MIXED-EFFECTS MODELS FOR A DIRECTIONAL RESPONSE:

A CASE STUDY WITH LOBLOLLY PINE MICROFIBRIL ANGLE

by

LEWIS C. JORDAN

(Under the Direction of Daniel B. Hall)

ABSTRACT

We propose methods for fitting mixed-effects regression models for an angular response and apply it to predict microfibril angle in loblolly pine (*Pinus taeda* L.) In this study, we generalize Presnell et al.'s (1998) spherically projected multivariate linear model, based on the angular normal distribution, by inclusion of random effects to account for the within-cluster correlation typical of repeated measures, longitudinal data and other clustered data structures. In the microbril angle example, the data are clustered due to the presence of multiple measurements of the response on each sample tree. This study suggests that mixed-effects angular response models are superior to population level models based on several statistical indices. Model comparisons were done based on several criteria including Likelihood ratio tests (LRT's), Akaike information criterion (AIC) and Bayesian information criterion (BIC).

INDEX WORDS: Circular data, clustered data, random effects, regression

MIXED-EFFECTS MODELS FOR A DIRECTIONAL RESPONSE: A CASE STUDY WITH LOBLOLLY PINE MICROFIBRIL ANGLE

by

LEWIS C. JORDAN

B.S.F.R., The University of Georgia, 2001

M.S., The University of Georgia, 2003

A Thesis Submitted to the Graduate Faculty of

The University of Georgia in Partial Fulfillment

of the

Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2005

MIXED-EFFECTS MODELS FOR A DIRECTIONAL RESPONSE:

A CASE STUDY WITH LOBLOLLY PINE MICROFIBRIL ANGLE

by

LEWIS C. JORDAN

Approved:

Major Professor: Daniel B. Hall

Committee: Richard F. Daniels

William P. McCormick

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2005

DEDICATION

To: Dr. Daniel B. Hall

"All his systems were adapted to fit them for the rough elements which they were destined to encounter. Frequent and severe punishments, which were expected to be born with more than Spartan fortitude, came to be considered less as inflictions of disgrace than as trials of obstinate endurance. To make his boys hardy, and to give them early forming habits, seemed to be his only aim."

- Charles Lamb's description of his former professor and mentor, William Wales, Mathematical Master at Christ's Hospital. Taken from Caroline Alexander's "The Bounty".

ACKNOWLEDGEMENTS

Thanks especially to Dr. Daniel B. Hall for his help, direction, and patience. My sincerest thanks are given to thank Dr. Richard F. Daniels for financial support, advice, and granting me the freedom to pursue various research interests these past 4 years. I would like to thank Dr. Bill McCormick for his helpful comments and suggestions. Thanks are also extended to the Faculty and Staff of the University of Georgia's Department of Statistics. I would also like to thank all of my friends who have encouraged me while pursuing this degree. Finally, I would like to express my deepest love, thanks, and appreciation to my parents, Mr. Charlie and Mrs. Chris Jordan.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION

Directional, or angular, response problems are commonly encountered in a multitude of scientific fields including: astronomy, biology, geology, and medicine. The directions are regarded as points on the circumference of a circle in two dimensions or on the surface of a sphere in three dimensions. In general, directions may be visualized as points on the surface of a hypersphere but observed directions are obviously angular measurements (Mardia 1975). The general theory of directional data has been covered extensively by Mardia (2000), Batschelet (1981) and most recently by Fisher (1993).

The theory of regression models when the response variable is angular is a quite neglected area of statistics, even though problems of this type are not uncommon (Fisher and Lee 1992). The general purpose of regression analysis is to model the conditional mean value of a response variable in terms of one or more explanatory variables by some relatively simple mathematical relationship. The most commonly used distributions in modeling directional data are the Fisher-von Mises, Wrapped Normal, Cardioid, and Wrapped Cauchy distributions. More recently, the angular normal distribution (Presnell et al 1998, Mardia and Jupp 2000, Watson 1983) has been used in the development of circular regression models.

Presnell et al. (1998) proposed a spherically projected multivariate linear (SPML) model based on the angular normal distribution to fit circular regression models. The SPML model treats directional observations as projections onto the unit circle or sphere of unobserved

responses from a multivariate linear model. Presnell suggests that the previously proposed models based on the Fisher von-Mises, Wrapped Normal, Cardioid, and Wrapped Cauchy are impractical in analyzing circular data. The flaws associated with these models include: implausibility of the fitted models; non-identifiably of parameters; multimodal likelihoods; and computational difficulties associated with the estimation algorithms.

In this project, mixed-effects models will be developed from Presnell et al's. (1998) spherically projected multivariate linear model. We apply the mixed-effects SPML model to describe how MFA changes within individual trees, from pith to bark for loblolly pine. Oftentimes, population level estimates are desired not for a single tree, but for a group of individuals. In addition, it is important to be able to make predictions at the individual tree level and to be able to account for between tree heterogeneity and within tree correlation appropriately in the model. The inclusion of tree-specific random effects in the SPML model accomplishes these goals. In addition, it gives the model the appropriate inference space: the population of all trees represented by the sample, rather than the sampled trees only, as would be the case for a purely fixed effects model that includes tree-specific parameters.

CHAPTER 2

LITERATURE REVIEW

ANGULAR DATA

Data that are in the form of angular or orientation measurements in the plane (circular data) or in space (spherical data) arise in numerous scientific fields. Circular data, or directional data, is the simplest case of this category of data, and is the focus of this study. Circular data are summarized as locations on the unit circle or as angles over a 360^o or 2π radians range, with the endpoints of each range corresponding to the same location on the circle. For directional data, unit lengths starting at the origin and pointing in the direction of their angle are used to represent individual observations. It is because of this representation that angular observations in two-dimensions are called circular.

Numerical representation as an angle is not necessarily unique since the angular value depends on the choice of what is labeled as the zero-direction [true East or true North ("azimuthal")] and the sense of rotation (clockwise or counter-clockwise). Two major approaches to directional statistics, namely the intrinsic approach (directions are considered as points on the circle itself) and the embedding approach (directions are considered as special points in the plane) are commonly used (Mardia and Jupp 2000). The embedding approach regards each point θ on the circle as the unit vector $X = (\cos\theta, \sin\theta)$ in the plane enables one to take the expectations. A single observation $A = \theta^o \left(0 < \theta^o < 360\right)$ represents the angle made by

the vector with the positive x-axis (the point (1,0) on the unit circle) in the counter-clockwise direction. The Cartesian coordinates of the vector are $(x, y) = (\cos\theta^o, \sin\theta^o)$.

The main difference between circular data and data measured on a linear scale is the wrap-around nature, with no minimum or maximum ($0 = 2\pi$), and the measurement is periodic with θ being the same as $\theta + 2c\pi$ for any integer *c*. Appropriate statistical measures and analysis of circular data differ from those of linear methods. Consider for example two angles 1^o and 359^o that are two degrees apart. If the values are viewed graphically in two dimensions, no problems are readily apparent, and the observations are centered on 0^o However, if linear statistical methodology is applied, the mean and standard deviation of the two values are 180^o and 179^o respectively.

An appropriate measure of the mean direction for a set of directions is obtained by treating the data as unit vectors and using the direction of their resultant vector. For a set of angular measurements, $\theta_1 \ldots \theta_n$, convert each observation to its rectangular form $(\cos\theta_i, \sin\theta_i)$, $i = 1 \ldots n$. The resultant vector, *R*, or these *n* unit vectors from the origin is obtained by summing them component wise to get $R = \left(\sum_{i=1}^{n} \cos\theta_i, \sum_{i=1}^{n} \sin\theta_i \right) = (\overline{C}, \overline{S})$. The length of the resultant vector is given as $R = \sqrt{(\overline{C}^2 + \overline{S}^2)}$. The direction of this resultant vector is known as the circular mean direction and is denoted as $\overline{\theta}$. A quadrant specific inverse tangent definition of the circular mean direction is

$$\overline{\theta} = \begin{pmatrix} \arctan(\overline{S} / \overline{C}), & \text{if } \overline{C} > 0 \\ \pi / 2, & \text{if } \overline{C} = 0 \text{ and } \overline{S} > 0 \\ -\pi / 2, & \text{if } \overline{C} = 0 \text{ and } \overline{S} > 0 \\ \arctan(\overline{S} / \overline{C}) + \pi, & \text{otherwise} \end{pmatrix}. \qquad (1)$$

Distributions of circular data are those whose total probability is concentrated on the circumference of a unit circle. The most frequently used circular distribution is the von Mises distribution, which is similar to the normal distribution on the line (Mardia 2000, Fisher 1993). The von Mises distribution is symmetric and unimodal, and is characterized by a mean direction μ, and concentration parameter quantifying the dispersion *k*. As k increases from zero, the von Mises density peaks higher about μ, and if *k* is zero the distribution is uniform with no preferred direction. If $k \geq 2$ the von Mises distribution VM(μ, *k*), can be approximated by the wrapped normal distribution WN(μ, *p*), $\left[p = \exp\left(-0.5\sigma^2\right)\right]$ which is a symmetric unimodal distribution obtained by wrapping a normal N(μ, σ^2) distribution around the circle (Mardia and Jupp 2000, Fisher 1993). Other models used to analyze symmetric unimodal circular data are the Cardioid and Wrapped Cauchy distributions. For a more detailed description of circular distributions see (Jammalamadaka and SenGupta 2001, Mardia and Jupp 2000, Fisher 1993).

