moran's I coefficient to a random sample of the possible permutations. The first approach is computationally cumbersome, since there are $49!$ ($6.082818 \text{ E}+62$) possible arrangements of the data for a plot without mortality. Taking a random sample from these $n!$ possible outcomes will provide the basis for an unbiased estimate of the p-value for the observed Moran's I coefficient. The number of random samples drawn needs to be large enough to resolve p-values at a desired level of precision. The observed Moran's I coefficient is then ranked among

these values and the p-value calculated from the number of permutations more extreme than the observed Moran's I coefficient. The randomization method was used to calculate the p-values for the derived Moran's I coefficients in this study. Ten thousand samples were randomly drawn from the $n!$ possible outcomes for a specified data set to obtain a p-value.

A plot of the Moran's I coefficients significant at the 0.10 level is given in Figure 25, revealing both positive and negative spatial autocorrelations across all ages. Plots with non-significant Moran's I coefficients account for the gap present in Figure 25. An initially decreasing trend with age is noticeable for the positively correlated stands with respect to the number and magnitude of significant indices. The initially decreasing trend with age is consistent with the hypothesized result of positive correlation at young ages with a shift towards an uncorrelated or negatively correlated stand at later ages. There is a total of 282 significant Moran's I coefficients across all plots and ages, inclusive of both the ground-line diameter and diameter at breast height data. A significant spatial correlation at the 0.10 alpha level for 13.4% ($\frac{282}{2112}$) of the plots across all ages was exhibited.

The number of Moran's I coefficients that are significant is presented in Table 12 by planting density, age, and diameter measurement. The results show that the number of plots with significant spatial correlation decreases with an increase in age for the data derived from ground-line diameter measurements. The decreasing trend in the number of significant Moran's I coefficients with age is not present for the data derived from breast height diameter measurements. To evaluate the impact of planting density on the number

Figure 25: Moran's I coefficients significant at an alpha level of 0.10 across all ages, for the residuals from a fitted model for ground-line diameters (d) and breast height diameters (D), with a mean equal to $\frac{-1}{n-1}$

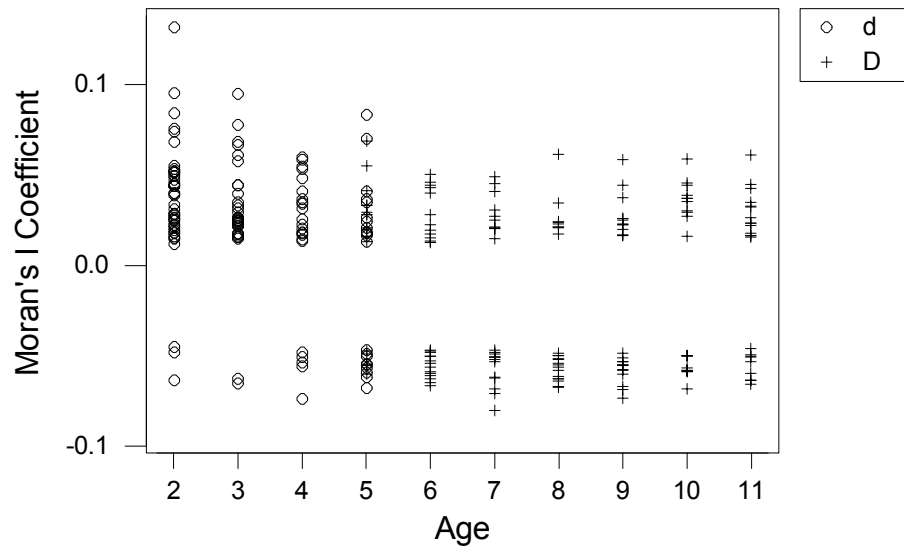


Table 12: Frequency of significant Moran's I coefficients (0.10 alpha level) across all planting densities, ages, and diameter measurements considered in this research

Diameter	Age	Planting density (TPA)									Sum
		302	453	605	680	907	1210	1361	1815	2722	
d	2	2	5	8	3	7	5	6	5	3	44
d	3	1	4	9	5	4	4	3	4	1	35
d	4	1	2	4	3	3	3	5	0	2	23
d	5	0	2	4	1	7	2	4	1	1	22
D	5	2	3	1	2	5	2	0	2	1	18
D	6	2	5	2	2	7	2	3	2	1	26
D	7	0	2	1	0	9	0	3	7	1	23
D	8	3	2	4	1	6	0	5	2	1	24
D	9	2	2	4	2	5	0	4	4	2	25
D	10	1	1	2	1	5	0	5	2	2	19
D	11	3	3	3	0	7	1	2	3	1	23
Sum		17	31	42	20	65	19	40	32	16	282
Prop (%)		12.9	11.7	15.9	15.2	12.3	14.4	15.2	12.1	12.1	

Where TPA = Trees per acre

d = Ground-line diameter

D = Diameter at breast height

Prop (%) = The proportion of rejections of the null hypothesis, expressed as a percentage

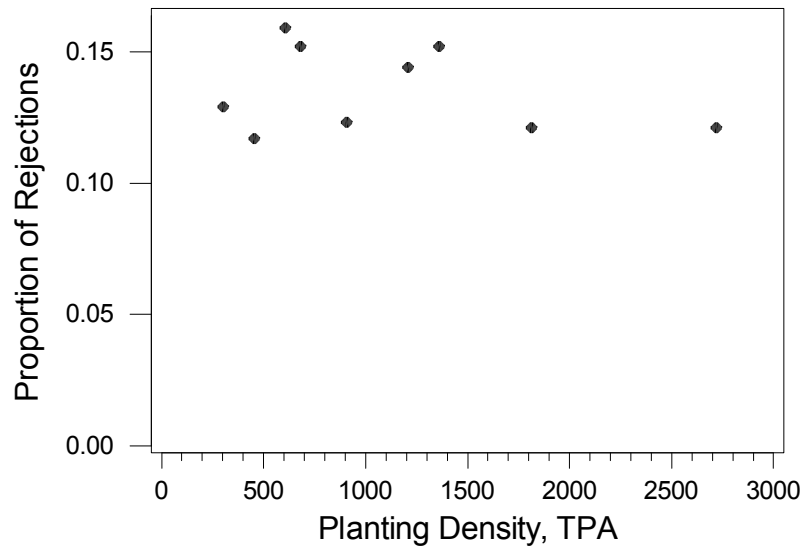
of plots with significant Moran's I coefficients, each planting density needs to be viewed with respect to the number of plots that were observed at that density. Table 2 gives the number of plots for the nine different planting densities established.

Given that the number of plots in the study varies with the planting density considered, a proportion of the number of plots with a significant Moran's I coefficient was computed. The proportion is given in Table 12 and displays the percentage of plots within a single planting density with significant spatial autocorrelation across ages 2 to 11. The percentage of rejections across planting densities averaged 13.5%, with a range of 11.7 to 15.9%. A graph of the average proportion of rejections by planting density is presented in Figure 26 (a) and does not show any trend in the proportion of rejections across planting density. The planting rectangularity of the plots does not seem to interact with the average proportion of rejection. Figure 26 (b) shows an initial negative trend between the average proportion of rejections of the null hypothesis and age. The trend levels off at age 5, and is fairly constant through age 11. Regression analysis indicated age as a significant predictor ($p\text{-value} < 0.001$) of Moran's I coefficient, while planting density did not aid in prediction ($p\text{-value} > 0.20$).

The initially decreasing trend may represent a planting density and age interaction. To elucidate any interactions between planting density and age, graphs of the proportion of rejections at each age over all planting densities were computed. For ground-line (Figure 27 (a)) and breast height (Figure 27 (b)) diameters, the proportion of rejections does not seem to have any discernable trends with the planting density and age of the plot. These comparisons indicate that there is not a strong interaction between the

Figure 26: Average proportion of significant Moran's I coefficients at an alpha level of 0.10 for (a) each planting density (trees per acre) and (b) over ages 2 to 5 for ground-line diameter (d), and over ages 5 to 11 for diameter at breast height (D)

(a)



(b)

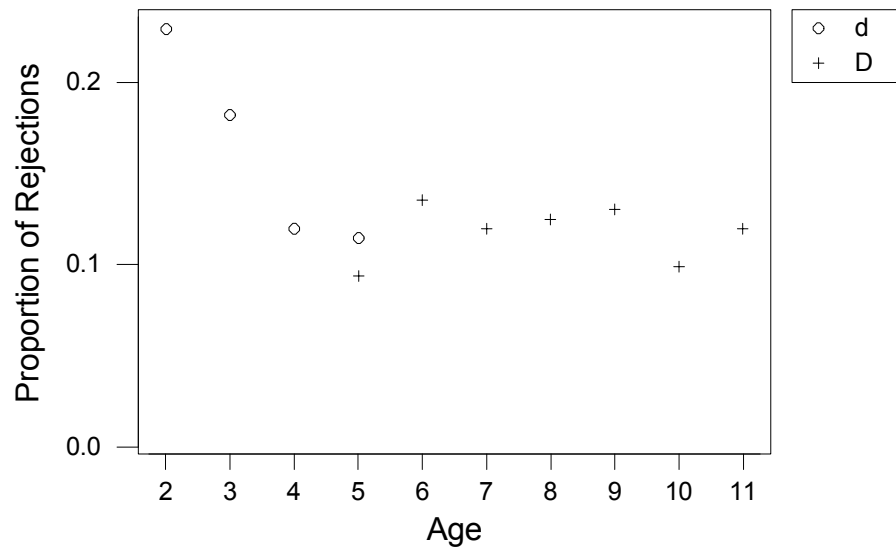
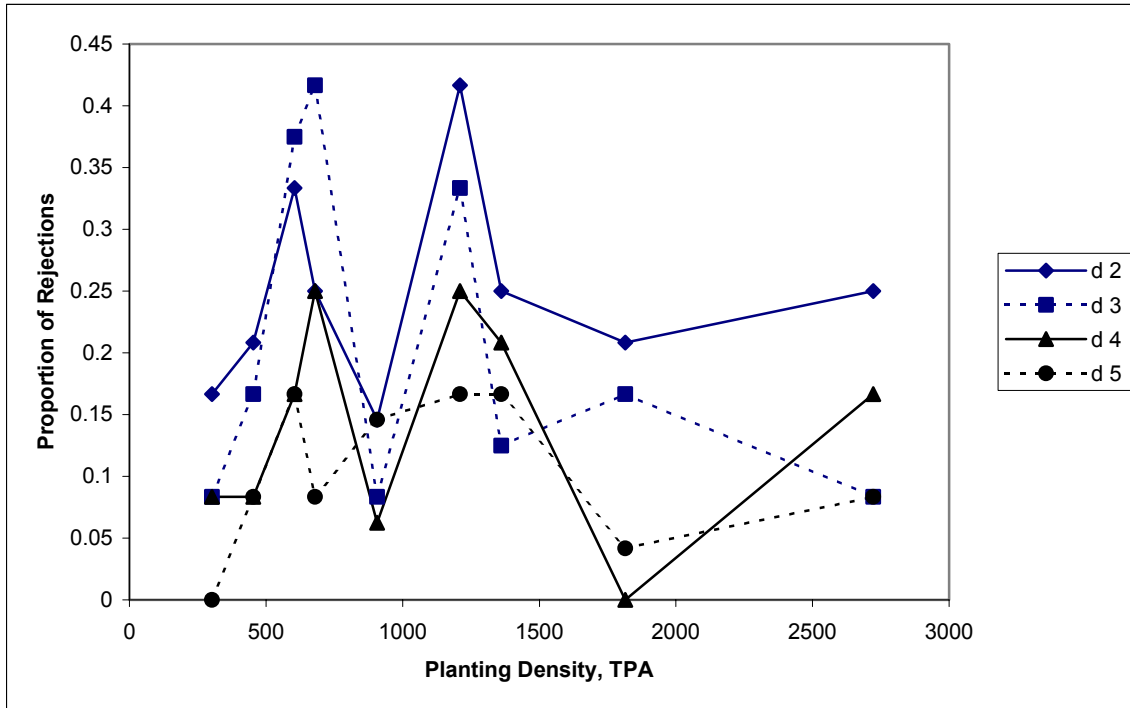
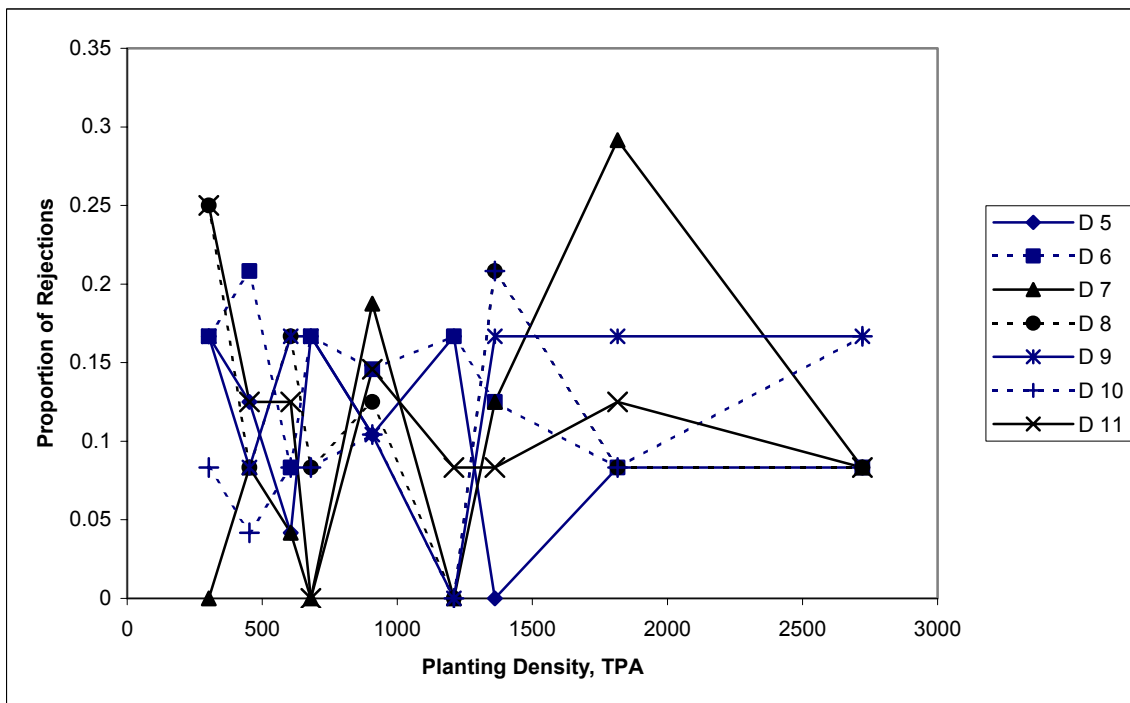


Figure 27: Proportion of significant Moran's I coefficients at an alpha level of 0.10 for each planting density (trees per acre) over (a) ages 2 to 5 for ground-line diameter (d), and (b) ages 5 to 11 for diameter at breast height (D)

(a)



(b)



age and planting density of juvenile loblolly pine with respect to the presence of significant spatial autocorrelation.

6.5. Geary's C Coefficient

Another means of testing for spatial autocorrelation is a statistic developed by Geary (1954). Termed Geary's C, the coefficient is:

$$C = \frac{n-1}{2 \sum_i \sum_j A_{i,j}} \frac{\sum_i \sum_j A_{i,j} (Z(s_i) - Z(s_j))^2}{\sum_i u_i^2} \quad (42)$$

where all terms are as previously defined.

Geary's C coefficient is asymptotically normal under the null hypothesis of no spatial autocorrelation and has a mean equal to one and variance equal to σ_C^2 . Positive spatial autocorrelation is indicated if $0 \leq C < 1$, and negative spatial autocorrelation is indicated if $C > 1$. The index of spatial autocorrelation is bounded due to the squared difference between measurements at spatial locations i and j. One would expect C to be less than one if stems close to one another are relatively close in observational values, and C should be greater than one if close observations differ considerably (see Cliff and Ord (1981) and Haining (1990) for further discussion of Geary's C coefficient). Calculation of the moments of Geary's C can be estimated using either the randomization or normality procedures discussed in the previous section. Geary's C coefficients were calculated and ranked for a random sample of ten thousand permutations from the n! possible outcomes. The randomization procedure resulted in a p-value estimate for the test of spatial autocorrelation by ranking the observed Geary's C coefficients among those derived from the ten thousand random permutations. The null hypothesis of no

spatial autocorrelation was rejected when the p-value estimated from the randomization scheme was less than an alpha level of 0.10.

The residuals from equation (39) were used as the measurements at each spatial location for estimating Geary's C coefficient. Figure 28 presents the Geary's C coefficients for all plots that have significant autocorrelation at the 0.10 alpha level. The derived Geary's C coefficients reveal positive and negative spatial autocorrelation over all ages considered in this research (Figure 28). Stands that did not have significant Geary's C coefficients constitute the gap present in Figure 28. Visual inspection discloses an initial decrease in the number of significant autocorrelations with an increase in age; these results are consistent with the hypothesized correlation patterns for young stands. There is a total of 268 significant Geary's C coefficients across all plots in the study, representing 12.7% ($\frac{268}{2112}$) of the plots.

Table 13 presents the number of significant spatial correlation coefficients by planting density, age, and diameter measurement. An initially decreasing trend in the number of plots with significant spatial autocorrelation with age was noted in Table 13. Table 13 also gives the proportion of plots with significant Geary's C coefficients by planting density. The proportions are relatively constant across all planting densities, averaging 12.9% with a range of 7.6 to 21.2%. A plot of the average proportion of rejections for each planting density is given in Figure 29 (a). No trend is visible in the average proportion of rejections with planting density. No interaction between the planting rectangularity of the plots and the average proportion of rejections was noted for

Figure 28: Geary's C coefficients significant at an alpha level of 0.10 across all ages for the residuals from a fitted model for ground-line diameters (d) and breast height diameters (D), with a mean equal to one

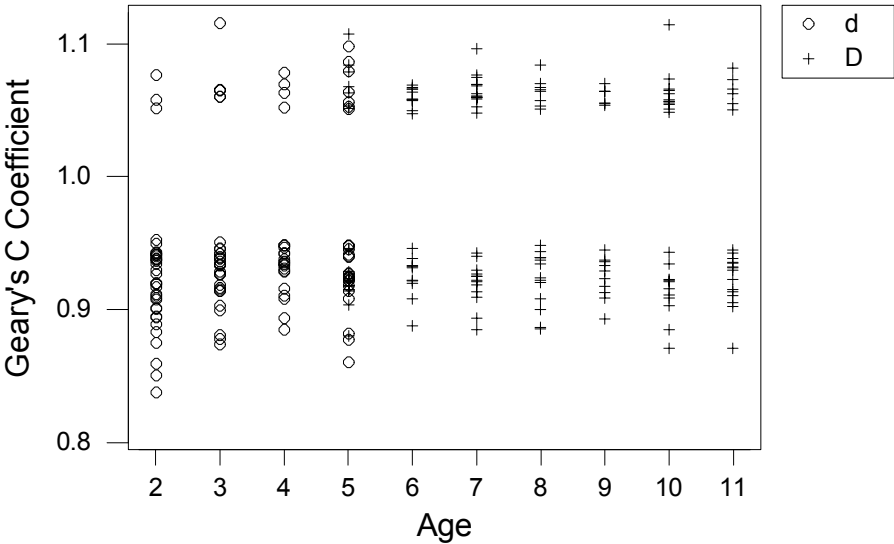


Table 13: Frequency of significant Geary's C coefficients (0.10 alpha level) across all planting densities, ages, and diameter measurements considered in this research

Diameter	Age	Planting density (TPA)									Sum
		302	453	605	680	907	1210	1361	1815	2722	
d	2	2	5	9	0	5	3	7	3	1	35
d	3	3	2	7	1	7	5	4	1	1	31
d	4	0	4	3	3	2	4	4	2	1	23
d	5	0	1	8	0	7	3	7	2	0	28
D	5	0	2	3	2	5	2	4	2	2	22
D	6	3	3	2	0	6	3	1	2	1	21
D	7	0	3	3	0	10	2	4	4	1	27
D	8	0	3	3	0	7	2	2	0	4	21
D	9	1	2	1	1	2	2	3	3	1	16
D	10	4	1	3	1	4	1	5	3	1	23
D	11	3	2	2	2	5	1	2	2	2	21
Sum		16	28	44	10	60	28	43	24	15	268
Prop (%)		12.1	10.6	16.7	7.6	11.4	21.2	16.3	9.1	11.4	

Where TPA = Trees per acre

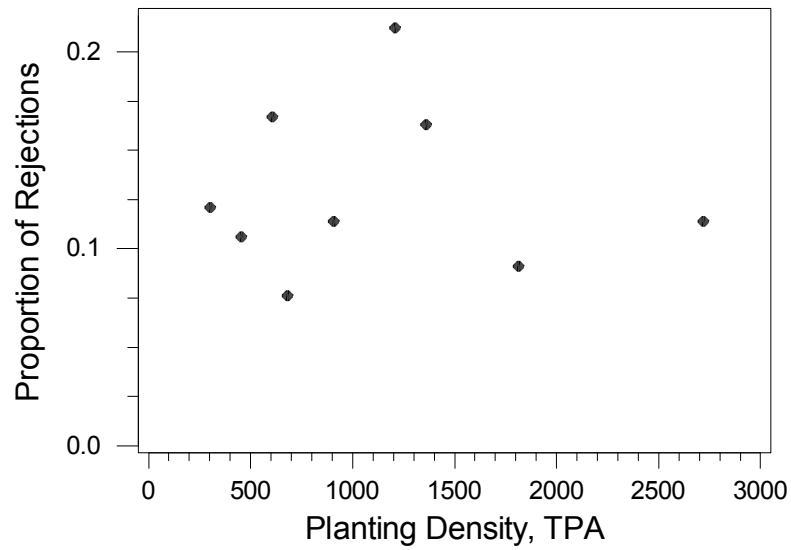
d = Ground-line diameter

D = Diameter at breast height

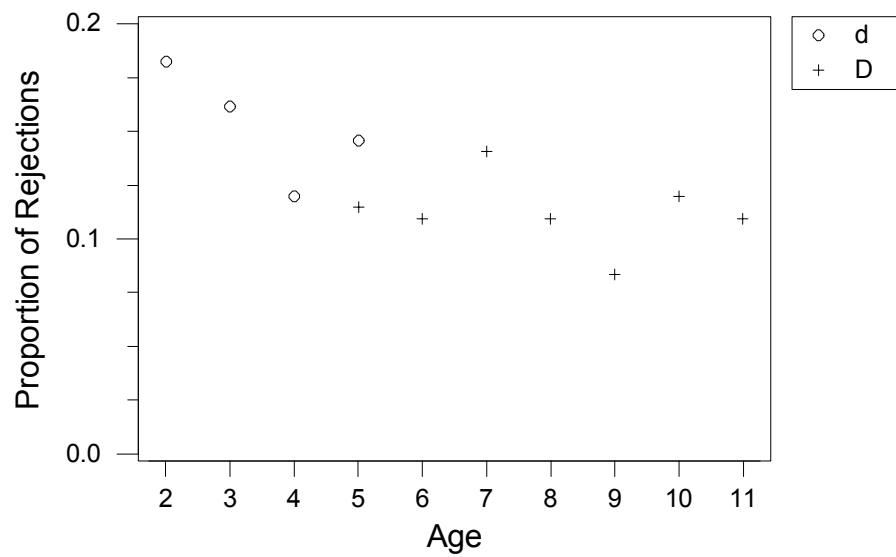
Prop (%) = The proportion of rejections of the null hypothesis, expressed as a percentage

Figure 29: Average proportion of significant Geary's C coefficients at an alpha level of 0.10 for (a) each planting density (trees per acre) and (b) over ages 2 to 5 for ground-line diameter (d), and over ages 5 to 11 for diameter at breast height (D)

(a)



(b)

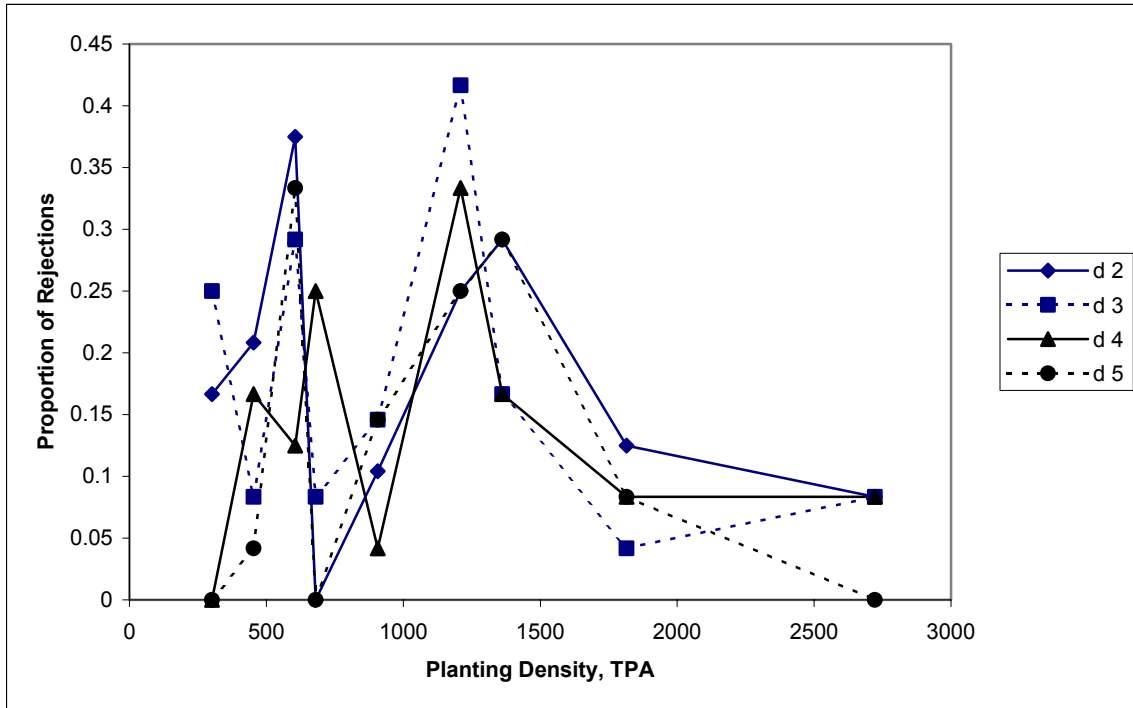


the significant Geary's C coefficients. The average proportion of significant Geary's C coefficients is given for ground-line and breast height diameters over each age in Figure 29 (b), revealing an initial decreasing trend in the proportion of rejections with an increase in age. Regression analysis showed age to be a significant predictor (p-value < 0.001) of Geary's C coefficient, while planting density did not aid in modeling (p-value = 0.194).

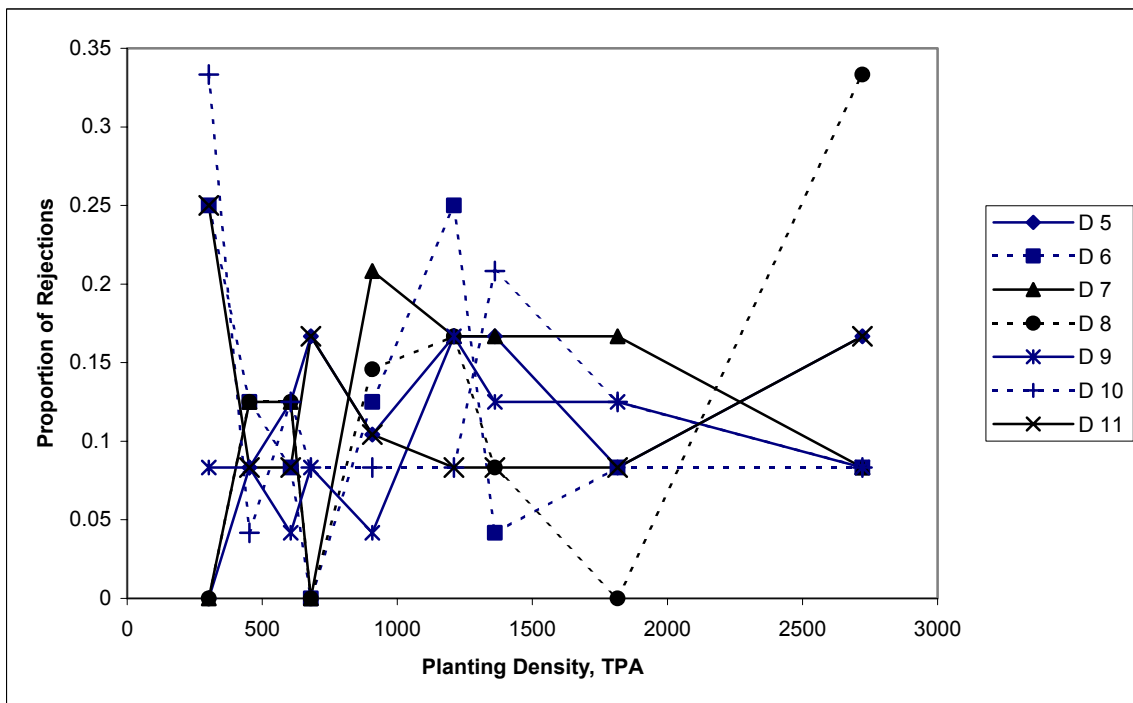
A graphical representation of the proportion of rejections by planting density and age is presented in Figure 30 (a) for ground-line diameters, and does not indicate any clear trends with planting density from ages 2 to 5. The proportion of significant Geary's C coefficients for each planting density over ages 5 to 11 for diameter at breast height is presented in Figure 30 (b), and does not expose any clear trends across the planting densities at each age. The aforementioned lack of distinguishable trends in the presence of spatial autocorrelation and planting density, as measured by Geary's C coefficient, indicates that the planting density of juvenile loblolly pine is not strongly related to the occurrence of significant spatial trends.

Figure 30: Proportion of significant Geary's C coefficients at an alpha level of 0.10 for each planting density (trees per acre) over (a) ages 2 to 5 for ground-line diameter (d), and (b) ages 5 to 11 for diameter at breast height (D)

(a)



(b)



6.6. Spatial Models

In utilizing a spatial model with lattice data there are at least two sources of variation, large- and small-scale. Large-scale variation was previously defined as the variation in any mean trends in the data, including that of explanatory variables. Small-scale variation is due to measurement error and spatial autocorrelation with neighbors.

The general form of the spatial regression model is given by:

$$Z = X\beta + \varepsilon, \quad (43)$$

where the error term is distributed by:

$$\varepsilon \sim N(0, \Sigma) \quad (44)$$

The variance-covariance matrix, Σ , is fitted using one of three possible covariance structures: conditional autoregressive (CAR), simultaneous autoregressive (SAR), or moving average (MA) models. The three structures differ in the choice of the variance-covariance matrix, Σ . A brief review of these structures is given.

A conditional autoregressive model defines the variance-covariance matrix as follows:

$$\Sigma = (I - \rho A)^{-1} M, \quad (45)$$

and assumes that given the spatial neighborhood, all of the information about s_i is known (Haining 1990). That is:

$$\begin{aligned} & \Pr(Z(s_i) | \{Z(s_j) : j \neq i\}) \\ &= \Pr(Z(s_i) | Z(s_1), \dots, Z(s_{i-1}), Z(s_{i+1}), \dots, Z(s_n)), \\ & \text{for all } s_i \in T \end{aligned} \quad (46)$$

Where the stochastic process is distributed as follows:

$$Z \sim N(\mu, \Sigma), \quad (47)$$

and the variance-covariance matrix is as defined above. The following conditions and definitions are given: $(I - \rho A)$ is invertible, $(I - \rho A)^{-1} M$ is symmetric and positive definite, M is a diagonal matrix with elements $(\sigma_1^2, \dots, \sigma_n^2)$, A is the previously defined *known* symmetric neighborhood matrix, ρ is the only spatial dependence parameter, and $\mu_i = E(Z(s_i))$. Note that neighborhood weights between sites s_i and s_j are related in a symmetric manner.

The simultaneous autoregressive model assumes the following variance-covariance form:

$$\Sigma = (I - B)^{-1} \Lambda (I - B')^{-1}, \quad (48)$$

where B is the spatial dependence matrix and does not have to be symmetric, and Λ is the diagonal variance matrix. The least-squares estimators of simultaneous autoregressive parameters may be inconsistent due to the residuals being correlated with neighboring data values (Cressie 1993); estimation by maximum likelihood techniques is a viable alternative.

The moving average model has a covariance structure of the form:

$$\Sigma = (I + G)\Omega(I + G'), \quad (49)$$

where G is the invertible spatial dependence matrix and Ω is the diagonal variance matrix. The moving average model corresponds to the time series moving average model (Cliff and Ord 1981).

All three covariance structures can be utilized to model data on spatial lattices. Further details are given by Cliff and Ord (1981), Haining (1990), Cressie (1993), and

Schabenberger and Pierce (2002) on each structure. The conditional autoregressive model was deemed appropriate for this study because of consistent least squares estimation and conditional relationship to the data values. The CAR model was the only structure considered.

6.7. CAR Parameter Estimation

The parameters of the CAR spatial model were estimated by maximizing the following likelihood function (Cressie 1993):

$$L(\beta, \Sigma; Z) = (2\pi)^{-\frac{n}{2}} |M|^{-\frac{1}{2}} |I - \rho A|^{\frac{1}{2}} \exp\left\{-\left(\frac{1}{2}\right)(Z - \mu)' M^{-1} (I - \rho A)(Z - \mu)\right\} \quad (50)$$

The linear equation follows where $\mu = X\beta$:

$$Z = X\beta + \varepsilon, \quad (51)$$

and Z is distributed by (47), X is the matrix of regressors, β is the vector of parameters, and ε is the error term. The error term is distributed as:

$$\varepsilon \sim N(0, \Sigma), \quad (52)$$

where Σ is as previously defined for the CAR model. The distributions employed in (47) and (52) require an assumption of normality. To validate the assumption of a Normal distribution and enable usage of the conditional autoregressive model, goodness-of-fit statistics were evaluated.

The fit of the Normal distribution to the juvenile diameter data was examined using the Anderson-Darling (equation (19)), Cramer-von Mises (equation (20)), and Kolmogorov-Smirnov (equation (21)) test statistics. The goodness-of-fit statistics were calculated on random samples of 50 plots for ground-line diameter at ages 2 and 5, and breast height diameters at ages 8 and 11 (Table 14). The results indicate that most plots

fail to reject the null hypothesis that the empirical data come from the Normal distribution, implying that the Normal distribution does a good job of characterizing the juvenile diameter distributions. The number of plots from the random samples that failed to reject the null hypothesis decreased as age increased (Table 14), indicating that the Normal distribution may not be appropriate in more mature stands.

The two-parameter Weibull distribution was shown to adequately fit the juvenile diameter distributions considered in this research (Section 4.3). The Normal distribution also appears appropriate for characterizing the diameter distributions of loblolly pine at juvenile ages. Example graphs of the fitted Normal and Weibull distributions are given for randomly selected plots in Figure 31 (a-b) for ground-line diameters at ages 2 and 5, and in Figure 32 (a-b) for breast height diameters at ages 8 and 11. In the example cases presented the goodness-of-fit statistics indicate both distributions adequately fit the data. To avoid switching distributional forms when forecasting diameter distributions, the two-parameter Weibull distribution was used for consistency in characterizing the diameter distributions. The normality assumption was made to facilitate the conditional autoregressive analysis on the juvenile diameter data.

The employed model is a first-order conditional autoregressive model due to the presence of only one spatial dependency parameter and neighborhood weight matrix. Homoskedasticity was conditionally assumed, which implies that $M = \sigma^2 I$. The stochastic error term has an expected value equal to zero, which does not depend on spatial location, and the covariance between any two points is only a function of the

Table 14: Results for the Anderson-Darling, Cramer-von Mises, and Kolmogorov-Smirnov goodness-of-fit statistics for fit to Normal distribution from 50 random plots for ground-line diameter (d) at ages 2 and 5, and breast height diameter (D) at ages 8 and 11

	d, Age 2	d, Age 5	D, Age 8	D, Age 11
Anderson-Darling				
NS	49	50	48	45
S	1	0	2	5
Cramer-von Mises				
NS	49	50	49	47
S	1	0	1	3
Kolmogorov-Smirnov				
NS	41	50	47	47
S	9	0	3	3

Where NS = The number of plots that are non-significant at the 0.20 alpha level

S = The number of plots that are significant at the 0.20 alpha level

Figure 31: The fitted Normal and two-parameter Weibull distributions for ground-line diameter at ages (a) 2 and (b) 5

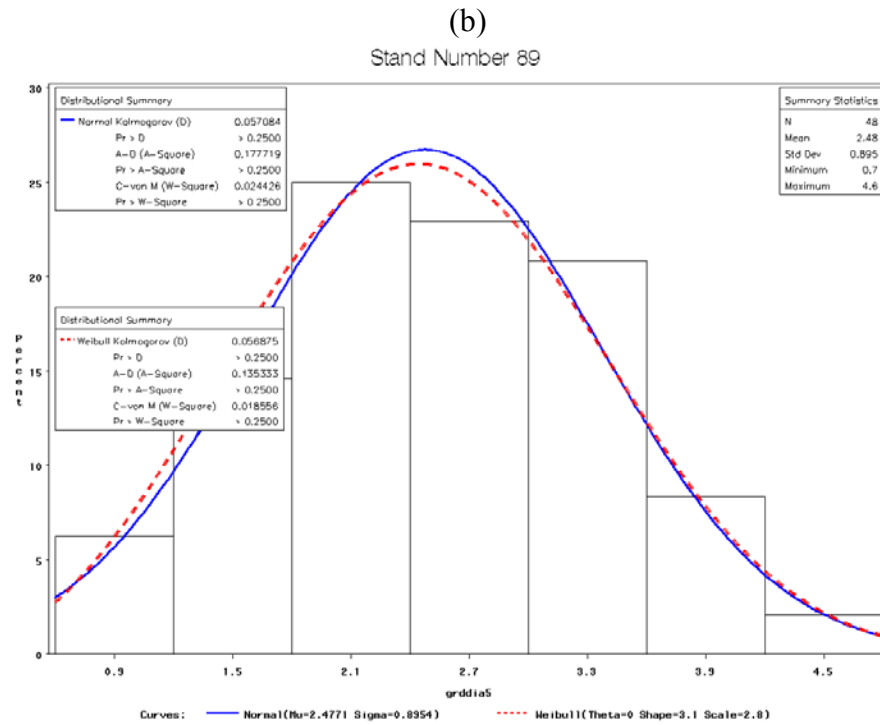
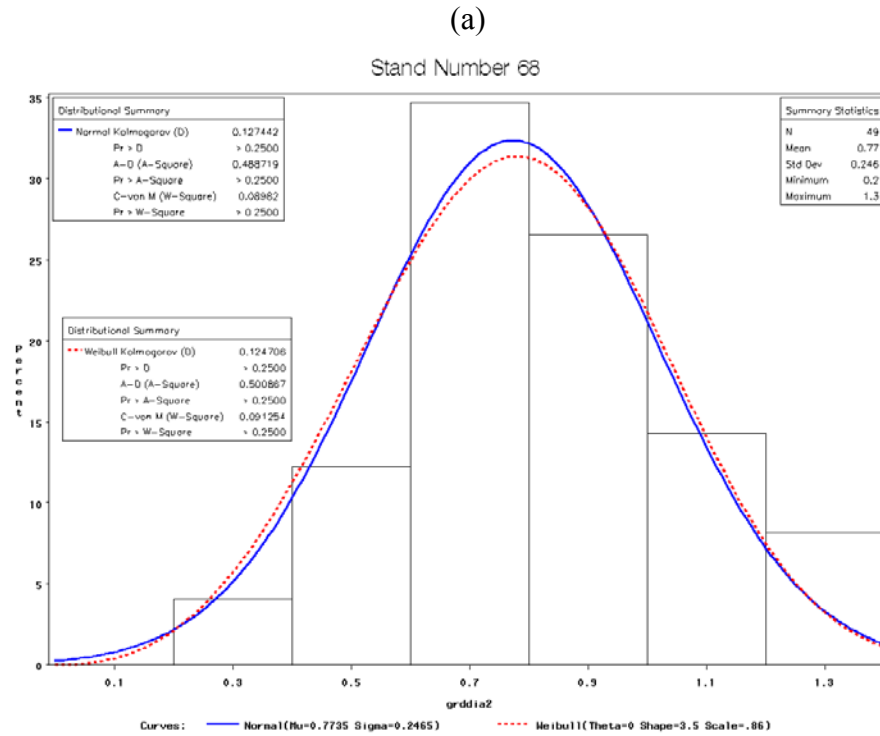
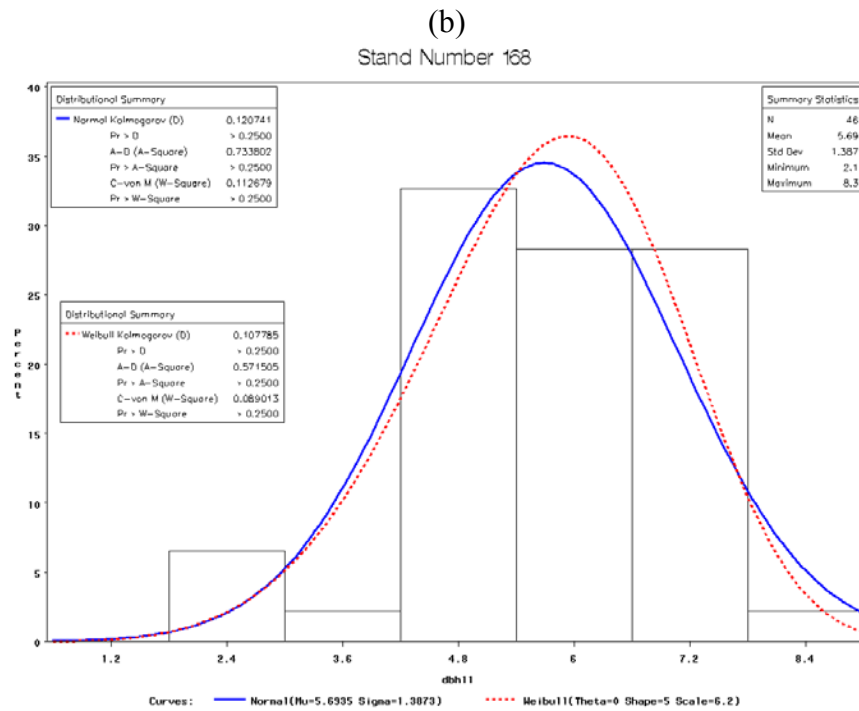
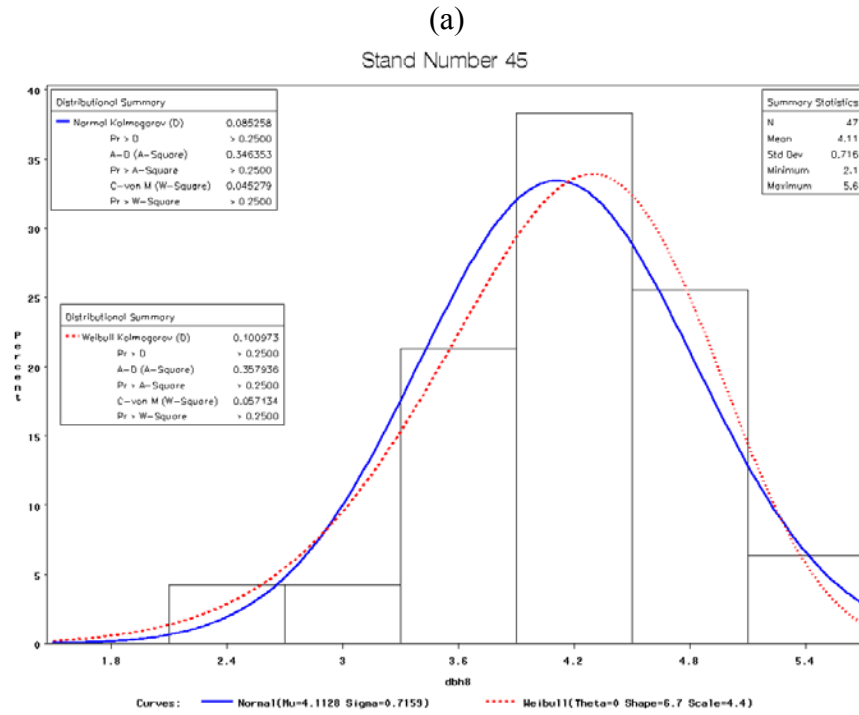


Figure 32: The fitted Normal and two-parameter Weibull distributions for breast height diameter at ages (a) 8 and (b) 11



spatial separation. These properties imply that the ε process is second-order stationary (Schabenberger and Pierce 2002).