MICROFIBRIL ANGLE FORMATION

Microfibril angle (MFA) refers to the mean helical angle between the cellulose fibrils and the longitudinal cell axis. A tree is sheathed by a thin cambial layer, which is composed of cells capable of repeated division. New cells produced on the inside of this sheath become new wood, while those moved to the outside become part of the phloem. In a process that may take several weeks to complete, the cell enlarges and the cell wall gradually thickens as biopolymers produced within the cells are progressively added to the inside of the wall. Eventually, the fluid filling the cell is lost and the cell has a thickened wall, consisting of primary and secondary wall layers and a hollow center (Haygreen and Bowyer 1996).

Megraw (1985) gives an excellent description of the structure of microfibrils.

> "Microfibrils are long strands of physically aggregated and more or less parallel polysaccharide chains. They contain a crystalline cellulose core surrounded by shorter chain hemicelluloses partly linked to this core, all encased into a rigid structure by the surrounding lignin. They are usually arranged in sheets or lamellae that lie parallel to the cell surface. In the primary wall they are loosely and more or less randomly interwoven. The secondary cell wall is comprised of three distinct layers: the S_1, S_2, and S_3 walls (Figure 2.1). The S_1 layer contains alternating lamellae of Z and S helical orientation at 50 to 70 degrees to the fiber axis. In the thick S_2 layer, they are usually oriented in a Z helix around the cell, but in some cases also occur in an S helix. In the S_2 layer microfibrils are highly parallel and steeply aligned to the fiber axis. In the S_3 layer they again occur in a very flat helical orientation, with angles 60 to 90 degrees to the cell axis. Because the S_2 layer is many times thicker than the other layers, its properties normally dominate, and the term fibril angle refers to the microfibular angle in the S_2 layer."

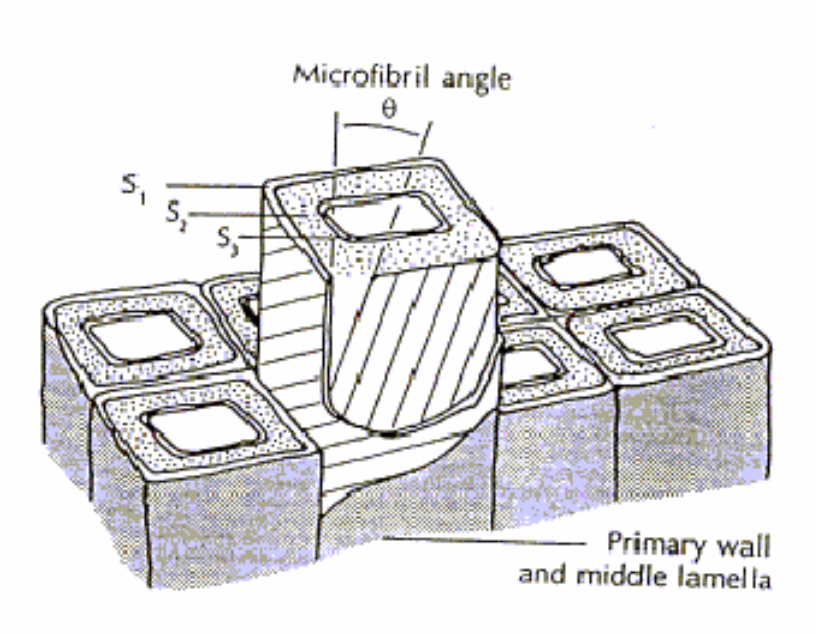

Figure 2.1. Orientation of microfibrils in the S2 layer of the tracheid cell wall (after Dickson and Walker 1997).

WITHIN TREE VARIATION OF MICROFIBRIL ANGLE

Variations in wood quality of any species can be attributed to variation within a tree, between trees in a particular stand, between different growing sites, and between different silvicultural regimes (Addis et al. 1995). MFA varies within each growth ring, from pith to bark, with height in the stem and among trees. Cave and Walker (1994) reported that the MFA decreases from the first earlywood cell to the last latewood cell. MFA in loblolly pine is large near the pith and decreases rapidly out to 10 or more rings from the pith, and then continues dropping, regardless of height, but at a much slower rate until such time as it essentially stabilizes. MFA also tends to increase as tracheid length decreases after sudden increases in growth (Waldrop 1951). Just as for tracheid length, the decrease in angle with age takes place at a slower rate near the base of the tree than it does at upper heights. This results in higher MFA's for a given number of rings from the pith at the butt and breast height regions than at several meters in height and above (Megraw 1985).

MFA varies considerably within the three zones of tree wood (juvenile, transition, and mature), thus affecting the mechanical properties of the wood. Within tree cells, the MFA in the S_2 part of the secondary wall is characteristically greater in juvenile wood. Pillow et al. (1953) found that MFA in the juvenile wood of open-grown loblolly pines averaged 20^o larger than those in a closely spaced natural stand. In mature loblolly pine MFA is small, averaging about 5^o to 10^o, while in juvenile wood the MFA is large, averaging 25^o to 35^o and often up to 50^o in rings next to the pith, then decreasing outward in the juvenile core (Larson et al. 2001). MFA has been found to decrease from 33^o at age 1 to 23^o at age 10, and to 17^o at age 22, in fast-grown loblolly pines (Ying et al. 1994). Bendtsen and Senft (1986) found that MFA within loblolly pine had stabilized at age 30.

MIXED EFFECT MODELS

It is often the case that data collected in a designed experiment are correlated. Correlated data occur under a variety of scenarios and go by several different names including multivariate observations, grouped data, repeated measurements, longitudinal data, and spatially correlated data (Verbeke and Molenberghs 2000). Forestry related data are typically collected from permanent plots or individual trees over time, e.g. height, diameter, basal area, volume, and trees per hectare. The assumption of independence of repeated measures in forestry is often violated by the repeated sampling of permanent plots, or in our case individual trees (Clutter 1961, Bailey and Clutter 1974, Lappi and Bailey 1988, Gregoire et al. 1995). Data of this structure lend themselves to analysis via mixed effects modeling techniques. Mixed effects models allow for the inclusion of multiple sources of correlation and/or heterogeneity, in addition to allowing for treatment and/or covariate effects with fixed effects parameters (Hall and Clutter 2004).

LINEAR MIXED EFFECTS MODELS

Linear mixed effects models include both fixed and random effects, which occur linearly in the model function. The incorporation of random effects can be regarded as additional error terms in a general linear model, accounting for correlation among observations within the same group. Using the notation of Pinheiro and Bates (2000), Vonesh and Chinchilli (1997), and Laird and Ware (1982), the general form of the linear mixed-effects model for data with one level of clustering can be written as:

$$\begin{aligned} \mathbf{y}_i &= \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i, \quad i = 1,\ldots,M \\ \mathbf{b}_i &\sim N(\mathbf{0}, \boldsymbol{\psi}), \boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \sigma^2\mathbf{I}) \end{aligned} . \tag{2}$$

Where,

$\mathbf{y}_i = \left(y_{i1}, y_{i2}, \ldots, y_{in_i}\right)^T$ = the n_i-dimensional response vector for the *i*th group,

$\mathbf{X}_i$ = the $n_i \times p$ fixed-effects design matrix,

$\boldsymbol{\beta}$ = the *p*-dimensional vector of fixed effects,

$\mathbf{Z}$ = the $n_i \times q$ random-effects design matrix,

$\mathbf{b}_i$ = the q-dimensional vector of random effects,

$\boldsymbol{\varepsilon}_i$ = the n_i-dimensional within-group error vector,

$\boldsymbol{\psi}$ = a $q \times q$ between-subject covariance matrix.

The random effects ($\mathbf{b}_i$) and within-groups ($\boldsymbol{\varepsilon}_i$) errors are assumed to be independent for different groups and independent of each other for the same group. The distribution of the random effects are assumed normal with mean **0**, thus it is characterized by its variance-covariance matrix $\boldsymbol{\Psi}$. No constraints other than assuming a positive-definite matrix is put on $\boldsymbol{\Psi}$. Assuming $\mathbf{b}_i$ and $\boldsymbol{\varepsilon}_i$ are independent, the mean and variance-covariance of $\mathbf{y}_i$ are given by

$$\begin{aligned} E\left(\mathbf{y}_i\right) &= \mathbf{X}_i\boldsymbol{\beta} \\ Var\left(\mathbf{y}_i\right) &= \boldsymbol{\Sigma}_\mathbf{i}\left(\boldsymbol{\theta}\right) = \mathbf{Z}_i\boldsymbol{\psi}\mathbf{Z}_i^T + \sigma^2\mathbf{I} \end{aligned} \tag{3}$$

The overall variance-covariance of $\mathbf{y}_i$ depends on the unknown parameter vector $\boldsymbol{\theta} = \left[\sigma^2, Vech(\boldsymbol{\psi})^T\right]^T$. From Equation (2) correlation is accounted for solely through the subject-specific random effects ($\mathbf{b}_i$).

The likelihood function for Equation (2) is the probability density for the data given the parameters, but regarded as a function of the parameters with the data fixed. Estimation of $\left(\boldsymbol{\beta}, \boldsymbol{\psi}, \sigma^2\right)$ can be carried out using normal theory maximum likelihood applied to the marginal distribution (Vonesh and Chinchilli 1997):

$$\mathbf{y}_i \sim N_{n_i}\left(\mathbf{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma}_i(\boldsymbol{\theta})\right). \tag{4}$$

Here, $\boldsymbol{\Sigma}_i(\boldsymbol{\theta}) = \mathbf{Z}_i\boldsymbol{\psi}\mathbf{Z}_i^T + \sigma^2\mathbf{I}$, and $\theta = \left[\sigma^2, vech(\boldsymbol{\psi})^T\right]^T$ is the known vector of distinct variance-covariance parameters. The maximum likelihood estimators of $\boldsymbol{\beta}$ and the random effects ($\mathbf{b}$) are given by:

$$\begin{aligned}\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}) &= \left(\sum_{i=1}^{M}\mathbf{X}_i^T\boldsymbol{\Sigma}_i(\boldsymbol{\theta})^{-1}\mathbf{X}^T\right)^{-1}\sum_{i=1}^{M}\mathbf{X}_i^T\boldsymbol{\Sigma}_i(\boldsymbol{\theta})^{-1}\mathbf{y}_i \\ \hat{\mathbf{b}}_i(\boldsymbol{\theta}) &= \boldsymbol{\psi}\mathbf{Z}_i^T\boldsymbol{\Sigma}_i(\boldsymbol{\theta})^{-1}\left(\mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})\right)\end{aligned} \tag{5}$$

where $\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})$ is the generalized least squares estimator of $\boldsymbol{\beta}$ assuming $\boldsymbol{\theta}$ is known and $\hat{\mathbf{b}}_i(\boldsymbol{\theta})$ is the best linear unbiased predictor of $\mathbf{b}_i$. If $\boldsymbol{\theta}$ is replaced with its maximum likelihood estimate $\hat{\boldsymbol{\theta}}$, then $\hat{\boldsymbol{\beta}}\left(\hat{\boldsymbol{\theta}}\right)$ is the corresponding maximum likelihood estimate of $\boldsymbol{\beta}$ and $\hat{\mathbf{b}}_i = \hat{\mathbf{b}}_i\left(\hat{\boldsymbol{\theta}}\right)$ is the estimated best linear unbiased predictor of $\mathbf{b}_i$.

The maximum likelihood estimate of $\boldsymbol{\theta}$ is obtained by maximizing the profile likelihood:

$$\hat{L}(\boldsymbol{\theta}) = -\frac{1}{2}\left\{N\log(2\pi) + \sum_{i=1}^{M}\left(\hat{\mathbf{e}}_i^T\,\Sigma_i^T\,\hat{\mathbf{e}} + \log|\boldsymbol{\Sigma}_i|\right)\right\}, \tag{6}$$

where, $N = \sum_{i=1}^{M} n_i$, $\hat{\mathbf{e}} = \mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}$, and all other variables previously defined.

Although fairly general, Equation (2) is somewhat restrictive because of the assumption $Var(\boldsymbol{\varepsilon}_i) = \sigma^2\mathbf{I}$, or conditional independence and homoskedasticity of the within-cluster errors. A more flexible form of Equation (2) can be specified where $\boldsymbol{\varepsilon}_i \sim N\left(\mathbf{0}, \sigma^2\boldsymbol{\Lambda}_i(\boldsymbol{\gamma})\right)$. Where, $\sigma^2\boldsymbol{\Lambda}_i(\boldsymbol{\gamma})$ is an $n_i \times n_i$ intrasubject covariance matrix with additional parameter $\boldsymbol{\gamma}$. Correlation can now be accounted for through the subject-specific random effects $\{\mathbf{b}_i\}$, or through specification of an

appropriate intrasubject covariance matrix $\left(\sigma^2 \mathbf{\Lambda}_i(\boldsymbol{\gamma})\right)$. For a more thorough review, readers are referred to Vonesh and Chinchilli (1997).

Maximum likelihood estimation of the linear mixed-effects model components presented above oftentimes produces, in general, no analytical expressions for the estimators (Searle et al. 1992). Generally, when no closed form expression for the estimators is possible, numerical optimization techniques must be used to maximize the log and profile log likelihood equations. The most commonly employed methods are procedures including Newton-Raphson (with or without Fisher scoring), the Marquardt method, and the EM algorithm. Descriptions and comparisons of the various estimation methods for linear mixed-effects models can be found in Searle et al. (1992) and Vonesh and Chinchilli (1997).

NONLINEAR MIXED EFFECTS MODELS

Nonlinear mixed-effects models (NLMM's) are mixed-effects models where some or all of the fixed and random effects occur nonlinearly in the model function (Pinheiro and Bates 2000). NLMM's are an extension of linear mixed-effects models where the conditional expectation of the response given the random effects is allowed to be a nonlinear function of the coefficients (Bates and Watts 1988).

Numerous authors discuss development of NLMM's including: Searle et al. (1992), Vonesh and Chinchilli (1997), and Pinheiro and Bates. Following Pinheiro and Bates (2000), the general formulation of a nonlinear mixed-effects model can be given as:

$$\begin{aligned} \mathbf{y}_i &= \mathbf{f}_i\left(\boldsymbol{\varphi}_i, \boldsymbol{\upsilon}_i\right) + \boldsymbol{\varepsilon}_i, \quad i = 1,\ldots,M \\ \mathbf{b}_i &\sim N\left(\mathbf{0}, \boldsymbol{\psi}\right), \boldsymbol{\varepsilon}_i \sim N\left(\mathbf{0}, \sigma^2 \mathbf{I}\right) \end{aligned} . \tag{7}$$

where, M is the number of groups, n_i is the number of observations on the ith group, and $\boldsymbol{\varepsilon}_i$ is a normally distributed within-group error term. Here, $\mathbf{f}_i$ is a real-valued, differentiable function of a vector-valued mixed effects parameters $\boldsymbol{\varphi}_i$ and a vector of covariates $\boldsymbol{\upsilon}_i$. No constraints other than assuming they are positive-definite symmetric matrices are put on $\boldsymbol{\psi}$. The mixed effects parameters $\boldsymbol{\varphi}_i$ take the form:

$$\boldsymbol{\varphi}_i = \mathbf{A}_i\boldsymbol{\beta} + \mathbf{B}_i\mathbf{b}_i. \tag{8}$$

Where,

$\mathbf{A}_i$ = the fixed effects design matrix,

$\boldsymbol{\beta}$ = the p-dimensional vector of fixed effects,

$\mathbf{B}_i$ = the random effects design matrix,

$\mathbf{b}_i$ = the q-dimensional vector of random effects.

The parameters of the NLMM are $\boldsymbol{\beta}$, σ^2 and $\boldsymbol{\psi}$, so the likelihood function based on all of the data can be written as $L\left(\boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi} \mid \mathbf{y}\right)$, where $\mathbf{y} = \left(\mathbf{y}_1^T, \ldots, \mathbf{y}_M^T\right)$ is the combined data vector from all groups. The likelihood function is equal to the joint probability density function:

$$p\left(\mathbf{y} \mid \boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi}\right) = \prod_{i=1}^{M} p\left(\mathbf{y}_i \mid \boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi}\right) \tag{9}$$

The conditional density $p\left(\mathbf{y}_i \mid \mathbf{b}_i, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi}\right)$ given $\mathbf{b}_i$ is a normal density, inherited from the normality vector of error terms $\boldsymbol{\varepsilon}_i$ [Equation (7)]. Conditional on $\mathbf{b}_i$, $\boldsymbol{\varphi}_i$ is non-random, so that $\mathbf{y}_i = \mathbf{f}_i\left(\boldsymbol{\varphi}_i, \boldsymbol{\upsilon}_i\right) + \boldsymbol{\varepsilon}_i$ has the usual form of a fixed effects nonlinear model, and $\boldsymbol{\varepsilon}_i \sim N\left(\mathbf{0}, \sigma^2\mathbf{I}\right)$ implies:

$$\mathbf{y}_i \mid \mathbf{b}_i \sim N\left(\mathbf{f}_i\left(\boldsymbol{\varphi}_i, \boldsymbol{\upsilon}_i\right), \sigma^2\mathbf{I}\right). \tag{10}$$

Therefore, $p(\mathbf{y}_i \mid \mathbf{b}_i, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi})$, is given by a multivariate normal density with mean $\mathbf{f}_i(\boldsymbol{\varphi}_i, \boldsymbol{\upsilon}_i)$ and variance-covariance matrix $\sigma^2\mathbf{I}$.

Because the random effects are unobserved quantities, maximum likelihood estimation in mixed-effects models is based on the marginal density of the responses $\mathbf{y}_i$, which is calculated as:

$$p(\mathbf{y}_i \mid \boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi}) = \int p(\mathbf{y}_i \mid \mathbf{b}_i, \boldsymbol{\beta}, \sigma^2) p(\mathbf{b}_i \mid \boldsymbol{\psi}) d\mathbf{b}. \tag{11}$$

Where, $p(\mathbf{y}_i \mid \boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi})$is the marginal density of $\mathbf{y}_i$, $p(\mathbf{y}_i \mid \mathbf{b}_i, \boldsymbol{\beta}, \sigma^2)$ is the conditional density of $\mathbf{y}_i$ given the random effects $\mathbf{b}_i$, and the marginal distribution of $\mathbf{b}_i$ is $p(\mathbf{b}_i \mid \boldsymbol{\psi})$. Using these results, we have that $p(\mathbf{y}_i \mid \boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi})$ is given by the integral of a product of multivariate normal densities. Substituting these normal densities and using Equation (9) we have

$$\begin{aligned} p(\mathbf{y} \mid \boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi}) = \prod_{i=1}^{M} \int (2\pi\sigma^2)^{-n_i/2} \exp\left[-\frac{1}{2\sigma^2} \left\| \mathbf{y}_i - \mathbf{f}_i(\boldsymbol{\varphi}_i, \boldsymbol{\upsilon}_i) \right\|^2 \right] \times \\ (2\pi)^{-q/2} |\boldsymbol{\psi}|^{-1/2} \exp\left[-\frac{1}{2} \mathbf{b}_i^T \boldsymbol{\psi}^{-1} \mathbf{b}_i \right] d\mathbf{b}_i \end{aligned} \tag{12}$$

For simplification, it is convenient to express $\boldsymbol{\psi}^{-1} = \sigma^{-2}\boldsymbol{\Delta}^T\boldsymbol{\Delta}$, where $\boldsymbol{\Delta}$ is a square-root matrix of $\sigma^2\boldsymbol{\psi}^{-1}$. This simplifies Equation (12) to:

$$p(\mathbf{y} \mid \boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi}) = \prod_{i=1}^{M} \frac{|\boldsymbol{\Delta}|}{(2\pi\sigma^2)^{(n_i+q)/2}} \int \exp\left[\frac{\left\| \mathbf{y}_i - \mathbf{f}_i(\boldsymbol{\beta}, \mathbf{b}_i) \right\|^2 + \left\| \boldsymbol{\Delta}\mathbf{b}_i \right\|^2}{-2\sigma^2} \right] d\mathbf{b}_i \tag{13}$$

In Equation (13) we have changed notation replacing $\mathbf{f}_i(\boldsymbol{\varphi}_i, \boldsymbol{\upsilon}_i)$ with $\mathbf{f}_i(\boldsymbol{\beta}, \mathbf{b}_i)$to make explicit the dependence of this quantity on the random effects vector $\mathbf{b}_i$. Replacing the left-hand side of Equation (13) with $L(\boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi} \mid \mathbf{y})$, yields an expression for the likelihood function.