The log-likelihood of (50) is given in equation (53), and was employed to obtain the MLEs of β and Σ .

$$\begin{aligned} \text{Log}L(\beta, \Sigma; Z) = & -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2) + \frac{1}{2} \log(|I - \rho A|) \\ & - \frac{1}{2\sigma^2} (Z - X\beta)'(I - \rho A)(Z - X\beta) \end{aligned} \quad (53)$$

The following estimating equations result:

$$\hat{\beta} = (X'(I - \rho A)X)^{-1} X'(I - \rho A)Z \quad (54)$$

$$\hat{\sigma}^2 = (Z - X\hat{\beta})'(I - \rho A)(Z - X\hat{\beta}) / (n - p) , \quad (55)$$

where $p = k + 1$, and k = the number of parameters estimated.

To solve for the MLEs, an iterative routine² was utilized. The solution space for ρ is limited by the minimum and maximum eigenvalues of the A matrix, $\lambda_{\min}$ and $\lambda_{\max}$, respectively. These eigenvalues $\lambda_1, \dots, \lambda_n$ also enable fast computation of the determinant of $(I - \rho A)$ (Cressie 1993). The determinant is given by:

$$|I - \rho A| = \prod_{i=1}^n (1 - \rho \lambda_i) . \quad (56)$$

For $\lambda_{\min} < 0 < \lambda_{\max}$, the spatial-dependence parameter ρ ranges between $\frac{1}{\lambda_{\min}} < \rho < \frac{1}{\lambda_{\max}}$,

if $0 \leq \lambda_{\min} < \lambda_{\max}$ then $\rho < \frac{1}{\lambda_{\max}}$, and if $\lambda_{\min} < \lambda_{\max} \leq 0$ then $\rho > \frac{1}{\lambda_{\min}}$. The maximization

of the log-likelihood function (53) can only occur within the bounds on the spatial dependency parameter. The output from the iterative routine is saved to a file to enable identification of the ρ value that maximizes the log-likelihood function with the estimates

² S-Plus 2000, Insightful Corp., Inc.

of β and σ^2 . The procedure presented was performed for the spatial linear model including the mean trend and that of the covariate total stem height. A likelihood ratio test was applied to test the significance of the spatial dependency parameter ρ . The likelihood ratio uses the log-likelihood given in equation (53) and the log-likelihood under the null hypothesis of no spatial autocorrelation to test the significance of ρ .

The significant spatial dependency parameters (p-values < 0.10) for the spatial linear model utilizing the covariate of total stem height are presented in Figure 33. Over all plots and ages there is a total of 351 significant spatial autocorrelation structures at an alpha level of 0.10, representing 17% of the data ($\frac{351}{2112}$). The separation between positive and negative spatial dependency parameters present in Figure 33 results from the exclusion of plots with non-significant spatial dependency parameters. The apportionment of plots with significant spatial autocorrelation by planting density, age, and diameter measurements is given in Table 15.

The proportion of rejections for each planting density was computed, averaging 15.8% with a range of 10.6 to 21.2%. A graphical representation of the average proportion of significant ρ values for each planting density over all ages is given in Figure 34 (a), and does not elucidate any significant trends in the average proportion of rejections with planting density. The planting rectangularity of the plots does not seem to aid in differentiating any differences in the oscillation of the average proportion of rejections of ρ values for the spatial linear model including the covariate of total stem height. Figure 34 (b) shows an increasing trend in the average proportion of rejections with age, continuing until age 9 where it levels off and no further change in the average

Figure 33: Derived ρ values significant at an alpha level of 0.10 across all ages for the spatial linear model including the covariate total stem height, for ground-line diameters (d) and breast height diameters (D)

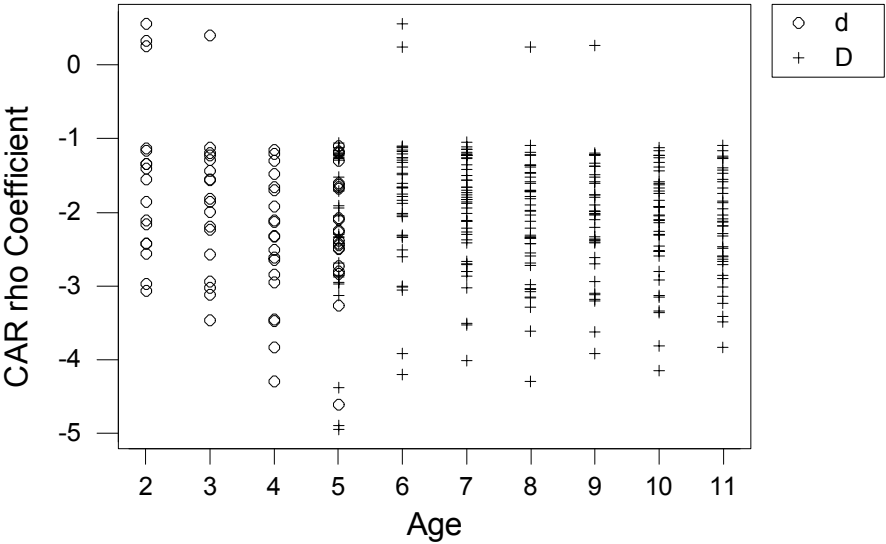


Table 15: Frequency of significant spatial dependency in spatial linear models with a covariate term (0.10 alpha level), across all planting densities, ages, and diameter measurements considered in this research

Diameter	Age	Planting density (TPA)									Sum
		302	453	605	680	907	1210	1361	1815	2722	
d	2	0	4	2	1	6	1	1	1	1	17
d	3	2	0	4	1	4	2	2	3	0	18
d	4	4	2	1	3	5	1	2	2	0	20
d	5	1	3	4	2	7	1	2	2	0	22
D	5	4	4	2	1	12	0	3	2	1	29
D	6	0	6	3	1	13	1	3	7	2	36
D	7	1	6	5	1	16	1	8	7	2	47
D	8	2	7	4	5	11	2	8	4	3	46
D	9	1	2	3	6	12	2	8	2	3	39
D	10	2	5	5	3	12	3	5	3	1	39
D	11	2	4	5	4	10	3	6	3	1	38
Sum		19	43	38	28	108	17	48	36	14	351
Prop (%)		14.4	16.3	14.4	21.2	20.5	12.9	18.2	13.6	10.6	

Where TPA = Trees per acre

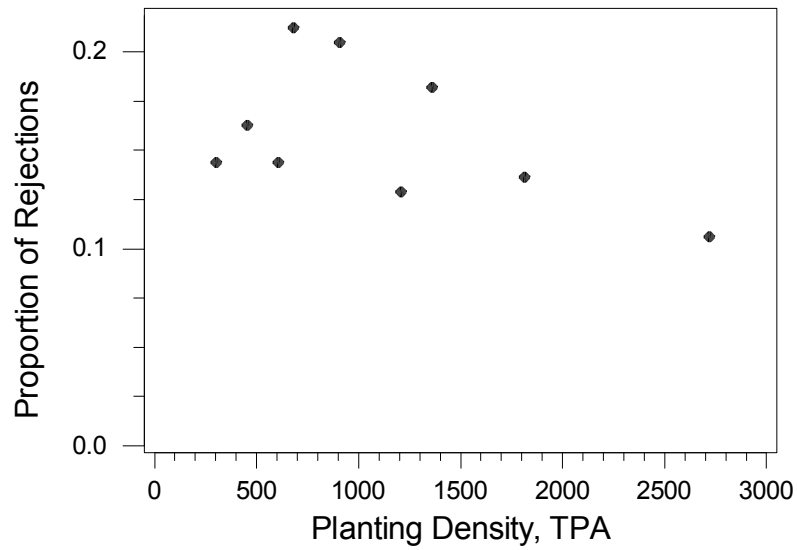
d = Ground-line diameter

D = Diameter at breast height

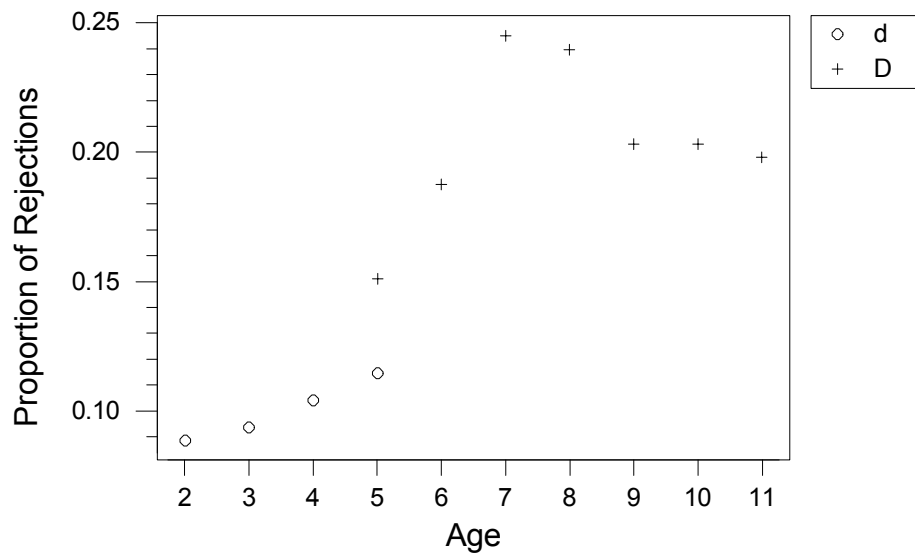
Prop (%) = The proportion of rejections of the null hypothesis, expressed as a percentage

Figure 34: Average proportion of significant p values at an alpha level of 0.10 for the spatial linear model including the covariate total stem height for (a) each planting density (trees per acre) and (b) over ages 2 to 5 for ground-line diameter (d), and over ages 5 to 11 for diameter at breast height (D)

(a)



(b)



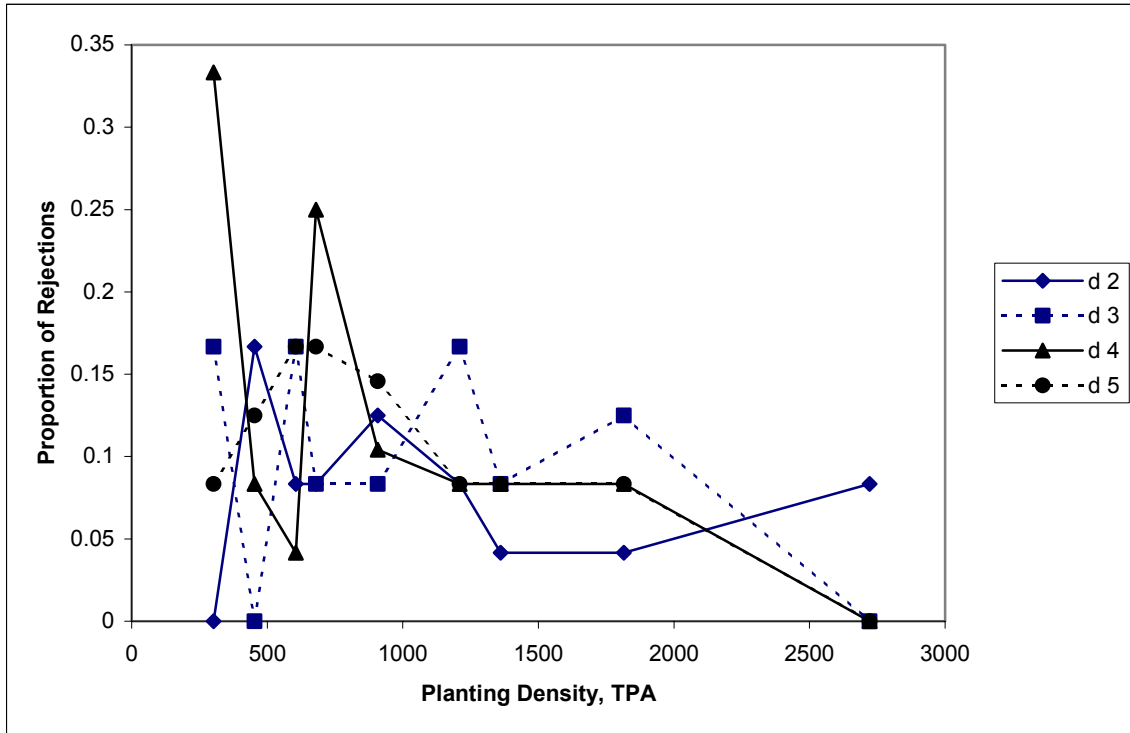
proportion of rejection is evident.

Regression analysis indicated that both age and planting density were significant predictors ($p\text{-values} \leq 0.01$) of the spatial dependence parameter. Graphs of the proportion of rejections were employed to investigate possible interactions between planting density and age. The proportion of rejection for each planting density over ages 2 to 5 is presented in Figure 35 (a) for the spatial linear model including the covariate of total stem height. No visible trend is present across planting densities or age for the proportion of significant p values derived from ground-line diameter measurements. The absence of any trends continues for the proportion of significant p values derived from the diameter at breast height measurements over ages 5 to 11 (Figure 35 (b)). For the spatial model including that of the covariate total stem height, no strong trends were noted between the planting density and the presence significance of spatial dependency on a plot. A positive trend was noticeable between the average proportion of rejections and age, with the trend strongest over ages 5 to 8 where the onset of competition occurred for most plots.

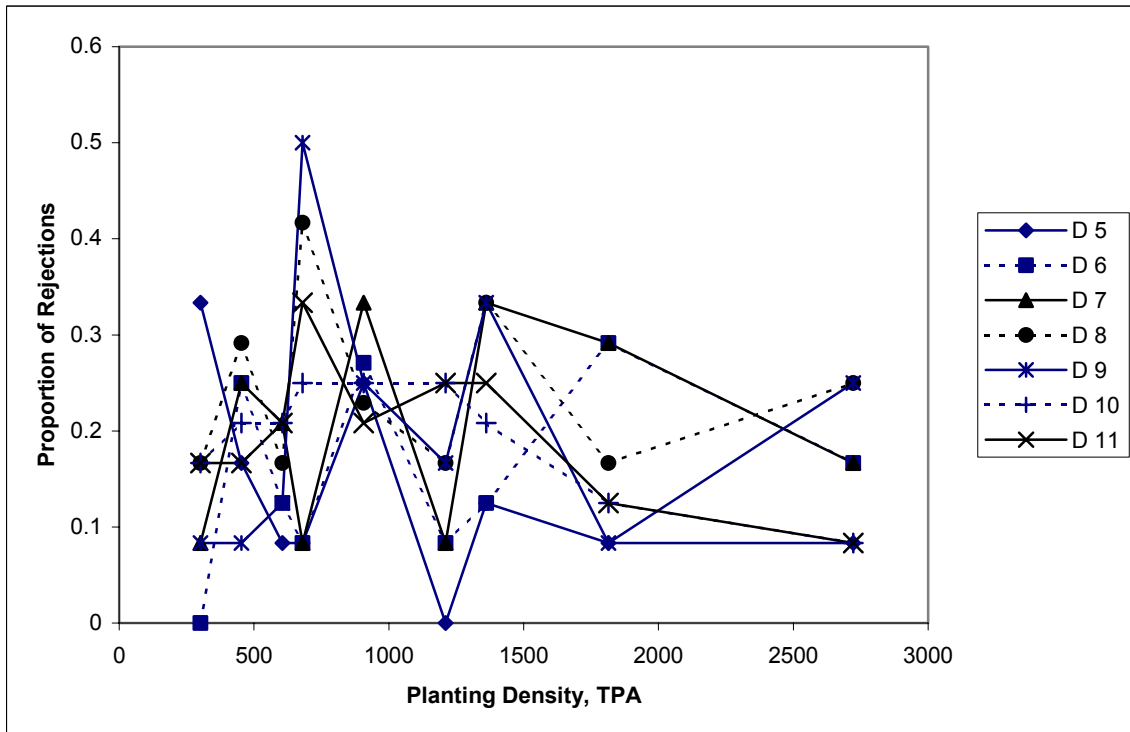
To investigate the effectiveness of the CAR model in accounting for any spatial dependency within the plots, the residuals from the spatial linear model were evaluated using Geary's C coefficient (equation (42)). If the CAR model adequately accounts for the spatial influence among stems, then the residuals from the spatial linear model should not have a significant spatial autocorrelation. Geary's C coefficient was calculated for the 351 plots that indicated significant inter-tree spatial dependencies as determined by the CAR model at an alpha level of 0.10.

Figure 35: Proportion of significant p values at an alpha level of 0.10 for the spatial linear model including the covariate total stem height for each planting density (trees per acre) over (a) ages 2 to 5 for ground-line diameter (d), and (b) ages 5 to 11 for diameter at breast height (D)

(a)



(b)



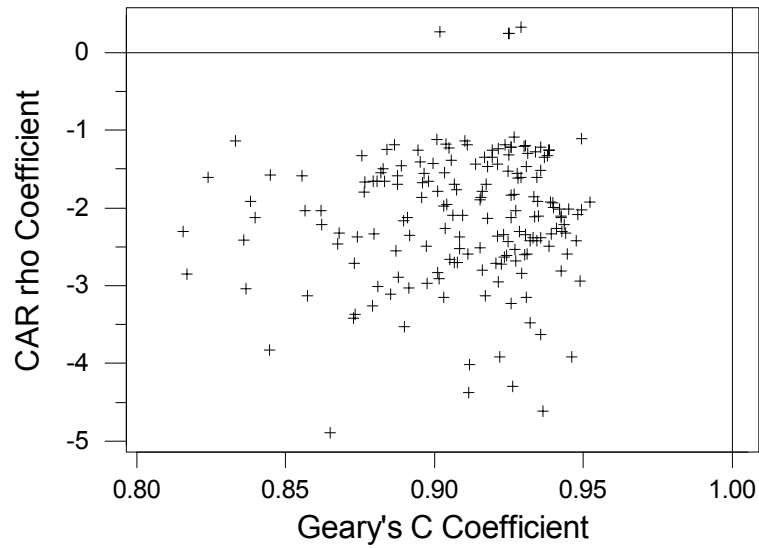
There were 177 plots with a significant Geary's C coefficient (alpha level = 0.10) derived from the spatial linear model residuals. The high number of rejections may indicate that the CAR model is not accounting for all of the inter-tree spatial influences. All of the significant Geary's C coefficients indicated a positive spatial autocorrelation structure among the residuals. A plot of the significant Geary's C coefficients derived from the residuals is presented in Figure 36 with the respective spatial dependency parameters. Figure 36 reveals that almost all of the plots with significant Geary's C coefficients have negative spatial dependency parameter estimates.

The shift in spatial dependency from negative to positive may be due to the CAR model overcorrecting for the inter-tree spatial influences. If the CAR model accounts for a negative autocorrelation structure and the resulting residuals from the spatial linear model are positively autocorrelated, then the CAR model may have overcorrected for the spatial structure in the stand. To address this concern, an investigation into the influence the spatial dependency parameter has on the correlations between observations was conducted.

6.8. Spatial Dependency Parameter Comparison

To determine the relationship that the known neighborhood weight matrix, A , and spatial dependency parameter, ρ , have with the spatial autocorrelation present in a stand, two stands were assigned similar ρ values. The stands examined were stand number 128 with a planting density of 302 TPA (12-by-12 ft spacing), and stand number 65 with a planting density of 2722 TPA (4-by-4 ft spacing). Both stands were established on a

Figure 36: Derived spatial dependency parameter values significant at an alpha level of 0.10 and corresponding significant (alpha level = 0.10) Geary's C coefficients evaluated from the residuals of the spatial linear model, with reference lines indicating no spatial autocorrelation for the CAR model ($\rho = 0$) and Geary's C coefficient ($C = 1$)



square spacing and were devoid of mortality at age 8, the stand age considered in this analysis. The neighborhood weight matrix employed was given by equation (37), where the weights are the inverse of the distance to neighboring stems.

The solution space for ρ is limited by the eigenvalues of the neighborhood weight matrix, hence the two stands may have different ranges for ρ due to the disparity in the weight matrices. Stand number 128 had a minimum eigenvalue of -0.1318 and a maximum of 1.4617 . The minimum and maximum eigenvalues for stand number 65 were -0.3954 and 4.3851 , respectively. Since zero is included in the range of possible eigenvalues for both stands, the ρ parameters are bounded by $\frac{1}{\lambda_{\min}} < \rho < \frac{1}{\lambda_{\max}}$. The range of possible spatial dependency parameters was $[-7.5869, 0.6841]$ and $[-2.5290, 0.2280]$ for stand numbers 128 and 65, respectively.

With no mortality present, each stand had 49 stem locations that were used to construct the neighborhood weight matrix where, as previously defined, all diagonal elements ($i = j$) were equal to zero and all off-diagonal elements ($i \neq j$) were equal to the inverse of the Euclidean distance between locations i and j . A spatial dependency parameter was chosen and a portion of the variance-covariance matrix was calculated from $(I - \rho A)^{-1} M$, resulting in a matrix with dimensions 49×49 . To minimize confounding effects the conditional homoskedasticity assumption was maintained and σ^2 was set equal to one ($M = \sigma^2 I = I$). The off-diagonal elements of the resulting matrix were the covariances between the i^{th} and j^{th} locations ($\text{Cov}(s_i, s_j)$), and the diagonal elements were the variances at each spatial location ($\text{Var}(s_i)$).

The linear correlation between each distinct spatial pair was calculated from:

$$\rho_{s_i, s_j} = \frac{Cov(s_i, s_j)}{\sqrt{Var(s_i) \cdot Var(s_j)}}, \quad (57)$$

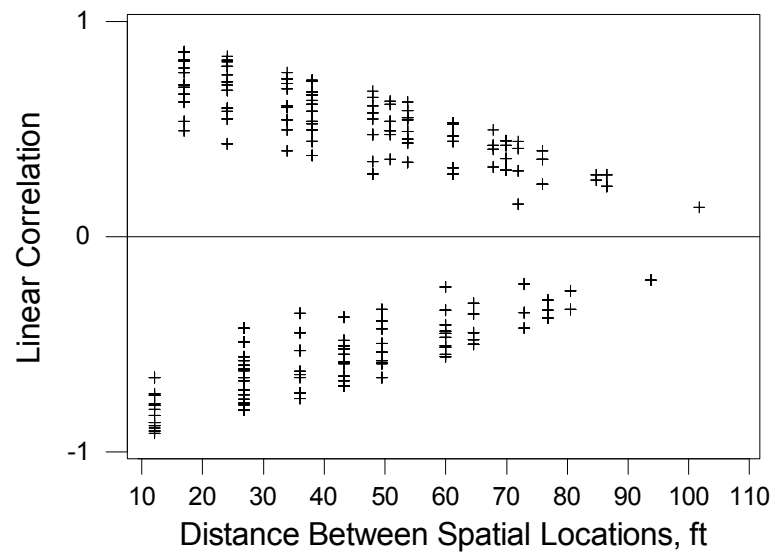
for every $s_{i,j}$ where $i \neq j$. For each example stand there were 1,176 linear correlations derived from all of the distinct spatial pairs. The linear correlations were then plotted against the distance of separation between the i^{th} and j^{th} locations to elucidate the underlying magnitude and direction of any spatial autocorrelation implied from the employed spatial dependency parameter.

Figure 37 (a-b) illustrates the linear correlations derived from spatial dependency parameters near the lower bound for each stand. The results indicate a very strong wave function, oscillating from negative to positive with an increase in the distance of separation. The magnitude of the oscillation present in Figure 37 (a-b) decreases to zero as the distance between two spatial locations increases. When the spatial dependency parameters were close to the upper bound for each example stand, strong positive linear correlations were present (Figure 38 (a-b)), indicating a strong positive autocorrelation between spatial locations. The strength of the positive linear correlation between spatial locations decreased with an increase in the distance of separation between locations.

Graphs of the linear correlations were constructed for identical spatial dependency parameter values for both example stands at intermediate positive and negative ρ values. Figure 39 (a-b) demonstrates that for a spatial dependency parameter value of -1.0 , both stands exhibit negative linear correlations, though these correlations do not oscillate between negative and positive linear correlations as was noted in Figure 37 (a-b). The trends in the linear correlations for a spatial dependency parameter value of 0.10 are

Figure 37: Linear correlations derived from spatial dependency parameters (ρ) near the lower bound of possible values, with a zero reference line, for (a) stand number 128, a 12-by-12 ft spacing, ρ value of -7.57 , and (b) stand number 65, a 4-by-4 ft spacing, ρ value of -2.51

(a)



(b)

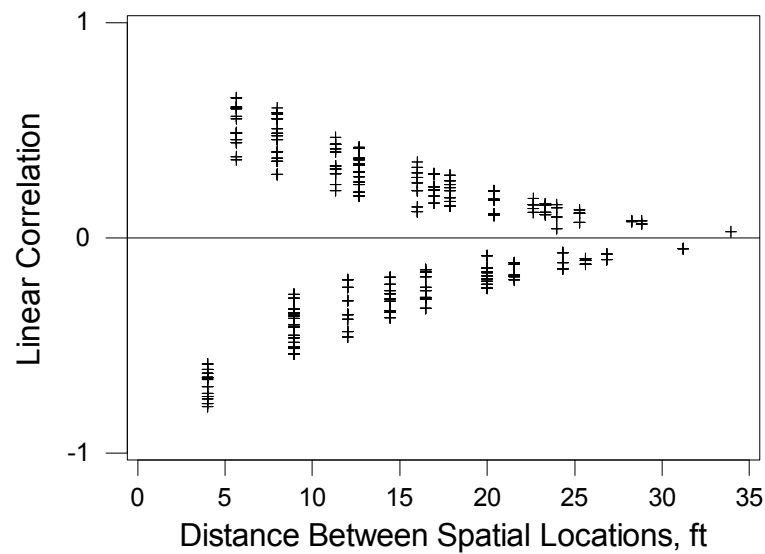
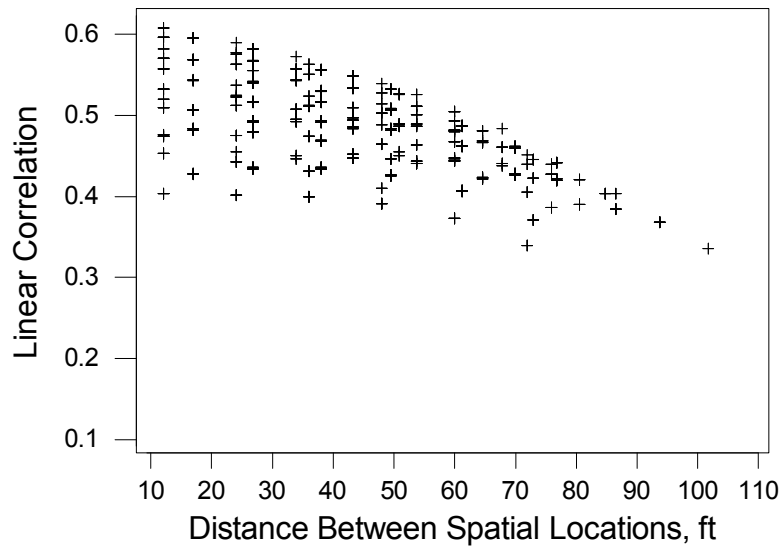


Figure 38: Linear correlations derived from spatial dependency parameters (ρ) near the upper bound of possible values, for (a) stand number 128, a 12-by-12 ft spacing, ρ value of 0.67, and (b) stand number 65, a 4-by-4 ft spacing, ρ value of 0.21

(a)



(b)

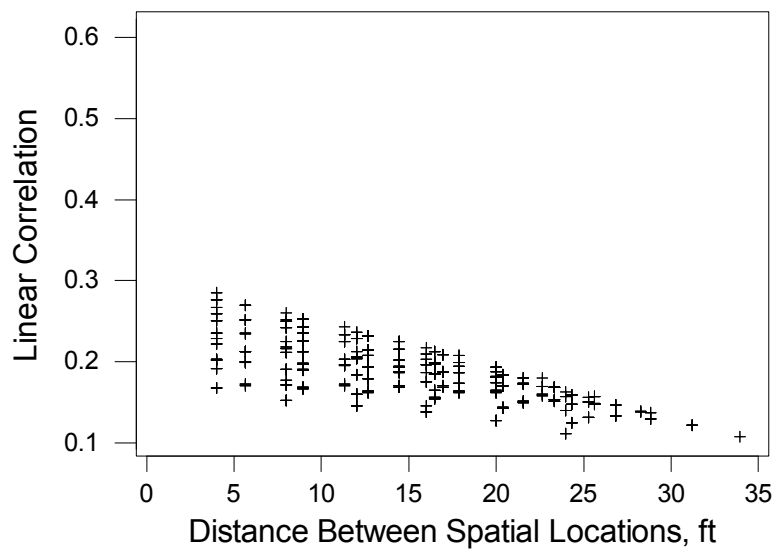
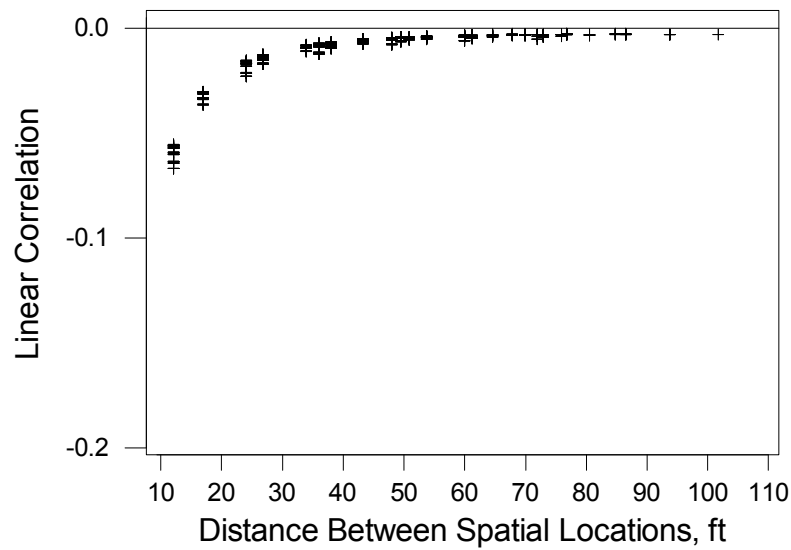
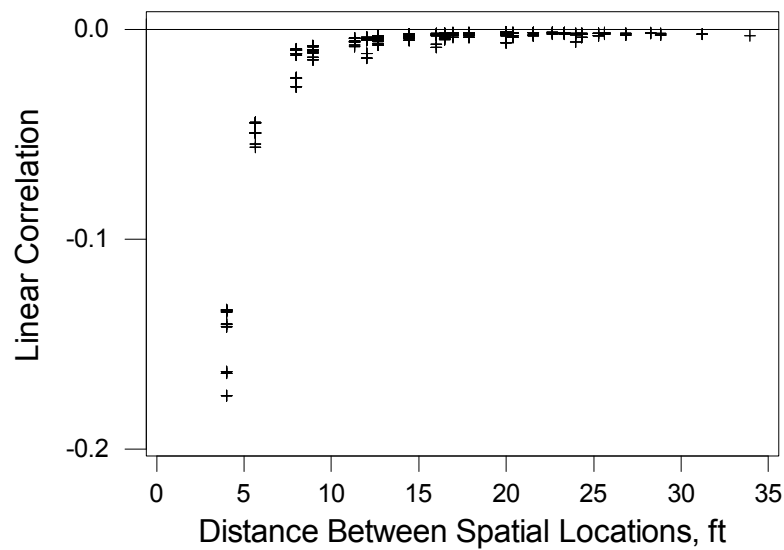


Figure 39: Linear correlations derived from spatial dependency parameters equal to -1.0 , with a zero reference line, for (a) stand number 128, a 12-by-12 ft spacing, and (b) stand number 65, a 4-by-4 ft spacing

(a)



(b)

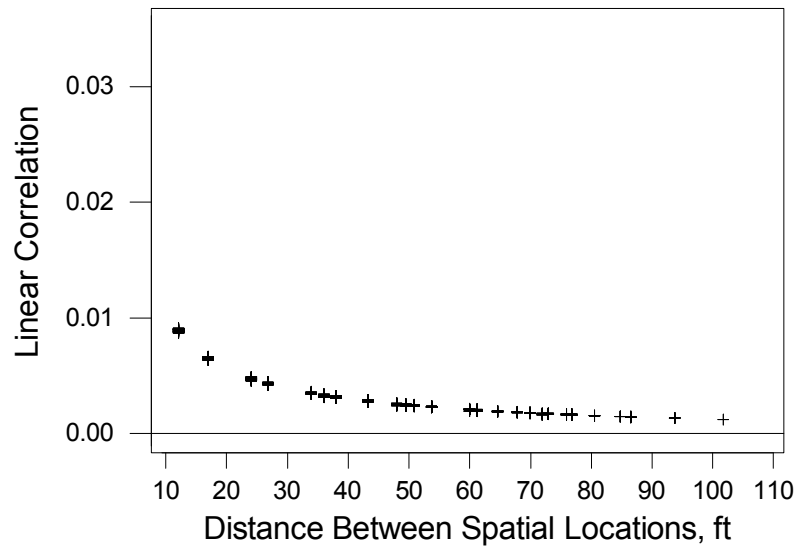


presented in Figure 40 (a-b) for both example stands. All of the derived correlations in Figure 40 (a-b) are positive, though close to zero. In both Figure 39 (a-b) and Figure 40 (a-b), the linear correlations quickly asymptote to zero with an increase in the distance between spatial locations, indicating a weak spatial relationship in each stand for the given spatial dependency parameter. For identical spatial dependency parameter values, the strength of the correlations present in the stands differed. The derived linear correlations should be viewed with respect to the relative size of the spatial dependency parameter within the range of possible values. In Figure 39 (a-b) and Figure 40 (a-b) the spatial dependency parameter values applied to the 12-by-12 ft spacing were more centrally located within the bounds of allowable values when compared to the 4-by-4 ft spacing, possibly accounting for some of the differences between the derived linear correlations.

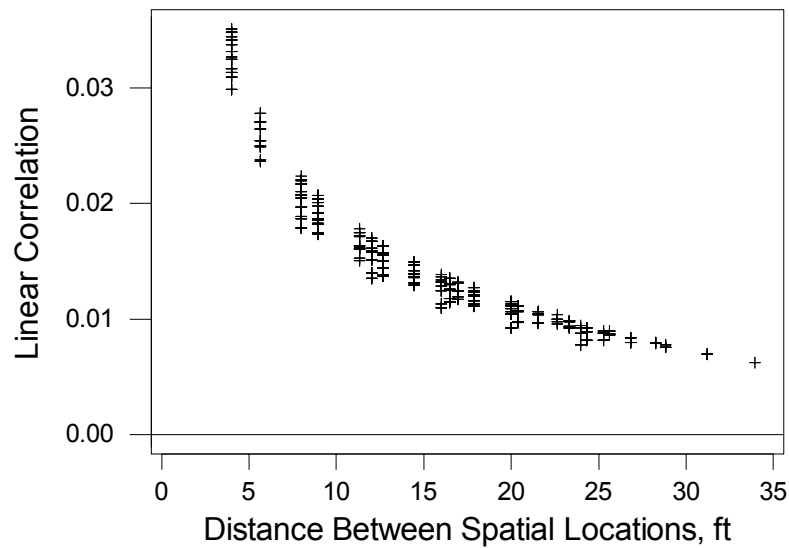
The comparisons conducted evaluated the impact specific spatial dependency parameters have on the autocorrelation within a stand, and can only be adequately interpreted through taking into consideration the bounds of the solution space for the possible ρ values. The bounds were derived from the eigenvalues of the neighborhood weight matrix and will vary if the spatial locations within a stand are altered (e.g. mortality). No actual diameter measurements were used to compare the effects of varying the spatial dependency parameter values, only the spatial locations of the stems in a stand and the neighborhood weight matrix was required. The spatial dependency parameter values considered near the upper and lower bounds for each stand indicated strong autocorrelation in the spatial locations. The comparisons made at identical values

Figure 40: Linear correlations derived from spatial dependency parameters equal to 0.10, with a zero reference line, for (a) stand number 128, a 12-by-12 ft spacing, and (b) stand number 65, a 4-by-4 ft spacing

(a)



(b)



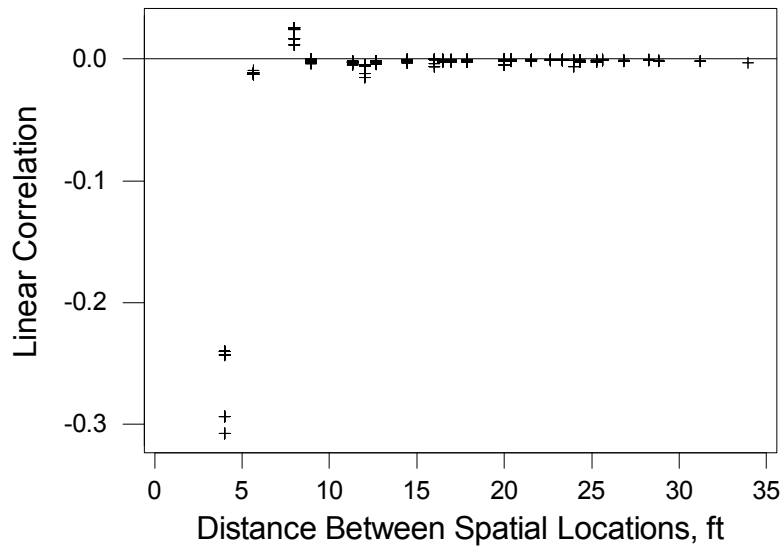
that were more centralized within the bounds did not reveal any strong autocorrelation among spatial locations.

To ascertain at what level the spatial dependency parameter invokes a wave function, linear correlations between distinct spatial pairs were derived for various levels of the spatial dependency parameter. In evaluating the spatial dependency parameters on the 4-by-4 ft and 12-by-12 ft spacings, it was determined that any comparisons made between stands with different spacings and weight matrices should take into account the relative location of the spatial dependency parameter with respect to the bounds on the solution space for ρ .

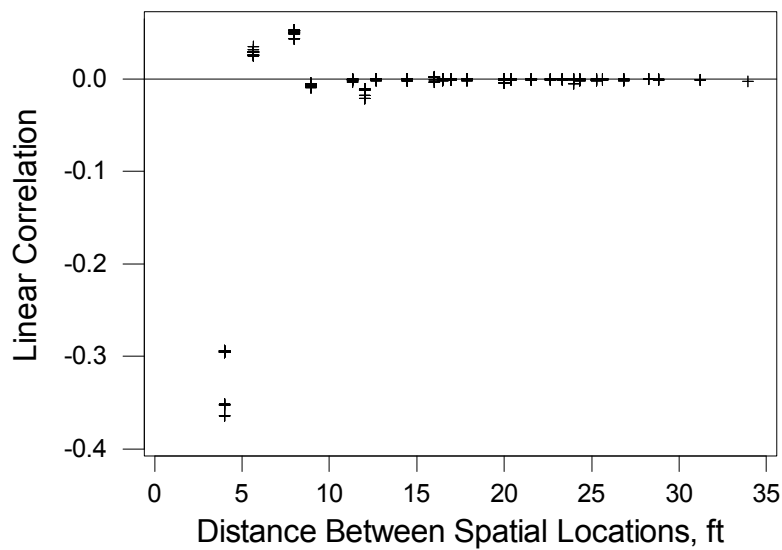
To determine at what level the spatial dependency parameter is associated with the wave function, the example stand presented earlier with a spacing of 4-by-4 ft was employed. In Figure 37 (b) the spatial dependency parameter (-2.51) used to estimate the linear correlations was very close to the lower bound for ρ (-2.5290), and a strong wave function was detected. Figure 39 (b) revealed a negative spatial autocorrelation among the locations that quickly asymptoted to zero. The spatial dependency parameter in Figure 39 (b) was approximately 40% of the lower bound ($\frac{-1.0}{-2.5290}$) for ρ . Linear correlations were derived for moderate ρ values at approximately 70 and 80% of the lower bound (Figure 41 (a-b)). The plots exhibited an initially oscillating trend in the linear correlations between locations at shorter distances of separation. The oscillation quickly asymptoted to zero with an increase in the distance of separation, indicating that only ρ values near the lower bound will result in strong oscillating trends between negative and positive spatial autocorrelation.

Figure 41: Linear correlations derived from spatial dependency parameters (ρ) near the lower bound of possible values, with a zero reference line, for stand number 65, a 4-by-4 ft spacing, with ρ values of (a) -1.7703 and (b) -2.0232

(a)



(b)

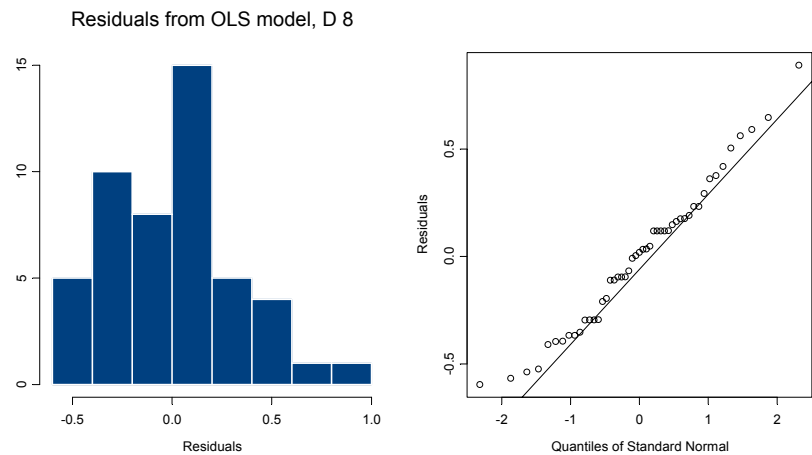


The CAR analysis in Section 6.7 indicated 351 plots with significant spatial dependency parameters. Geary's C coefficient was calculated on the residuals from the estimated spatial linear model for plots with significant spatial trends; a positive spatial autocorrelation was found among 177 plots. The previous analysis determined the effect the spatial dependency parameter had on the linear correlations between data values separated by various distances. Only those plots with spatial dependency parameters near the lower bound resulted in a strong oscillating function between negative and positive spatial autocorrelations. The presence of significant positive spatial autocorrelations among the residuals from the spatial linear model may be a result of the CAR model overcorrecting for negative autocorrelation. The following example using stand number 65 and the diameter at breast height measurements at age 8 will aid in elucidating the excessive correction for spatial dependence by the CAR model.

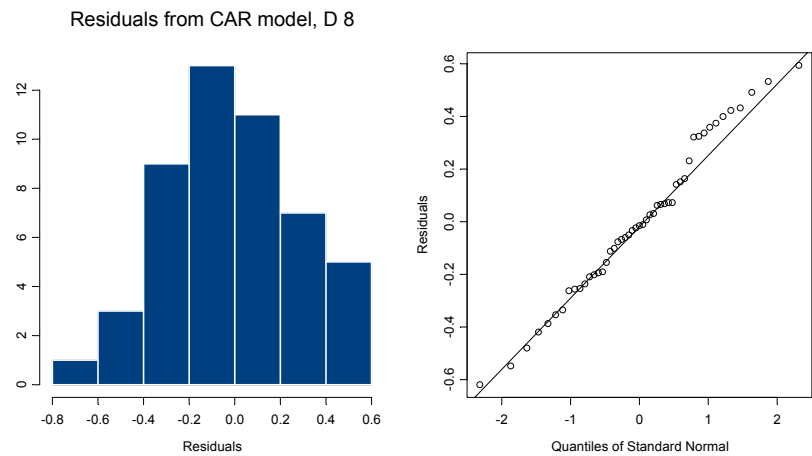
Using the spatial linear model given by equation (51) with a covariate of total stem height, the CAR analysis indicated that stand number 65 had a significant negative spatial dependency parameter ($p\text{-value} = 0.0442$). From the ordinary least squares (OLS) analysis, Geary's C coefficient indicated significant negative spatial autocorrelation (permutation $p\text{-value} = 0.9584$) among the residuals. Histograms and qq plots of the residuals are presented in Figure 42 (a-b) for the OLS and CAR models. Further analysis employing Geary's C coefficient was performed on the residuals from the CAR spatial linear model. The resulting Geary's C coefficient indicated significant positive spatial autocorrelation (permutation $p\text{-value} = 0.034$) among the residuals from the spatial linear model.

Figure 42: Histograms and quantile-quantile plots of the residuals for the (a) OLS and (b) CAR models fitted to diameter at breast height (D) at age 8

(a)



(b)



The presence of positive spatial autocorrelation among the residuals from the CAR spatial linear model may be due to an overcorrection with respect to the estimated spatial dependency parameter. Figure 43 illustrates the empirical semivariograms for the residuals from the OLS and CAR models. An oscillating trend between negative and positive autocorrelation can be viewed in the empirical semivariogram for the OLS residuals, while a positive autocorrelation is visible in the empirical semivariogram for residuals from the CAR analysis. The derived spatial dependency parameter had a ρ value of -1.4532 , approximately 57% of the lower bound. Linear correlations between spatial locations were derived for stand number 65 using the estimated ρ value of -1.4532 (Figure 44). The trend visible in Figure 44 indicates predominately negative linear correlations between points separated by close distances. Further, the trend in the linear correlations asymptotes to zero at further distances, and does not exhibit a strong oscillating trend. The CAR analysis detected and accounted for a spatial autocorrelation in the data, but may have overcorrected for a negative spatial autocorrelation structure. Analyses performed on the residuals of the spatial linear model detected a positive spatial autocorrelation structure in the residuals.

Figure 36 presented plots where the CAR model had a significant spatial dependency parameter, but the residuals from the CAR model still contained a significant spatial structure. The results indicated that the CAR model might overcorrect for a negative spatial dependency parameter in the model. Care should be taken when implementing the CAR model; the residuals should be checked to determine if any strong spatial structure is present.

Figure 43: Semivariograms for the OLS and CAR residuals from modeling diameter at breast height (D) at age 8

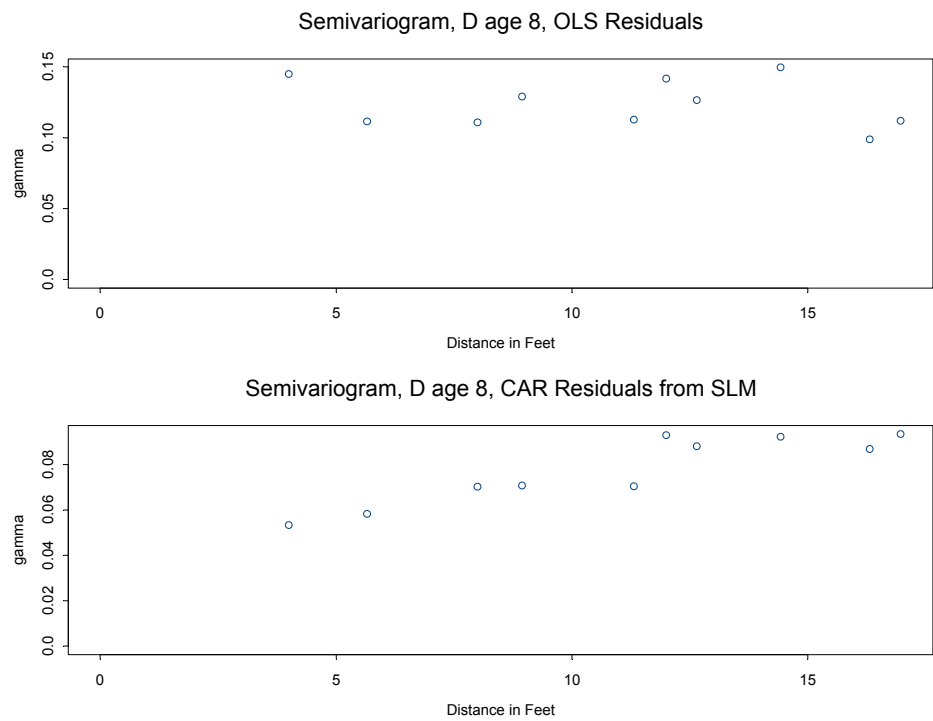
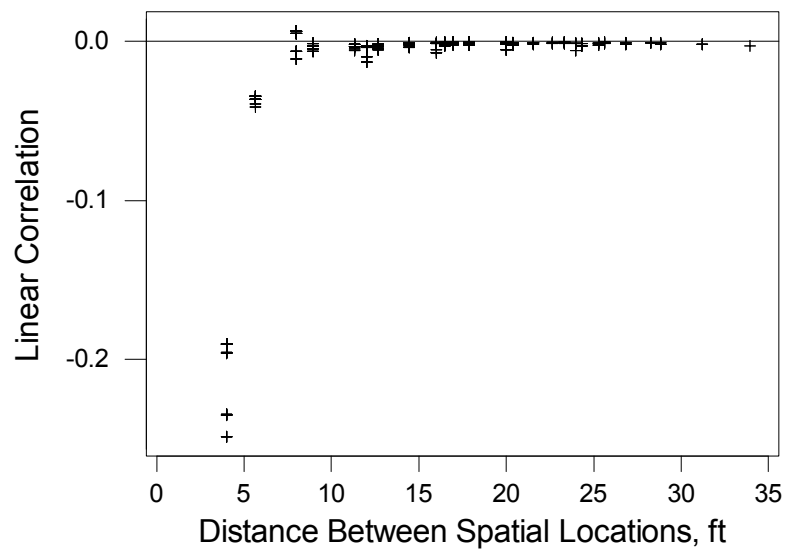


Figure 44: Linear correlations derived from the estimated spatial dependency parameter (-1.4532), with a zero reference line, for stand number 65, a 4-by-4 ft spacing at age 8



6.9. Spatial Summary

Though earlier analysis conducted using the Weibull distribution (Sections 4 and 5) assumed the data to be independent observations, an investigation into the presence of inter-tree spatial relationships was conducted. Using the methods described herein and various spatial data analysis tools (empirical semivariogram, Moran's I coefficient, Geary's C coefficient, and the CAR model), inter-tree relationships were estimated. Accounting for the overall effect of the spatial dependency between stems will aid in interpreting stand dynamics of juvenile loblolly pine trees. The spatial relationships that manifest themselves in the significant spatial dependency estimates are microsite and competition effects. Any genetic differences within the plots are not accounted for and are assumed to be negligible (all planted seedlings came from the same growing stock raised in the same nursery).

The empirical semivariograms were derived for both the raw diameter data and residuals from a fitted model to the diameter data. Neither application of the semivariograms revealed any strong trends in the covariance between observations and the distances separating stems on a plot. The near horizontal trend in most empirical semivariogram plots indicates that a nugget model may be appropriate and no adjustment for inter-tree spatial dependencies on a plot necessary for the juvenile loblolly pine data considered in this research. A nugget model would not take into account the distance between stems, but would be a constant correction factor to account for the white noise (measurement error or micro-scale variation) in the data.

Moran's I coefficients were derived to estimate the amount of spatial autocorrelation present within the residuals from a fitted model to the diameter

measurements on each plot. The null hypothesis of no spatial autocorrelation was rejected for 282 of 2112 plots over all ages. A decreasing trend in the average proportion of rejections of the null hypothesis with age was initially present, but leveled off by age 5. No trend between the proportion of rejections and planting density was noted.

The derived Geary's C coefficients from the residuals of a fitted model to the diameter measurements indicated significant spatial autocorrelation in 268 of 2112 plots across all ages. No strong trend was noted between the proportion of rejections and planting density. Age was initially strongly related to the average proportion of rejections with a negative trend, but this trend leveled off by age 5; these results concur with those from Moran's I index.

The conditional autoregressive spatial linear model including the covariate total stem height discerned significant spatial influences in 351 plots. A trend was noted between the presence of spatial influences and age; the trend initially increased but leveled off by age 9. No trends were detected with respect to spatial dependencies and either planting density or planting rectangularity. The residuals from the spatial linear model were evaluated with Geary's C coefficient to ascertain if any spatial dependencies were present. The results indicated that in some cases the CAR model overcorrected for a negative spatial autocorrelation structure, and a positive spatial autocorrelation was imposed on the plot.

For the spatial analysis tools employed, the null hypothesis of no spatial autocorrelation was tested using an alpha-level of 0.10 for the 2112 independent tests conducted for each spatial index. The high number of independent tests increases the

chance of at least one type I error occurring. Though one cannot say with certainty, the overall rejection rate may be overestimated.

The variance-covariance matrix of the conditional autoregressive model was evaluated for two example stand spacings (4-by-4 and 12-by-12 ft spacings) to compare the effects the spatial dependency parameter value has on the spatial autocorrelation structure. Values that were near the upper bounds of permissible spatial dependency parameter values indicated strong positive spatial autocorrelations, while values that were near the lower bounds demonstrated a wave function between negative and positive autocorrelations. Those values that were further from the solution space bounds did not indicate the presence of strong spatial autocorrelation among the spatial locations. Therefore, any comparisons of spatial dependency parameter estimates across different neighborhood weight matrices should be examined as a function of the relative location within the permissible bounds for the spatial dependency parameter.

Modeling was conducted to determine if the estimates from the various spatial indices derived for each plot would assist in explaining the variation in the scale and shape parameter estimates derived previously from a two-parameter Weibull distribution. The estimates of spatial influence (Moran's I, Geary's C, and CAR spatial dependencies) did not adequately describe the variation in the scale and shape parameter estimates. Even modeling conducted on only those plots having significant indices did not produce sufficient R-square values. Further modeling was performed on the estimates of spatial dependency to see if they could be adequately modeled as a function of age, planting density, and relative spacing. Modeling of Moran's I and Geary's C coefficients was unsuccessful, with R-square values less than 0.14. Modeling of the CAR derived spatial

dependency parameters achieved R-square values of approximately 35%, but due to the varying nature of the neighborhood matrices, the utility of any model was questionable.

7. SPATIAL STRUCTURE SIMULATION

The use of simulation techniques to generate tree characteristics has been in use for many years. The main goal of a simulation is to adequately describe a situation that could naturally occur given a set of input parameters. Assuming that the simulator adequately represents the empirical data from which it was estimated, the simulator could then be used to evaluate the impact of specific changes in the system. Simulators can also be used to develop projections for future outcomes from the system. In forestry, a common simulator output is a set of breast height diameters derived for a specific set of inputs (e.g. stand age, planting density, site quality). Simulation of stand diameter distributions facilitates inventory procedures, operational planning, and forecasting.

A common technique used to simulate stand diameters *randomly* draws a sample of n diameters from a diameter distribution function estimated for a given set of stand-level inputs. These diameters can then be used to obtain summary statistics for the resulting stand. If the simulated diameters are to be used to predict future stand characteristics, a methodology that accounts for growth and competition processes occurring within the stand is needed. A spatial representation of the generated diameters is necessary when using distance dependent methods to account for growth and competition processes. In one individual-stem growth and yield simulator for loblolly pine, PTAEDA2, the diameters are randomly assigned to specified spatial locations (Burkhart et al. 1987). The random assignment of diameters to spatial locations does not take into account any micro-site variability or distance-based competition effects. While these effects may not be influential if the stand is generated at planting (or shortly thereafter), there may be a significant spatial autocorrelation among the stems in a stand

generated at later ages. To accommodate this, a methodology that accounts for inter-tree spatial dependencies at the time of stand simulation was developed.

7.1. Stand Simulation

Simulation of individual stem diameters that have an imposed spatial autocorrelation structure will account for the initial inter-tree dependencies in the stand. The stands in this research were analyzed as lattice data. To simulate spatial lattice data the spatial locations (s_i) must be known a priori. In conjunction with the spatial locations, an estimate of the spatial dependency parameter, ρ , and an estimate of the variance, σ^2 , are needed for the stand to be simulated. The spatial dependency parameter and variance estimates come from previous spatial analysis on stands with similar characteristics. The spatial process is assumed to follow the detrended distribution below.

$$U \sim N(0, \Sigma) \quad (58)$$

The spatial process (equation (58)) is assumed to be Normally distributed; a Normal distribution was shown to adequately characterize the juvenile diameter data considered in this research (Section 6.7). A conditional autoregressive model given by equation (45) characterizes the variance-covariance matrix for the spatial process. The previously estimated $\hat{\rho}$ and $\hat{\sigma}^2$ values for the juvenile loblolly pine stands are employed in the stand simulation. The Cholesky decomposition method (Cressie 1993) allows the variance-covariance matrix to be decomposed as:

$$\Sigma = LL' \quad (59)$$

where L is the $n \times n$ lower triangular matrix. From this decomposition the data can be simulated so that the spatial process follows equation (58). The following equation enables the simulation of the detrended spatially dependent vector:

$$U = L\varepsilon \quad (60)$$

where $\varepsilon = (\varepsilon(s_1), \dots, \varepsilon(s_n))'$, a vector of uncorrelated random variables distributed by:

$$\varepsilon \sim N(0,1). \quad (61)$$

The realizations of the diameters can then be generated from the following equation (Haining 1990):

$$Z(s_i) = \mu_i + u_i = \mu_i + \sum_{j=1}^n l_{ij} \varepsilon(s_j). \quad (62)$$

From equation (62) a spatially dependent diameter is derived for each spatial location. The newly simulated stand should have spatial dependencies comparable to stands with similar characteristics. Over many iterations, the average stand spatial characteristics will approach those characteristics applied in the simulation techniques.

Other methods are available to simulate spatially dependent variates that do not require an assumption of Normality. Simulated annealing is an optimization technique that can be used to solve combinatorial optimization problems (Laarhoven and Aarts 1987). The simulated annealing technique takes an initial set of data at specific locations and perturbs them until a target optimization point is achieved through an objective function. The objective function measures the difference between the current realization and the target point (Goovaerts 1997). The perturbation would then switch two random points and either accept or reject the swap based on a decision rule. The process is then

repeated until the target optimization point is achieved (possibly within a certain tolerance level) or the objective function fails to be significantly reduced.

7.2. Simulation Application

The Cholesky decomposition method was used to simulate realizations of a stand with spatially dependent diameters. An example follows using stand number 145, one of the example stands presented in the Appendices. Stand number 145 was chosen because of the highly significant spatial dependency parameter (p-value = 0.0158) in the spatial linear model.

The focus of the simulation in a lattice structure is on the data values at specific locations and not on the actual placement of the locations. Given that the spatial locations for stand number 145 are assumed known, these same locations were used for the simulated data sets. The planting density for stand number 145 is 2722 TPA with a planting rectangularity of 4-by-4 ft (1:1). The previously derived CAR model for the example stand was employed in the simulator. The conditional homoskedasticity assumption was maintained such that $M = \sigma^2 I$. Since the same spatial locations were assigned to the simulated data sets, the neighborhood matrix A is unchanged. The variance-covariance matrix used in the Cholesky decomposition was:

$$\Sigma_{sim} = (I - \hat{\rho}_{sim} A)^{-1} \hat{\sigma}_{sim}^2 \quad (63)$$

where Σ_{sim} = the variance-covariance matrix for the simulated data set

$\hat{\rho}_{sim} = -1.474545$, the estimated spatial dependency parameter

$\hat{\sigma}_{sim}^2 = 0.7882264$, the estimated variance

The Cholesky decomposition given by equation (59) was calculated in accordance with the estimated spatial dependency parameter. The uncorrelated random variables (ε) were generated from a $N(0,1)$ distribution. The random vector of spatially dependent simulated diameter observations results from equation (62), where the mean trend was added to the generated values to obtain the diameter realizations.

Comparisons were made between a simulated spatially dependent data set, the median of 100 realizations of the simulator, and the original data set to ascertain the effectiveness of the simulator. The spatial linear mean trend model was fit to the original data and the data from random realizations of the simulator. Table 16 gives the estimated spatial dependency parameter ($\hat{\rho}$), the p-value for the likelihood ratio test, and the estimated intercept for each respective data set.

The relative closeness of the significant spatial dependency parameters (alpha level = 0.10) in both the original and simulated data sets (Table 16) indicates that the stands have similar spatial dependencies. The nearness of the estimated intercept terms for the spatial linear models suggests that the diameter data have similar mean trends. For the original data and one realization of the simulator, a two-sample t-test failed to reject the null hypothesis that the difference in the means equals zero (p-value = 0.8103). A two-sample Kolmogorov-Smirnov goodness-of-fit test performed on the original and simulated data set failed to reject the null hypothesis that there are no differences in the two cdfs (p-value = 0.9602). Frequency histograms of the diameters for the original data

Table 16: Spatial index comparisons of the original data (breast height diameters at age 8) from stand number 145, a random realization of the spatial simulator, and the median of 100 random realizations of the spatial simulator

	Original Data, Stand Number 145	One Random Realization of Simulated Data	Median of 100 Random Realizations of Simulated Data
$\hat{\rho}$	-1.475	-1.998	-1.423
Likelihood ratio test p-value	0.0158	0.0007	0.0193
Intercept	2.6683	2.7377	2.6629
Moran's I Coefficient	-0.047	-0.057	-0.047
p-value for I	0.961	0.998	0.958
Geary's C Coefficient	1.055	1.043	1.030
p-value for C	0.978	0.935	0.822

Where $\hat{\rho}$ = The estimated spatial dependency parameter

set and one realization of the spatial simulator are presented in Figure 45 and show similar diameter distributions. To facilitate the visual comparison between the original diameter data set and a realization of the simulator, the spatial arrangements of the diameters on the lattice are given (Figure 46 (a-b)).

The spatial location of stems in the simulated data sets were the same as that of the original data set, and any mortality that was present is represented in the same location (e.g. the upper right hand corner of Figure 46 (a-b)). The diameter distributions are comparable for the data sets, but, as expected, the diameter realizations at specific spatial locations differ. To further test the likeness of any spatial dependencies between the original data, a single realization of the simulator, and the median of 100 realizations of the simulator, several methods were evaluated.

Empirical semivariograms were computed for the original ($\hat{\gamma}(h)_{orig}$) and simulated ($\hat{\gamma}(h)_{sim}$) data sets. Since the same spatial locations were used in the simulated data sets, the distance bins and number of distinct pairs in each bin is the same for all of the empirical semivariograms. To further investigate the similarities of the data sets, Moran's I and Geary's C coefficients were computed for the original and simulated data sets. The resulting coefficients and p-values are given in Table 16. The estimated coefficients for both Moran's I and Geary's C indices indicate a negative spatial autocorrelation present in both the original and simulated stands. The p-values for Moran's I coefficient, determined by the randomization method, indicated significant negative spatial autocorrelations. The median Moran's I and Geary's C coefficients from the 100 simulations exhibit a high degree of agreement with the results from the original

Figure 45: Frequency histogram for diameters (D at age 8) from the original data from stand number 145 and a random realization from the spatial simulation

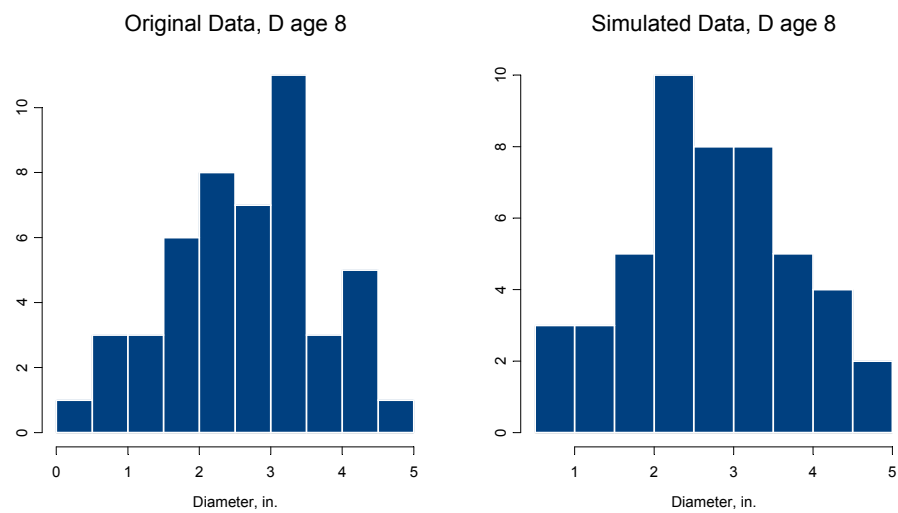
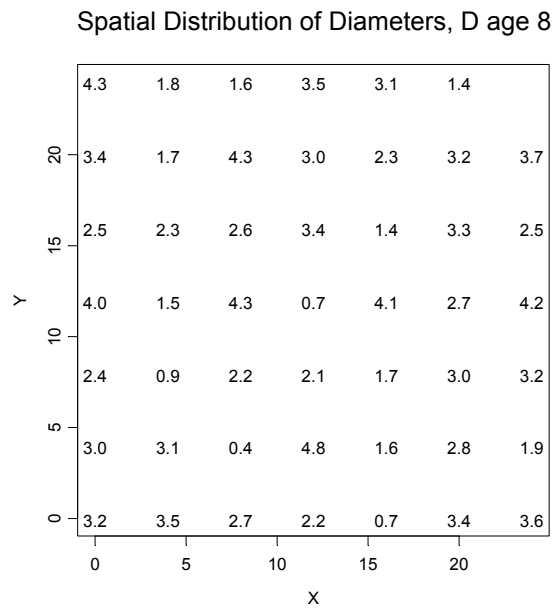
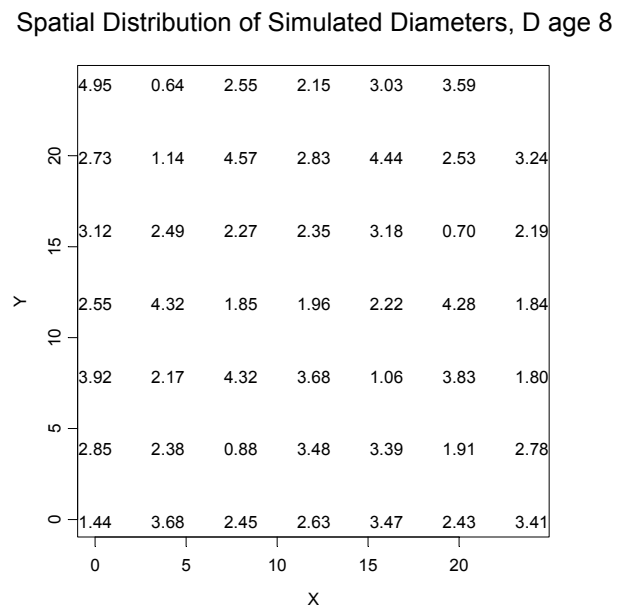


Figure 46: Spatial distribution of diameters for (a) the original diameter at breast height data (D at age 8) from stand number 145, and (b) a random realization from the spatial simulation

(a)



(b)



data. The agreement in these indices exemplifies, as expected, that over the long run the spatial characteristics of the original data can be replicated.

Though the original data can be replicated through averaging many simulations, this may not be beneficial if the objective is to predict future outcomes of a system. The randomness invoked in a single simulation assists in capturing some of the possible occurrences that could take place. Therefore, only a single realization of the simulator is necessary to obtain an adequate representation of the system.

8. SUMMARY AND RECOMMENDATIONS

A two-parameter Weibull distribution was fitted to the juvenile stand data collected in a loblolly pine spacing trial, employing maximum likelihood estimation techniques. Goodness-of-fit tests did not find many significant differences between the empirical distribution function and the derived two-parameter Weibull function. The Anderson-Darling test rejected the null hypothesis (0.20 alpha level for all goodness-of-fit tests) that the data come from the specified distribution in 0.8% of the stands, while the Cramer-von Mises test rejected 0.7%, and the Kolmogorov-Smirnov test rejected 3.1% of the fitted distributions. The lack of significant differences indicates that the two-parameter Weibull distribution adequately characterizes the empirical diameter distributions of juvenile loblolly pine.

The Normal distribution was also fitted to the juvenile diameter data considered in this research. A Normality assumption was required to perform analyses utilizing a conditional autoregressive model and to enable the simulation of spatially dependent data employing the Cholesky decomposition method. Goodness-of-fit statistics indicated that the Normal distribution sufficiently characterized the juvenile diameter data, but an increasing trend in the rejection of the Normal was present with age; the Normal distribution may not provide an appropriate fit to the diameter data at later ages. Therefore, to avoid switching distributions at later ages, the two-parameter Weibull distribution was fitted to enable diameter distribution forecasting of juvenile loblolly pine.

The parameter estimates from the fitted two-parameter Weibull diameter distributions were compared for several stand-level characteristics to detect any trends

present. A rectangularity effect was not discerned among the derived two-parameter Weibull estimates, skewness coefficients, or kurtosis coefficients for plots with the same planting density of 907 trees per acre; these findings corroborate the results given by Sharma et al. (in review). The planting rectangularity of the plots compared was 1:3 (4-by-12 ft) and 1:1.33 (6-by-8 ft). The lack of any significant differences between the different planting rectangularities indicates that stems may be planted over a range of rectangularities without any noticeable difference in their diameter distributions in the juvenile stage, increasing the amount of operational flexibility for planting regimes.

A significant temporal effect was found for the estimated scale parameter from the two-parameter Weibull diameter distribution over ages 5 to 11. The relationship between stand age and the estimated scale parameter was positive, indicating that an increase in age corresponds to an increase in the range of the diameters in the estimated distribution for a stand. A significant planting density effect was noted at each age with the scale parameter estimates. The planting density effect was negatively associated with the estimated scale parameter, implying that with an increase in planting density there was a decrease in the estimated scale parameter.

A significant temporal effect was detected for the shape parameter estimates over ages 5 to 11 for stands with planting densities ranging from 302 to 907 trees per acre. Plots with higher planting densities, from 1210 to 2722 trees per acre, did not have significantly different shape parameter estimates over ages 5 to 11. Across all planting densities at age 5, no significant difference in the shape parameter estimates was noted, implying that there are no differences in the skewness of the estimated diameter distributions. At age 6, and continuing through age 11, a significant difference in the

estimated shape parameter was found across all planting densities. The trend between the shape parameter and planting density was negative; suggesting an increase in planting density was associated with a decrease in the shape parameter estimate. Plots that have lower planting densities would then be more likely to have diameter distributions that are more left skewed.

Stand relative spacing was computed for each plot at each remeasurement period to ascertain stand competitiveness levels and detect any trends present that would aid in describing the scale or shape parameter estimates. Relative spacing was employed as a potentially more meaningful non-chronological measure to evaluate stand competition status. The individually derived relative spacings did not elucidate any trends in either the scale or shape parameter estimates that were not discerned by age or planting density for the juvenile loblolly pine data considered in this investigation. The wide dispersion of the relative spacing values over the range of the relative spacing bins used to group similar stands inhibited trend recognition.

A reparameterization, aimed at reducing the amount of parameter-effects nonlinearity, of the two-parameter Weibull distribution was carried out to reduce the bias and variance of the estimated shape parameters. The reparameterization did not have an extensive impact on the dispersion of the shape parameter. The natural variation in the empirical diameter distributions across all of the plots is reflected in the inability of the reparameterization to reduce the dispersion of the shape parameters.

Models were fitted to the estimated scale and shape parameters from the two-parameter Weibull distribution. These models utilize linearized forms of the regressors to improve fit statistics. The regressors employed in the final model were stand age,

planting density (trees per acre), and the average height of the dominant stems (in ft). Other factors were considered in model construction to aid in predicting the scale and shape parameters. Relative spacing did not aid in describing the trends present for the scale or shape parameters. A term representing the current density of the stand did not outperform a planting density term. The model presented to predict the scale parameter, equation (26), has an R-square value of 0.940 and mean square error of 0.006. The shape parameter prediction equation (33) has an R-square value of 0.407 and mean square error of 0.045. These prediction equations enable estimates of two-parameter Weibull diameter distributions for juvenile loblolly pine.

The spatial dependencies between trees in the juvenile loblolly pine stands considered in this research were quantified using multiple spatial analysis tools. Empirical semivariograms were derived for the original diameter data and the residuals from a linear model fitted to the data. The empirical semivariograms did not reveal any strong spatial trends as a function of distance between pairs of observations. No lasting positive or negative spatial trend was noted for either data form, for both the ground-line diameter data from ages 2 to 5, and the breast height diameter data from ages 5 to 11. The lack of visible spatial trends indicates that a nugget model would be adequate to act as a correction for the white noise process present in the data.

Moran's I coefficient was employed to detect the magnitude and direction of any spatial trends in the residuals from a linear model fitted to the diameter data. Significant spatial autocorrelations (0.10 alpha level for all spatial autocorrelation tests) were found in 13.4% of the plots over all ages. Rejection of the null hypothesis of no spatial autocorrelation was initially strongly related to age with a negative trend for Moran's I

coefficient, but this trend leveled off by age 5. Though the overall number of rejections initially decreased with age, there was no trend with the planting density of the plots. Additionally, Geary's C coefficient was used to evaluate spatial autocorrelation in the residuals from a fitted model to the diameter data. Over all ages considered in this study, Geary's C coefficient indicated a significant spatial autocorrelation in 12.7% of the plots. Age was initially negatively related to the number of rejections of the null hypothesis for the juvenile loblolly pine stands in this investigation, but the relationship leveled off by age 5. The planting density was not strongly associated with the occurrence of null hypothesis rejection.

The results of Moran's I and Geary's C coefficients in this investigation are similar to those found in Reed and Burkhardt (1985) and Liu and Burkhardt (1994). Reed and Burkhardt (1985) employed a binary weighting scheme where only trees whose tree area polygons shared common borders were considered neighbors. Moran's I and Geary's C indices were evaluated on stem basal area and were not strongly related to age, site index, or density of loblolly pine stands. Liu and Burkhardt (1994) modeled diameter with a second-order trend surface; the error terms were evaluated for spatial autocorrelation using Moran's I coefficient. A binary weight matrix was employed where only stems in the same column were considered neighbors. In general Liu and Burkhardt (1994) found that the derived Moran's I coefficients went from positively significant to non-significant with an increase in age.

Conditional autoregressive models that specify the variance-covariance matrix given a known neighborhood structure were employed to derive spatial dependency parameters that characterize the presence and level of spatial influence in stands. A

spatial linear model including a covariate of total stem height was evaluated to account for the large-scale variation in the diameter data. Across ages 2 to 5 for ground-line diameters and ages 5 to 11 for breast height diameters, the conditional autoregressive model found significant spatial dependency parameters in 17% of the plots. No trend was discernable between the planting density of a stand and the presence of significant spatial influence. An initial positive trend was noted between stand age and the existence of significant spatial dependencies, but this trend leveled off by age 9. Geary's C coefficient was used to evaluate the presence of any spatial autocorrelation among the residuals from spatial linear models. Of the plots with significant spatial dependency parameters, 50% had significant positive spatial autocorrelations in the residuals. It was determined that this was due to an overcorrection in the conditional autoregressive model for negative autocorrelation.

The effect the spatial dependency parameter value has on the spatial autocorrelation structure was determined for two stands with square spacings (4-by-4 and 12-by-12 ft). Linear correlations between spatial locations were derived from the variance-covariance matrix of the conditional autoregressive model applied to each stand. Strong spatial autocorrelation was exhibited when the spatial dependency parameter values were near the limits of the solution space. When the spatial dependency parameter values were further within the bounds of the solution space, the linear correlations did not reveal any strong spatial autocorrelation among the locations.

In modeling the scale and shape parameter estimates for the two-parameter Weibull distribution, a term that accounted for the amount of spatial autocorrelation present in the stand was non-significant. The estimated Moran's I coefficient, Geary's C

coefficient, and spatial dependency parameter from the conditional autoregressive model did not significantly aid in explaining the variation in the scale or shape parameter estimates.

A simulator was employed to explicitly derive tree diameters in a juvenile forest stand. Each stem was given a specific spatial location and the Cholesky decomposition method was applied to simulate realizations of the diameter data. A spatial dependency was imposed on the simulated stand to more accurately describe the inter-tree correlations. The simulation technique presented derived a spatially dependent diameter for each spatial location of the simulated stand, enabling more spatially accurate representations of distance-dependent diameters within a stand. The utilization of a spatial dependency in stand simulation may further aid stand-level forecasting.

Further research is needed to adequately understand the implications of forecasting with stands where the initial tree sizes are assigned randomly or with an imposed spatial relationship. A stand simulator such as PTAEDA2 (Burkhardt et al. 1987) could be employed to determine the effect the competition processes invoked in the simulator have on the generated stands. Comparisons could include stands that were generated with various spatial structures in the tree sizes and stands that were generated with randomly assigned initial diameters; the effects of any spatial trends present at stand generation could then be evaluated at several points in stand development.

Future work is needed using diameter increment to evaluate spatial trends in forest stands; diameter increment may better represent inter-tree spatial influences within a stand. Work comparing the effects various neighborhood weight matrices have on the

estimates of spatial autocorrelation would demonstrate the consequences of under- and over-specification of neighbors.

Finally, the author extends a note of caution in extrapolating the results presented beyond the range of rectangularities, planting densities, and ages considered in this investigation for juvenile loblolly pine.

LITERATURE CITED

- Amateis, R.L., H.E. Burkhart, and S.M. Zedaker. 1988. Experimental design and early analyses for a set of loblolly pine spacing trials. *USDA For Serv Gen Tech Rep NC-120 North Cent For Exp Stn.* pp. 1058-1065.
- Avery, T.E. and H.E. Burkhart. 2002. *Forest Measurements*. 5th ed. McGraw-Hill, N.Y.
- Bailey, R.L. 1980. Individual tree growth derived from diameter distribution models. *Forest Science* 26:626-632.
- Bailey, R.L. and T.R. Dell. 1973. Quantifying diameter distributions with the Weibull function. *Forest Science* 19:97-104.
- Bennett, F.A. and J.L. Clutter. 1968. Multiple-product yield estimates for unthinned slash pine plantations - pulpwood, sawtimber, gum. *U S For Serv Res Pap SE-35*. 21 p.
- Bickel, P.J. and K.A. Doksum. 1977. *Mathematical Statistics*. Prentice-Hall, N.J.
- Biging, G.S. and M. Dobbertin. 1992. A comparison of distance-dependent competition measures for height and basal area growth of individual conifer trees. *Forest Science* 38:695-720.
- Borders, B.E. and W.D. Patterson. 1990. Projecting stand tables - A comparison of the Weibull diameter distribution method, a percentile-based projection method, and a basal area growth projection method. *Forest Science* 36:413-424.
- Borders, B.E., R.A. Souter, R.L. Bailey, and K.D. Ware. 1987. Percentile-based distributions characterize forest stand tables. *Forest Science* 33:570-576.
- Brooks, J.R., B.E. Borders, and R.L. Bailey. 1992. Predicting diameter distributions for site-prepared loblolly and slash pine plantations. *Southern Journal of Applied*

- Forestry* 16:130-133.
- Burk, T.E. and J.D. Newberry. 1984. A simple algorithm for moment-based recovery of Weibull distribution parameters. *Forest Science* 30:329-332.
- Burkhart, H.E. 1971. Slash pine plantation yield estimates based on diameter distribution: An evaluation. *Forest Science* 17:452-453.
- Burkhart, H.E., K.D. Farrar, R.L. Amateis, and R.F. Daniels. 1987. Simulation of individual tree growth and stand development in loblolly pine plantations on cutover, site-prepared areas. FWS-1-87. *Virginia Polytechnic Institute and State University, School of Forestry and Wildlife Resources, Blacksburg, VA.* 47 p.
- Cao, Q.V. and H.E. Burkhart. 1984. A segmented distribution approach for modeling diameter frequency data. *Forest Science* 30:129-137.
- Cliff, A.D. and J.K. Ord. 1981. *Spatial Processes: Models and Applications*. Pion, London
- Clutter, J.L. and F.A. Bennett. 1965. Diameter distributions in old field slash pine plantations. *Ga For Res Counc Rep 13, Macon, Ga.* 12 p.
- Cohen, A.C. 1965. Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples. *Technometrics* 7:579-588.
- Cressie, N.A.C. 1993. *Statistics for Spatial Data*. rev ed. John Wiley & Sons, Inc., N.Y.
- Daniels, R.F. 1976. Simple competition indices and their correlation with annual loblolly pine tree growth. *Forest Science* 22:454-456.
- Daniels, R.F., H.E. Burkhart, and T.R. Clason. 1986. A comparison of competition measures for predicting growth of loblolly-pine trees. *Canadian Journal of Forest Research* 16:1230-1237.

- Droessler, T.D. and T.E. Burk. 1989. A test of nonparametric smoothing of diameter distributions. *Scandinavian Journal of Forest Research* 4:407-415.
- Dubey, S.D. 1967. Some percentile estimators for Weibull parameters. *Technometrics* 9:119-129.
- Ek, A.R., J.N. Issos, and R.L. Bailey. 1975. Solving for Weibull diameter distribution parameters to obtain specified mean diameters. *Forest Science* 21:290-292.
- Garcia, O. 1981. Simplified method-of-moments estimation for the Weibull distribution. *New Zealand Journal of Forestry Science* 11:304-306.
- Gates, D.J. and M. Westcott. 1981. Negative skewness and negative correlations in spatial competition models. *Journal of Mathematical Biology* 13:159-171.
- Geary, R.C. 1954. The contiguity ratio and statistical mapping. *Incorporated Statistician* 5:115-145.
- Goovaerts, P. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, N.Y.
- Gove, J.H. and S.E. Fairweather. 1989. Maximum-likelihood estimation of Weibull function parameters using a general interactive optimizer and grouped data. *Forest Ecology and Management* 28:61-69.
- Gove, J.H. and G.P. Patil. 1998. Modeling the basal area-size distribution of forest stands: A compatible approach. *Forest Science* 44:285-297.
- Haara, A., M. Maltamo, and T. Tokola. 1997. The k-nearest-neighbour method for estimating basal-area diameter distribution. *Scandinavian Journal of Forest Research* 12:200-208.
- Hafley, W.L. and H.T. Schreuder. 1977. Statistical distributions for fitting diameter and

- height data in even-aged stands. *Canadian Journal of Forest Research* 7:481-487.
- Hafley, W.L., W.D. Smith, and M.A. Buford. 1982. A new yield prediction model for unthinned loblolly pine in plantations. *N.C. State Univ. Sch. For. Res., Raleigh, NC. Bioecon. Modeling Proj. Tech. Rep. 1.* 65 p.
- Haining, R.P. 1990. *Spatial Data Analysis in the Social and Environmental Sciences.* University Press, Cambridge, Great Britain
- Hardin, J.W., D.J. Leopold, and F.M. White. 2001. *Harlow and Harrar's Textbook of Dendrology.* 9th ed. McGraw-Hill, N.Y.
- Harms, W.R., C.D. Whitesell, and D.S. Debell. 2000. Growth and development of loblolly pine in a spacing trial planted in Hawaii. *Forest Ecology and Management* 126:13-24.
- Hyink, D.M. and J.W. Moser. 1983. A generalized framework for projecting forest yield and stand structure using diameter distributions. *Forest Science* 29:85-95.
- Journel, A.G. and C.J. Huijbregts. 1978. *Mining Geostatistics.* Academic Press, London
- Kaluzny, S.P., S.C. Vega, T.P. Cardoso, and A.A. Shelly. 1998. *S+SpatialStats : User's Manual for Windows and UNIX.* Springer-Verlag, N.Y.
- Kangas, A. and M. Maltamo. 2000. Calibrating predicted diameter distribution with additional information. *Forest Science* 46:390-396.
- Kilikki, P., M. Maltamo, R. Mykkanen, and R. Paivinen. 1989. Use of the Weibull function in estimating the basal area DBH-distribution. *Silva Fennica* 23:311-318.
- Knoebel, B.R. and H.E. Burkhart. 1991. A bivariate distribution approach to modeling forest diameter distributions at two points in time. *Biometrics* 47:241-253.
- Laarhoven, P.J.M.v. and E.H.L. Aarts. 1987. *Simulated Annealing: Theory and*

- Applications*. D. Reidel Publishing Company, Dordrecht, Holland
- Leak, W.B. 1964. An expression of diameter distribution for unbalanced, uneven-aged stands and forests. *Forest Science* 10:39-50.
- Leak, W.B. 1965. The J-shaped probability distribution. *Forest Science* 11:405-409.
- Lee, Y. and J.D. Lenhart. 1998. Influence of planting density on diameter and height in East Texas pine plantations. *Southern Journal of Applied Forestry* 22:241-244.
- Lemin, R.C. and H.E. Burkhart. 1983. Predicting mortality after thinning in old-field loblolly pine plantations. *Southern Journal of Applied Forestry* 7:20-23.
- Lenhart, J.D. 1988. Diameter-distribution yield-prediction system for unthinned loblolly and slash pine plantations on non-old-fields in East Texas. *Southern Journal of Applied Forestry* 12:239-242.
- Lin, C. and P.M. Morse. 1975. A compact design for spacing experiments. *Biometrics* 31:661-671.
- Liu, J. and H.E. Burkhart. 1994. Spatial characteristics of diameter and total height in juvenile loblolly pine (*Pinus taeda* L.) plantations. *Forest Science* 40:774-786.
- Lorimer, C.G. and A.G. Krug. 1983. Diameter distributions in even-aged stands of shade-tolerant and midtolerant tree species. *The American Midland Naturalist* 109:331-345.
- Magnussen, S. 1990. Application and comparison of spatial models in analyzing tree-genetics field trials. *Canadian Journal of Forest Research* 20:536-546.
- Magnussen, S. 1994. A method to adjust simultaneously for spatial microsite and competition effects. *Canadian Journal of Forest Research* 24:985-995.
- Maltamo, M. 1997. Comparing basal area diameter distributions estimated by tree species

- and for the entire growing stock in a mixed stand. *Silva Fennica* 31:53-65.
- Maltamo, M., J. Puumalainen, and R. Paivinen. 1995. Comparison of beta and Weibull functions for modelling basal area diameter distribution in stands of *Pinus sylvestris* and *Picea abies*. *Scandinavian Journal of Forest Research* 10:284-295.
- Matheron, G. 1963. Principles of geostatistics. *Economic Geology* 58:1246-1266.
- Matney, T.G., J.R. Ledbetter, and A.D. Sullivan. 1987. Diameter distribution yield systems for unthinned cutover site-prepared slash pine plantations in Southern Mississippi. *Southern Journal of Applied Forestry* 11:32-36.
- McTague, J.P. and R.L. Bailey. 1987. Compatible basal area and diameter distribution models for thinned loblolly pine plantations in Santa Catarina, Brazil. *Forest Science* 33:43-51.
- Moeur, M. 1993. Characterizing spatial patterns of trees using stem-mapped data. *Forest Science* 39:756-775.
- Moran, P.A.P. 1950. Notes on continuous stochastic phenomena. *Biometrika* 37:17-23.
- Moser, J.W. 1976. Specification of density for the inverse J-shaped diameter distribution. *Forest Science* 22:177-180.
- Murray, D.M. and K. von Gadow. 1991. Relationships between the diameter distributions before and after thinning. *Forest Science* 37:552-559.
- Penttinen, A., D. Stoyan, and H.M. Henttonen. 1992. Marked point-processes in forest statistics. *Forest Science* 38:806-824.
- Pienaar, L.V. and B.D. Shiver. 1993. Early results from an old-field loblolly pine spacing study in the Georgia Piedmont with competition control. *Southern Journal of Applied Forestry* 17:193-196.