As in Equation (3), Equation (8) is restricted by within-subject errors that are conditionally independent $N(\mathbf{0}, \sigma^2\mathbf{I})$ random variables. One can relax the assumption of a spherical variance-covariance structure on the error terms. A more flexible form of Equation (8) can be specified where $\boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \sigma^2\boldsymbol{\Lambda}_i(\boldsymbol{\gamma}))$. Where, $\sigma^2\boldsymbol{\Lambda}_i(\boldsymbol{\gamma})$ is an $n_i \times n_i$ intrasubject covariance matrix with parameters $[\sigma^2\boldsymbol{\gamma}^T]$. Interested readers are referred to Vonesh and Chinchilli (1997) for a more thorough review.

Unlike the linear mixed-effects model, there is generally no closed form solution to the maximum likelihood estimates of the nonlinear mixed-effects model (Pinheiro and Bates 2000, Vonesh and Chinchilli 1997). The presence of an integral in the marginal distribution of $\mathbf{y}$ makes the maximum likelihood procedure cumbersome, especially for multidimensional random effects (Gregoire et al. 1997). A variety of numerical optimization routines for maximizing the likelihood function have been proposed. Some of these methods consist of taking a first-order Taylor expansion of the model function $\mathbf{f}$ around the expected value of the random effects (Sheiner and Beal 1980, Vonesh and Carter 1992) or around the conditional modes of the random effects (Lindstrom and Bates 1990, Pinheiro and Bates 2000). The use of Gaussian quadrature has also been used (Davidian and Giltinan 1995). For a more thorough review of parameter estimation in nonlinear mixed-effects models, the reader is referred to: Searle et al. 1992; Davidian and Giltinan 1995; Vonesh and Chinchilli 1997; Pinheiro and Bates 2000.

EXISTING MICROFIBRIL ANGLE MODELS

In spite of its great importance, almost no existing literature deals with developing MFA prediction models. This is partly due to the inconsistent patterns of MFA found within and

between trees, which make parametric modeling difficult. Jordan et al. (2005) developed a multilevel nonlinear mixed effects model using a modified three parameter Logistic function for modeling earlywood and latewood MFA. The Jordan model takes:

$$f(Ring) = \frac{\beta_0}{1 + e^{\beta_1 Ring}} + \beta_2 \quad (14).$$

Where *Ring* is ring number from pith, $f(Ring)$ is the mean response function for MFA, β_0 corresponds to an initial value parameter, β_1 is the rate parameter, and β_2 is the lower asymptote. If β_1 is positive, as $Ring \to \infty$, $f(Ring) \to \beta_2$. By making the parameters of the Logistic function linear functions of height, a three-dimensional model was developed which describes the changes of earlywood and latewood microfibril angle within the tree. Jordan et al. found that a first order autoregressive correlation structure, and an assumption of error variance proportional to the mean in which the proportionality constant differs according to the level of wood type, best identified the within-group correlation and variance structures.

He et al. (2005) proposed a nonparametric smoothing spline analysis of variance (SS ANOVA) approach for modeling earlywood and latewood MFA. These authors that since the patterns of MFA at 1.4 meters differ substantially with increasing height, it is difficult to pinpoint the transitional position, making it difficult to develop sensible parametric models. Both earlywood and latewood MFA were modeled as functions of ring number from pith and height. The SS ANOVA method decomposes a multivariate function into additive components within a reproducing kernel Hilbert space (RKHS), and it requires the predictor variables, ring number and disk height to be scaled into a [0, 1] interval. The scaled ring number and disk height values are derived by dividing all ring numbers and disk heights by their corresponding maximum value. Let *B* be the population from which trees are drawn and *P* be the sampling

distribution. MFA at any specified position within a given tree is determined by ring number from pith and disk height. The He model takes:

$$y_{ijk} = m(i, h_{ij}, ring_{ijk}) + e_{ijk} \quad (15),$$

where, y_{ijk} = earlywood or latewood MFA of the k^{th} ring number of the j^{th} disk height from the i^{th} tree, $i \in B$ with i =1...n, $h_{ij} \in [0,1]$ = the scaled height of the j^{th} disk of the i^{th} tree with j=1… h_i and h_i = the number of disks in the i^{th} tree, $ring_{ijk} \in [0,1]$ = the k^{th} scaled ring number of the j^{th} disk from the i^{th} tree with k=1... r_{ij} , r_{ij} = the number of rings of the j^{th} disk from the i^{th} tree, and e_{ijk} = random error.

He et al.'s MFA models determine the within tree MFA variation and allow for a straightforward interpretation as an ANOVA model. Within tree MFA spatial autocorrelation was accounted for indirectly via correlated growth year and disk random effects. The within tree variation pattern of MFA was described by a tensor product of a cubic spline of ring number from pith and a cubic spline of ring height above ground. The fit statistics from He et al.'s models were found to be comparable to those of Jordan et al., and the models were found to be unbiased at all height levels and rings.

ANGULAR RESPONSE MODELS

Let, θ and ω represent a random angle and its angular mean, both measured by their angular displacement in radians from the zero direction. If $\mathbf{u} = (\cos(\theta), \sin(\theta))^T$ is the random direction in $\Re^2$, then its mean direction is the unit vector $\boldsymbol{\eta} = E(\mathbf{u})/\rho$, where $\rho = \| E(\mathbf{u}) \|$. Here, $E(\mathbf{u})$ represents the usual component wise expectation for vectors, and $\|.\|$ represents the usual

Euclidian norm. The parameter ρ, $0 \le \rho \le 1$, is called the mean resultant length and serves as a measure of concentration for directional distributions. For unimodal distributions, larger values of ρ indicate higher concentrations. Angles differing by multitudes of 2π are equivalent, and $\mathbf{u}$ and $\boldsymbol{\eta}$ are given by $\mathbf{u}^T = (\cos(\theta), \sin(\theta))$ and $\boldsymbol{\eta}^T = (\cos(\omega), \sin(\omega))$.

The theory of regression models when the response variable is angular is a quite neglected area of statistics, even though problems of this type are not uncommon (Fisher and Lee 1992). The general purpose of regression analysis is to model the mean value of a response variable in terms of one or more explanatory variables by some simple/complex mathematical relationship. Numerous distributions have been used in modeling directional data including: the Fisher-von Mises (FVM), Wrapped Normal, Cardioid and Wrapped Cauchy distributions. More recently, the angular normal (AN) distribution (Presnell et al 1998, Mardia and Jupp 2000, Watson 1983) has been used in the development of circular regression models.

Of the distributions described above, the most commonly used distribution in modeling directional data is the Fisher-von Mises (FVM) distribution. The distribution is referred to as von-Mises for the circular case, and the Fisher distribution for spherical data. The density of the FVM distribution is given by:

$$f(\theta; \omega, k) = [2\pi I_0(k)]^{-1} \exp[k\cos(\theta - \omega)] \quad -\pi < \theta, \omega \le \pi, k \ge 0 \tag{16}$$

Where, ω is the mean direction, *k* is the concentration parameter,

$I_0(k) = (2\pi)\int_0^{2\pi} \exp[k\cos(\theta - \omega)]d\theta$ is the modified Bessel function of order zero.

In the regression setting, let $(\mathbf{x}_1, \mathbf{u}_1), \ldots (\mathbf{x}_n, \mathbf{u}_n)$ be independent observations, where $\mathbf{x}_i$ is a vector of covariate values and $\mathbf{u}_i$ is the corresponding directional response, with mean

direction $\boldsymbol{\eta}_i$ and mean resultant length ρ_i. For circular data, θ_i and ω_i are the angular representations of $\mathbf{u}_i$ and $\boldsymbol{\eta}_i$.

Most of the existing literature on regression for an angular response is based upon the von-Mises distribution. Gould (1969), Johnson and Wehrly (1978) and Fisher and Lee (1992) all assume θ_i follows a von-Mises distribution, with either the angular mean (ω_i), or the concentration parameter (k_i), or both depend on $\mathbf{x}_i$. Gould (1969) introduced the "barber's pole" model, where the mean angular response, conditional on a linear variate, is a curve winding in an infinite number of spirals up the surface of an infinite cylinder. The Gould model takes the form: $\omega_i = \omega_0 + \mathbf{x}_i\boldsymbol{\beta}$, and $k_i = k$ for all $i = 1,\ldots,n$, where ω_0, $\boldsymbol{\beta}$, and k are unknown parameters.