- Radtke, P.J. and H.E. Burkhart. 1999. Basal area growth and crown closure in a loblolly pine spacing trial. *Forest Science* 45:35-44.
- Ratkowsky, D.A. 1983. *Nonlinear Regression Modeling*. Marcel Dekker, Inc., N.Y.
- Reed, D.D. and H.E. Burkhart. 1985. Spatial autocorrelation of individual tree characteristics in loblolly pine stands. *Forest Science* 31:575-587.
- Rennolls, K., D.N. Geary, and T.J.D. Rollinson. 1985. Characterizing diameter distributions by the use of the Weibull distribution. *Forestry* 58:57-66.
- Reynolds, M.R., T.E. Burk, and W.C. Huang. 1988. Goodness-of-fit tests and model selection procedures for diameter distribution models. *Forest Science* 34:373-399.
- Ripley, B.D. 1976. The second-order analysis of stationary point processes. *Journal of Applied Probability* 13:255-266.
- Schabenberger, O. and F.J. Pierce. 2002. *Contemporary Statistical Models for the Plant and Soil Sciences*. CRC Press, Boca Raton, FL.
- Scott, D.W. 1985. Averaged shifted histograms: Effective nonparametric density estimators in several dimensions. *Annals of Statistics* 13:1024-1040.
- Sharma, M., H.E. Burkhart, and R.L. Amateis. (in press). Modeling the effect of density on the growth of loblolly pine trees. *Southern Journal of Applied Forestry*.
- Sharma, M., H.E. Burkhart, and R.L. Amateis. (in review). Spacing rectangularity effect on the growth of loblolly pine plantations. *Canadian Journal of Forest Research*.
- Shiver, B.D. 1988. Sample sizes and estimation methods for the Weibull distribution for unthinned slash pine plantation diameter distributions. *Forest Science* 34:809-814.
- Somers, G.L., R.G. Oderwald, W.R. Harms, and O.G. Langdon. 1980. Predicting mortality with a Weibull distribution. *Forest Science* 26:291-300.

- Stoyan, D. and A. Penttinen. 2000. Recent applications of point process methods in forestry statistics. *Statistical Science* 15:61-78.
- Strub, M.R. and H.E. Burkhart. 1975. A class-interval-free method for obtaining expected yields from diameter distributions. *Forest Science* 21:67-69.
- Terrell, G.R. and D.W. Scott. 1985. Oversmoothed nonparametric density estimates. *Journal of the American Statistical Association* 80:209-214.
- Zarnoch, S.J. and T.R. Dell. 1985. An evaluation of percentile and maximum likelihood estimators of Weibull parameters. *Forest Science* 31:260-268.
- Zhang, S., H.E. Burkhart, and R.L. Amateis. 1996. Modeling individual tree growth for juvenile loblolly pine plantations. *Forest Ecology and Management* 89:157-172.
- Zutter, B.R., R.G. Oderwald, R.M. Farrar, and P.A. Murphy. 1982. WEIBUL, a program to estimate parameters of forms of the Weibull distribution using complete, censored, and truncated data. FWS-3-82. *Virginia Polytechnic Institute and State University, School of Forestry and Wildlife Resources, Blacksburg, VA*. 17 p.
- Zutter, B.R., R.G. Oderwald, P.A. Murphy, and R.M. Farrar. 1986. Characterizing diameter distributions with modified data types and forms of the Weibull distribution. *Forest Science* 32:37-48.

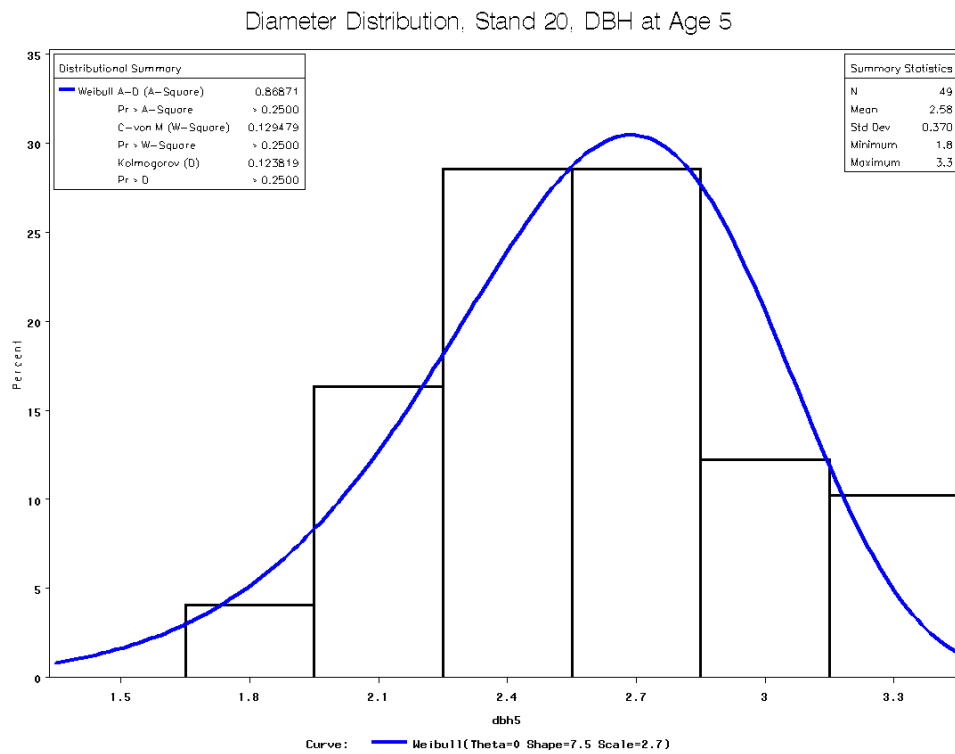
APPENDIX

Appendix I: Histograms for diameter at breast height from ages 5 to 11 for an example plot from each planting density (trees per acre) with the derived two-parameter Weibull distribution overlaid

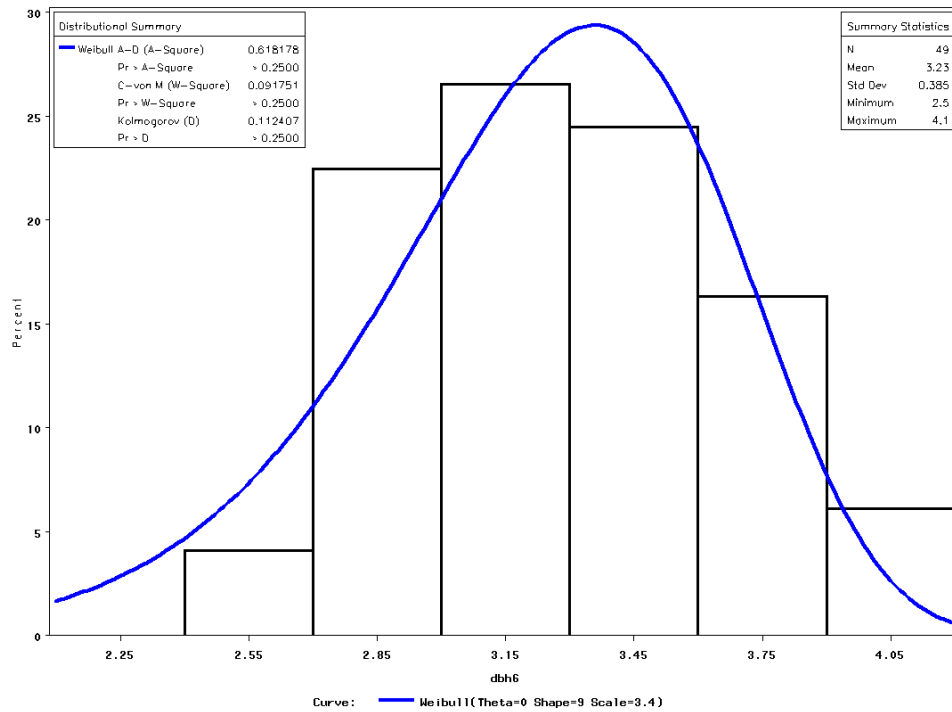
Appendix II: Empirical semivariograms for ground-line and breast height diameters for an example plot from each planting density (trees per acre)

Appendix III: Summary tables for an example plot from each planting density (trees per acre) containing all significant (alpha level of 0.10) spatial indices analyzed in this study

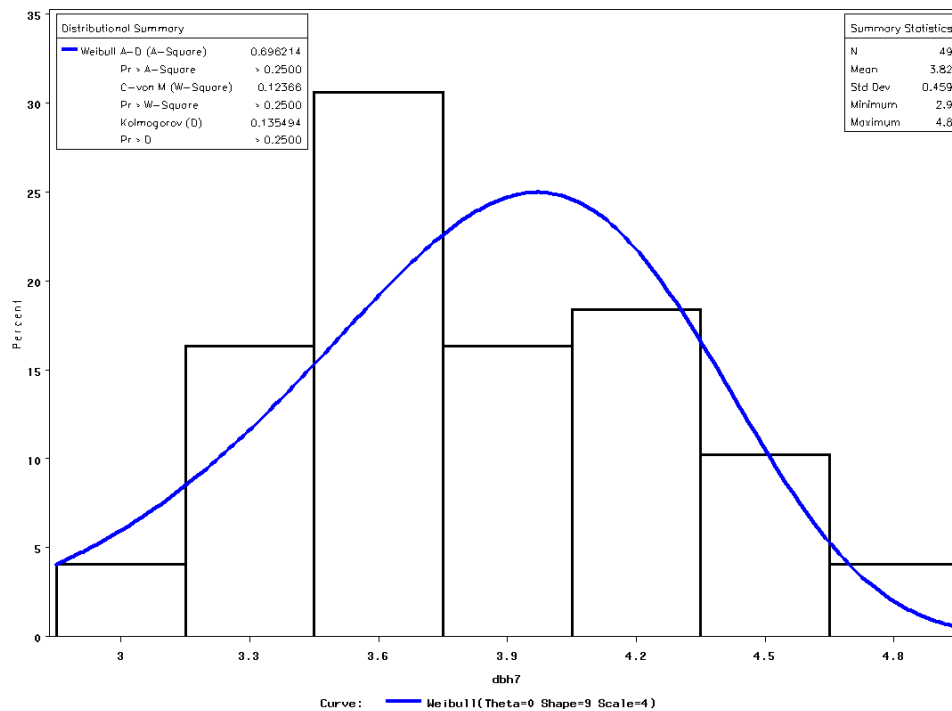
Appendix I: Histograms of an example plot from each planting density (trees per acre) with the derived two-parameter Weibull distribution overlaid. The breast height diameter distributions are given for ages 5 to 11. Summary statistics are in the upper right-hand corner for the diameter data (N = sample size, and Std Dev = standard deviation). A summary of the goodness-of-fit statistics is provided in the upper left-hand corner for the fit between the empirical distribution function and the derived two-parameter Weibull distribution function (A-D = Anderson-Darling, C-von M = Cramer-von Mises, Kolmogorov = Kolmogorov-Smirnov, and Pr = the respective p-value). The estimated scale and shape parameters of the two-parameter Weibull distribution are given below the graph for each fitted distribution (rounded to one decimal for graphics).



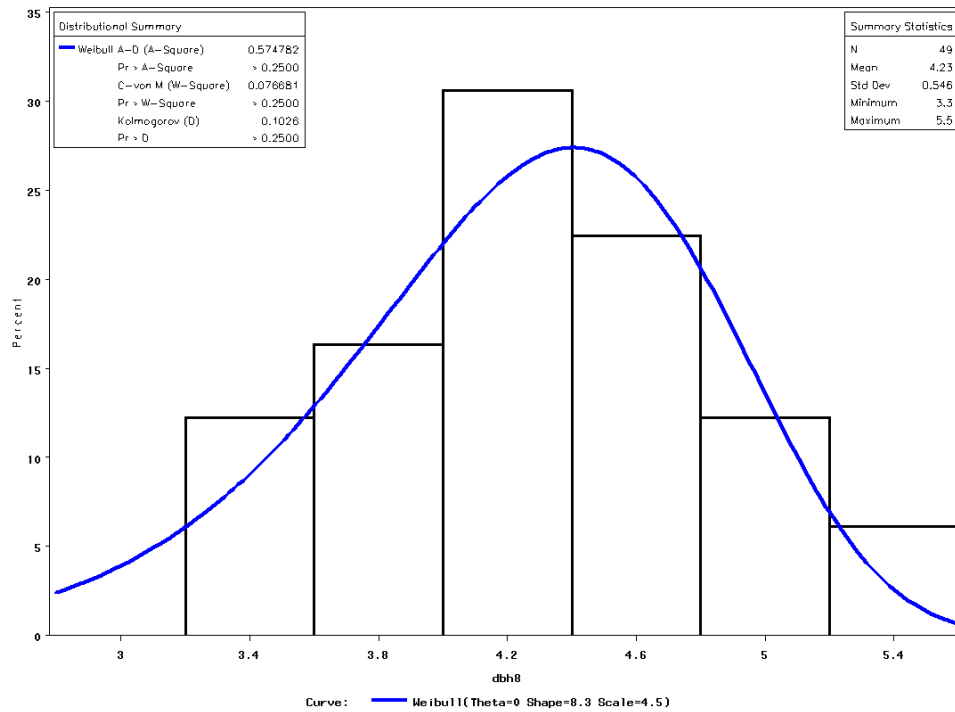
Diameter Distribution, Stand 20, DBH at Age 6



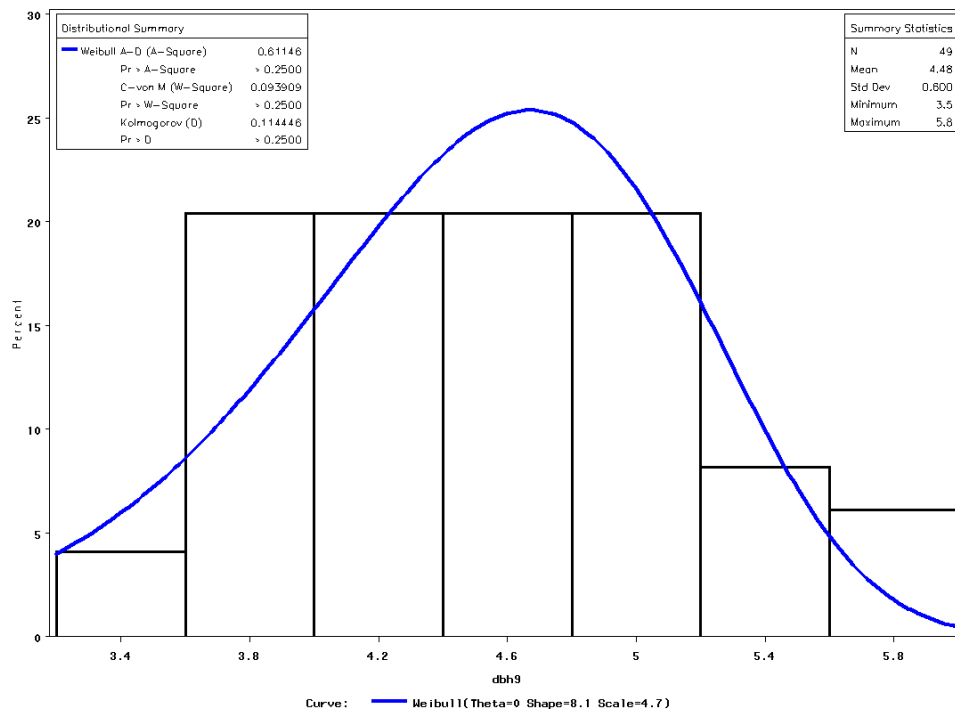
Diameter Distribution, Stand 20, DBH at Age 7



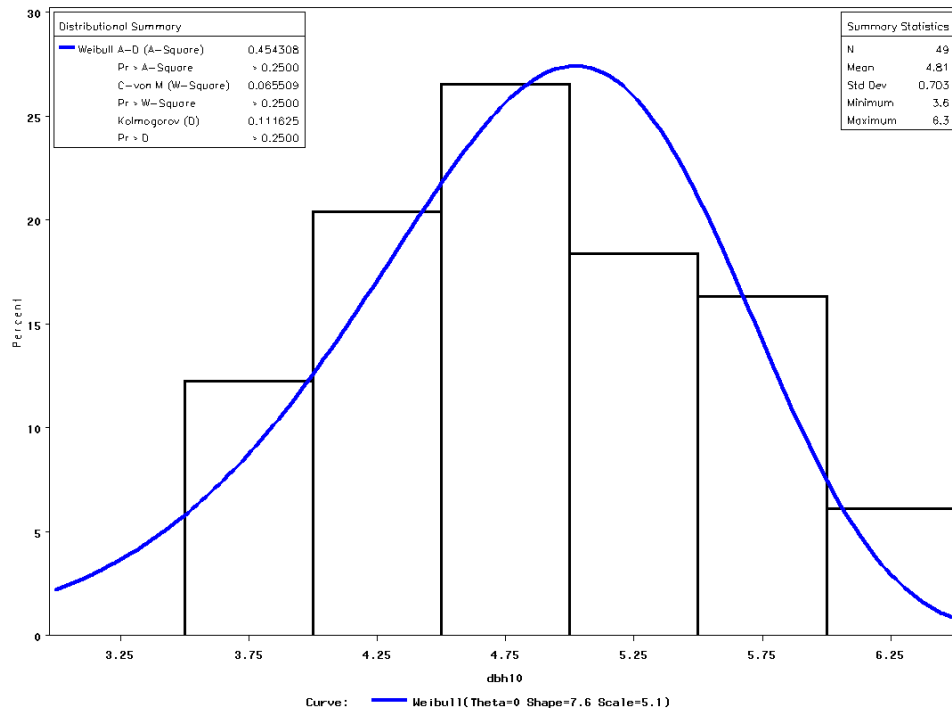
Diameter Distribution, Stand 20, DBH at Age 8



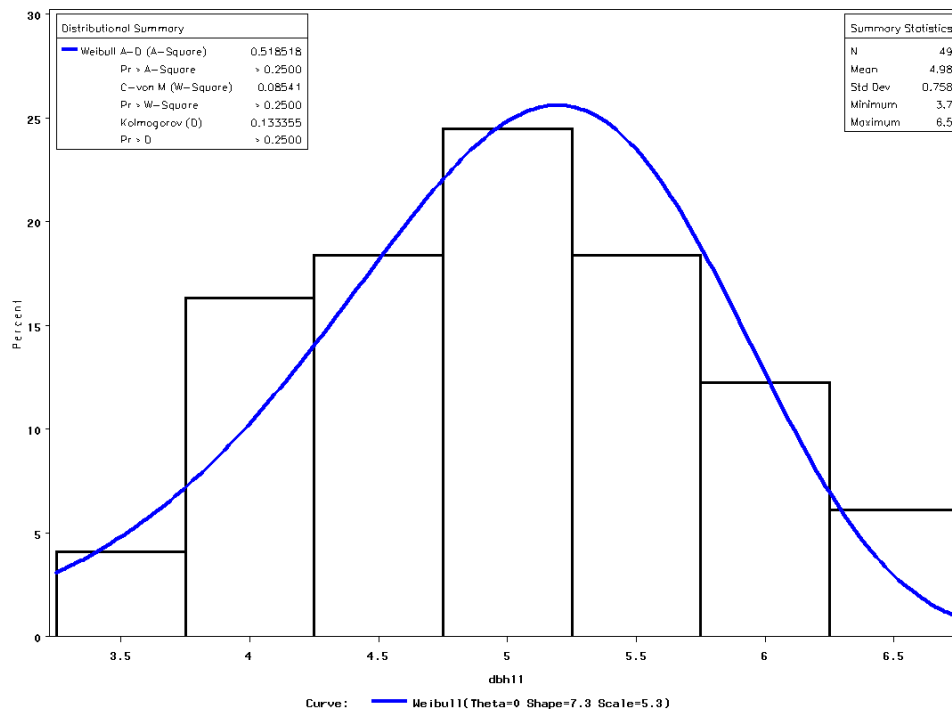
Diameter Distribution, Stand 20, DBH at Age 9



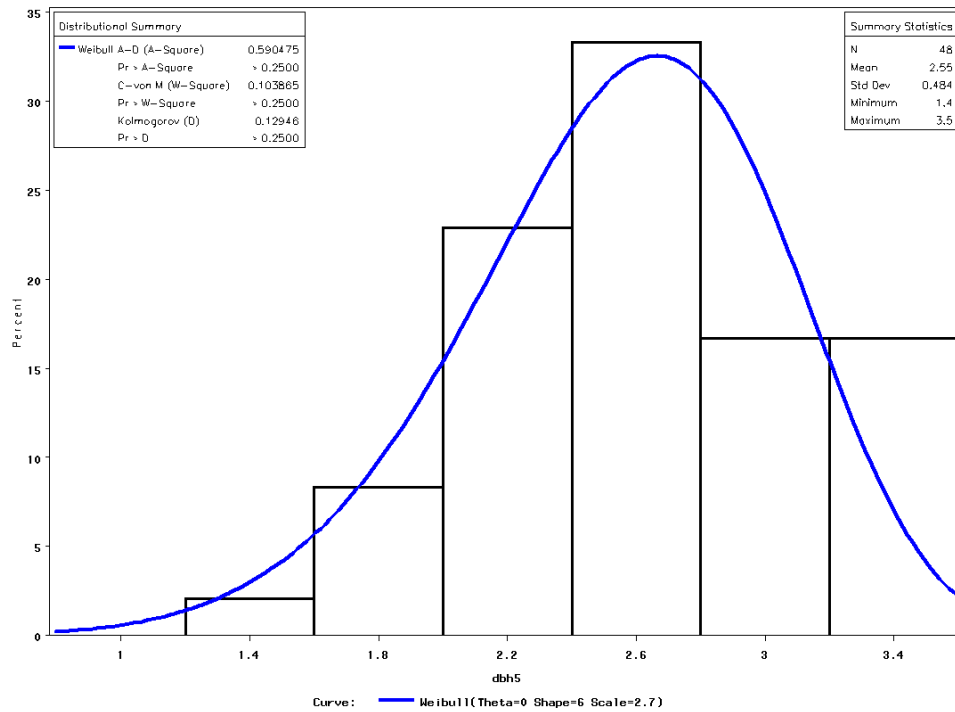
Diameter Distribution, Stand 20, DBH at Age 10



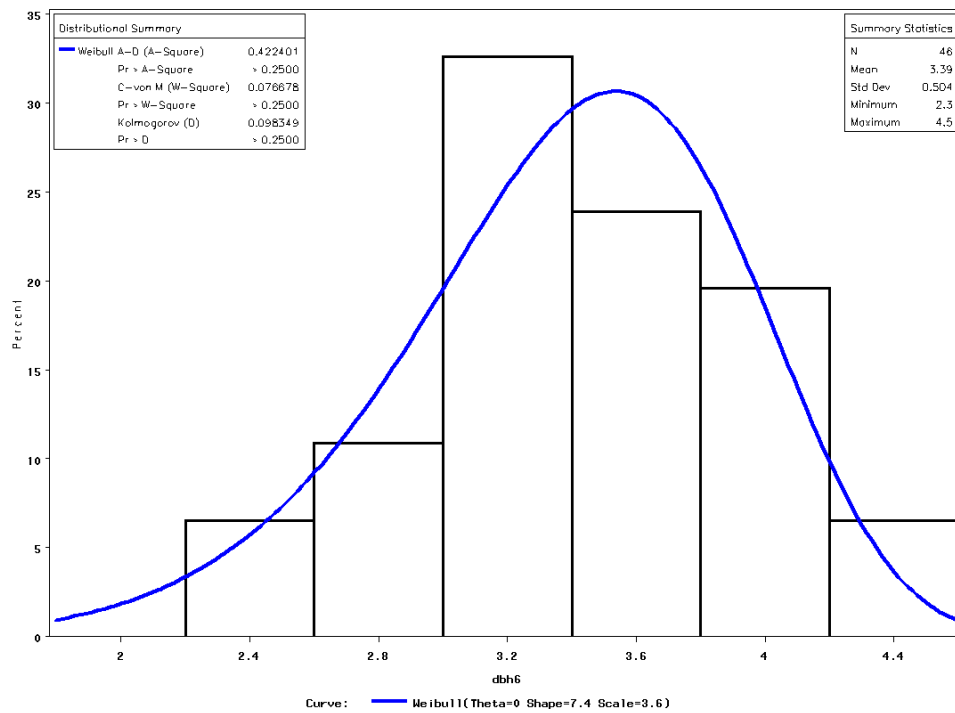
Diameter Distribution, Stand 20, DBH at Age 11



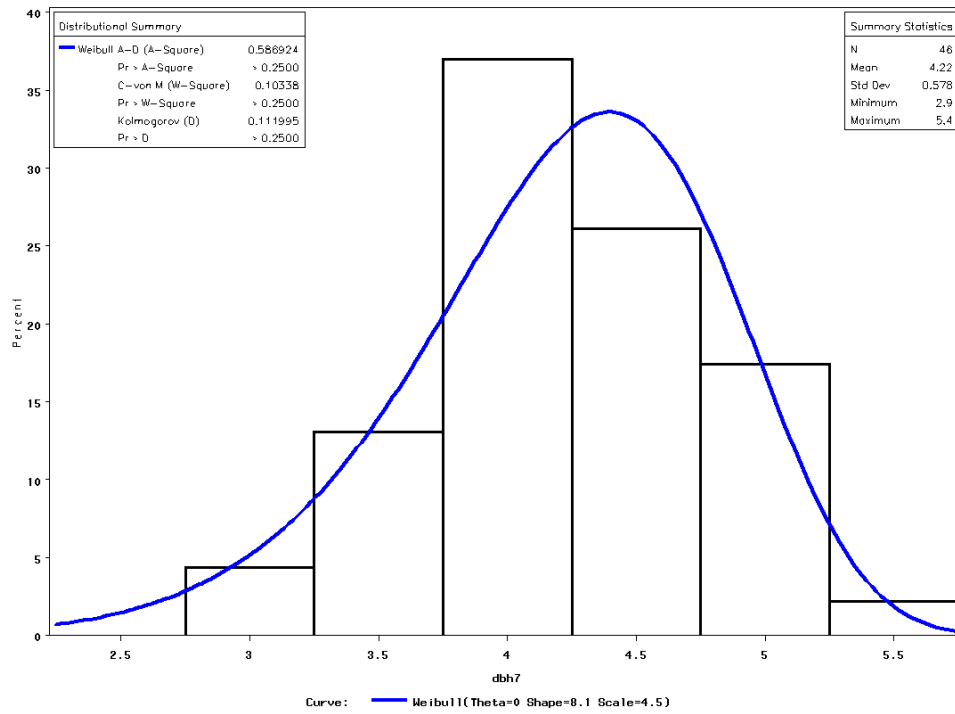
Diameter Distribution, Stand 44, DBH at Age 5



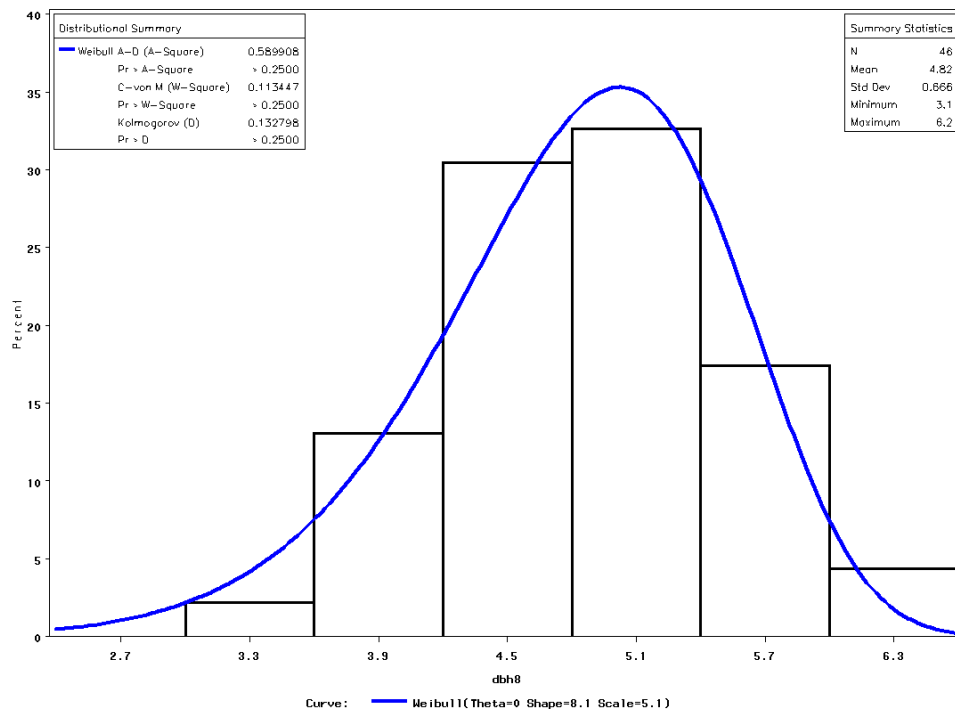
Diameter Distribution, Stand 44, DBH at Age 6



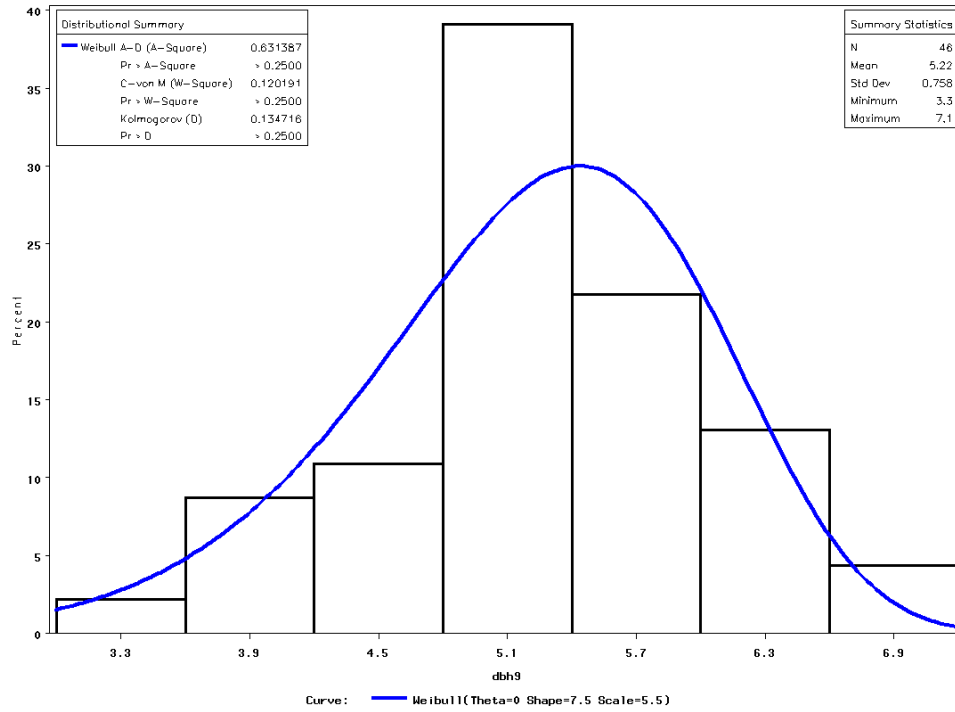
Diameter Distribution, Stand 44, DBH at Age 7



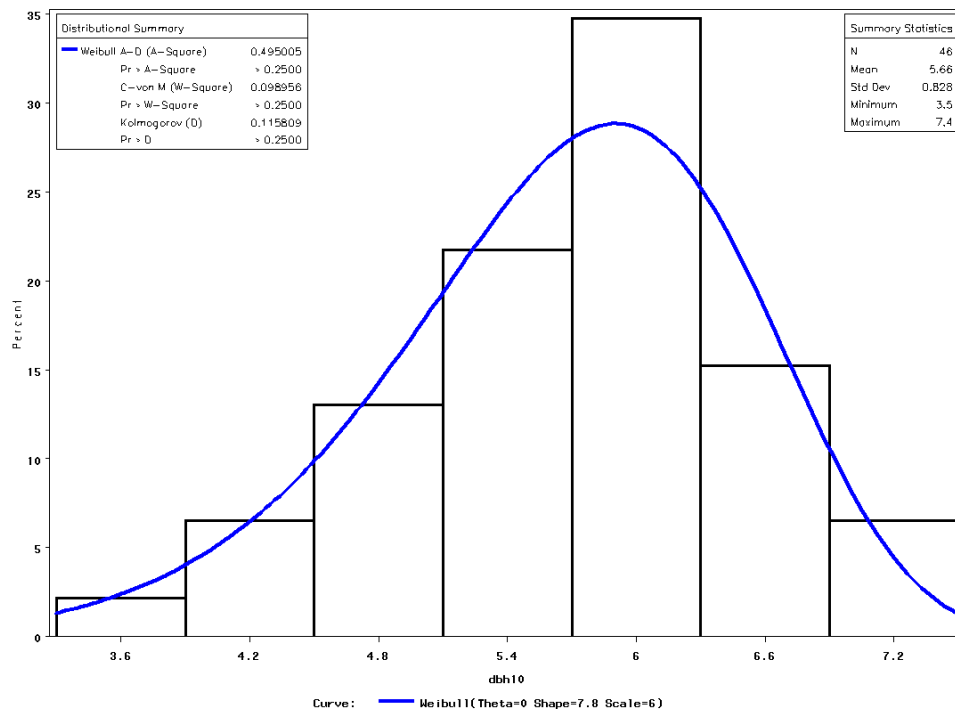
Diameter Distribution, Stand 44, DBH at Age 8



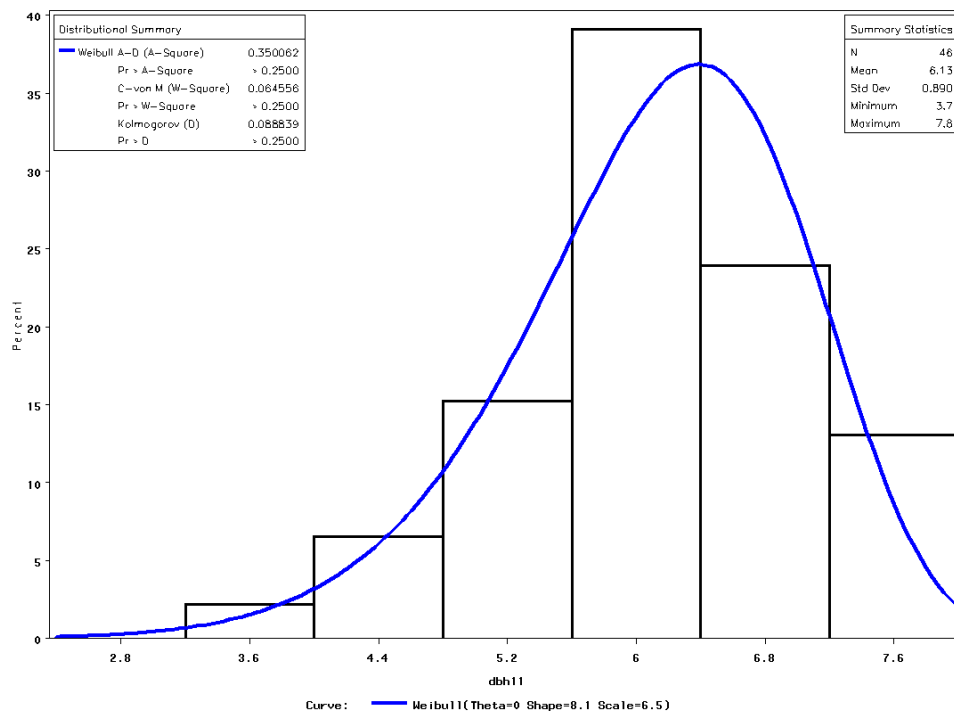
Diameter Distribution, Stand 44, DBH at Age 9



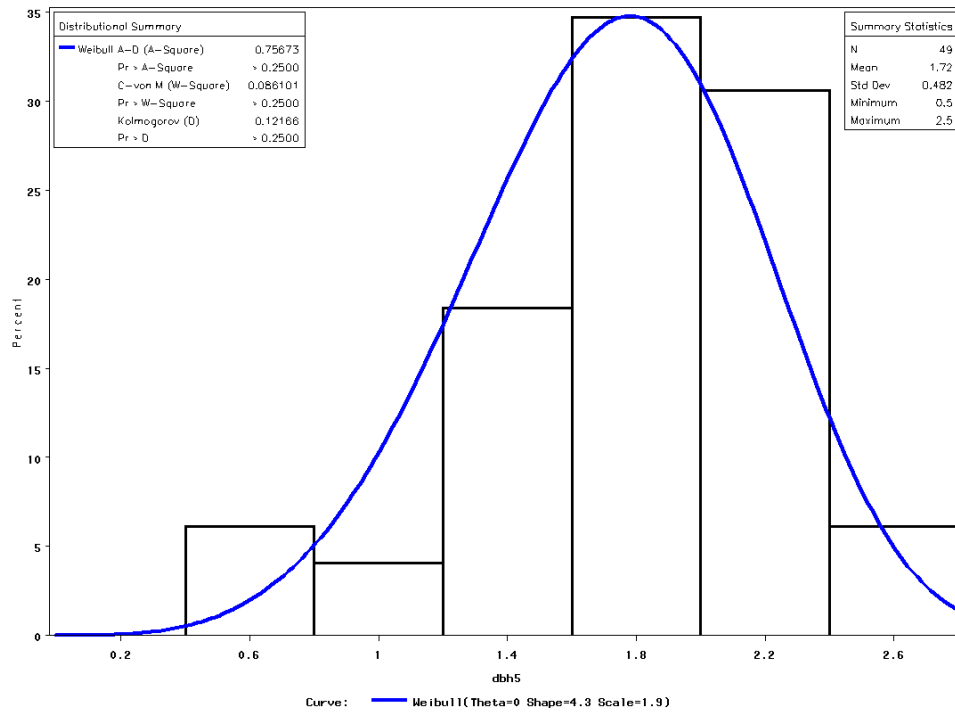
Diameter Distribution, Stand 44, DBH at Age 10



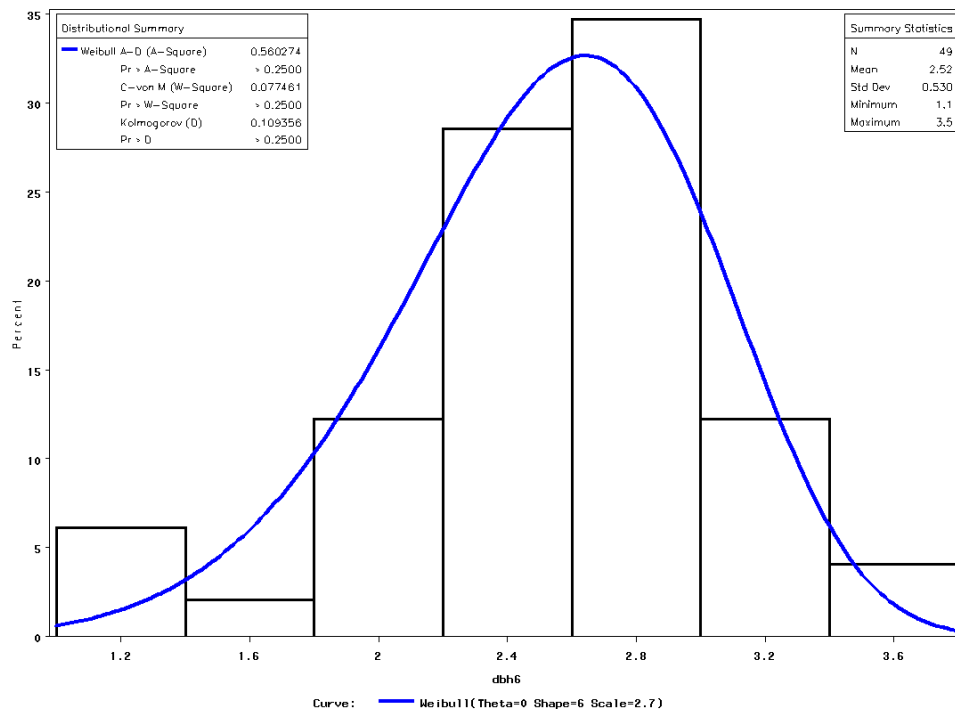
Diameter Distribution, Stand 44, DBH at Age 11



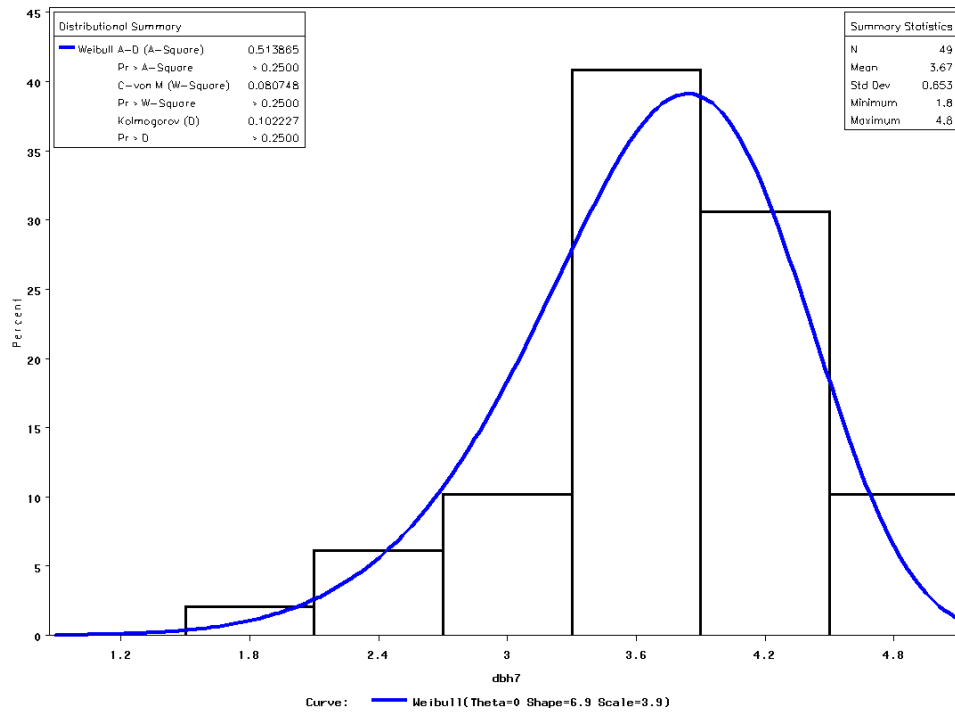
Diameter Distribution, Stand 59, DBH at Age 5



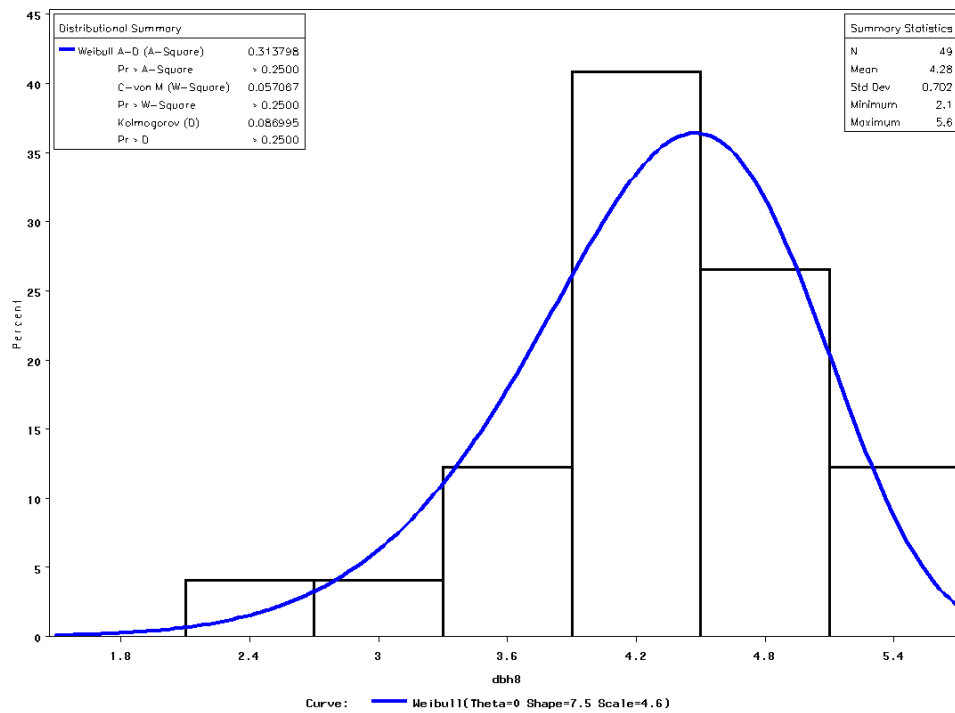
Diameter Distribution, Stand 59, DBH at Age 6



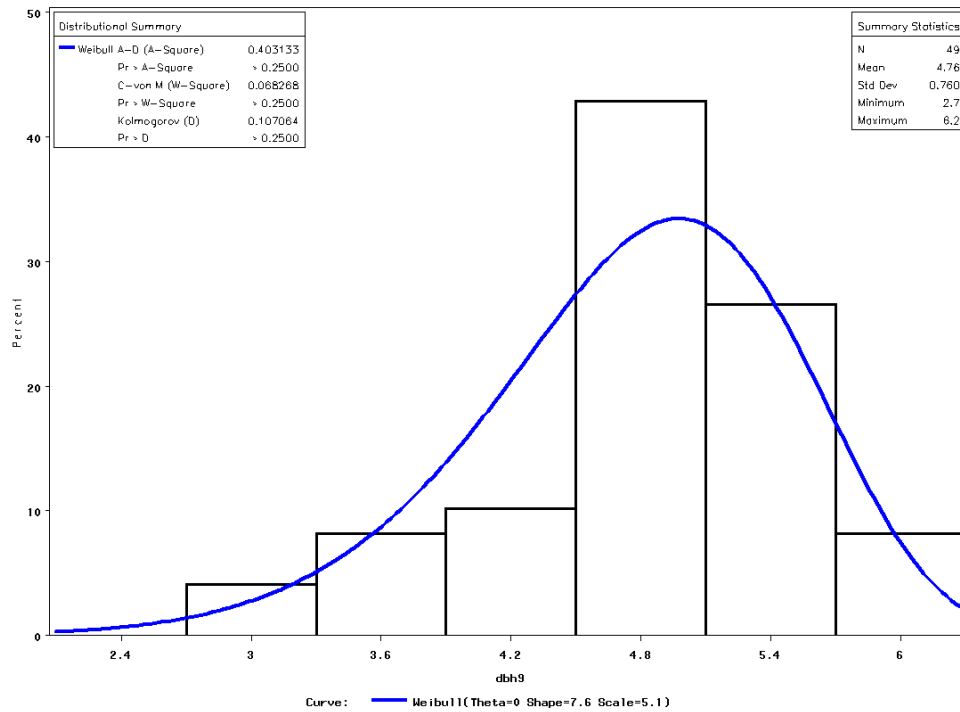
Diameter Distribution, Stand 59, DBH at Age 7



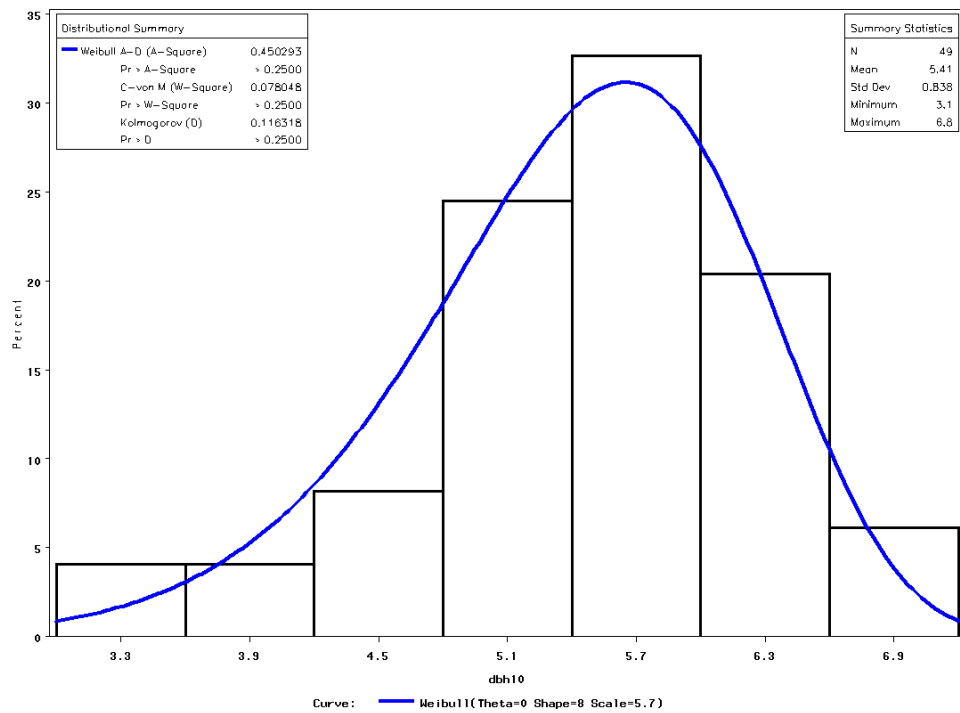
Diameter Distribution, Stand 59, DBH at Age 8



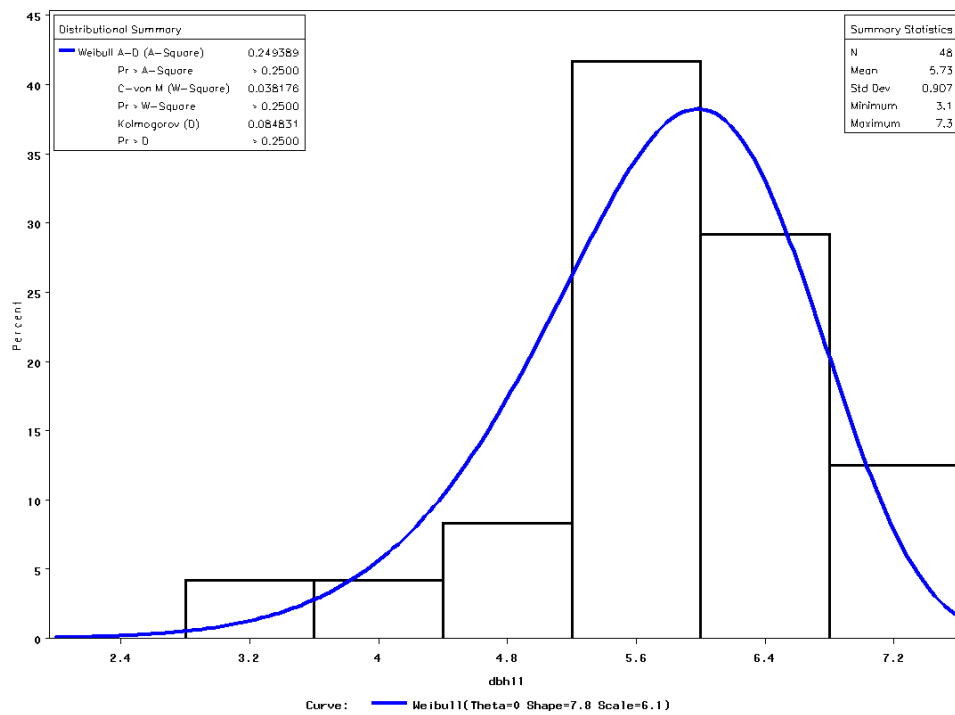
Diameter Distribution, Stand 59, DBH at Age 9



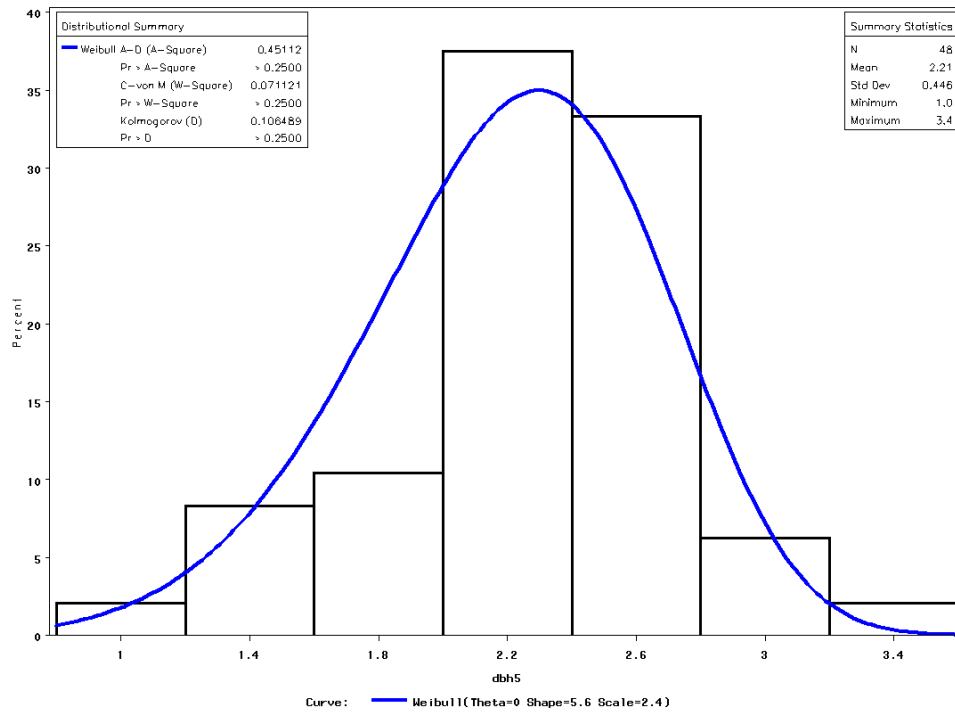
Diameter Distribution, Stand 59, DBH at Age 10



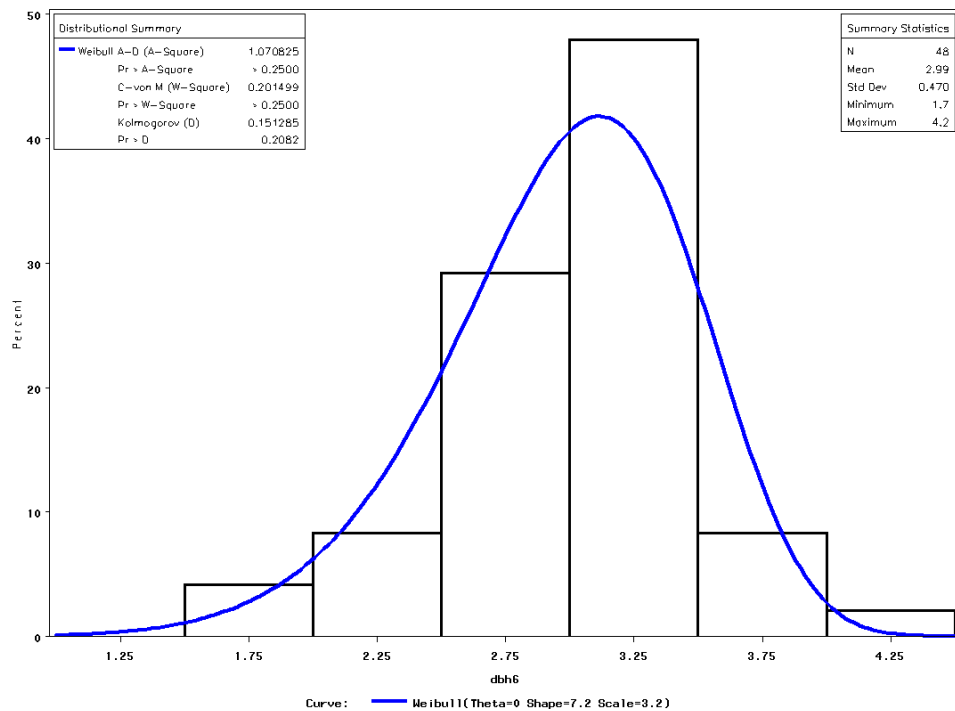
Diameter Distribution, Stand 59, DBH at Age 11



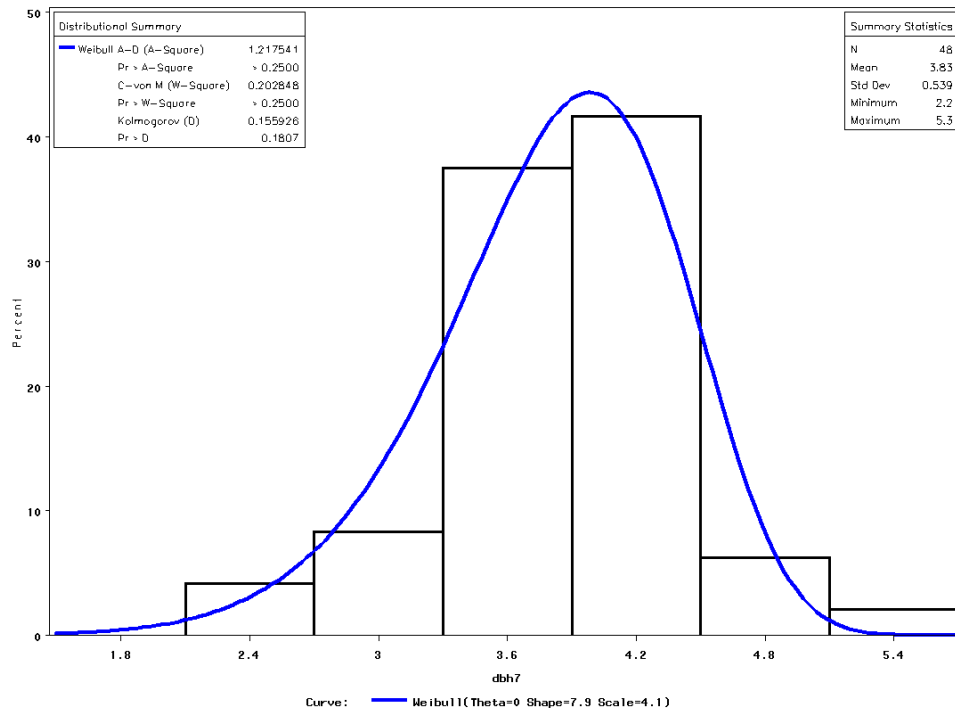
Diameter Distribution, Stand 86, DBH at Age 5



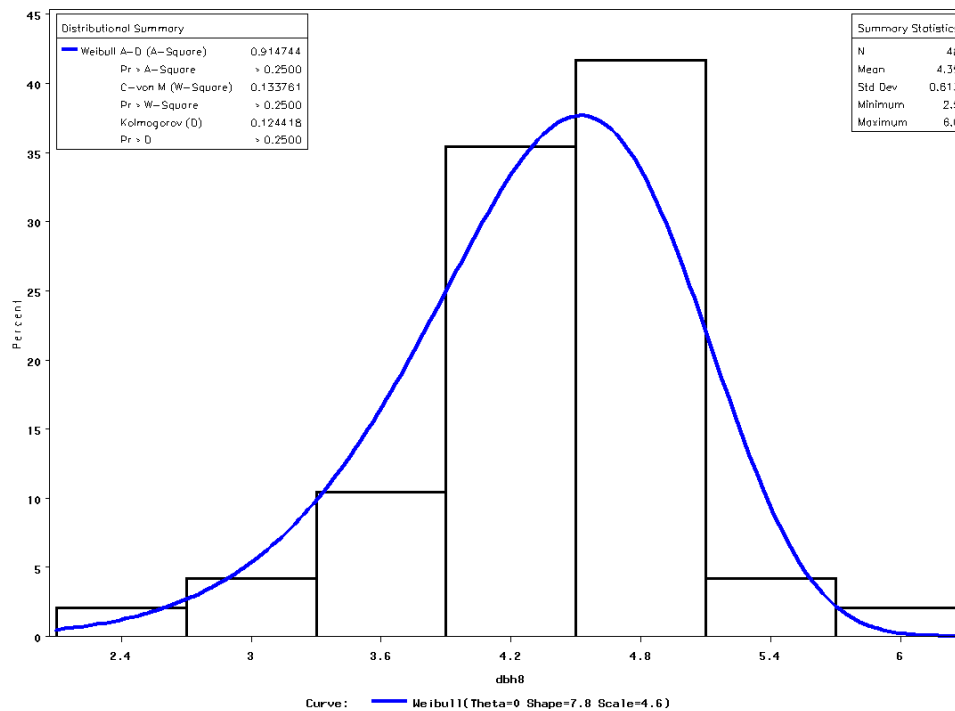
Diameter Distribution, Stand 86, DBH at Age 6



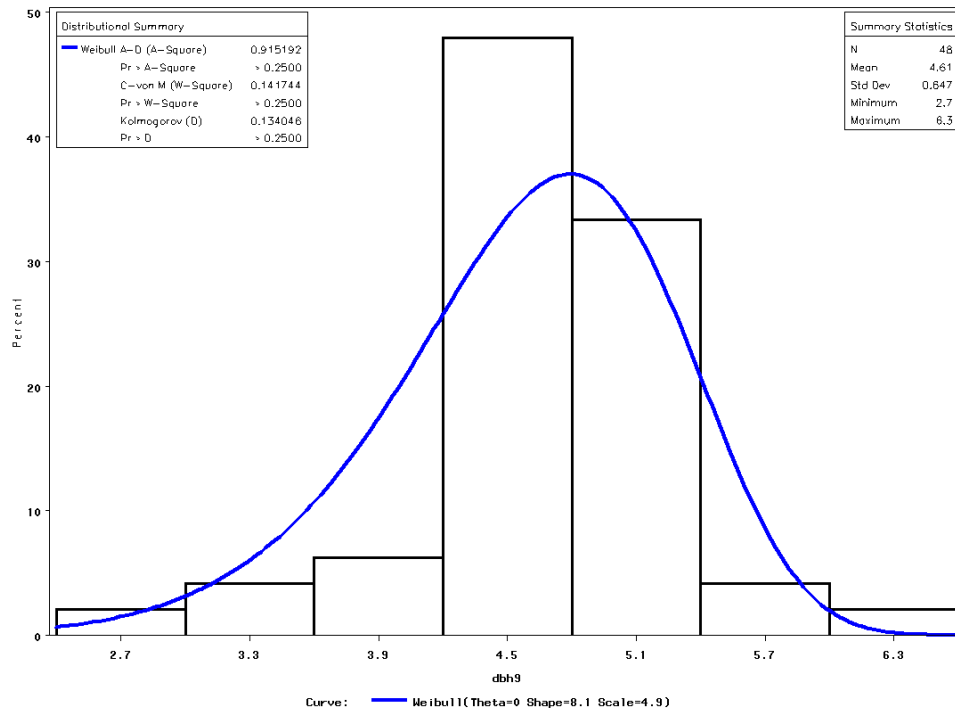
Diameter Distribution, Stand 86, DBH at Age 7



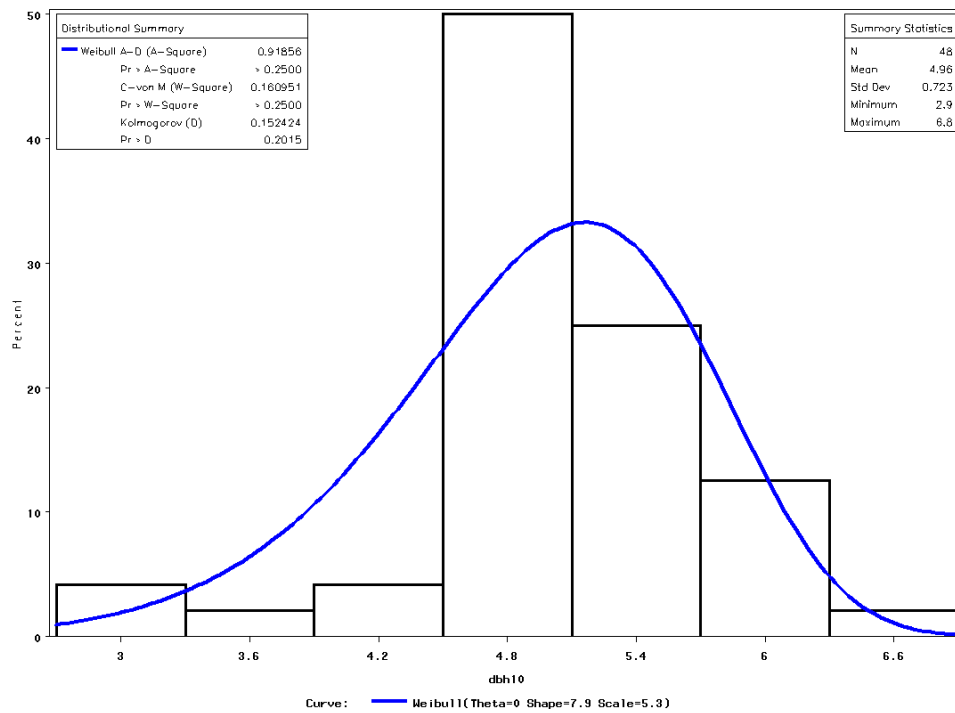
Diameter Distribution, Stand 86, DBH at Age 8



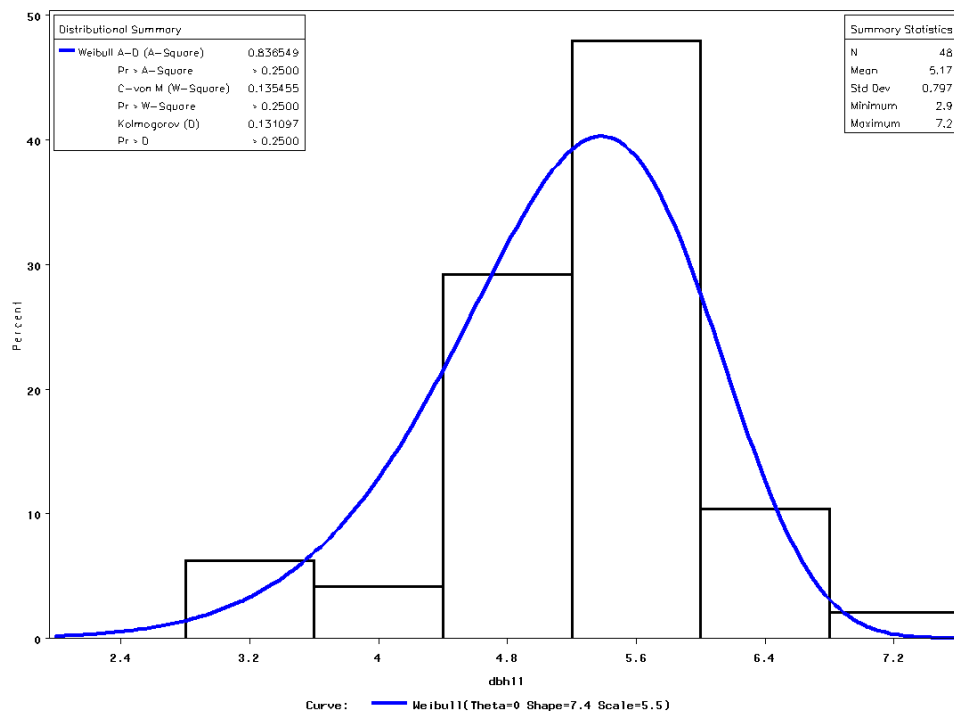
Diameter Distribution, Stand 86, DBH at Age 9



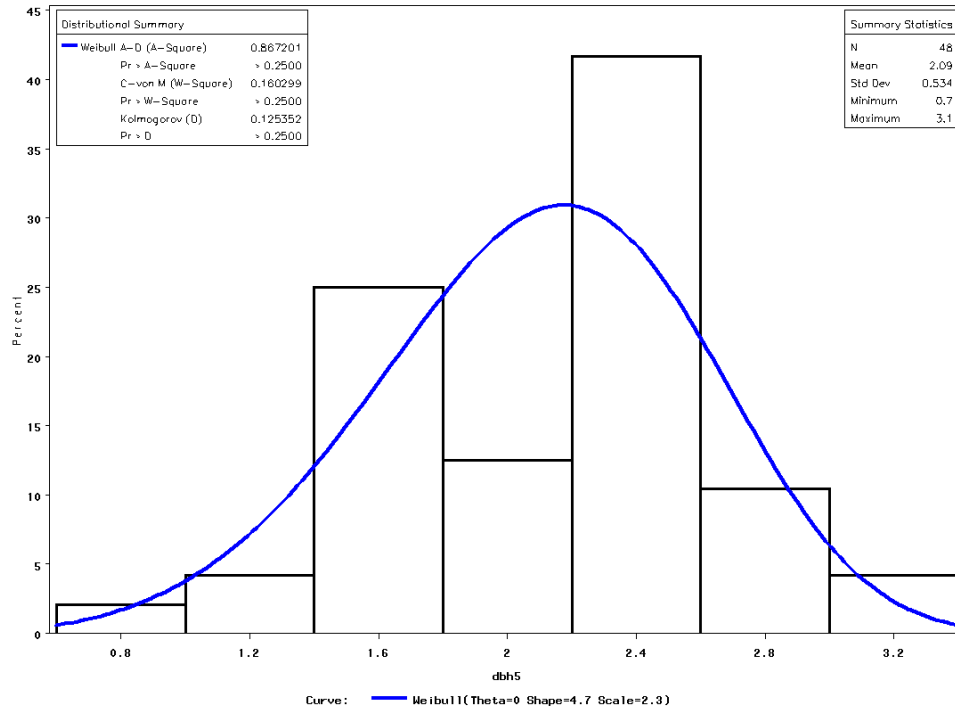
Diameter Distribution, Stand 86, DBH at Age 10



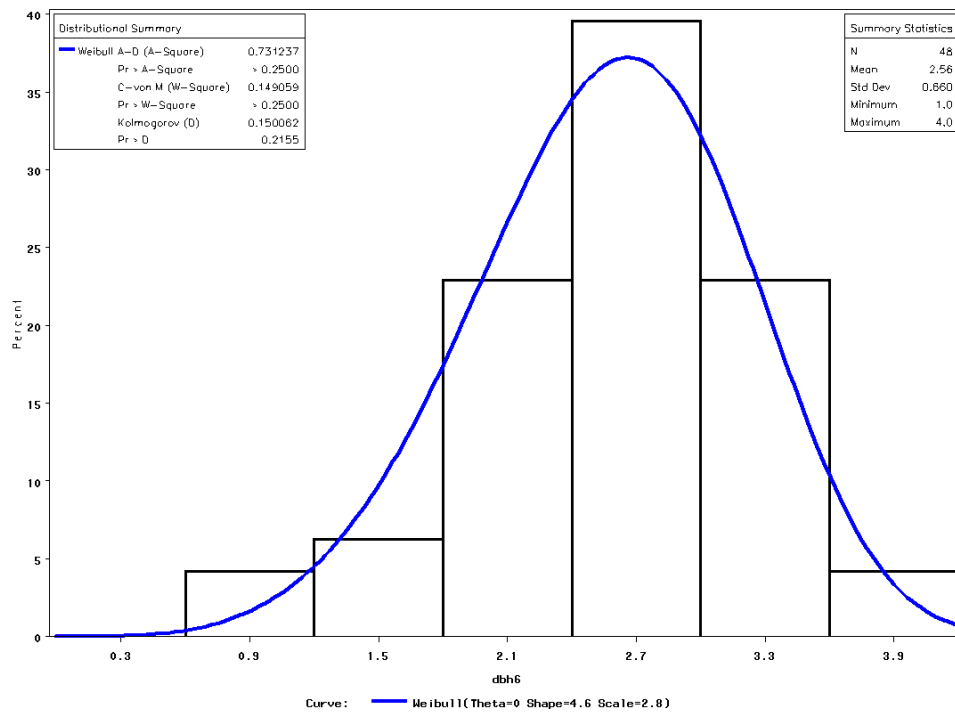
Diameter Distribution, Stand 86, DBH at Age 11



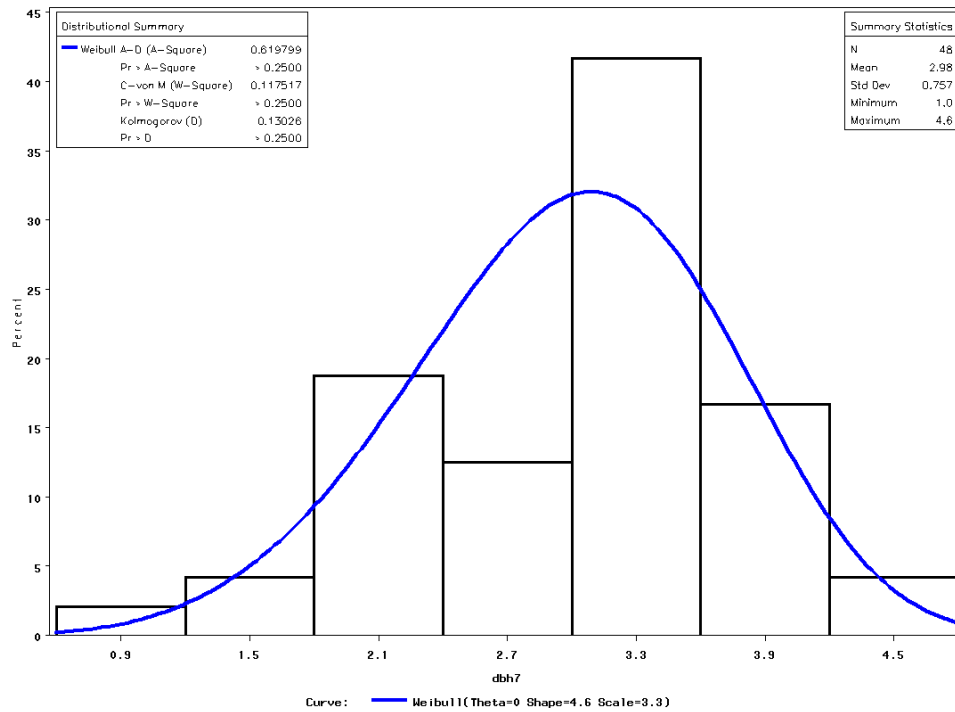
Diameter Distribution, Stand 105, DBH at Age 5



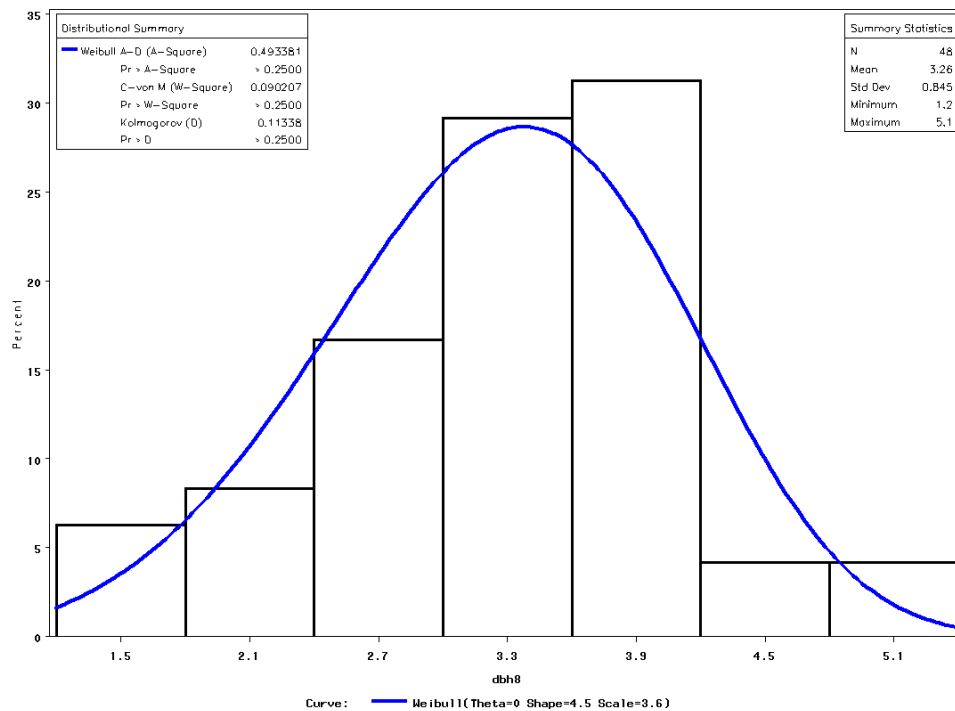
Diameter Distribution, Stand 105, DBH at Age 6



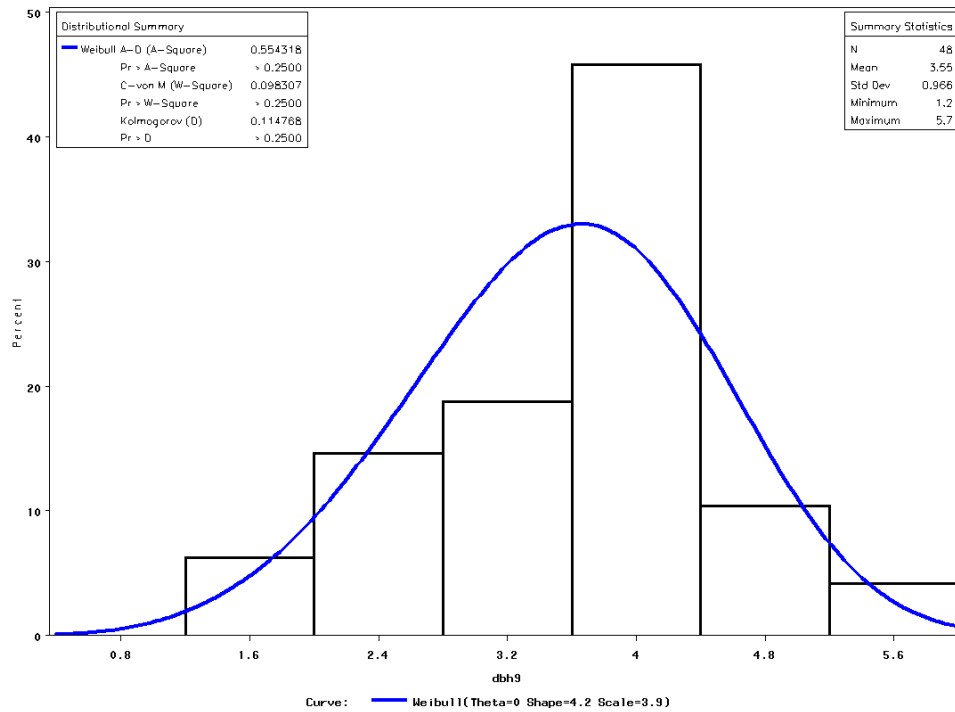
Diameter Distribution, Stand 105, DBH at Age 7



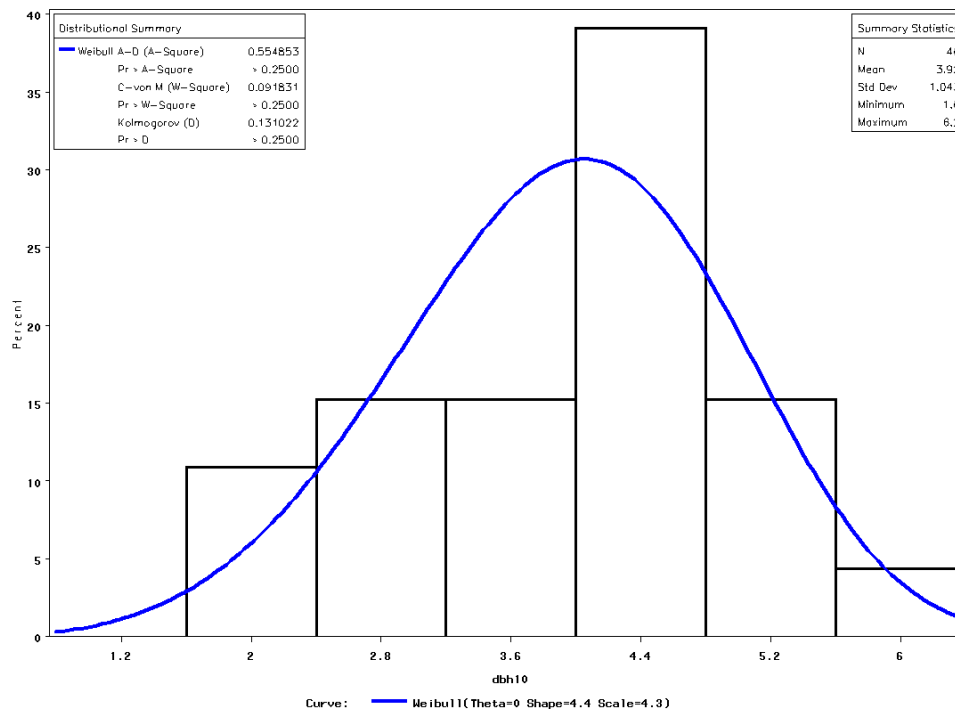
Diameter Distribution, Stand 105, DBH at Age 8



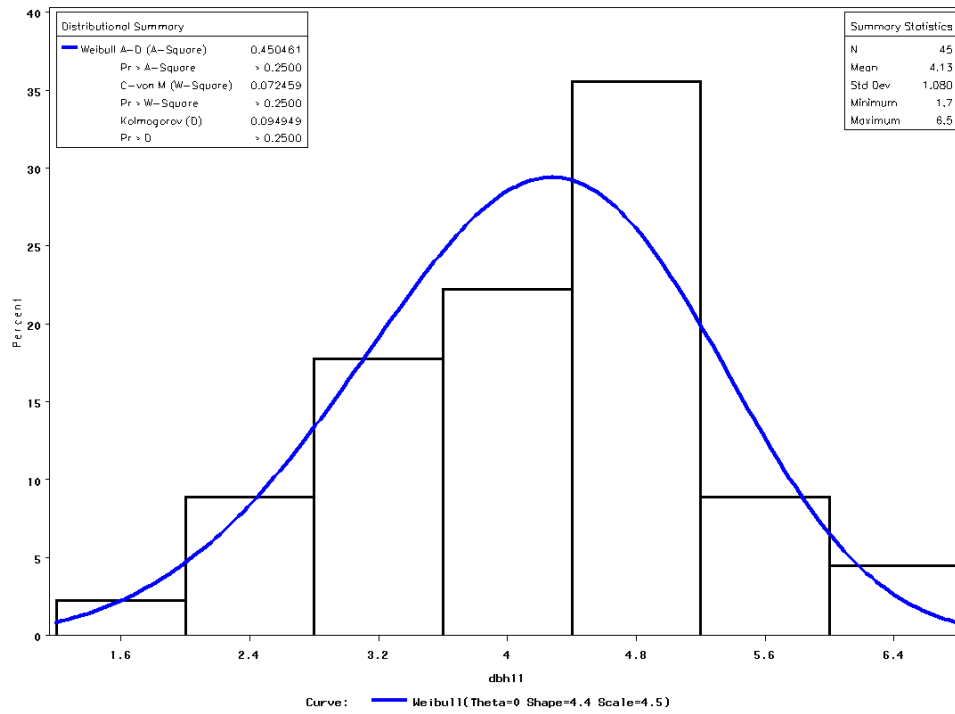
Diameter Distribution, Stand 105, DBH at Age 9



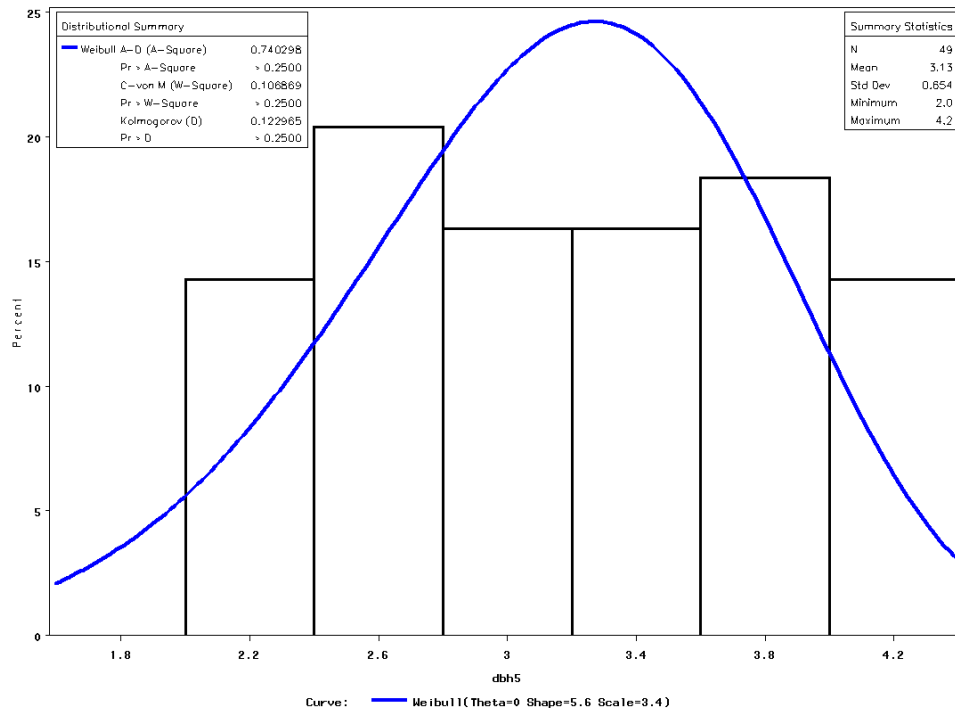
Diameter Distribution, Stand 105, DBH at Age 10