Johnson and Wehrly (1978) expanded upon Gould's model by modeling the joint distribution of the angular response and a linear variate, which allows direct estimation of the model parameters. Johnson and Wehrly proposed two models, one for location and one for concentration. The location model takes the form: $\omega_i = \omega_0 + 2\pi F(\mathbf{x}_i)$, and $k_i = k$ for all $i = 1,\ldots,n$, where ω_0 and k are unknown parameters, and F is a known continuous cumulative distribution function.

Fisher and Lee (1992) gave three generalizations of the Johnson and Wehrly models. The Fisher-Lee mean model takes the form: $\omega_i = \omega + g(\mathbf{x}_i^T\boldsymbol{\beta})$, and $k_i = k$ for all $i = 1,\ldots,n$, where ω, $\boldsymbol{\beta}$, and k are unknown parameters, and $g(.)$ is a monotone link function. The purpose of the function g is to map the real line to the circle, having the property that as x ranges from $(-\infty,\infty)$, $g(\mathbf{x}_i^T)$ ranges from $(-\pi,\pi)$. The Fisher-Lee concentration model takes $\omega_i = \omega$ and $k_i = h(\mathbf{x}_i)$, where ω is unknown, and h is some function mapping $\Re^k$ to $[0,\infty)$. Fisher and Lee suggested the use of $h(\mathbf{x}_i) = \exp(\mathbf{x}_i^T\boldsymbol{\lambda})$, where $\boldsymbol{\lambda}$ is an unknown coefficient vector to be

estimated. Fisher and Lee proposed fitting the above models using maximum likelihood estimation. Difficulties in the maximization of the likelihood arise with multiple maxima occurring in the likelihood function. Specifically, the likelihood is multimodal and increases as $\boldsymbol{\beta}_j \to \pm\infty$. To overcome this difficulty, Fisher and Lee propose graphical evaluation of the likelihood function.

Presnell et al. (1998) proposed a spherically projected multivariate linear (SPML) model based on the angular normal distribution to overcome the shortcomings of the models presented above. These authors suggest that the previously proposed models are impractical in analyzing such data. The flaws associated with these models include: implausibility of the fitted models; non-identifiably of parameters; multimodal likelihoods; and computational difficulties associated with the estimation algorithms. The SPML model treats directional observations as projections onto the unit circle of unobserved responses from a multivariate linear model. The basis of the distribution is simple. A sample of angles on the unit circle is generated by the angular normal distribution when we consider the projections onto the circle of a sample of points in the plane having a unit bivariate normal distribution (Scapini 2002).

Following Presnell et al. (1998), let $\mathbf{u}_i = \mathbf{y}_i / R_i$, $R_i = \|\mathbf{y}_i\|$, where the random vectors $\mathbf{y}_1, \ldots, \mathbf{y}_n$ are independent bivariate normal random vectors with common covariance matrix $\boldsymbol{\Sigma}$ and mean vectors $\boldsymbol{\mu}_i = \boldsymbol{\beta}^T \mathbf{x}_i$, with $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$. The components of $\boldsymbol{\mu}_i$ are $\mu_{ij} = \mathbf{x}_i^T \boldsymbol{\beta}_j$, with the first element of each $\mathbf{x}_i$ is 1 to allow for an intercept term. The parameters $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ are to be estimated but are not identifiable without further constraints, since for any $c > 0$, taking $\boldsymbol{\beta}^* = c\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}^* = \boldsymbol{\Sigma} / c$ does not alter the distribution of the observed directions. Presnell et al. suggest this issue could be addressed by restricting the determinant of $\boldsymbol{\Sigma}$ to equal 1. A simpler

alternative is to take $\boldsymbol{\Sigma} = \mathbf{I}$, the identity matrix. Under this restriction, the distribution of $\mathbf{u}_i$ has been called the "offset normal" (Mardia 1972), "displaced normal" (Kendall 1974), and an "angular normal" (Watson 1983).

Dropping the subscript *i*, but noting that $\mathbf{u}$, $\boldsymbol{\mu}$, $\boldsymbol{\eta}$, and γ depend on *i*, the density of $\mathbf{u}$ given by Watson (1983) is:

$$f(\mathbf{u}) = \frac{\int_0^\infty r^{d-1}\phi(r - \gamma \mathbf{u}^T \boldsymbol{\eta})dr}{(2\pi)^{d/2}\exp(\gamma^2/2)\phi(\gamma \mathbf{u}^T \boldsymbol{\eta})}, \tag{17}$$

where, $\gamma = \|\boldsymbol{\mu}\|$ is a concentration parameter, $\boldsymbol{\eta} = \boldsymbol{\mu}/\gamma$ (the unit vector) is the mean direction, and ϕ is the standard normal density. For the circular case, $\mathbf{u}^T\boldsymbol{\eta} = \cos(\theta - \omega)$, and the density of θ can be written as:

$$f(\theta) = \frac{\exp(-\gamma^2/2)}{2\pi}\left[1 + \frac{\gamma\cos(\theta-\omega)\Phi(\gamma\cos(\theta-\omega))}{\phi(\gamma\cos(\theta-\omega))}\right], \tag{18}$$

where Φ is the standard normal distribution function. Any FVM distribution is closely approximated by an angular normal distribution of the form above with the same angular mean and mean resultant length (Presnell et al. 1998).

For the circular case, which appears to be the most important and practical for applications, the log-likelihood function for the SPML model is:

$$l(\boldsymbol{\beta}) = \log(L(\boldsymbol{\beta})) = -\frac{1}{2}\sum_{i=1}^{n}\boldsymbol{\mu}_i^T\boldsymbol{\mu}_i + \sum_{i=1}^{n}\psi(\mathbf{u}_i^T\boldsymbol{\mu}_i) - n\log(2\pi), \tag{19}$$

where, $\boldsymbol{\mu}_i = \boldsymbol{\beta}^T\mathbf{x}_i$, $\mathbf{u}^T\boldsymbol{\mu}$ replaces $\gamma\cos(\theta-\omega)$ in equation (18), and

$$\psi(t) = \log\left[1 + \frac{t\Phi(t)}{\phi(t)}\right]. \tag{20}$$

In regression notation, let $\mathbf{X} = [\mathbf{x}_1 \ldots \mathbf{x}_n]^T$, $\mathbf{c} = [C_1 \ldots C_n]^T$, $\mathbf{s} = [S_1 \ldots S_n]^T$, and $\mathbf{u}_i = (C_i, S_i) = (\cos(\theta_i), \sin(\theta_i))$. The score function is:

$$\dot{l}(\boldsymbol{\beta}) = \begin{pmatrix} -\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}_1 + \mathbf{X}^T\mathbf{M}\mathbf{c} \\ -\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}_2 + \mathbf{X}^T\mathbf{M}\mathbf{s} \end{pmatrix}, \tag{21}$$

where $\mathbf{M}$ is the diagonal matrix $\mathbf{M} = \mathbf{M}(\boldsymbol{\beta}) = diag\left[\dot{\psi}(\mathbf{u}_1^T\boldsymbol{\beta}\mathbf{x}_1), \ldots, (\mathbf{u}_n^T\boldsymbol{\beta}\mathbf{x}_n)\right]$, with

$$\dot{\psi}(t) = t + \frac{\Phi(t)}{\phi(t) + t\Phi(t)}.$$

The likelihood equations can be written in matrix form as:

$$\mathbf{X}^T\mathbf{X}\boldsymbol{\beta} = \mathbf{X}^T\mathbf{M}\mathbf{U}, \tag{22}$$

where $\mathbf{U} = [\mathbf{c}, \mathbf{s}]$.

No closed form solution exists, but the maximum likelihood estimates can be obtained using an iterative procedure such as the Newton-Raphson or EM algorithms.

One representation of the fitted angular mean at *x* for the circular case is:

$$\hat{\omega} = \left[\tan^{-1}(\boldsymbol{\mu}_2 / \boldsymbol{\mu}_1) + \pi I(\boldsymbol{\mu}_1 < 0)\right] \bmod 2\pi, \tag{23}$$

where, $\boldsymbol{\mu}_1 = \boldsymbol{\beta}_1^T\mathbf{x}_i$, and $\boldsymbol{\mu}_2 = \boldsymbol{\beta}_2^T\mathbf{x}_i$. The range of $\tan^{-1}$ for real *x* is $(-\pi/2, \pi/2)$, and $\tan^{-1}(z)$ corresponds to an angle in the range $(-90, 90)$. If $(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$ is a point in the plane, then the corresponding direction, or the angle between the horizontal axis and the vector pointing from the origin to $(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$ is given by $\tan^{-1}(\boldsymbol{\mu}_2 / \boldsymbol{\mu}_1)$ as long as the point is in the first or fourth quadrants (as long as $\boldsymbol{\mu}_1 \geq 0$). When $(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$ is in the second or third quadrants ($\boldsymbol{\mu}_1 < 0$), the direction of the vector is actually opposite from the angle, and the addition of π (180 degrees) remedies this.

For example, the points $(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2) = (1,1)$ and $(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2) = (-1,-1)$ both give $\tan^{-1}(1/1) = \tan^{-1}(-1/-1) = \tan^{-1}(1) = (\pi/4) = 45^0$, which is the correct angle for $(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2) = (1,1)$, but not for $(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2) = (-1,-1)$, which corresponds to $(\pi/4) + \pi = 5\pi/4 = 225^0$. So in general, we get the correct angle if we take $\tan^{-1}(\boldsymbol{\mu}_2 / \boldsymbol{\mu}_1)$ when $\boldsymbol{\mu}_1 \geq 0$, and $\tan^{-1}(\boldsymbol{\mu}_2 / \boldsymbol{\mu}_1) + \pi$ when $\boldsymbol{\mu}_1 < 0$. Which leads to: $\tan^{-1}(\boldsymbol{\mu}_2 / \boldsymbol{\mu}_1) + \pi I(\boldsymbol{\mu}_1 < 0)$. The $(\bmod 2\pi)$ function just gives a range from $(0, 2\pi) = (0^0, 360^0)$ rather than $(-\pi/2, 3\pi/2) = (-90^0, 270^0)$. That is, it rewrites angles from $(-\pi/2, 0)$ as angles from $(3\pi/2, 2\pi)$, but this is an arbitrary choice.