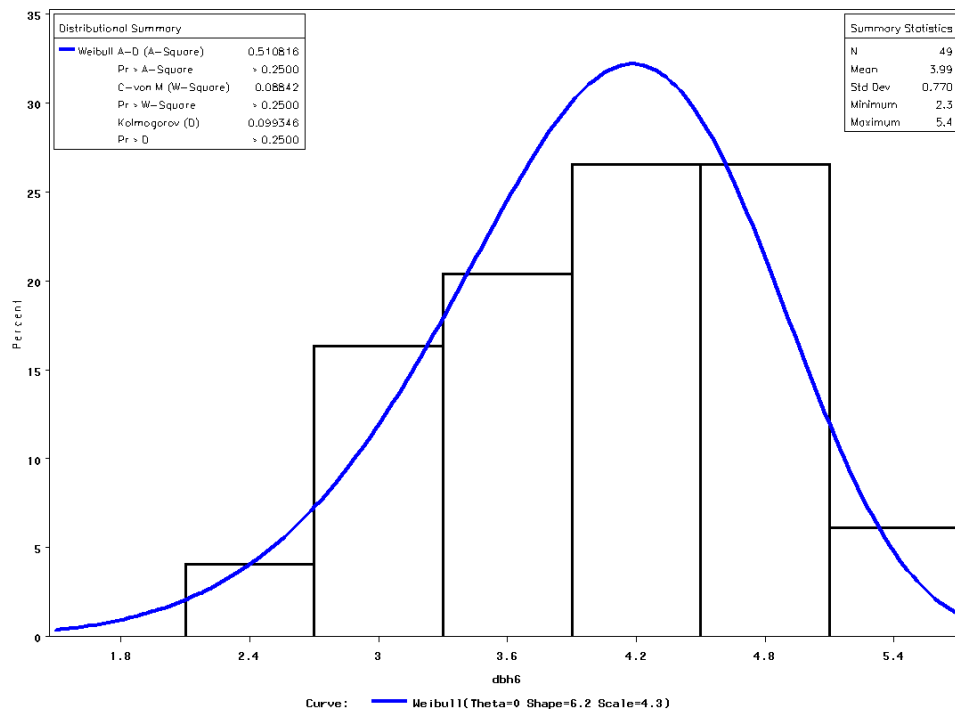
Diameter Distribution, Stand 105, DBH at Age 11



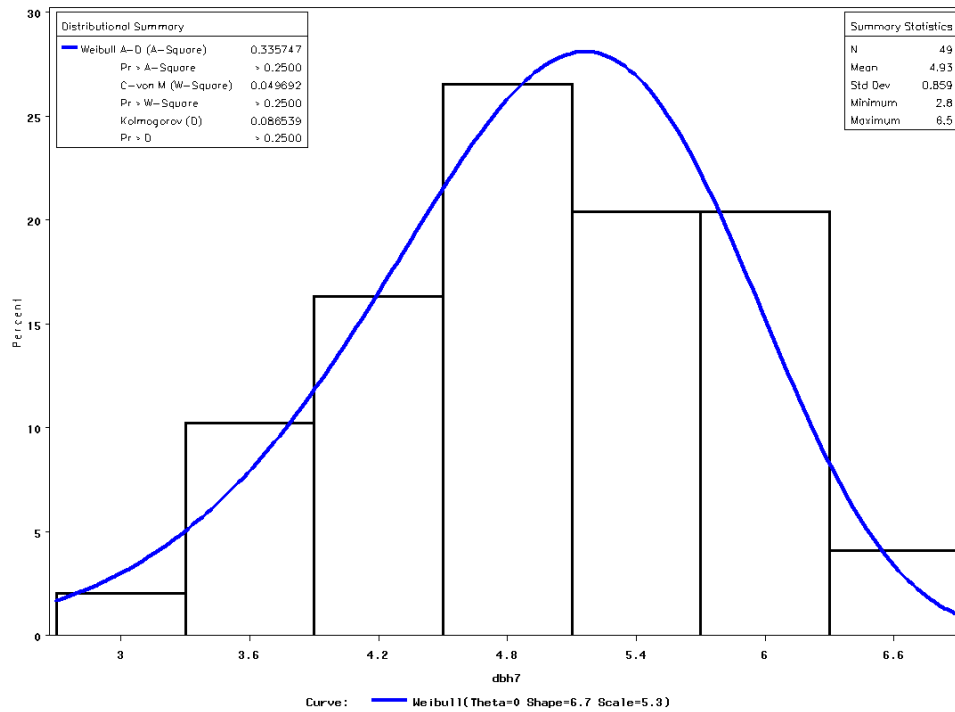
Diameter Distribution, Stand 128, DBH at Age 5



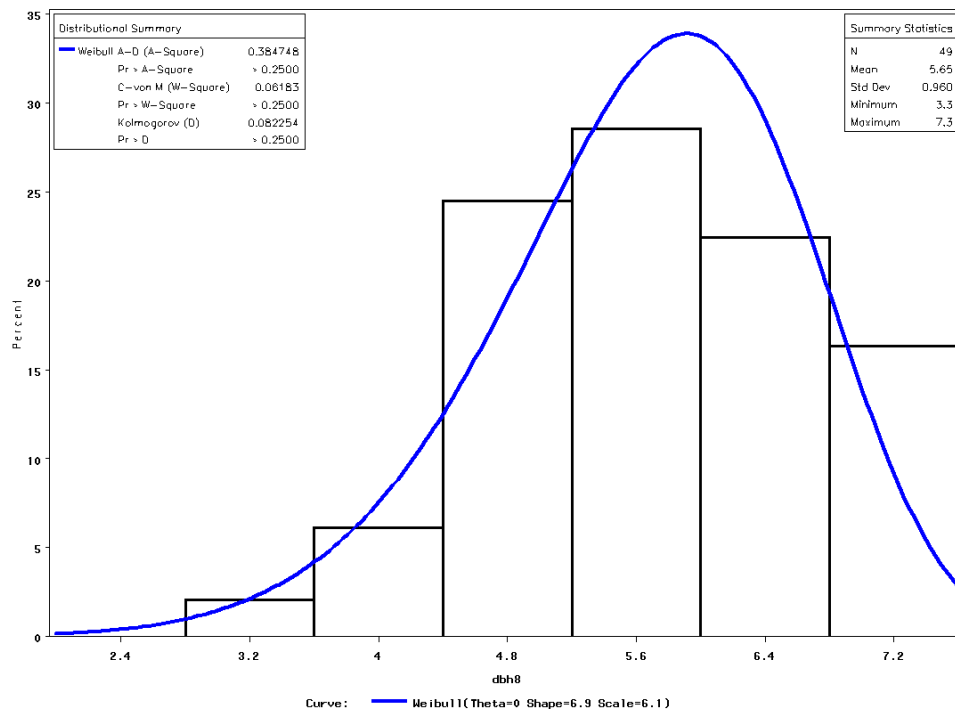
Diameter Distribution, Stand 128, DBH at Age 6



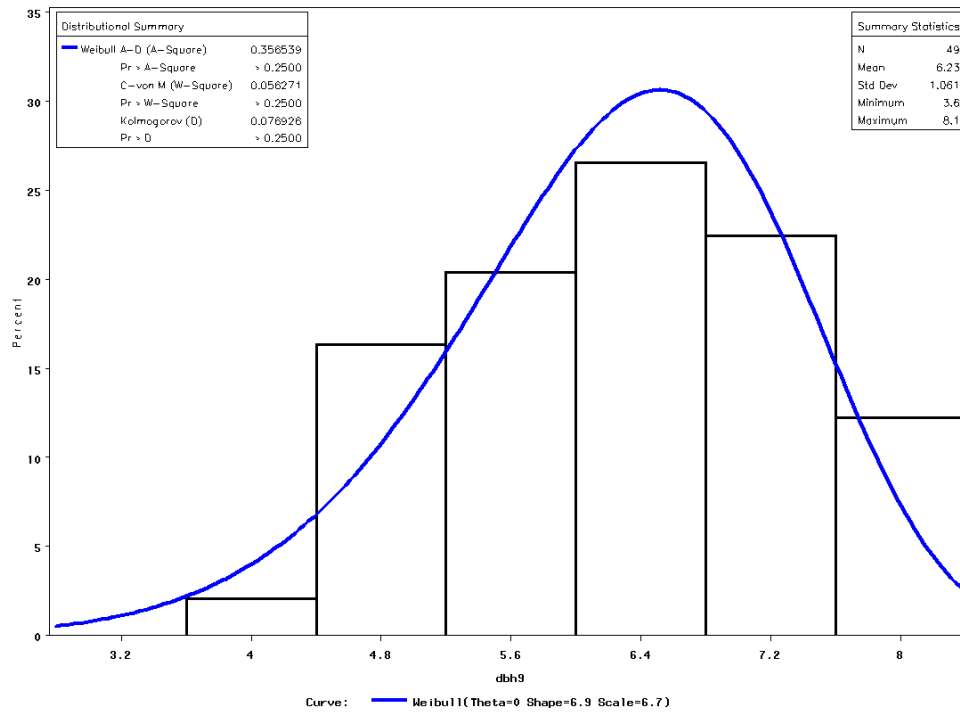
Diameter Distribution, Stand 128, DBH at Age 7



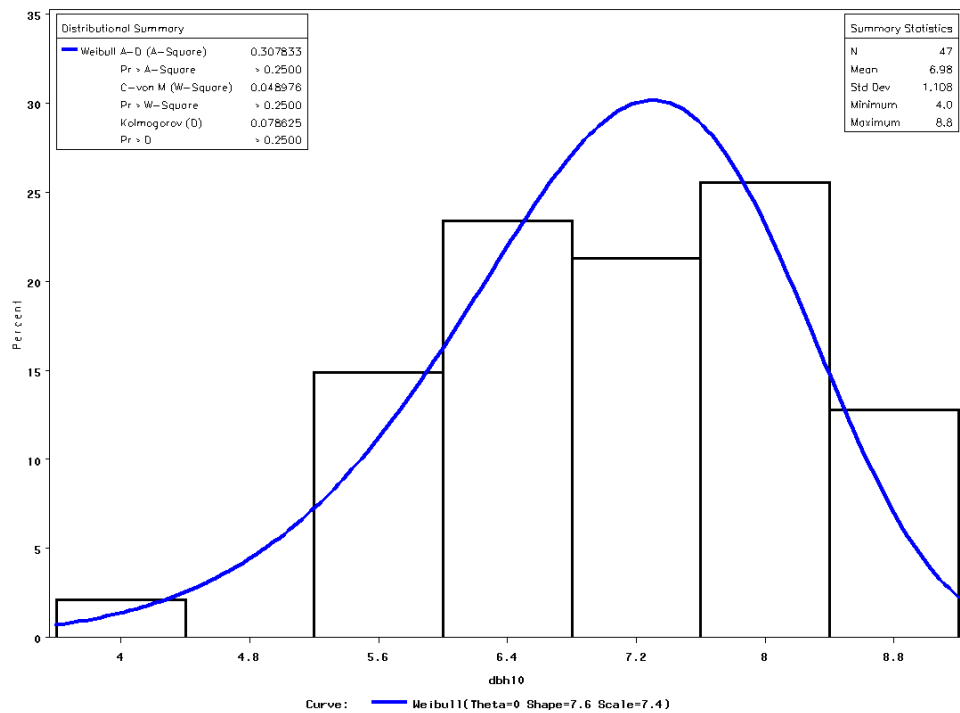
Diameter Distribution, Stand 128, DBH at Age 8



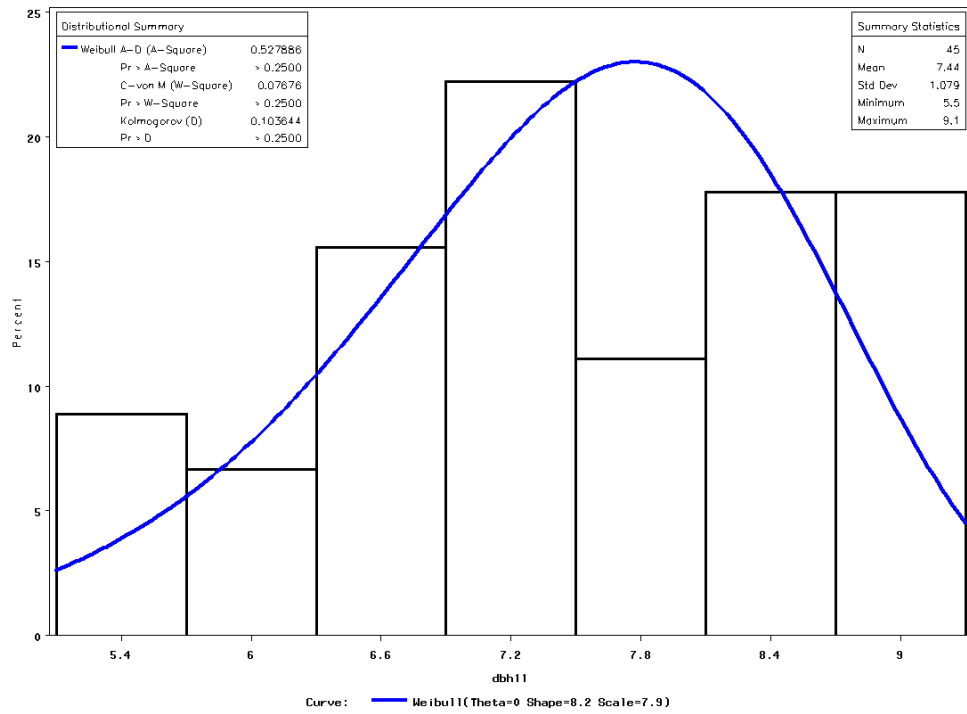
Diameter Distribution, Stand 128, DBH at Age 9



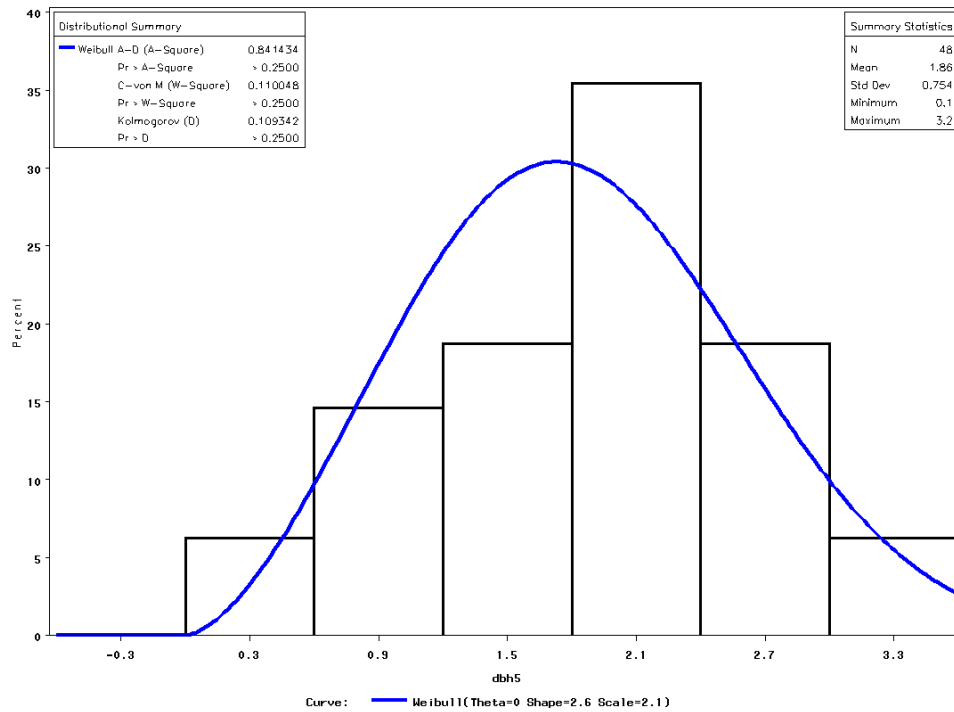
Diameter Distribution, Stand 128, DBH at Age 10



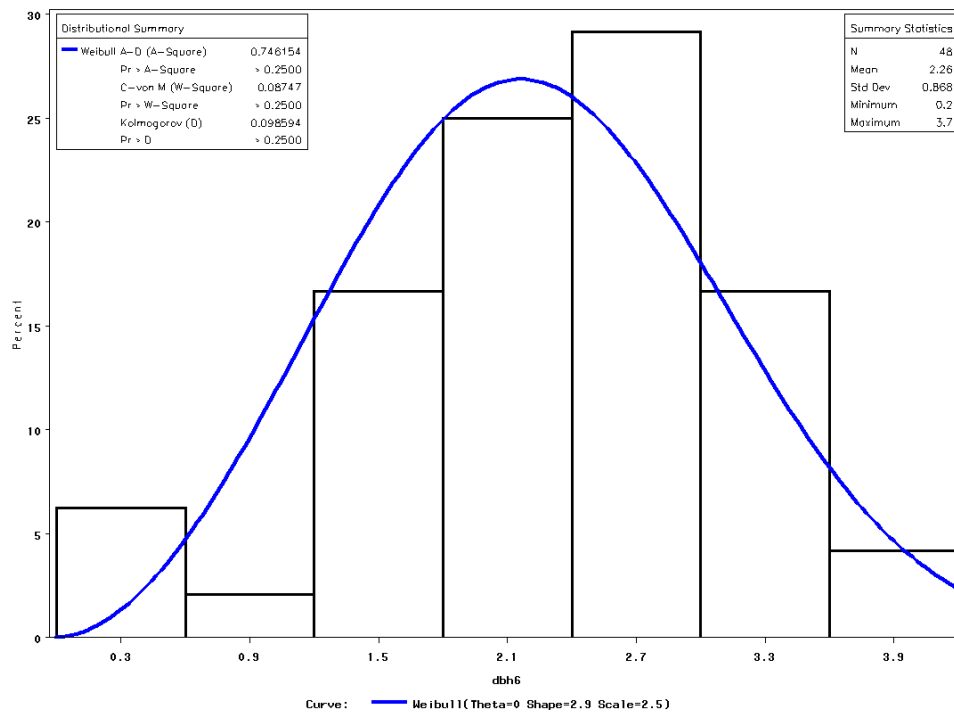
Diameter Distribution, Stand 128, DBH at Age 11



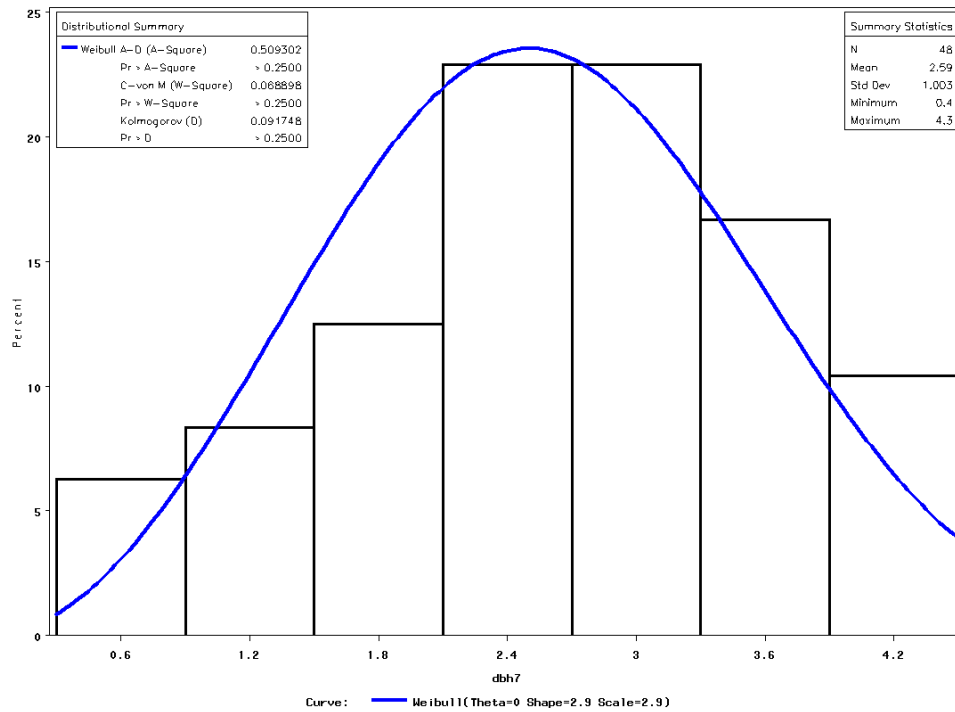
Diameter Distribution, Stand 145, DBH at Age 5



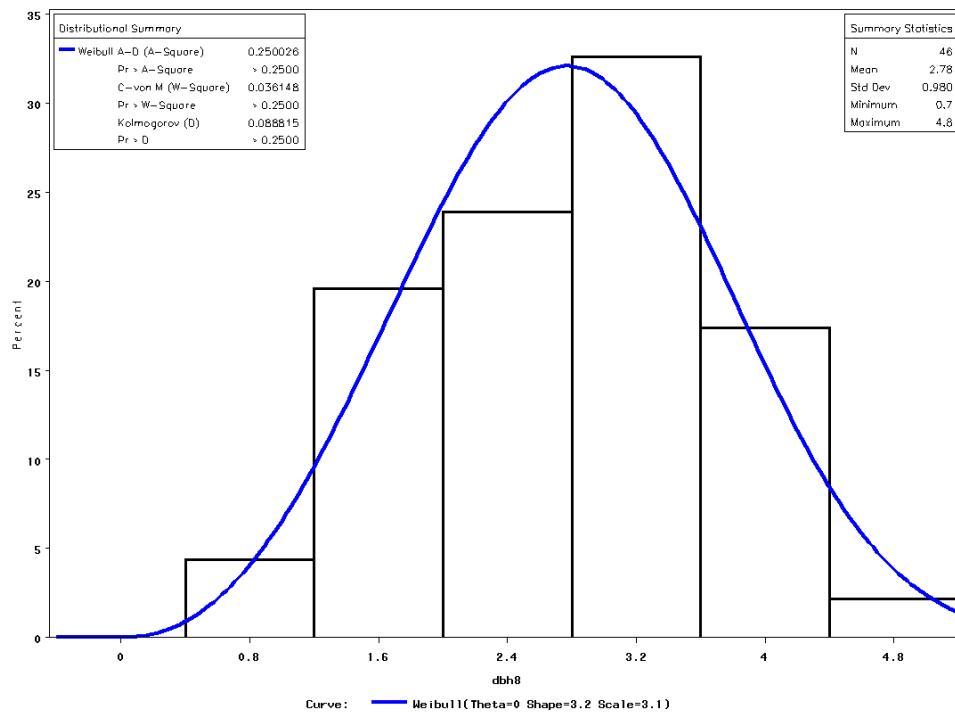
Diameter Distribution, Stand 145, DBH at Age 6



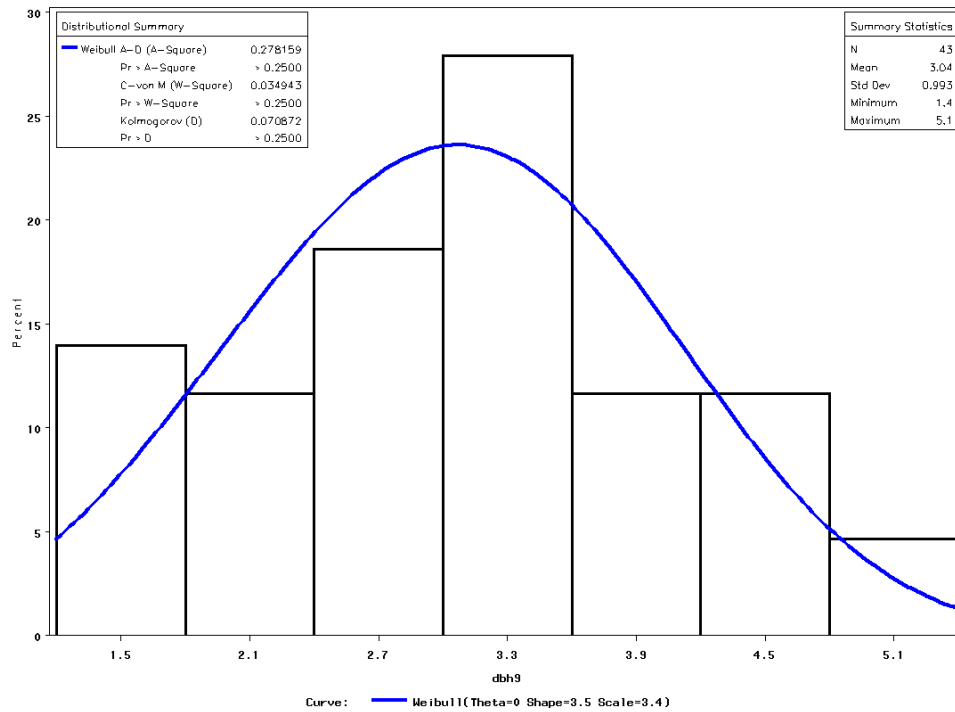
Diameter Distribution, Stand 145, DBH at Age 7



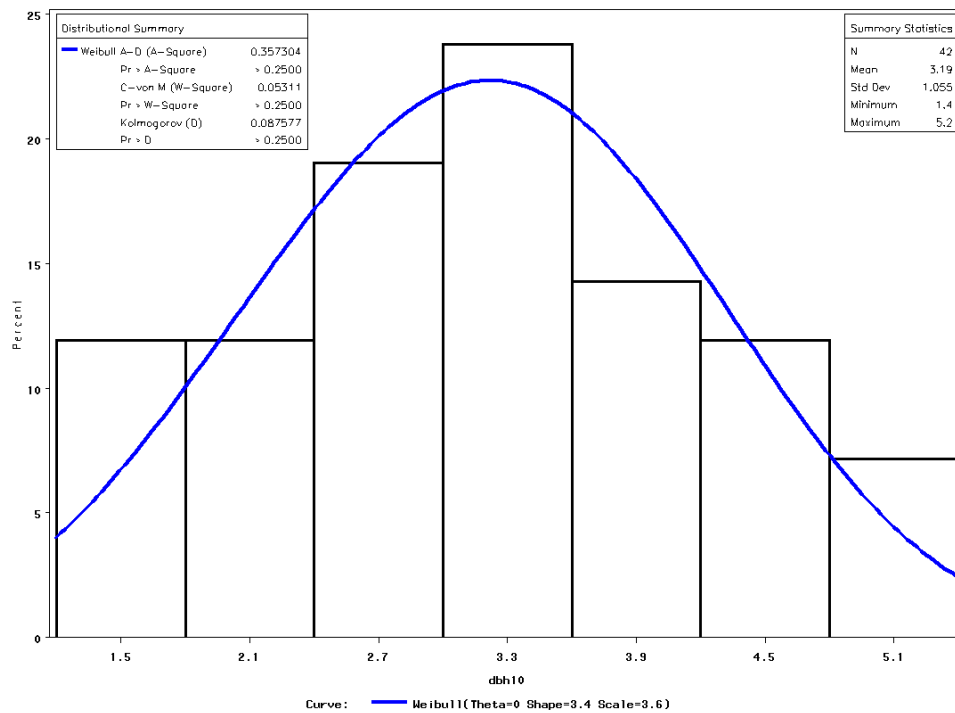
Diameter Distribution, Stand 145, DBH at Age 8



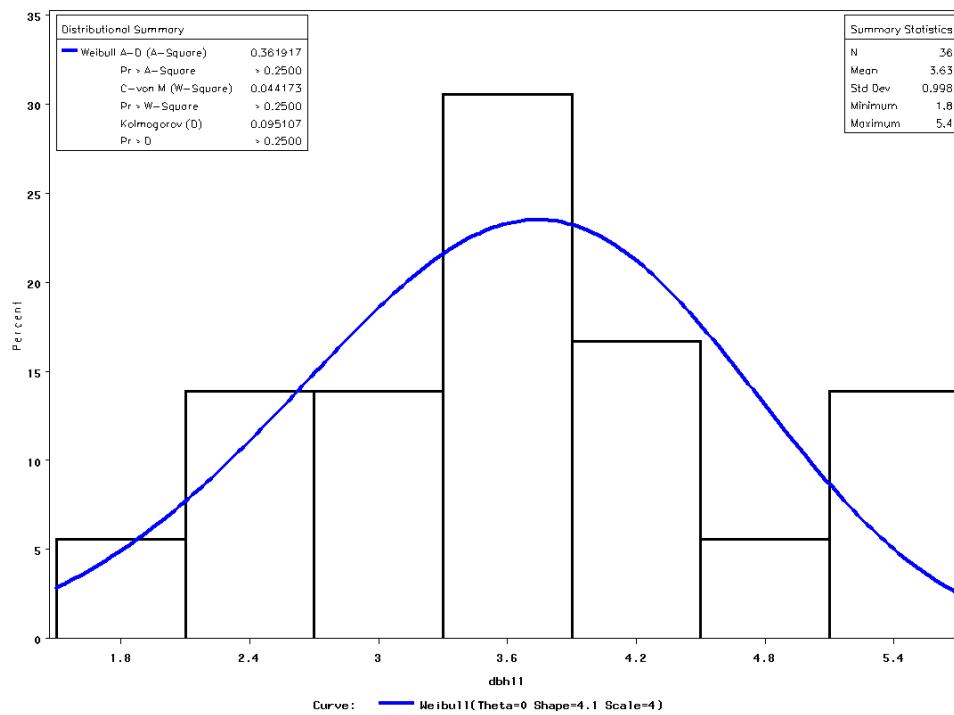
Diameter Distribution, Stand 145, DBH at Age 9



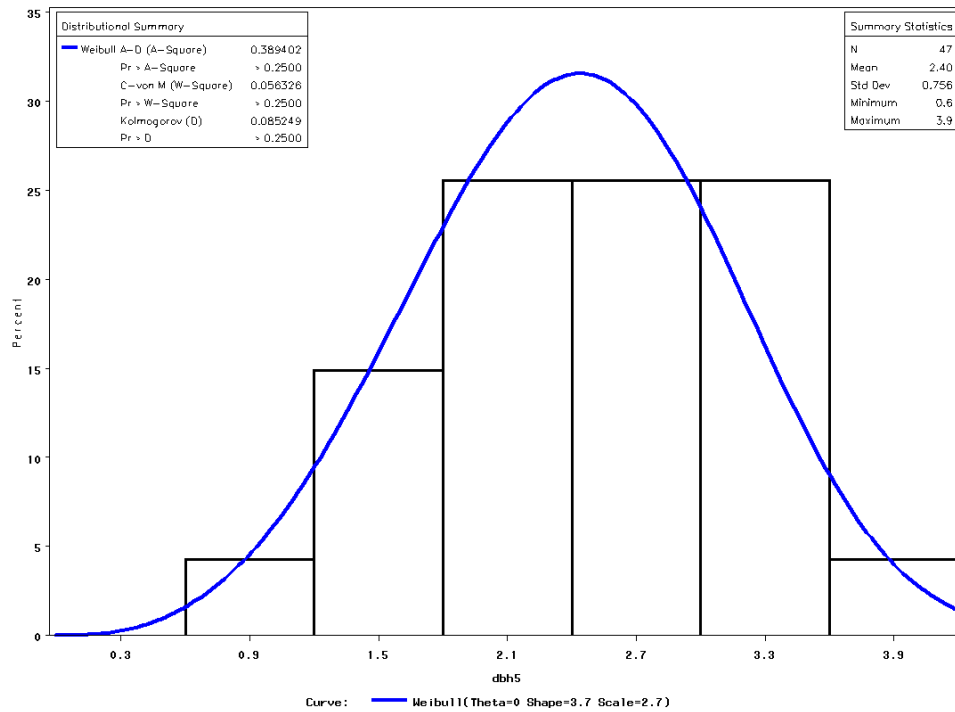
Diameter Distribution, Stand 145, DBH at Age 10



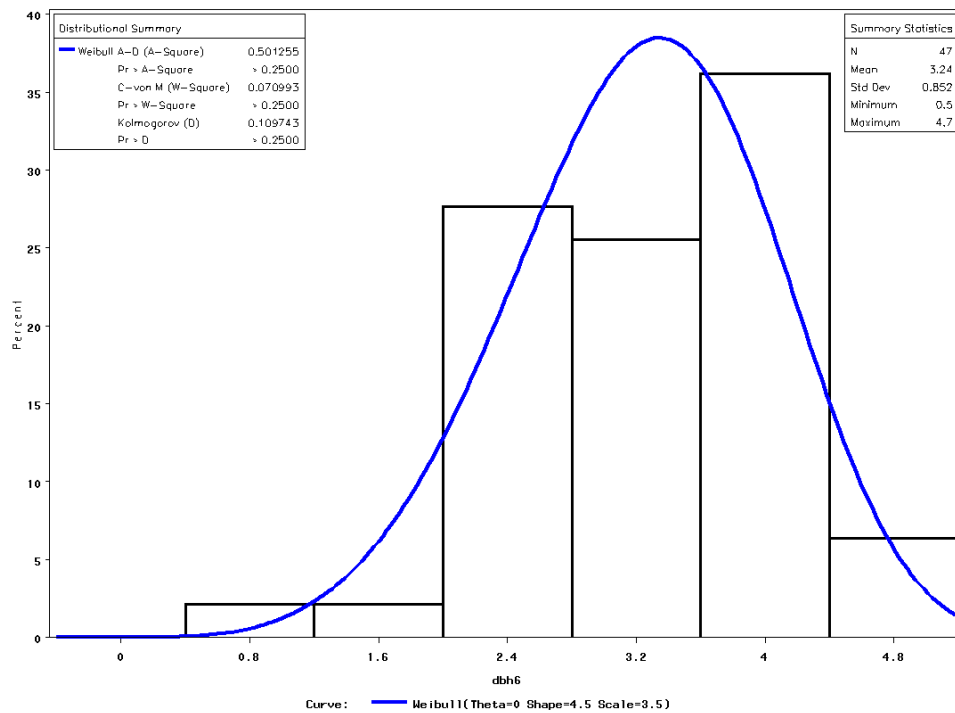
Diameter Distribution, Stand 145, DBH at Age 11



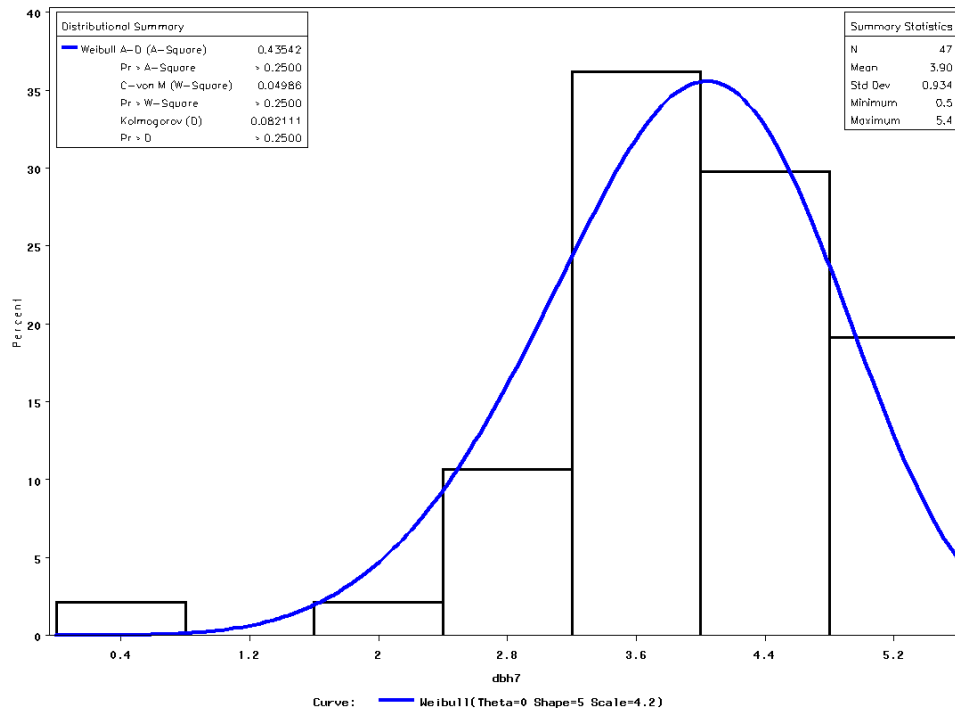
Diameter Distribution, Stand 158, DBH at Age 5



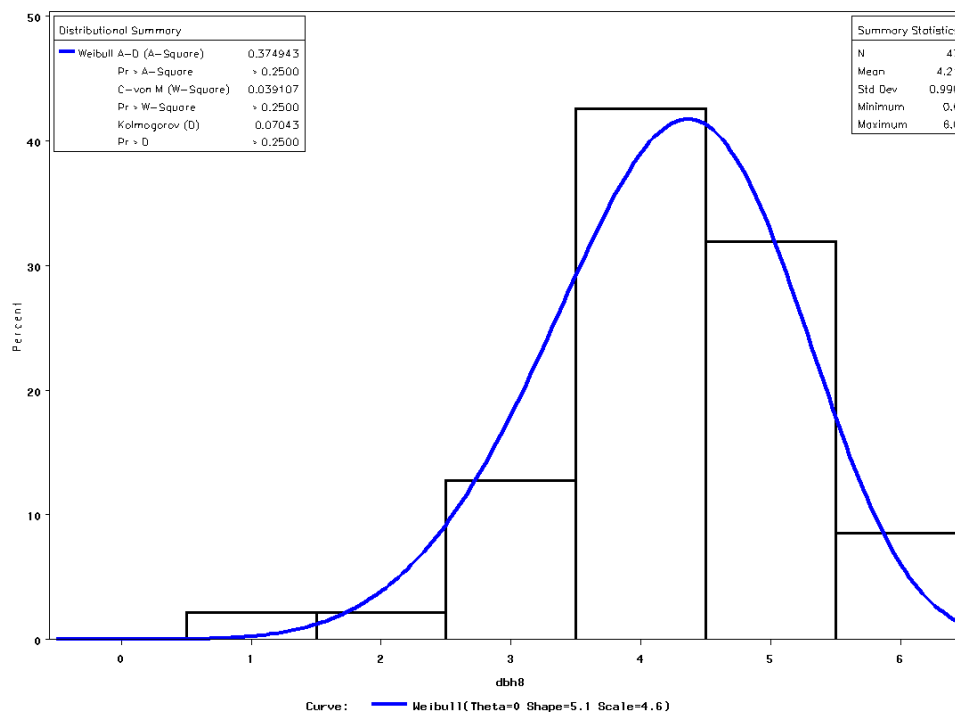
Diameter Distribution, Stand 158, DBH at Age 6



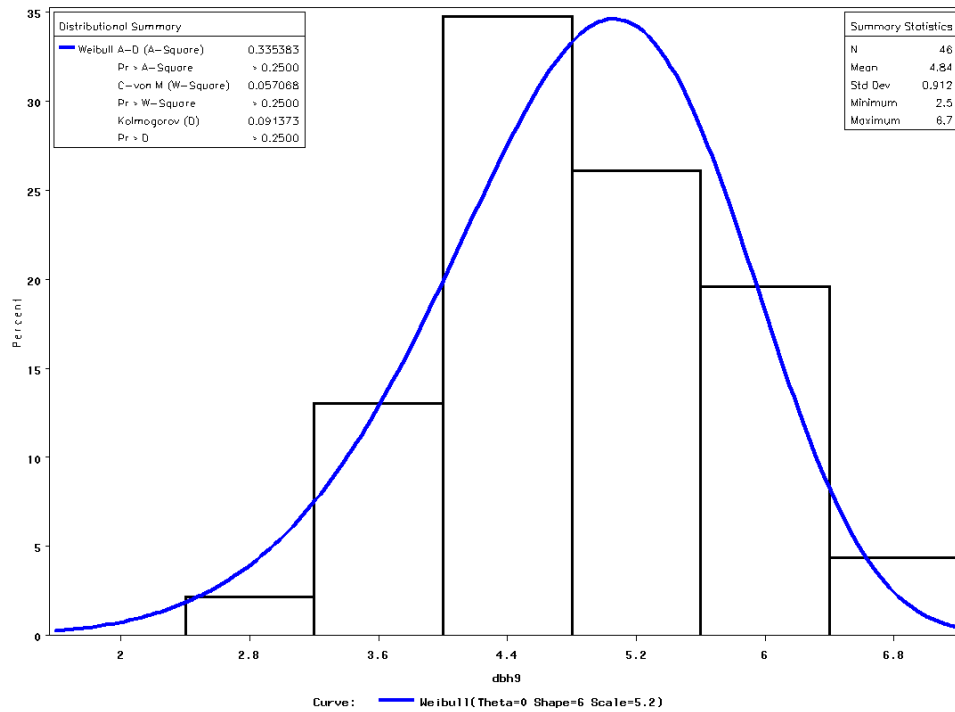
Diameter Distribution, Stand 158, DBH at Age 7



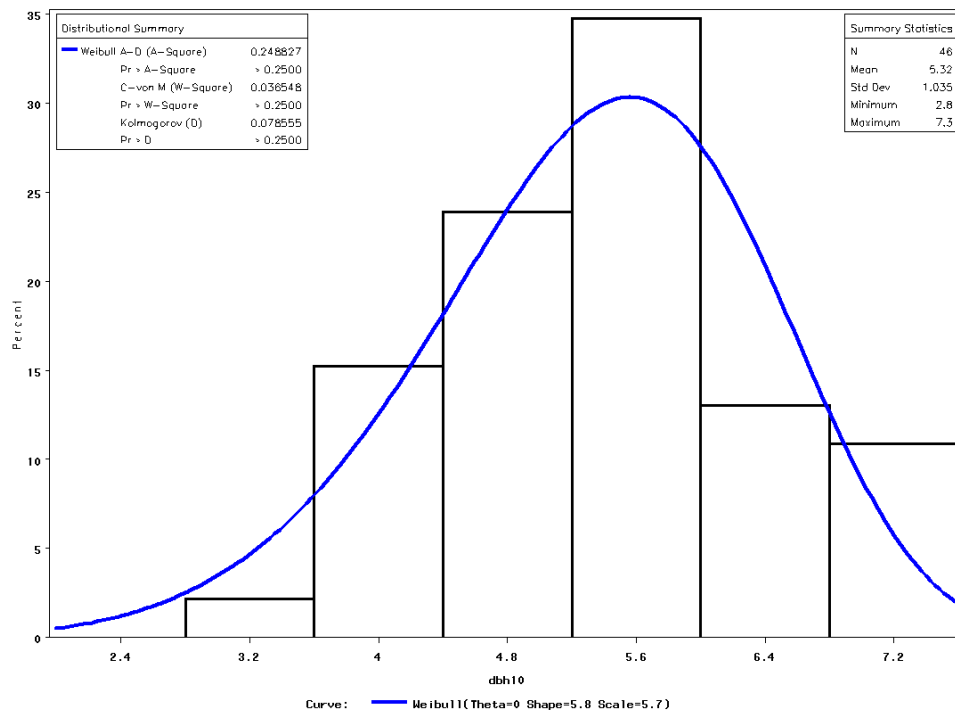
Diameter Distribution, Stand 158, DBH at Age 8