Presnell et al. applied the SPML model to describe the movements of blue periwinkle butterflies. The fit of the SPML model was found to be comparable to those of Fisher (1993) presented above, both having absolute residual values of 27 degrees. Both the SPML and Fisher models were found to have similar resultant lengths and both captured the trend of increasing concentration. More recently, Scapini et al. (2002) successfully applied the SPML model in the analysis of variance setting to predict the movements of sandhoppers. Scapini analyzed the movements of two different sandhopper species, at two levels of landscape, two levels of season, and sex. Scapini's model assumes that the position of the mean of the bivariate normal is a function of the explanatory variables. When all of the explanatory variables are qualitative (factors), the model is analogous to a multiway analysis of variance model with or without interactions. A model with the full set of interactions specifies a different angular normal for each combination of the levels of the factors (Scapini et al 2002).

COMPUTATIONAL METHODS

Although the Newton-Raphson algorithm can be used to fit the SPML, the expectation-maximization (EM) algorithm is particularly convenient for this purpose because the model fitting can be cast as an incomplete data problem. Presnell et al. (1998) recommend fitting the SPML model using the EM algorithm (Dempster et al. 1977) by treating the unobserved $R_i = \|\mathbf{y}_i\|,\ i = 1,\ldots,n$ as missing data. The complete data model is the multivariate normal regression model for $\mathbf{Y} = (\mathbf{y}_1,\ldots \mathbf{y}_n)^T$, which has sufficient statistic given by $\mathbf{X}^T\mathbf{Y} = \mathbf{X}^T\mathbf{R}\mathbf{U}$, where $\mathbf{R} = diag(R_1,\ldots,R_n)$ is the diagonal matrix of unobserved lengths, and $\mathbf{U} = [\mathbf{c},\mathbf{s}]$.

The *E*-step, or the expectation of the missing data given $\boldsymbol{\beta}, \mathbf{X}$, and the observed data $\boldsymbol{\theta}$ is given by $E(\boldsymbol{X}^T\boldsymbol{Y} \mid \boldsymbol{\theta}) = \boldsymbol{X}^T\boldsymbol{M}\boldsymbol{U}$, where again $\mathbf{M}$ is the diagonal matrix $\mathbf{M} = \mathbf{M}(\boldsymbol{\beta}) = diag[\psi(\mathbf{u}_1^T\boldsymbol{\beta}\mathbf{x}_1),\ldots,(\mathbf{u}_n^T\boldsymbol{\beta}\mathbf{x}_n)]$. The *M*-step consists of maximizing the log-likelihood evaluated at the expected value of the missing data taken at the E step, which leads to the closed form solution $\boldsymbol{\beta} = (\mathbf{X}^T\mathbf{X}^{-1})E(\mathbf{X}^T\mathbf{Y} \mid \boldsymbol{\theta})$. Given starting values, the EM algorithm proceeds by iterating $\mathbf{M}^{(k)} = \mathbf{M}(\hat{\boldsymbol{\beta}}^{(k)})$ and $\hat{\boldsymbol{\beta}}^{(k+1)} = (\mathbf{X}^T\mathbf{X}^{-1})\mathbf{X}^T\mathbf{M}^{(k)}\mathbf{U}$, $k = 0,1,\ldots,$ until $\hat{\boldsymbol{\beta}}$ converges. Presnell et al. (1998) found that the EM and NR algorithms are both usable approaches to fitting the SPML model and lead to the same MLEs. They found that although the NR algorithm converged in fewer iterations, the EM algorithm was much simpler to implement.

To obtain estimates of parameters in mixed-effects models for repeated measures data, in all but the simplest cases, requires iterative methods (Ware 1985, Lindstrom and Bates 1988). The two most commonly used algorithms for estimating parameters in mixed-effects models are

the EM (Dempster et al. 1977) and NR algorithms (Laird and Ware 1985, Lindstrom and Bates 1988, Laird et al 1987, Pinheiro and Bates 2000).

The EM iterations for mixed-effects models are based on regarding the random effects as unobserved data. At iteration *k*, the *E*-step consists of computing the expected value of the complete data log-likelihood given the observed response vector $\mathbf{y}$ and $\boldsymbol{\theta}^{(k)}$, the estimate of the variance-covariance parameter vector at the current step. This involves taking an expectation with respect to the distribution of $\mathbf{b} \mid \mathbf{y}, \boldsymbol{\theta}^{(k)}$. The *M*-step consists of maximizing this expectation with respect to $\boldsymbol{\theta}$ to produce $\boldsymbol{\theta}^{(k+1)}$. Each iteration results in an increase of the likelihood function (Navidi 1997, Pinheiro and Bates 2000).

The NR algorithm uses a first-order expansion of the gradient of the likelihood function around the current estimate of $\boldsymbol{\theta}^{(k)}$ to produce the next estimate of $\boldsymbol{\theta}^{(k+1)}$. Each NR iteration requires calculation of the score function and it derivative (the Hessian matrix of the log-likelihood) (Pinheiro and Bates 2000), and is computationally more expensive. The EM algorithm has an advantage over the NR algorithm in the fact that each iteration can be computed more quickly (Lindstrom and Bates 1988). However, the number of iterations required for convergence is generally much smaller for the NR than EM algorithm. The NR algorithm can also be adapted to handle most extensions of mixed-effects models, such as accounting for serial correlation in longitudinal data, where the EM algorithm cannot be easily generalized to handle these models. A more comprehensive comparison of the NR and EM algorithms in mixed-effects models is given by Lindstrom and Bates (1988).

The log-likelihood functions of the fixed-effects [Equation (19)] and mixed-effects models can be maximized using the EM algorithm combined with Gaussian quadrature to fit the model. In principle, the SAS PROC NLMIXED procedure can be used to fit the mixed-effects

SPML model without the EM algorithm using any one of several optimization algorithms, including the NR and various quasi-Newton methods combined with Gaussian quadrature, adaptive Gaussian quadrature, or one of the other integration approximation methods (Hall and Berenhaut 2002). The NLMIXED procedure allows for the specification of a conditional distribution of the data (given the random effects) having any general distribution. PROC NLMIXED fits mixed-models by maximizing an approximation to the likelihood integrated over the random effects. The limitations of using the NLMIXED procedure is that only low dimensional random effects models can be fit (up to three dimensions) and NLMIXED can not handle nested or crossed random effects (Wolfinger 1999).

The models in this paper were fit using the S-PLUS NLME library and the SAS (1999) PROC NLMIXED procedure.

CHAPTER 3

DATA

Eighteen trees representing six stands were selected from Southeast Texas for MFA analysis (Figure 3.1). The stands were located on land owned by forest products companies, and included only stands with similar silvicultural history: 1) site preparation with no herbaceous weed control; 2) no fertilization at planting except phosphorus on phosphorus-deficient sites; 3) stand density of at least 617 trees per hectare at the time of sampling. Trees larger than 12.7 centimeters in diameter were inventoried on three 0.04-hectare plots to determine stand density and diameter distribution. A sample of three trees proportional to the diameter distribution of each stand, to represent a range of tree sizes in the stand was chosen for MFA analysis. Diameter (cm) ranged from 14.0 to 29.2 and averaged 20.1. Height (m) ranged from 11.4 to 21.7 and averaged 17.2. Age (years) ranged from 20 to 27 and averaged 22. MFA (degrees) ranged from 8.1 to 39.8 and averaged 16.5.

The trees selected for analysis were felled and cross-sectional disks 2.54 centimeters thick were cut at 1.4 meters, and then at 3-meter intervals to a height of 13.7 meters. A radial strip 1.27 centimeters square in cross-section, extending from pith to bark was cut from each disk, dried, glued to core holders and sawn into two strips at the pith. The strip used for MFA analysis, was dried at 122^{o} Celsius, and analyzed by Silviscan® using x-ray diffraction at 1 millimeter intervals on the radial surface. A densitometer was used to determine specific gravity. The densitometer was calibrated to express specific gravity on an air dried basis and a specific

gravity value of 0.53 was used to separate earlywood and latewood. Traditionally, a specific gravity value of 0.48 is used to distinguish between earlywood and latewood specific gravity (Clark et al. 2004) based on green volume and dry weight. However, specific gravity values analyzed by Silviscan® are based upon dry volume and dry weight resulting in a reduction of volume on the order of 10 percent, thus resulting in higher specific gravity values. Separation of rings was accomplished using Silviscan's Analyse2001 program.