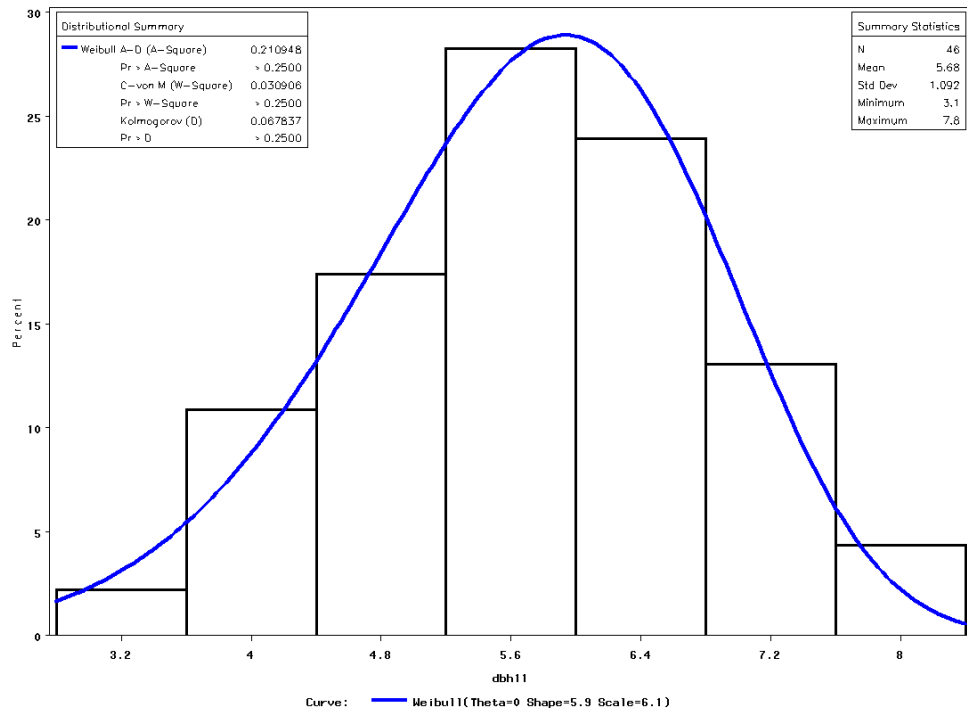
Diameter Distribution, Stand 158, DBH at Age 9



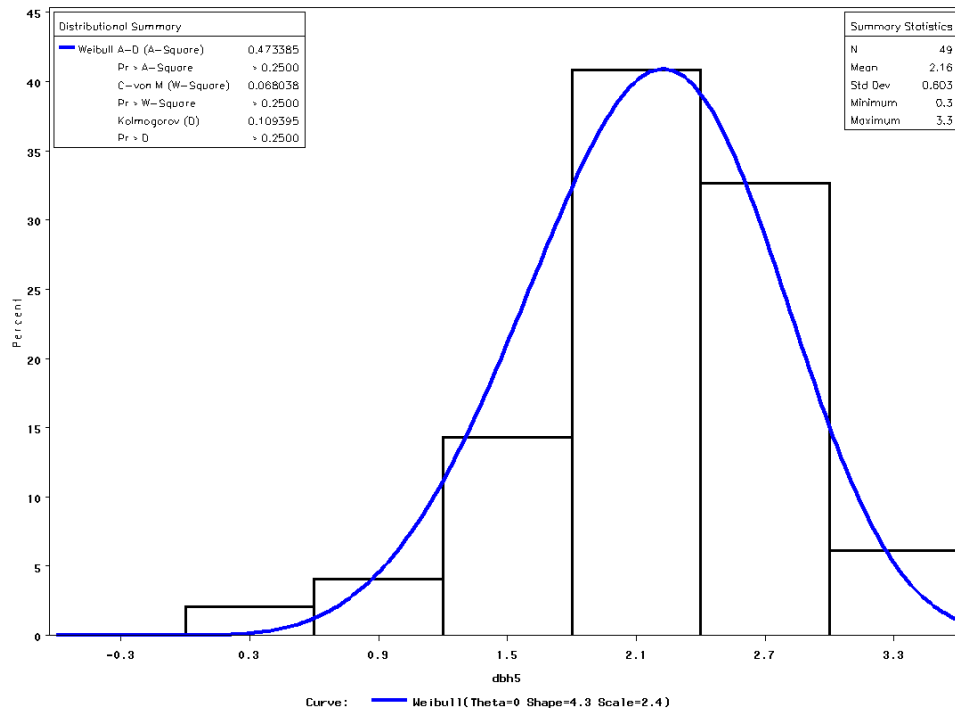
Diameter Distribution, Stand 158, DBH at Age 10



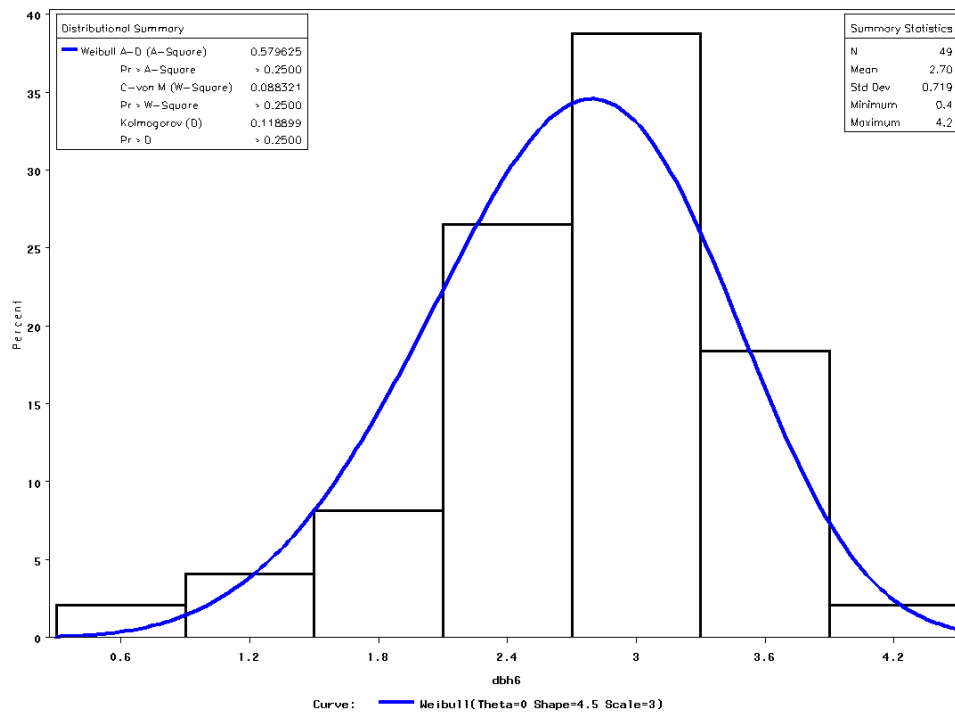
Diameter Distribution, Stand 158, DBH at Age 11



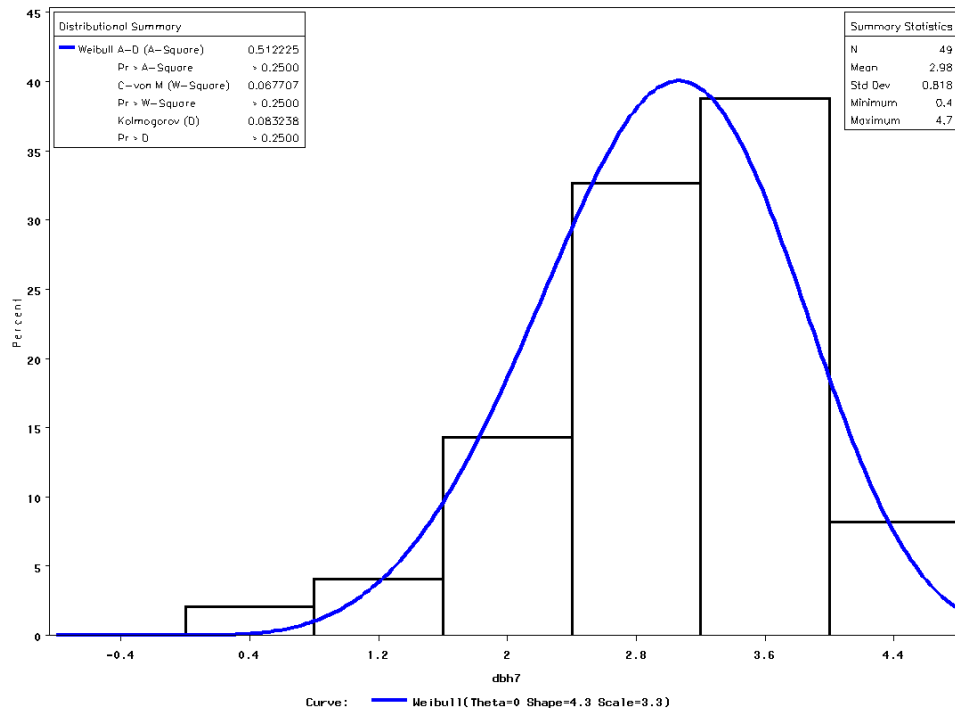
Diameter Distribution, Stand 181, DBH at Age 5



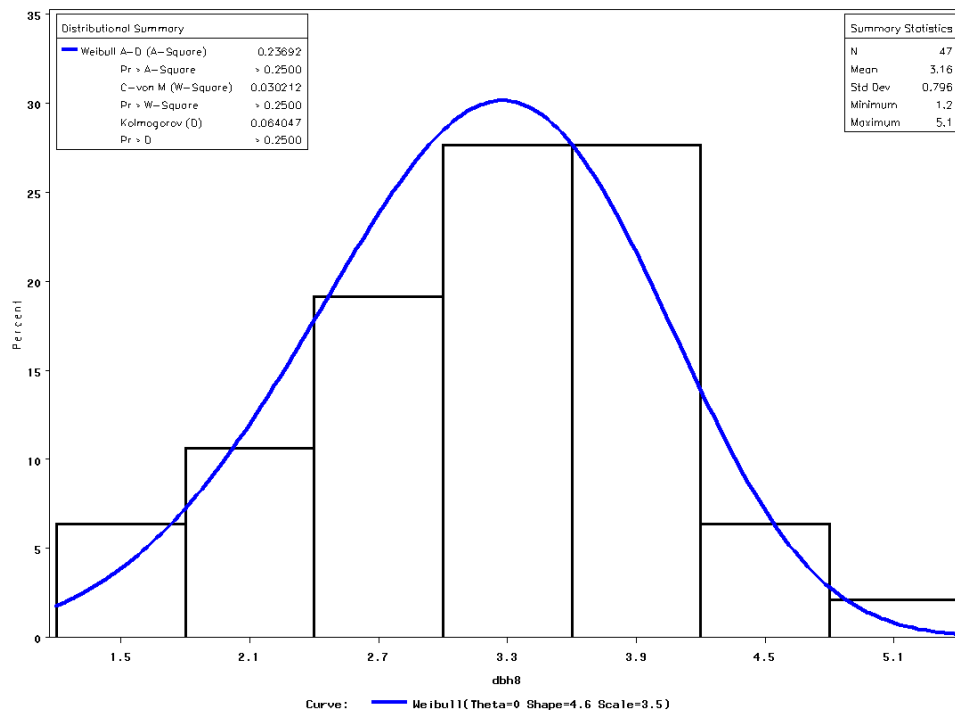
Diameter Distribution, Stand 181, DBH at Age 6



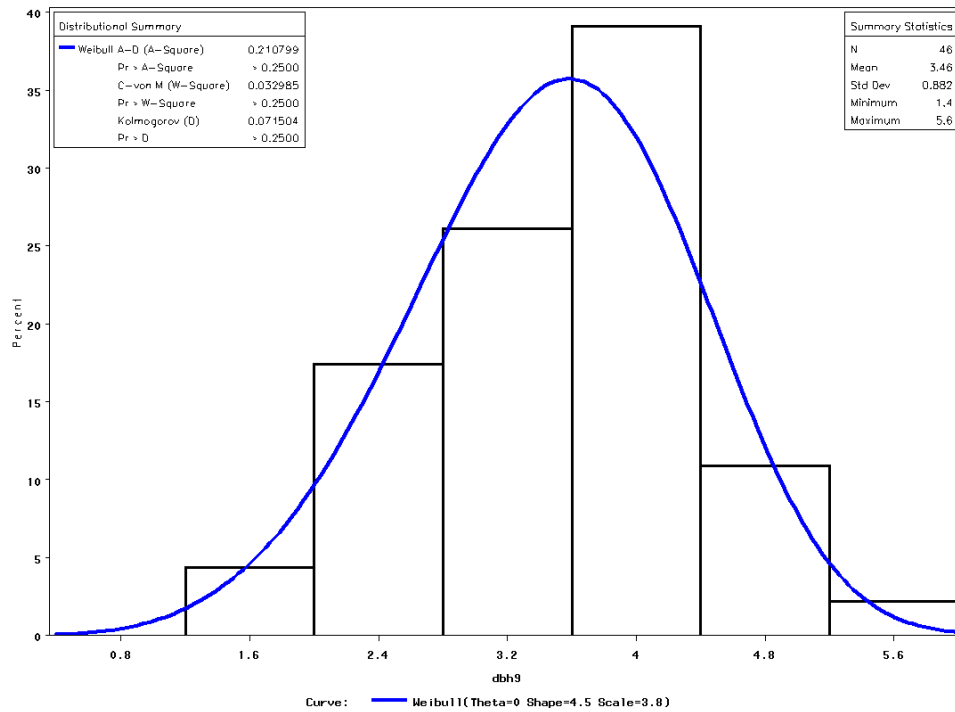
Diameter Distribution, Stand 181, DBH at Age 7



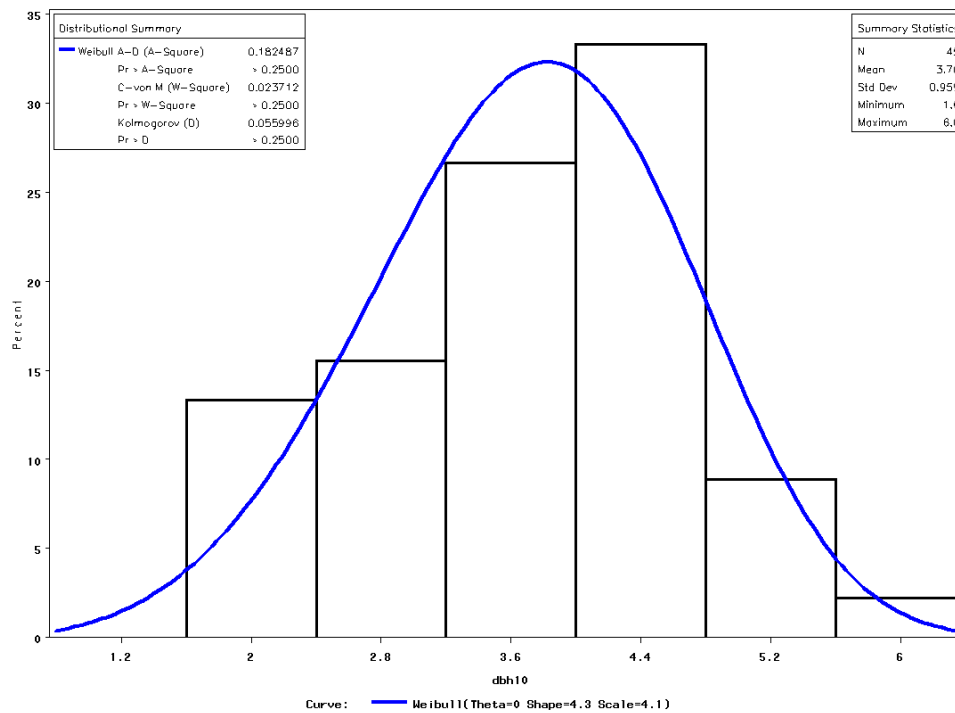
Diameter Distribution, Stand 181, DBH at Age 8



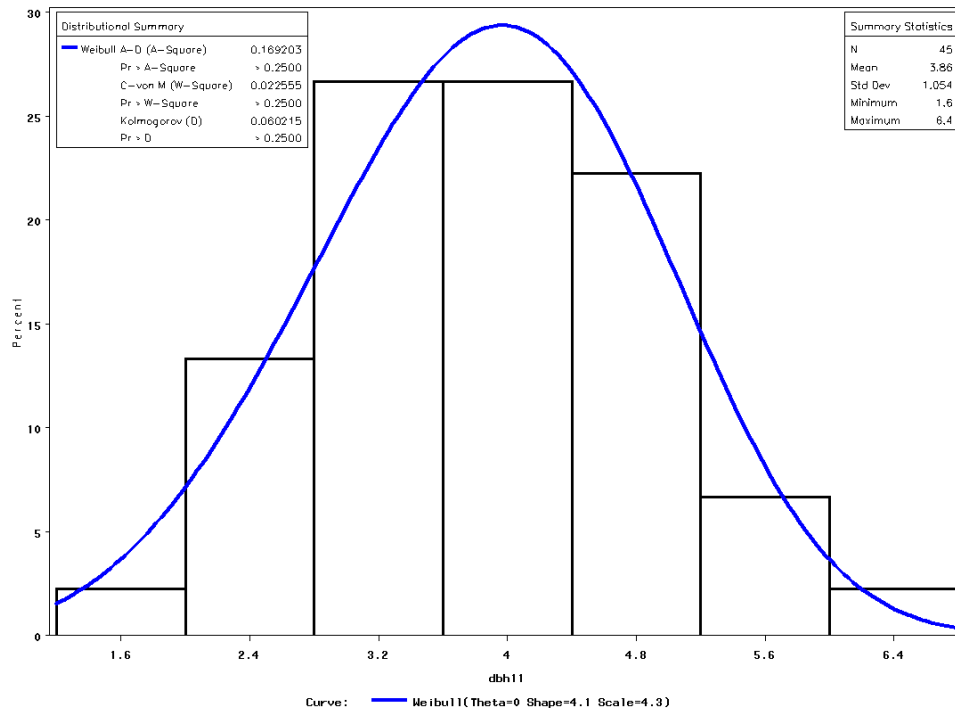
Diameter Distribution, Stand 181, DBH at Age 9



Diameter Distribution, Stand 181, DBH at Age 10

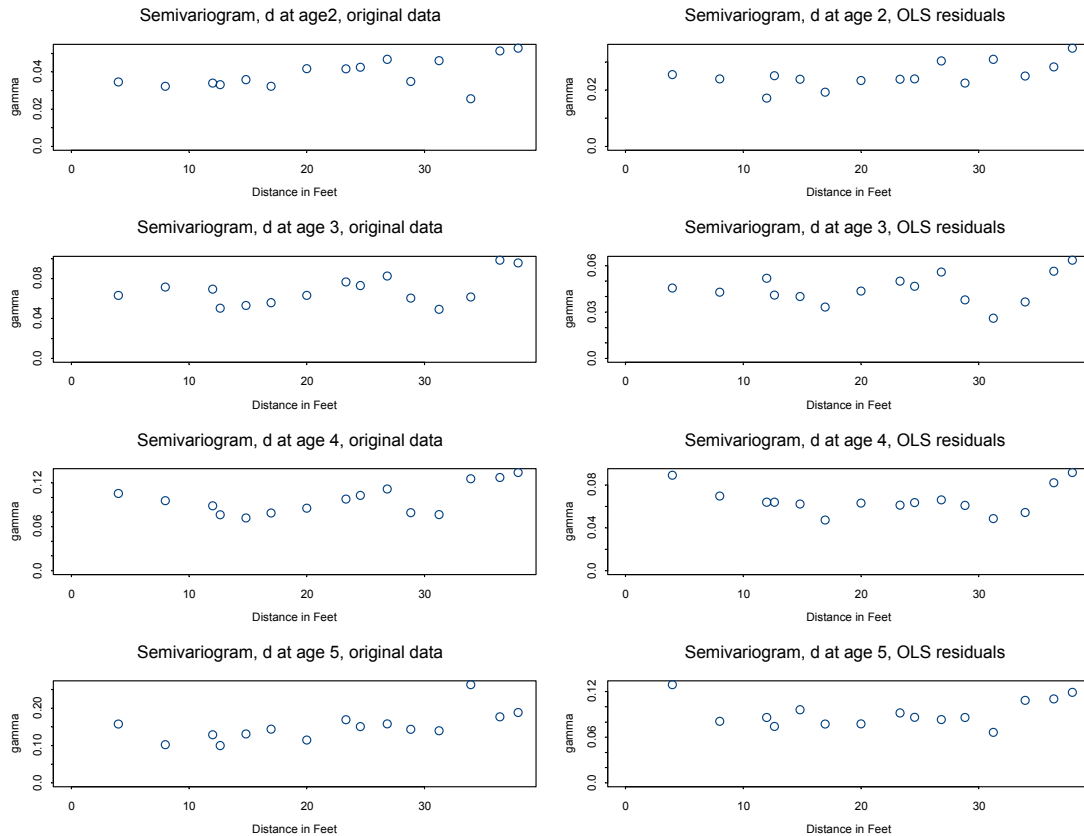


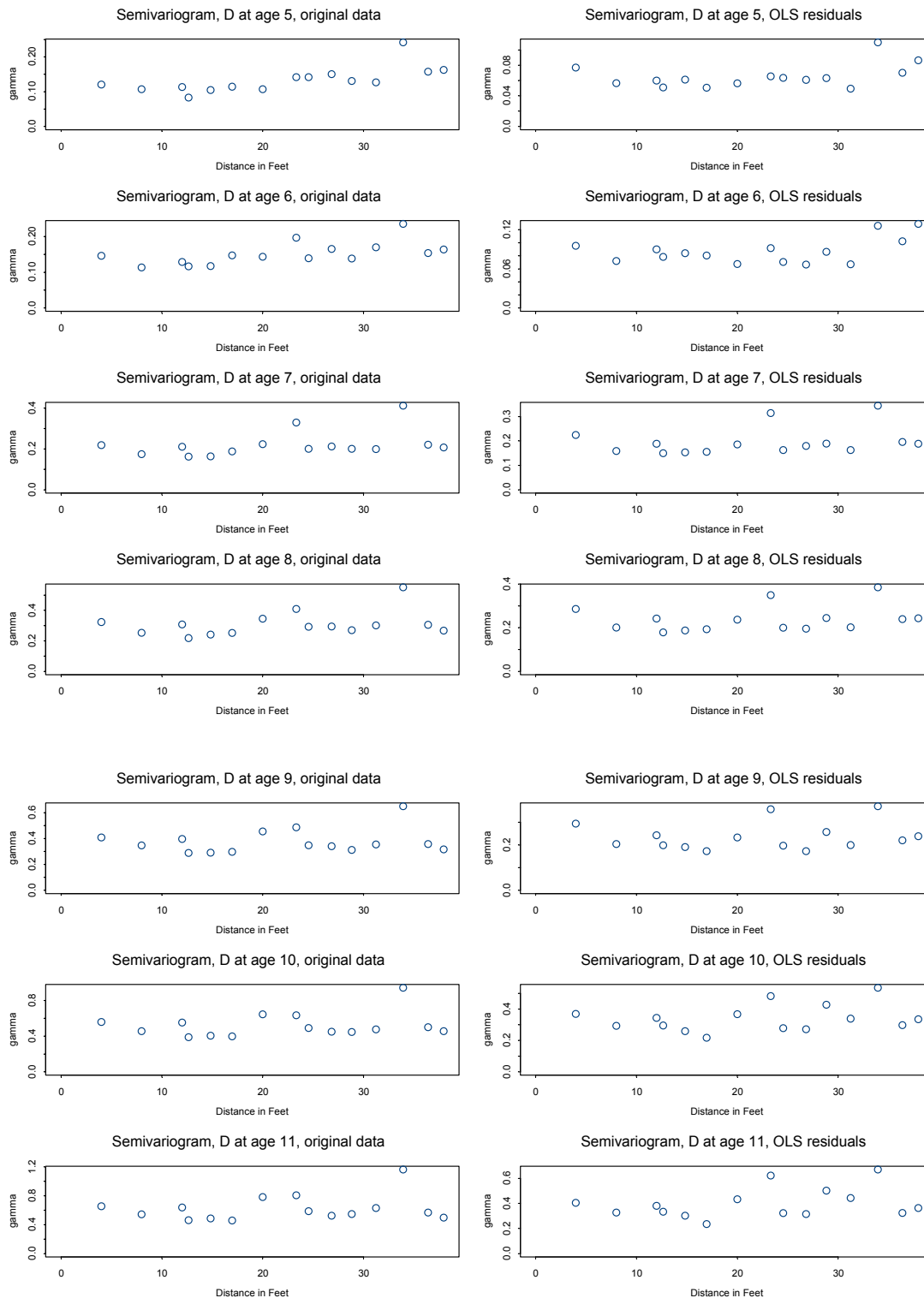
Diameter Distribution, Stand 181, DBH at Age 11



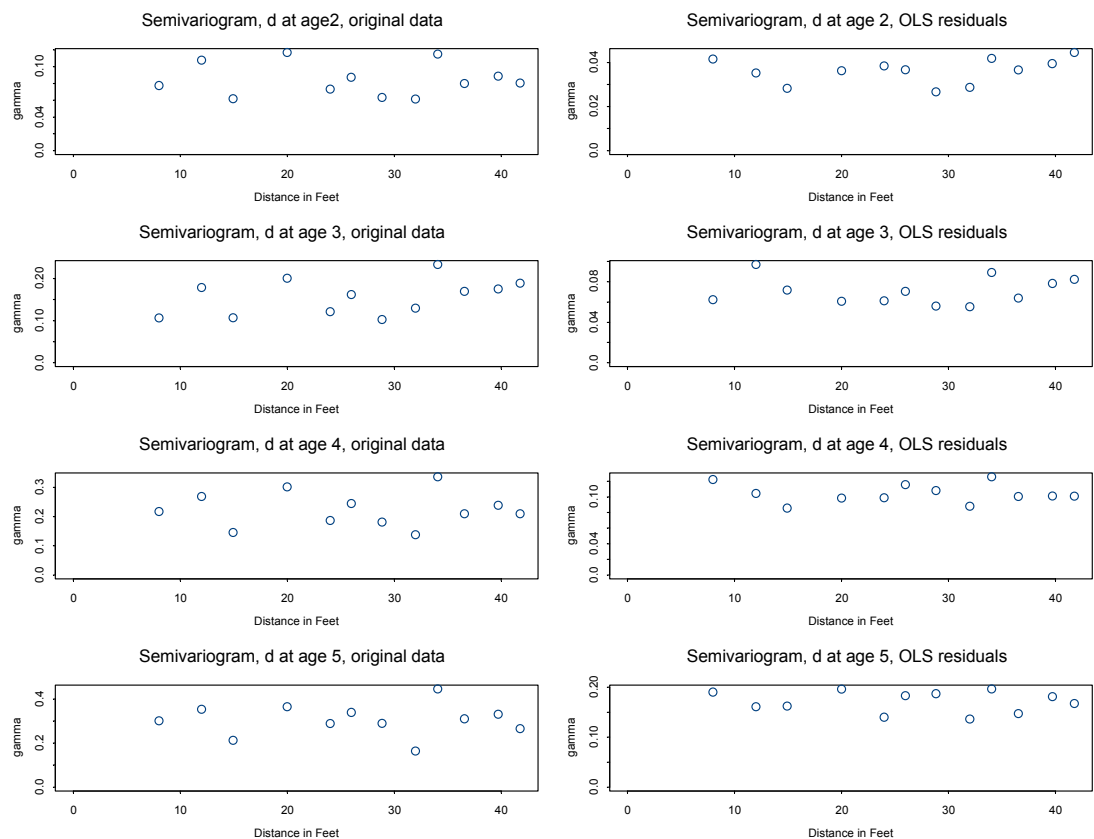
Appendix II: Empirical semivariograms for ground-line diameters (d) from ages 2 to 5, and breast height diameters (D) from ages 5 to 11. An example is presented from each planting density (trees per acre (TPA)) employed in this research. The empirical semivariograms are derived for the original data and the residuals from a fitted linear model.

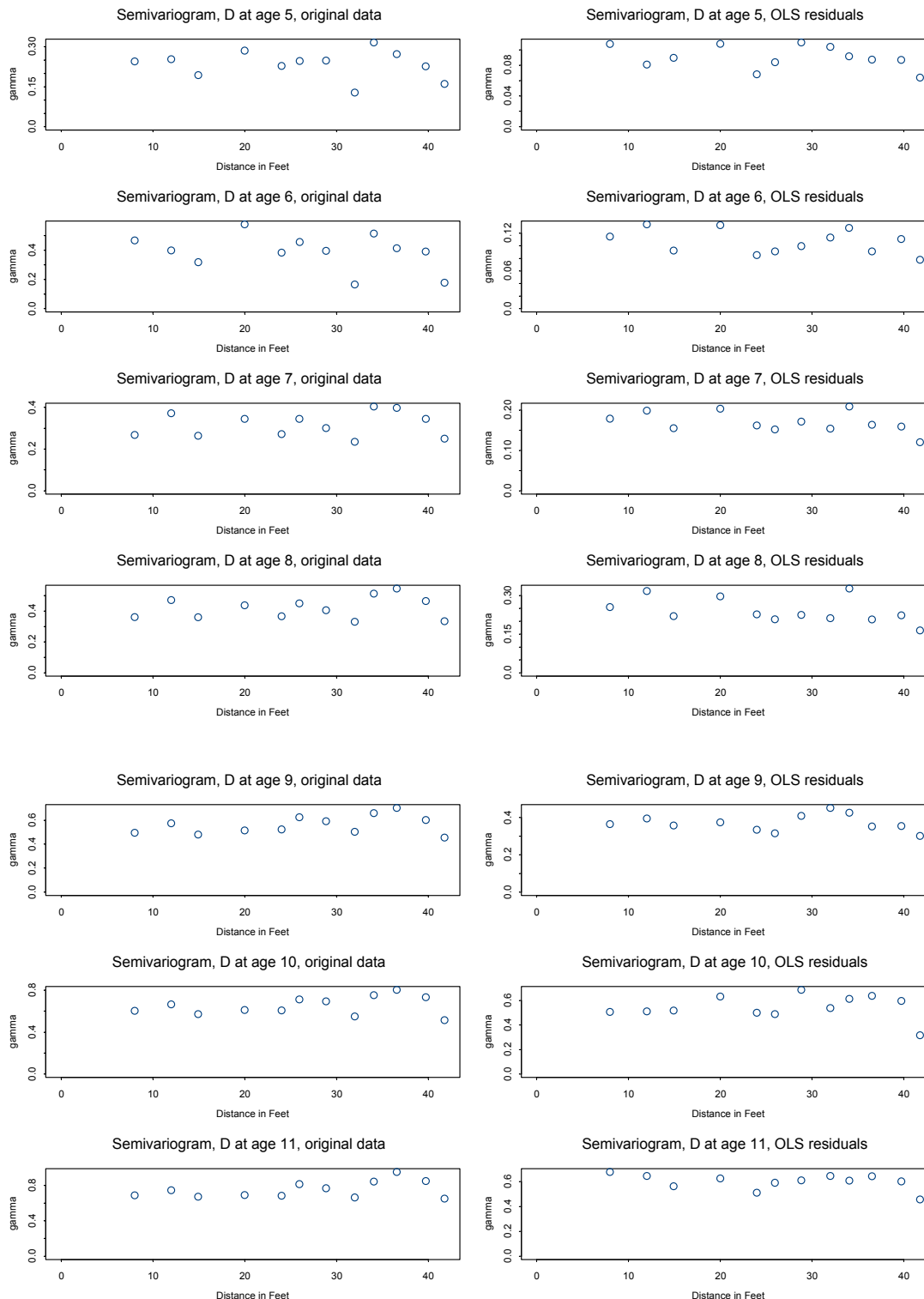
Stand number 20, 907 TPA



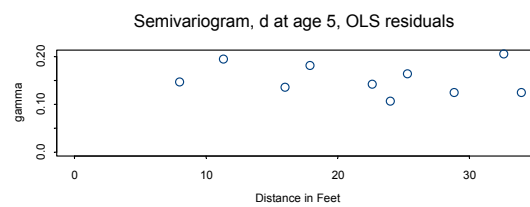
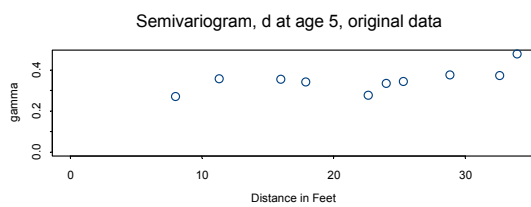
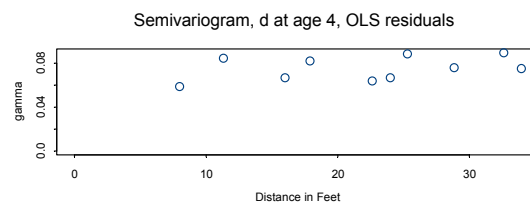
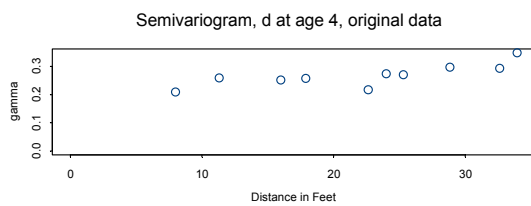
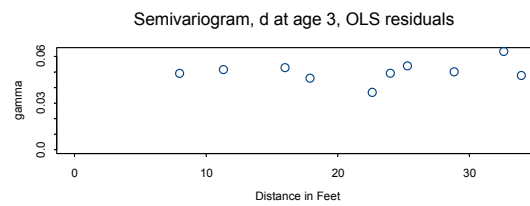
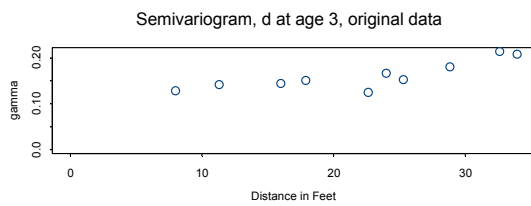
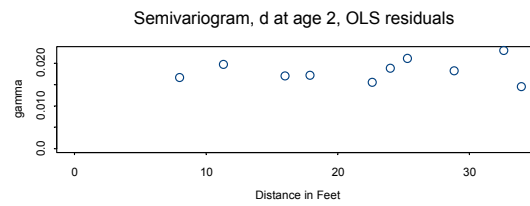
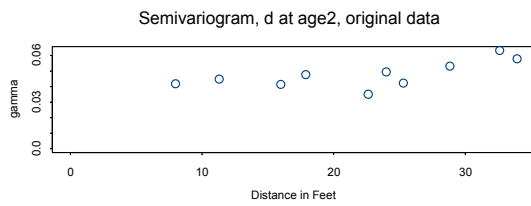