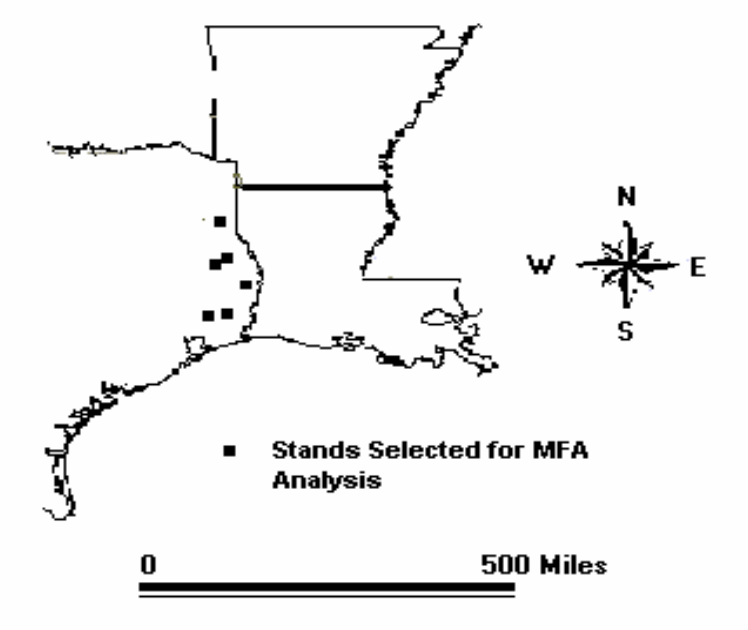

Figure 3.1. Location of loblolly pine stands sampled for microfibril angle analysis in Southeast Texas.

CHAPTER 4

METHODS

MODEL DEVELOPMENT

1. THE SPML MODEL

Expanding on Presnell et al. (1998), let $\mathbf{y}_{ij} = \mathbf{u}_{ij} \left\| \mathbf{y}_{ij} \right\|$ be the bivariate random vector for the *j*th, $j = 1, \ldots, n_i$ observation in the *i*th tree, $i = 1, \ldots, M$, $\mathbf{u}_{ij} = \left(\cos\left(\theta_{ij}\right), \sin\left(\theta_{ij}\right)\right)$ where θ_{ij} is the corresponding angle measured in radians from the zero direction, and $\left\| \mathbf{y}_{ij} \right\|$ is an unobserved latent vector, such that $\mathbf{y}_{ij} \sim N_2\left(\boldsymbol{\mu}_{ij}, \boldsymbol{\Sigma} = \mathbf{I}\right)$. We assume that $\mathbf{y}_{ij}$ follows a linear mixed-effects model as follows: Let, $\mathbf{y}_i = \left(\mathbf{y}_{i1}^T \ldots \mathbf{y}_{in_i}^T\right)^T$ be the $n_i \times 2$ latent response matrix for the *i*th cluster, such that

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{B}_i + \mathbf{E}_i, \tag{1}$$

where, $\mathbf{X}_i$ is an $n_i \times p$ covariate matrix, $\boldsymbol{\beta}$ is a $p \times 2$ matrix of regression coefficients, $\mathbf{Z}_i$ is a known random effects matrix of size $n_i \times q$, $\mathbf{B}_i$ is a $q \times 2$ matrix of random coefficients specific to tree *i*, and $\mathbf{E}_i$ is a $n_i \times 2$ matrix of errors. Where the transposed rows of $\mathbf{E}_i$ are independent, identically distributed $N\left(\mathbf{0}, \boldsymbol{\Sigma} = \mathbf{I}\right)$, and $\text{Vec}\left(\mathbf{B}_i\right) \sim N_{2q}\left(\mathbf{0}, \mathbf{D}\right)$, with **D** being a positive-definite matrix.

Equivalently, we may write the model in vector form by stacking the columns of the terms. This leads to

$$\mathrm{Vec}(\mathbf{y}_i) = [\mathbf{I}_2 \otimes \mathbf{X}_i]\mathrm{Vec}(\boldsymbol{\beta}) + [\mathbf{I}_2 \otimes \mathbf{Z}_i]\mathrm{Vec}(\mathbf{B}_i) + \mathrm{Vec}(\mathbf{E}_i), \quad (2)$$

where $\mathrm{Vec}(\mathbf{y}_i)$ is $2n_i \times 1$, and $\mathrm{Vec}(\mathbf{E}_i) \sim N(\mathbf{0}, \boldsymbol{\Sigma} = \mathbf{I} \otimes \mathbf{I}_{n_i})$. In this work we concentrate on the random intercept case, where $q = 1$. In this case, the model for $\mathbf{y}_{ij}$ can be written as

$$\mathbf{y}_{ij} = [\mathbf{I}_2 \otimes \mathbf{x}_{ij}^T]\mathrm{Vec}(\boldsymbol{\beta}) + \mathbf{b}_i + \mathbf{E}_{ij}, \quad (3)$$

where $\mathbf{y}_{ij}$ is 2×1, $[\mathbf{I}_2 \otimes \mathbf{x}_{ij}^T]$ is $2 \times 2p$ and $\mathbf{x}_{ij}^T$ is the *j*th row of $\mathbf{X}_i$, $\mathrm{Vec}(\boldsymbol{\beta})$ is $2p \times 1$, $\mathbf{E}_{ij}$ is 2×1 (the transpose of the *j*th row of $\mathbf{E}_i$), and $\mathbf{b}_i \overset{i.i.d}{\sim} N_2(\mathbf{0}, \mathbf{D})$ with $\mathbf{D}$ being a positive-definite variance-covariance matrix of size 2×2.

It is convenient to utilize an unconstrained parameterization for $\mathbf{D}$, (Pinheiro and Bates 1996, Hall and Berenhaut 2002). This leads to a slight modification of the model, where Equation (3) can now be rewritten as:

$$\mathbf{y}_{ij} = [\mathbf{I}_2 \otimes \mathbf{x}_{ij}^T]\mathrm{Vec}(\boldsymbol{\beta}) + \mathbf{D}^{T/2}\mathbf{b}_i^* + \mathbf{E}_{ij}. \quad (4)$$

Where, $\mathbf{b}_i^*$ are i.i.d. bivariate normal with mean $\mathbf{0}$ and identity variance-covariance matrix, and $\mathbf{D}^{T/2}(\alpha)$ is the transpose of $\mathbf{D}^{1/2}(\alpha)$, the upper triangular Cholesky factor $\mathbf{D}$. We parameterize the variance-covariance matrix of $\mathbf{b}_i^*$ through an unconstrained vector-valued parameter α that contains the nonzero elements of $\mathbf{D}^{1/2}$. In this work we concentrate on random intercept models, specifically:

$$\mathbf{b}_i^* = \begin{pmatrix} \mathrm{b}_{i1}^* & \mathrm{b}_{i2}^* \end{pmatrix}^T, \quad \mathbf{D}^{T/2} = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ 0 & \alpha_{22} \end{pmatrix}. \quad (5)$$

Equation (4) implies that conditional on $\mathbf{b}_i^*$, $(\mathbf{y}_{ij} \mid \mathbf{b}_i^*) \sim N_2(\boldsymbol{\mu}_{ij}, \boldsymbol{\Sigma} = \mathbf{I})$, where $\boldsymbol{\mu}_{ij} = [\mathbf{I}_2 \otimes \mathbf{x}_{ij}^T]\mathrm{Vec}(\boldsymbol{\beta}) + \mathbf{b}_i^* = \begin{pmatrix} \mathbf{x}_{ij}^T\boldsymbol{\beta}_1 + \mathrm{b}_{i1}^* & \mathbf{x}_{ij}^T\boldsymbol{\beta}_2 + \mathrm{b}_{i2}^* \end{pmatrix}^T$. Conditional on $\mathbf{b}_i^*$, the density of θ_{ij} is

$$f\left(\theta_{ij} \mid \mathbf{b}_i^*\right) = \frac{\exp\left(-\boldsymbol{\mu}_{ij}^T \boldsymbol{\mu}_{ij} / 2\right)}{2\pi}\left[1 + \frac{\mathbf{u}_{ij}^T \boldsymbol{\mu}_{ij}\, \Phi\left(\mathbf{u}_{ij}^T \boldsymbol{\mu}_{ij}\right)}{\phi\left(\mathbf{u}_{ij}^T \boldsymbol{\mu}_{ij}\right)}\right], \tag{6}$$

and the joint conditional density of $\boldsymbol{\theta}_i = \left(\theta_{i1} \ldots \theta_{in_i}\right)^T$ is

$$f\left(\boldsymbol{\theta}_i \mid \mathbf{b}_i^*\right) = \prod_{j=1}^{n_i} f\left(\theta_{ij} \mid \mathbf{b}_i^*\right). \tag{7}$$

The log-likelihood of the model is

$$\begin{aligned} \sum_{i=1}^{M} \log f\left(\boldsymbol{\theta}_i\right) &= \sum_{i=1}^{n} \log \int f\left(\boldsymbol{\theta}_i, \mathbf{b}_i^*\right) d\mathbf{b}_i^* \\ &= \sum_{i=1}^{n} \log \int f\left(\boldsymbol{\theta}_i \mid \mathbf{b}_i^*\right) f\left(\mathbf{b}_i^*\right) d\mathbf{b}_i^* \end{aligned} \tag{8}$$

$$= \sum_{i=1}^{M} \log \int \prod_{j=1}^{n_i} \frac{\exp\left(-\boldsymbol{\mu}_{ij}^T \boldsymbol{\mu}_{ij} / 2\right)}{2\pi}\left[1 + \frac{\mathbf{u}_{ij}{}^T \boldsymbol{\mu}_{ij}\, \Phi\left(\mathbf{u}_{ij}{}^T \boldsymbol{\mu}_{ij}\right)}{\phi\left(\mathbf{u}_{ij}{}^T \boldsymbol{\mu}_{ij}\right)}\right]\left[\frac{\exp\left(-\mathbf{b}_i^{*T} \mathbf{b}_i^* / 2\right)}{2\pi}\right] d\mathbf{b}_i$$

where $\boldsymbol{\mu}_{ij} = \left(\mathbf{x}_{ij}^T \boldsymbol{\beta}_1 + \mathrm{b}_{i1}^* \quad \mathbf{x}_{ij}^T \boldsymbol{\beta}_2 + \mathrm{b}_{i2}^*\right)^T$.