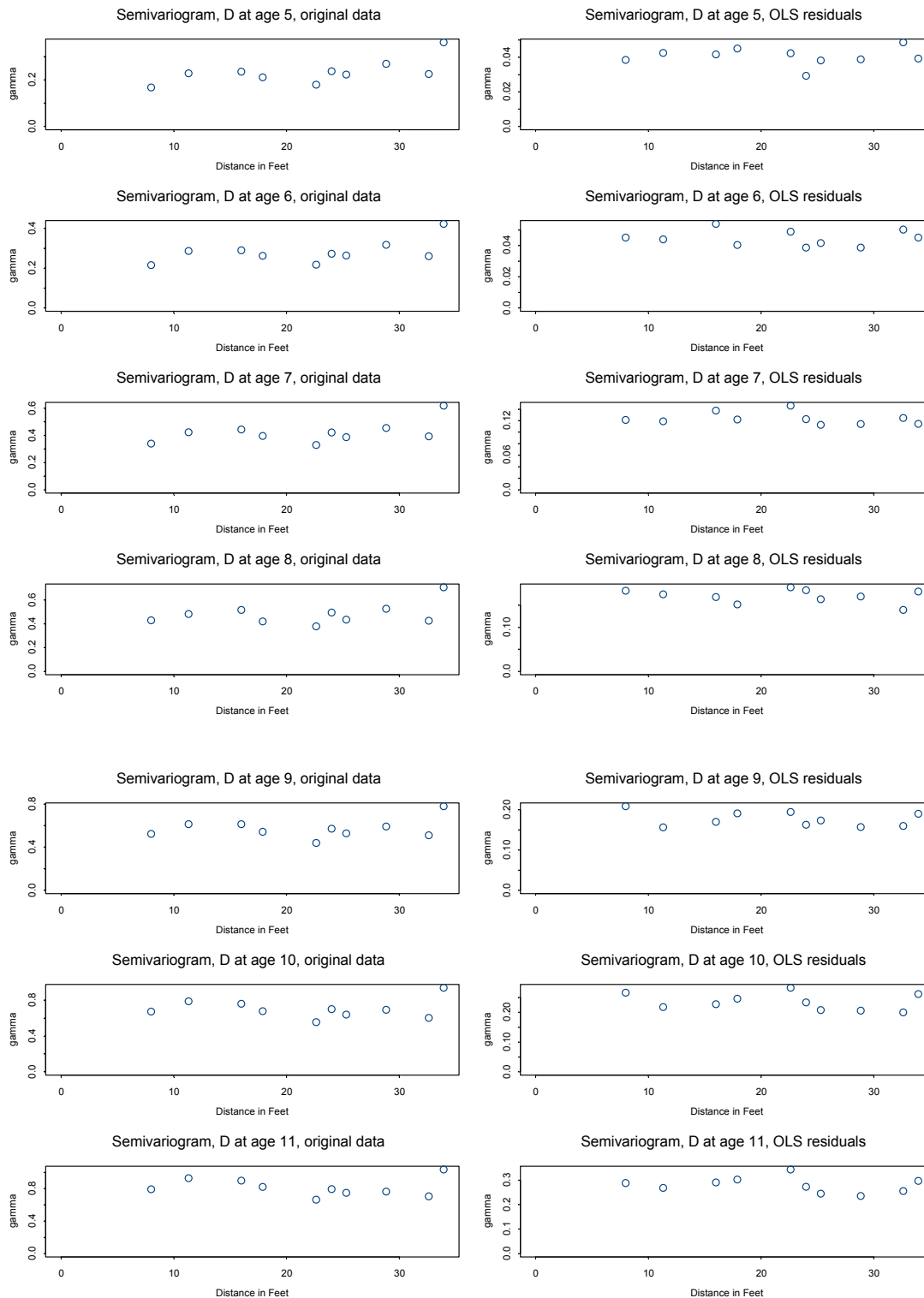
Stand number 44, 453 TPA



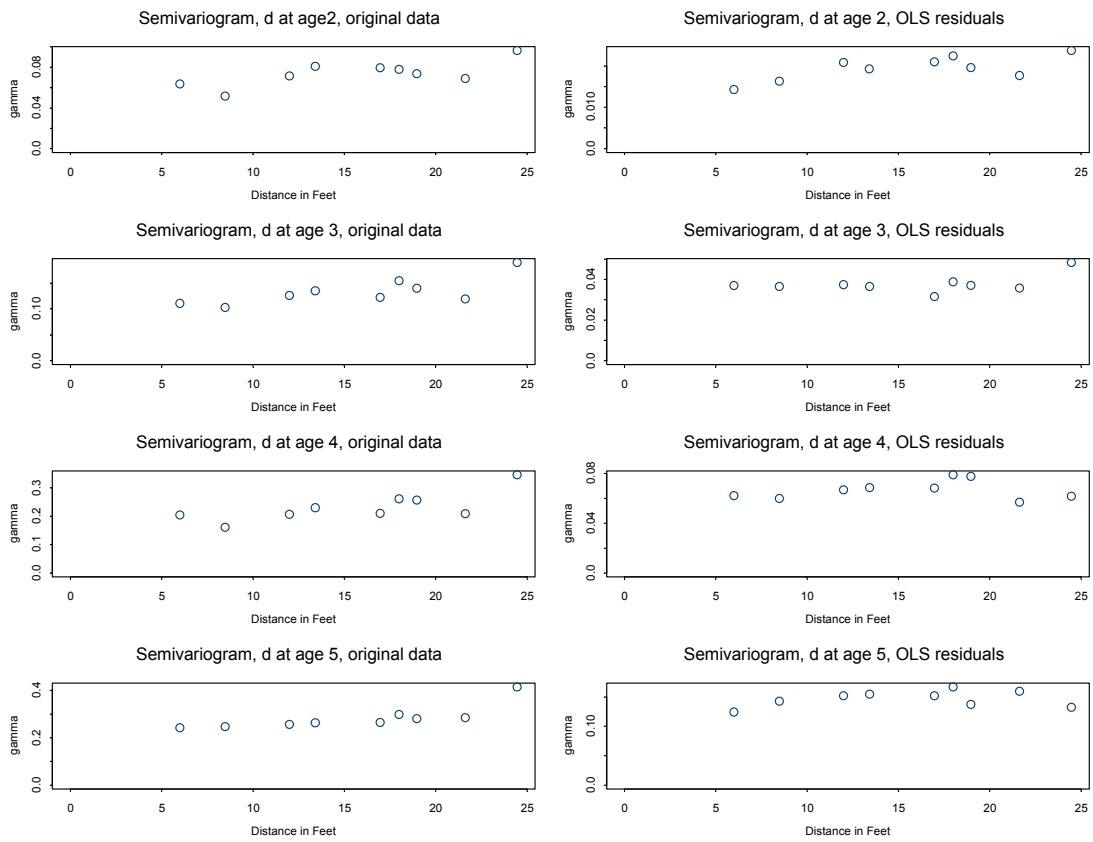


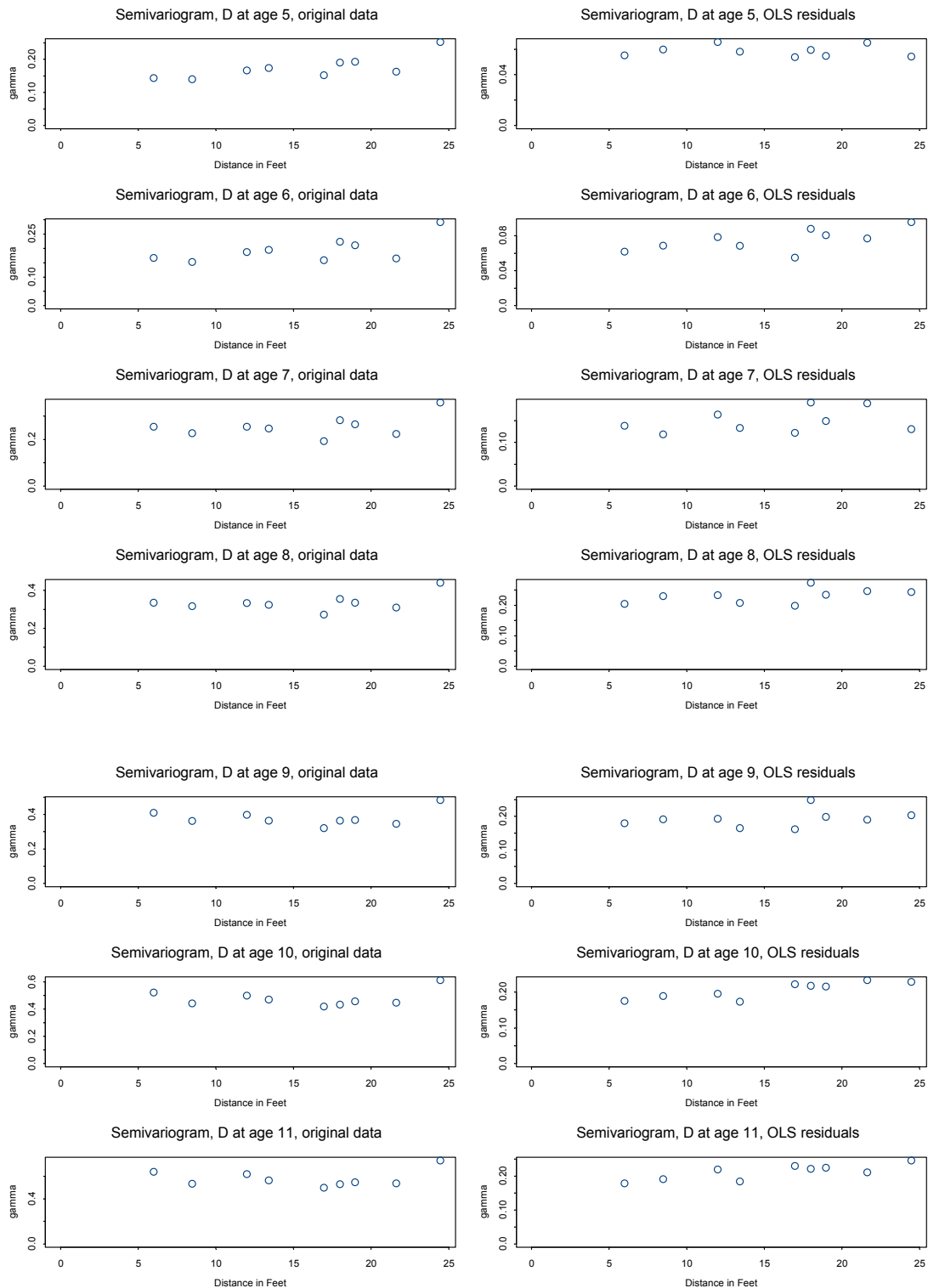
Stand number 59, 680 TPA



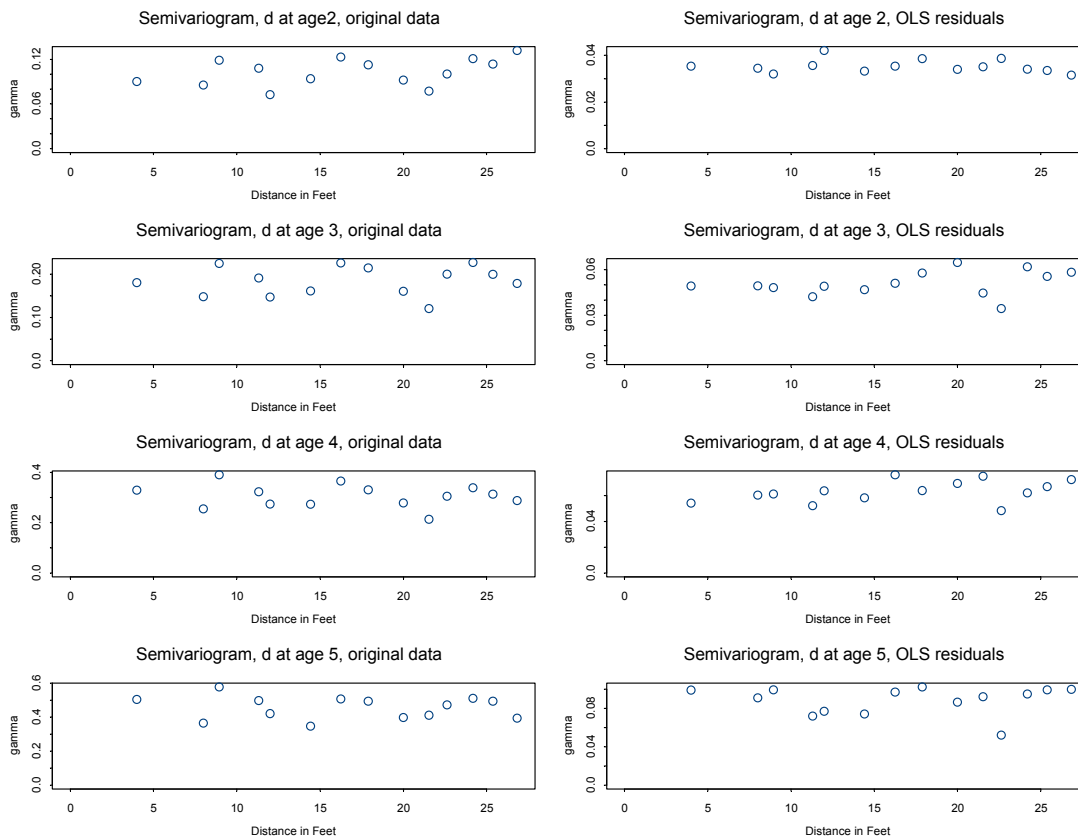


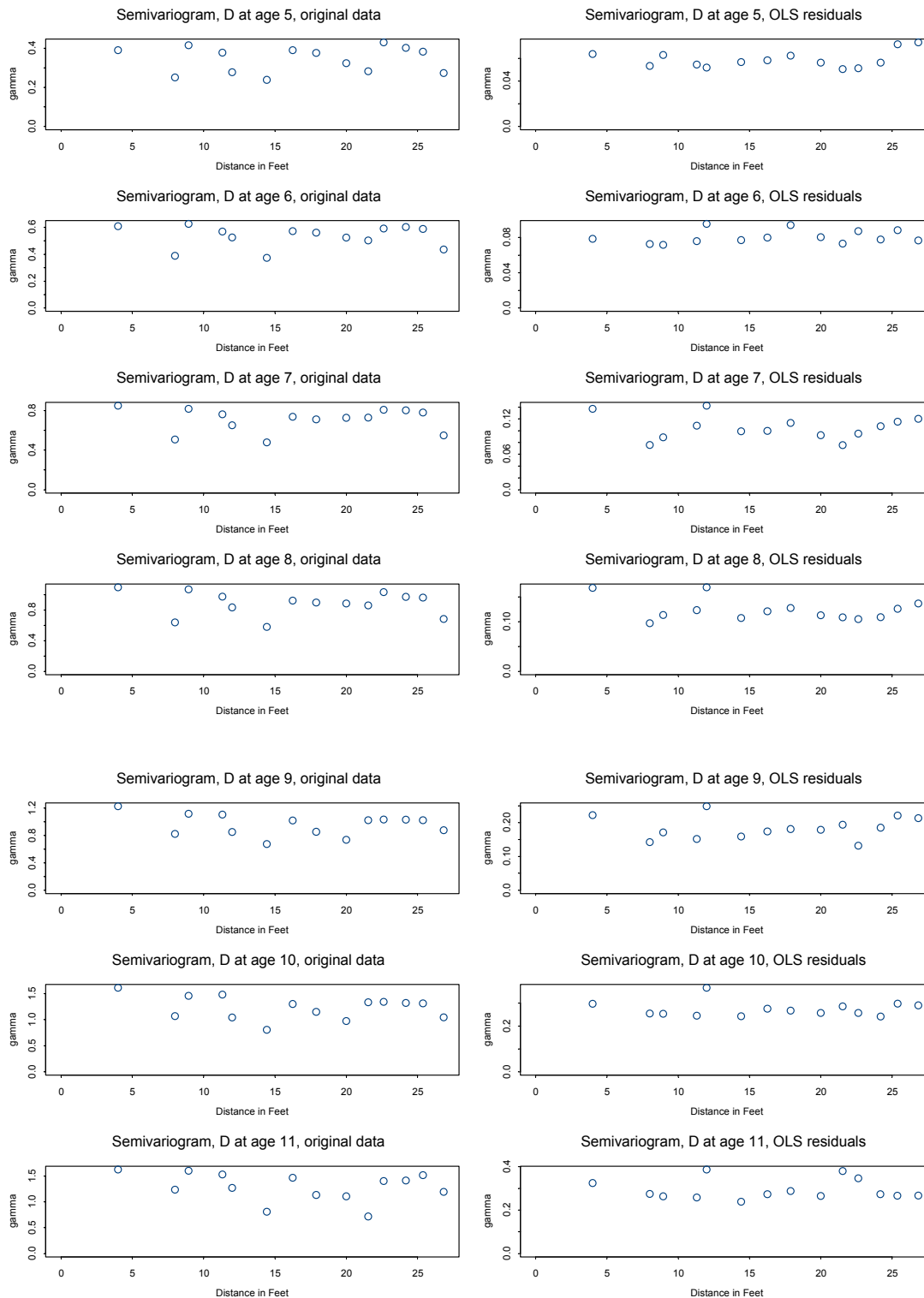
Stand number 86, 1210 TPA



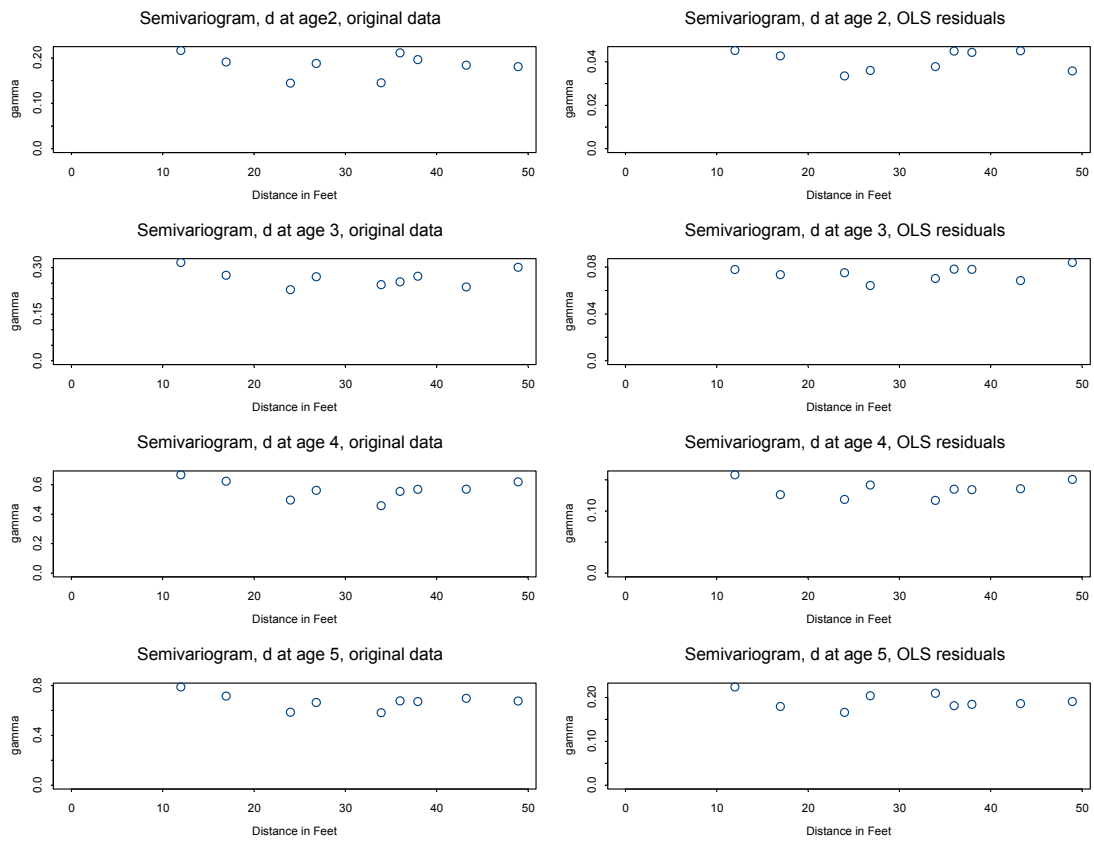


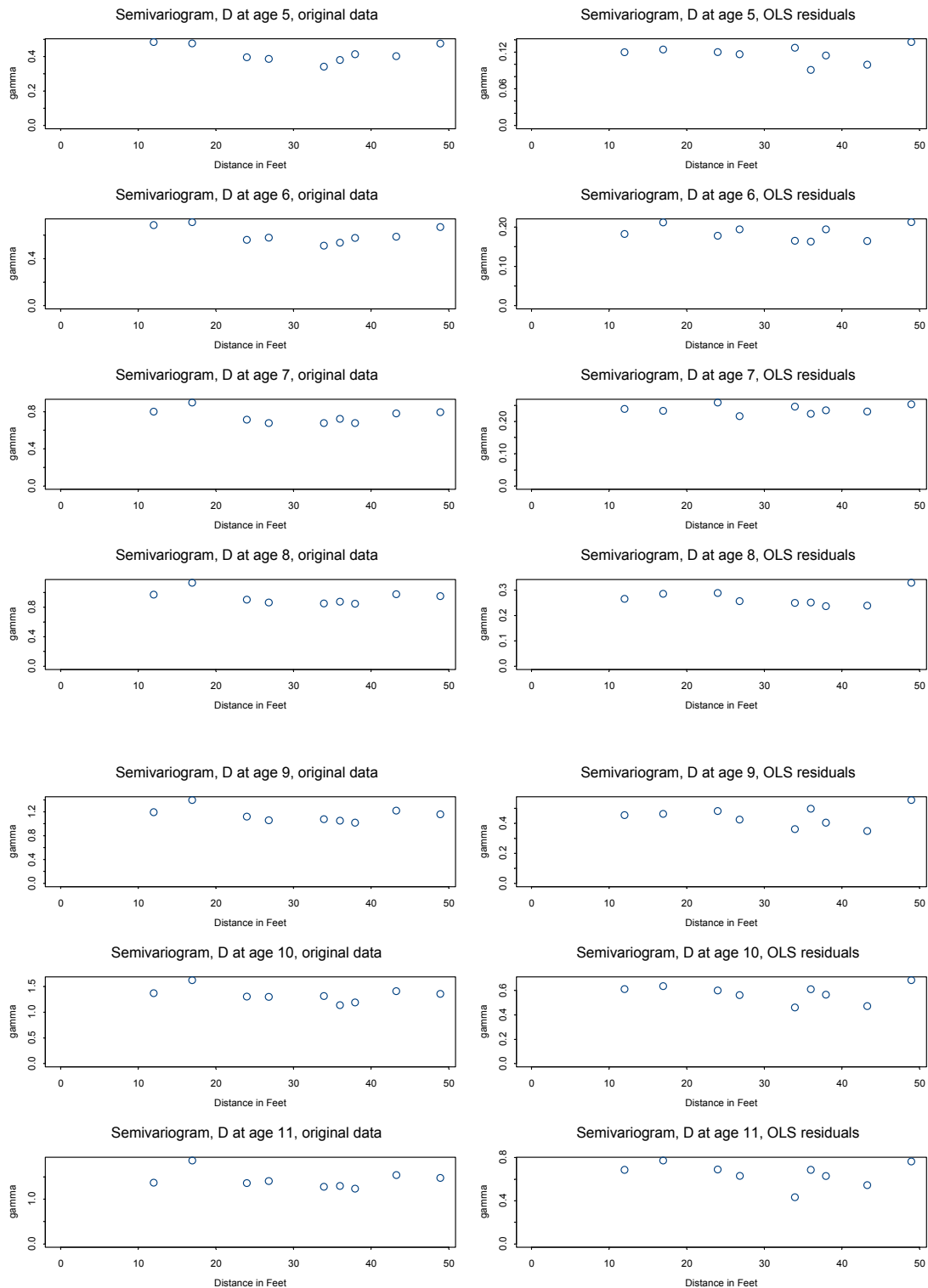
Stand number 105, 1361 TPA



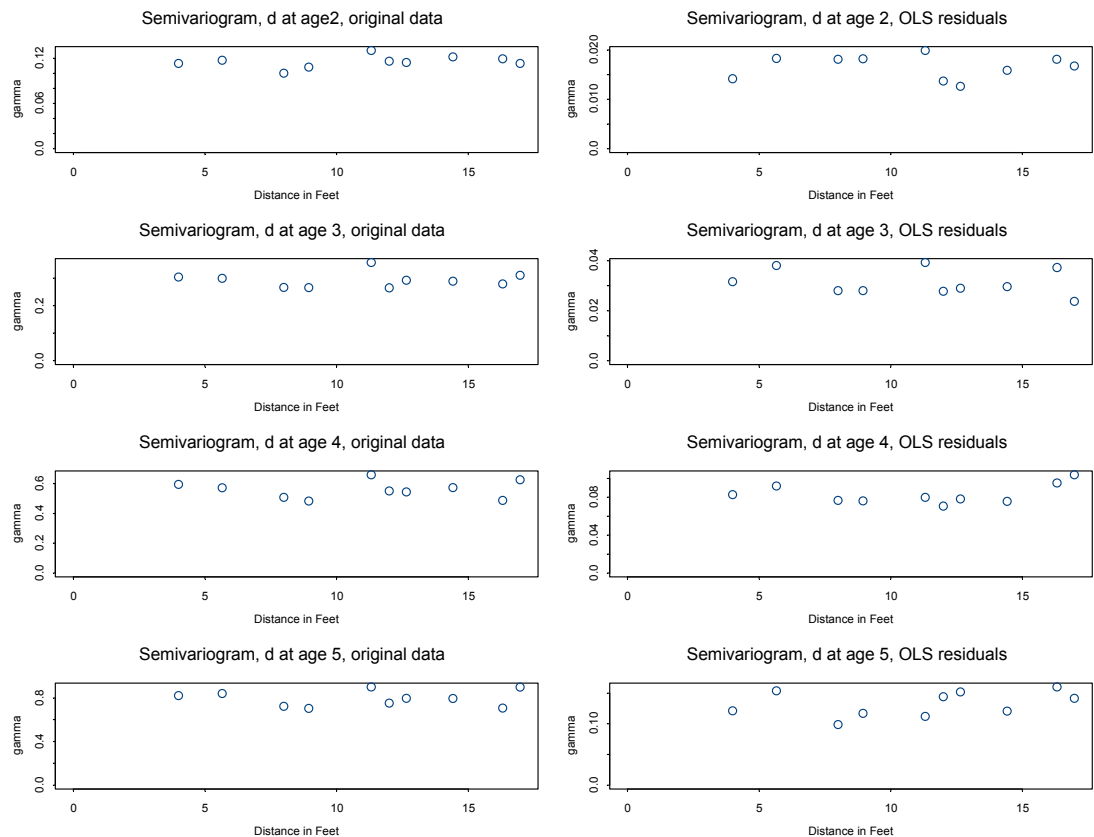


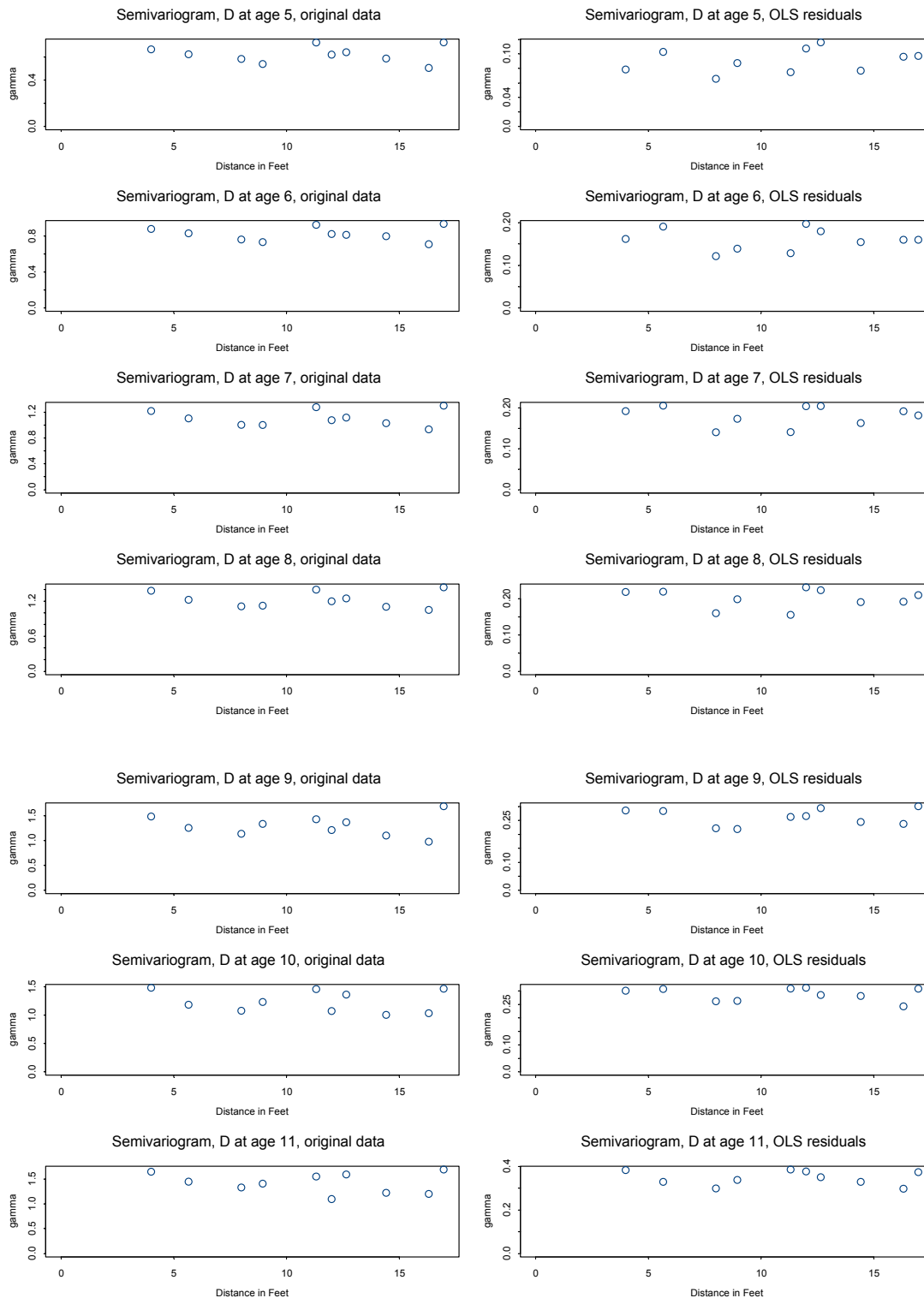
Stand number 128, 302 TPA



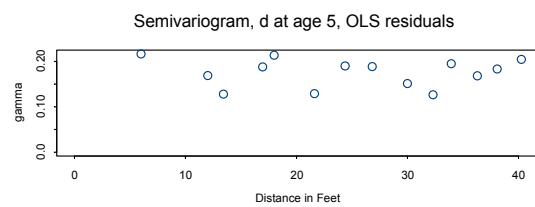
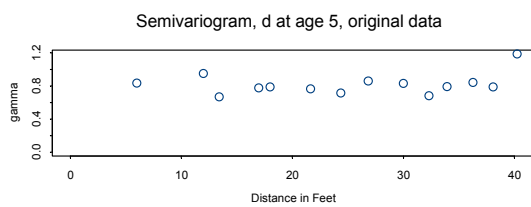
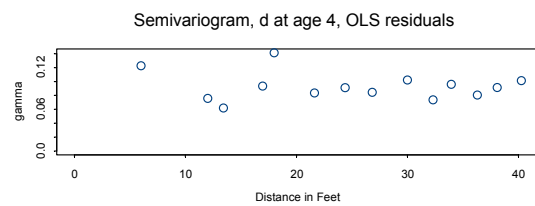
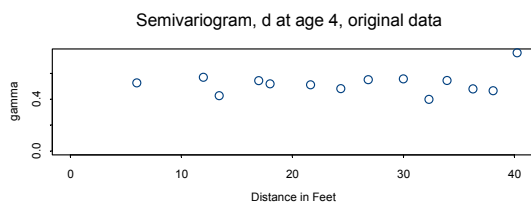
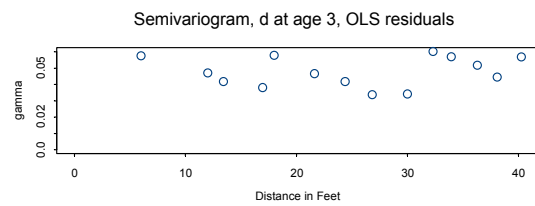
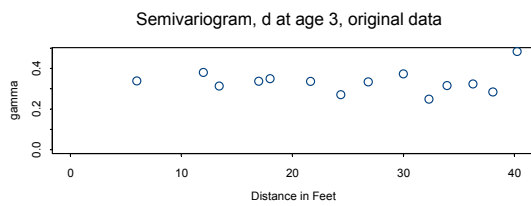
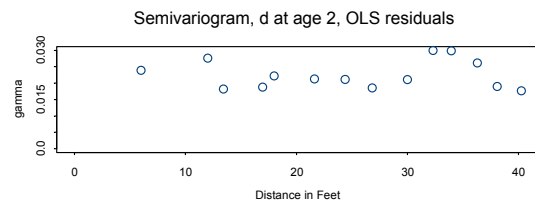
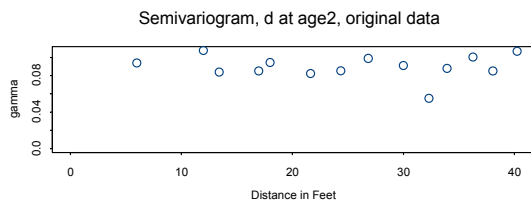


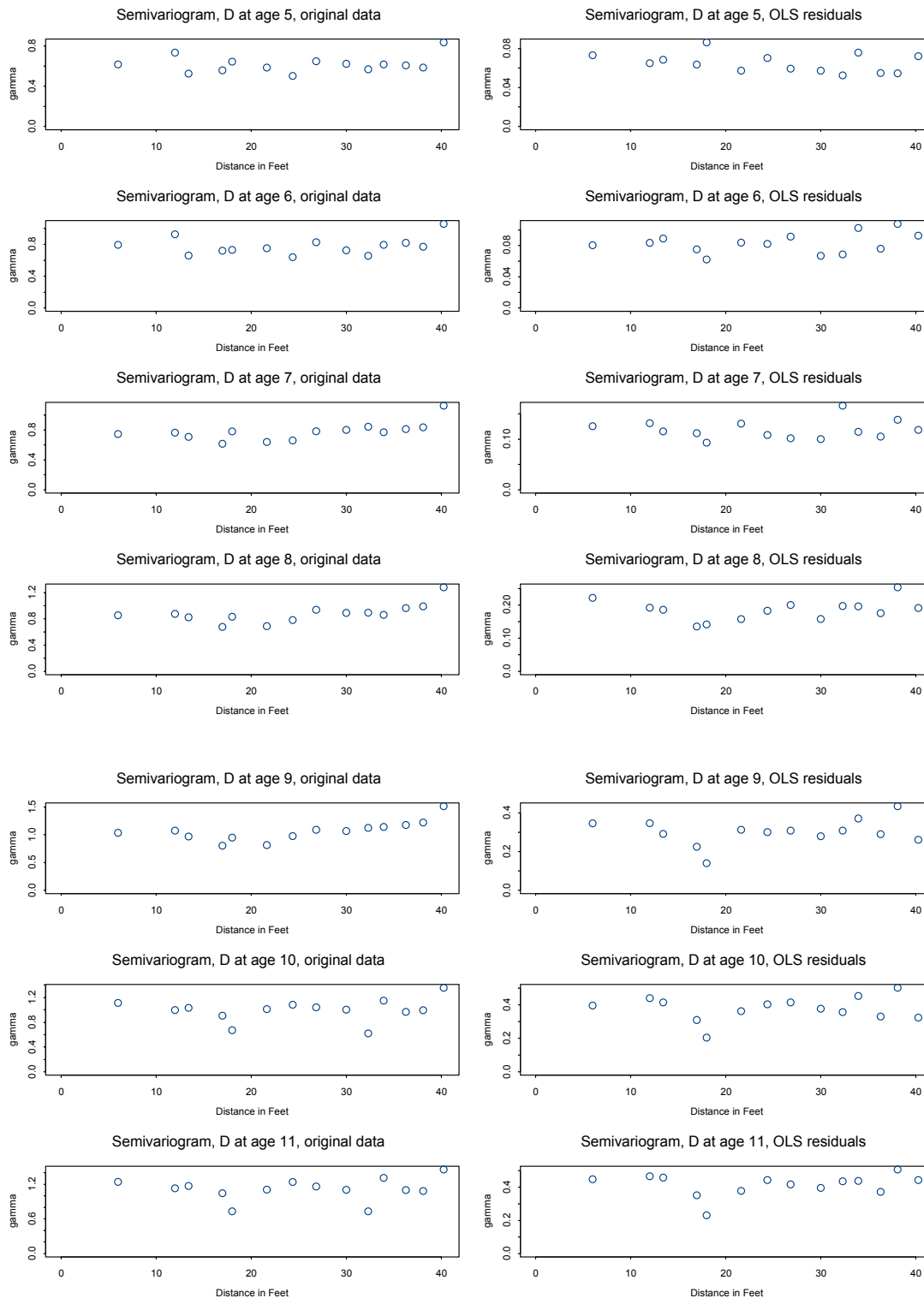
Stand number 145, 2722 TPA



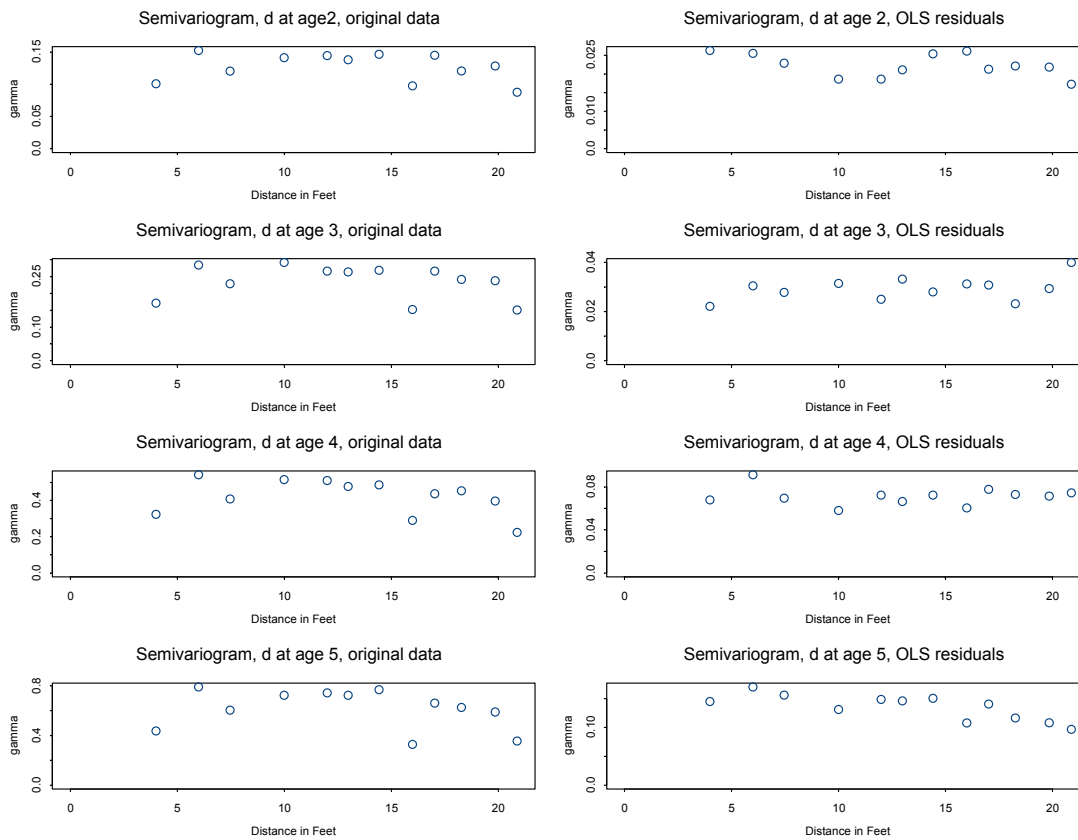


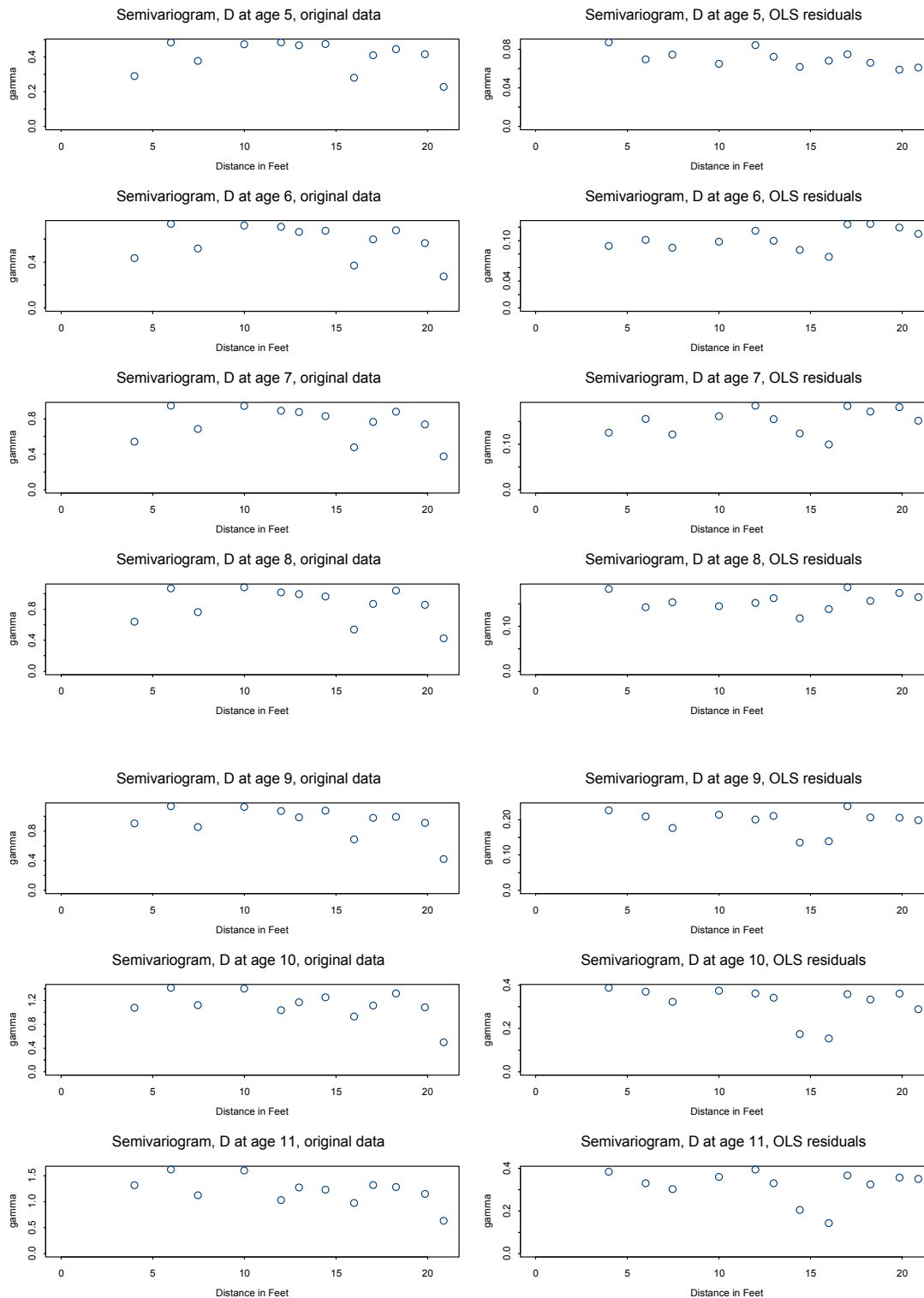
Stand number 158, 605 TPA





Stand number 181, 1815 TPA





Appendix III: Summary tables for an example plot from each planting density (trees per acre (TPA)) containing all significant (alpha level of 0.10) spatial indices analyzed in this study. The spatial indices considered were Moran's I coefficient, Geary's C coefficient, and a conditional autoregressive (CAR) spatial linear model including a covariate of total stem height. These indices were computed for ground-line diameters from ages 2 to 5, and for breast height diameters from ages 5 to 11.

Stand number 20, 907 TPA, all values significant at an alpha level of 0.10

Age	Diameter	Moran's I	Geary's C	CAR
2	d	--	--	--
3	d	--	--	--
4	d	--	--	--
5	d	-0.061694	--	-1.609409
5	D	--	--	--
6	D	--	--	-1.262584
7	D	--	--	-1.142690
8	D	-0.063866	--	-1.586941
9	D	-0.068658	--	-1.793538
10	D	--	--	-1.433394
11	D	--	--	-1.468215

CAR = The spatial linear model including the covariate of total stem height

d = Ground-line diameter

D = Diameter at breast height

Stand number 44, 453 TPA, all values significant at an alpha level of 0.10

Age	Diameter	Moran's I	Geary's C	CAR
2	d	--	--	--
3	d	--	--	--
4	d	--	--	--
5	d	--	--	-2.246373
5	D	-0.048988	0.956400	-3.128423
6	D	--	--	--
7	D	--	--	--
8	D	--	--	-2.421060
9	D	--	--	--
10	D	--	--	--
11	D	--	--	-2.608241

CAR = The spatial linear model including the covariate of total stem height

d = Ground-line diameter

D = Diameter at breast height

Stand number 59, 680 TPA, all values significant at an alpha level of 0.10

Age	Diameter	Moran's I	Geary's C	CAR
2	d	--	--	--
3	d	--	--	--
4	d	0.018057	0.932969	--
5	d	--	--	--
5	D	--	--	--
6	D	--	--	--
7	D	--	--	--
8	D	-0.049695	--	-3.284077
9	D	-0.048367	--	-3.205923
10	D	-0.049799	--	-3.336009
11	D	--	--	-2.287073

CAR = The spatial linear model including the covariate of total stem height

d = Ground-line diameter

D = Diameter at breast height

Stand number 86, 1210 TPA, all values significant at an alpha level of 0.10

Age	Diameter	Moran's I	Geary's C	CAR
2	d	0.04519812478972742	0.90045654659215002	--
3	d	0.01506530789689211	0.92874321137452731	--
4	d	--	--	--
5	d	--	--	--
5	D	--	--	--
6	D	0.01357313964026051	0.93307587041926654	--
7	D	--	--	--
8	D	--	--	--
9	D	--	--	--
10	D	--	--	--
11	D	--	--	--

CAR = The spatial linear model including the covariate of total stem height

d = Ground-line diameter

D = Diameter at breast height

Stand number 105, 1361 TPA, all values significant at an alpha level of 0.10

Age	Diameter	Moran's I	Geary's C	CAR
2	d	--	--	--
3	d	--	--	--
4	d	--	--	--
5	d	--	--	--
5	D	--	--	--
6	D	--	--	--
7	D	--	--	--
8	D	--	--	-1.386471
9	D	--	--	-1.192180
10	D	--	--	--
11	D	--	--	-1.394109

CAR = The spatial linear model including the covariate of total stem height

d = Ground-line diameter

D = Diameter at breast height

Stand number 128, 302 TPA, all values significant at an alpha level of 0.10

Age	Diameter	Moran's I	Geary's C	CAR
2	d	--	--	--
3	d	--	--	--
4	d	--	--	-3.478343
5	d	--	--	--
5	D	--	--	-2.972033
6	D	--	--	--
7	D	--	--	--
8	D	--	--	--
9	D	--	--	--
10	D	--	--	-3.809378
11	D	--	--	-3.230938

CAR = The spatial linear model including the covariate of total stem height

d = Ground-line diameter

D = Diameter at breast height

Stand number 145, 2722 TPA, all values significant at an alpha level of 0.10

Age	Diameter	Moran's I	Geary's C	CAR
2	d	--	--	--
3	d	--	--	--
4	d	--	--	--
5	d	--	--	--
5	D	--	--	--
6	D	--	--	--
7	D	--	--	-1.351017
8	D	--	--	-1.524804
9	D	-0.053280	--	-1.979537
10	D	-0.058641	1.056812	-2.099031
11	D	-0.065825	1.072769	-2.318762

CAR = The spatial linear model including the covariate of total stem height

d = Ground-line diameter

D = Diameter at breast height

Stand number 158, 605 TPA, all values significant at an alpha level of 0.10

Age	Diameter	Moran's I	Geary's C	CAR
2	d	--	--	--
3	d	-0.062889	--	-2.944672
4	d	--	--	--
5	d	--	--	--
5	D	--	--	--
6	D	--	--	--
7	D	--	--	-2.412055
8	D	-0.062884	--	-3.035055
9	D	-0.054808	--	-2.411758
10	D	--	--	--
11	D	--	--	--

CAR = The spatial linear model including the covariate of total stem height

d = Ground-line diameter

D = Diameter at breast height

Stand number 181, 1815 TPA, all values significant at an alpha level of 0.10

Age	Diameter	Moran's I	Geary's C	CAR
2	d	--	--	--
3	d	--	--	--
4	d	--	0.948656	--
5	d	--	1.050626	--
5	D	--	--	-1.270393
6	D	0.046116	0.932947	--
7	D	0.049060	0.921130	--
8	D	0.023204	--	--
9	D	0.025890	--	--
10	D	--	--	--
11	D	--	--	--

CAR = The spatial linear model including the covariate of total stem height

d = Ground-line diameter

D = Diameter at breast height

VITA

Bronson P. Bullock was born on June 13, 1973, in Cumberland, Pennsylvania. He graduated from Steinert High School, Hamilton Square, New Jersey, in June 1991. After completion of an Honors Thesis, he was awarded a Bachelor of Science degree in Natural Resource Management – Forestry, with honors from Rutgers, The State University of New Jersey – Cook College in May 1996. The following summer included an internship with Westvaco Corporation in Summerville, South Carolina. Entering the Forestry graduate program at Virginia Polytechnic Institute and State University in the fall of 1996, he acquired a Master of Science degree in Forest Biometry in December 1998 and continued his studies at Virginia Tech towards the Doctor of Philosophy degree in Forest Biometry. While working towards his Ph.D., he obtained a Master of Science degree in Statistics in May 2000. During his studies at Virginia Tech he served as both a Graduate Teaching Assistant and Graduate Research Assistant. The author is a member of the Society of American Foresters, American Society for Photogrammetry and Remote Sensing, American Statistical Association, and the Virginia Academy of Science. Upon completion of his Ph.D., he was employed as an Assistant Professor of Forest Biometrics and Timber Management by the Department of Forestry, College of Natural Resources, at North Carolina State University.

Bronson P. Bullock