With the general specification of the SPML model completed, we will now proceed to describe the development of an appropriate nonlinear mixed-effects model based on Jordan et al.'s (2005) Logistic model.

2. THE JORDAN MODEL

As presented above, the Jordan et al. (2005) model is based on the following form for the conditional expectation function:

$$f\left(Ring\right) = \frac{\beta_0}{1 + e^{\beta_1 Ring}} + \beta_2. \tag{9}$$

Using the notation of Pinheiro and Bates (2000), let y_{ij} = the response at the *j*th measurement of the *i*th tree. The model can be expressed as

$$y_{ij} = f\left(\boldsymbol{\beta}_{ij}, \boldsymbol{\upsilon}_{ij}\right) + \varepsilon_{ij}$$

$$i = 1,...,M, \quad j = 1,...,n_i \ , \tag{10}$$

where M is the number of trees, n_i is the number of observations on the ith tree, and ε_{ij} is a normally distributed within-group error term. Here, f is a real-valued, differentiable function of a vector-valued mixed effects parameters $\boldsymbol{\beta}_{ij}$ and a vector of covariates $\boldsymbol{\upsilon}_{ij}$. The mixed effects parameters $\boldsymbol{\beta}_{ij}$ take the form

$$\boldsymbol{\beta}_{ij} = \mathbf{A}_{ij}\boldsymbol{\beta} + \mathbf{B}_{ij}\mathbf{b}_i, \tag{11}$$

where, $\mathbf{b}_i$ is a random effects vector of size $q \times 1$ and $\mathbf{B}_{ij}$ is the associated random effects design matrix. The fixed effects design matrix and parameter vectors are $\mathbf{A}_{ij}$ and $\boldsymbol{\beta}$, respectively. We assume $\mathbf{b}_i \sim N(\mathbf{0}, \boldsymbol{\Psi})$, with no constraints other than assuming a positive-definite symmetric matrix $\boldsymbol{\Psi}$. It may be useful to restrict $\boldsymbol{\Psi}$ to have a simplified and more parsimonious form for the purposes of computational stability and speed; for example, this could be done by assuming the random effects are independent of each other, which would make $\boldsymbol{\Psi}$ a diagonal matrix. Hall and Clutter (2004) state that often there is no *a priori* reason for assuming the random effects parameters are uncorrelated, and that random effects pertaining to distinct response variables measured on the same unit will typically be correlated.

By allowing the β parameters in Equation (9) to be modeled in terms of both fixed and random effects, the predicted random effects can be plotted by height, indicating an appropriate function, i.e., linear, quadratic, or some higher order polynomial function. For development of an appropriate height structure, independent random effects were assumed implying a diagonal random effects variance-covariance matrix with zero off-diagonal covariance elements. After determination of the height structure, the assumption of independent random effects will be

relaxed, and a correlated variance-covariance structure will be applied to account for potential correlation among the random effects. All models fit in this study used the natural logarithm of disk height. Taking the natural logarithm of disk height rescales the slope of the height structure paying dividends in computational time and model stability. Numerous height structures were fit to the data and the best model, based on that model selection criteria AIC and BIC, and likelihood ratio test's (LRT's) takes the form:

$$MFA_{ij} = \frac{\beta_{0i}}{1+e^{\beta_{1i}Ring}} + \beta_{2i}, \tag{12}$$

where,

$$\beta_{ij} = \begin{pmatrix} \beta_{0i} \\ \beta_{1i} \\ \beta_{2i} \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} b_{0i} \\ b_{1i} \\ b_{2i} \end{pmatrix} = \begin{pmatrix} (\beta_{00} + b_{0i}) + \beta_{01}\ln(ht) + \beta_{02}[\ln(ht)]^2 \\ (\beta_{10} + b_{1i}) + \beta_{11}\ln(ht) \\ (\beta_{20} + b_{2i}) + \beta_{21}[\ln(ht)]^2 \end{pmatrix}. \tag{13}$$

Since the basic model structure has been specified, we relax the assumption of independence among the random effects estimates. This leads to fitting a general positive-definite variance-covariance structure in combination with the model given by (12) and (13):

$$Var(b_i) = Var\begin{pmatrix} b_{0i} \\ b_{1i} \\ b_{2i} \end{pmatrix} = \mathbf{\Psi} = \begin{pmatrix} \psi_{00} & \psi_{01} & \psi_{02} \\ \psi_{01} & \psi_{11} & \psi_{12} \\ \psi_{02} & \psi_{12} & \psi_{22} \end{pmatrix} \tag{14}$$

Results indicated that a correlated variance-covariance structure improved the model based on the statistical indices previously described above. The addition of the height covariates in the model may change the variance-covariance correlation of the estimated random effects. This suggests that some of the random effects may be dropped, or the variance-covariance structure could be simplified by, for example, assuming some block diagonal structure. Updating Equation (14), we allow:

$$Var(b_i) = Var\begin{pmatrix} b_{0i} \\ b_{1i} \\ b_{2i} \end{pmatrix} = \boldsymbol{\Psi} = \begin{pmatrix} \psi_{00} & \psi_{01} & 0 \\ \psi_{01} & \psi_{11} & 0 \\ 0 & 0 & \psi_{22} \end{pmatrix} \tag{15}$$

In Equation (15), we assume b_{2i} is independent of b_{0i} and b_{1i} due to low correlation, but allow correlation between b_{0i} and b_{1i}. The model presented above was fit using the S-PLUS NLME library.

COMPARISON CRITERIA

The comparison of the models was based on graphical analysis of the residuals, and four statistical indices: coefficient of determination (R^2)[1], root mean square error (*RMSE*), mean residual (*MR*), and mean absolute residual (*MAR*) (Loague and Green 1991, Mayer and Butler 1993). These criteria are given below as:

$$R^2 = 1 - \frac{\sum_{i=1}^{M}\sum_{j=`}^{n_i}\left(Y_{ij} - \hat{Y}_{ij}\right)^2}{\sum_{i=1}^{M}\sum_{j=`}^{n_i}\left(Y_{ij} - \overline{Y}_{ij}\right)^2} \tag{16}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{M}\sum_{j=1}^{n_i}\left(Y_{ij} - \hat{Y}_{ij}\right)^2}{N - p}} \tag{17}$$

$$MR = \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{n_i}\left(Y_{ij} - \hat{Y}_{ij}\right) \tag{18}$$

$$MAR = \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{n_i}\left|Y_{ij} - \hat{Y}_{ij}\right| \tag{19}$$

[1] In the forestry literature starting with Schlaegel (1982), one often sees the term 'fit index' for Equation (6) applied to nonlinear models, but in the statistical literature it is still commonly referred to as the coefficient of determination or R^2 [see, e.g., Kvålseth (1985)].

Where, Y_{ij}, $\hat{Y}_{ij}$, and $\overline{Y}$ are the actual, predicted, and average values of the independent variable, N is the total number of observations used to fit the model, and p is the number of parameters in the model. Since MFA is constrained between 0^0 and 90^0 $\left(0 \text{ and } \pi/2\right)$ use of the statistical indices above is appropriate, and calculations involving mean values are approximately equivalent to methods used in calculating means of directional data often cited in the directional statistics literature (Mardia 1972). All reported statistical indices were calculated with respect to residual MFA values in degrees for comparison of models and since the transformation from radians to degrees is just a multiplicative rescaling: $\theta_{ij}^{D} = \theta_{ij}^{R}\left(180/\pi\right)$. The models evaluated in this paper were also examined visually by plotting the residuals against the estimated values and independent variables. Plots of the fitted curves for the SPML and Jordan models were also examined, including plots of the fixed-effects and tree level responses.

To truly assess performance of the models fit in this study, validation of the models with an independent data set would be the most desirable approach, because quality of fit does not always insure the quality of prediction (Huang et al. 2003, Kozak and Kozak 2003). The commonly used validation methods of data splitting and cross-validation, as shown by Kozak and Kozak (2003), do not provide any additional information on model performance compared to the statistics obtained from models fit to the entire data set. Models validated with an independent data set prove that either the data are from the same population and will perform as per se validation utilizing data splitting, or the data are from a different population entirely, in which case the models should be refit to obtain more appropriate parameter estimates. Due to the scarcity of data resembling that used herein, the models fit in this study were evaluated using the goodness-of-fit statistics and graphical analyses described above.

CHAPTER 5

RESULTS

RESULTS

1. THE SPML MODEL

Before employing the mixed-effects SPML model presented above, we need to specify an appropriate fixed effects model. The predictor variables used in this study consist of ring number, disk height, and transformations of the two. Presnell et al. (1998) suggests that nested models can be compared using likelihood ratio tests (LRT's), citing that the chi-squared approximation to the null distribution of the test statistic is accurate for samples as small as n = 20. However, the choice of a final model was based on the Akaike Information criterion (AIC), defined by $AIC = -2l + 2p$, where l is the log-likelihood of the fitted model and p is the number of parameters in the model. Table 1 contains a series of models with an increasing number of predictors. From Table 1, it can be seen that using the natural logarithm of ring and disk height improves, AIC, and BIC. According to the statistical criterion, Model 6 is preferred with the lowest AIC. Table 2 contains results of LRT's of the nested models. According to Table 2, Model 6 provides a significantly better fit to the data than do Models 4 or 5.

Wald tests indicated that not all of the parameters in Model 6 were significant at the 0.05 level. Model 6 was then reduced by removing the most insignificant variable then refitting the model until all parameters were significant at the 0.05 level. The final fixed-effects model takes the form:

$$\begin{pmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{\beta}_1^T\mathbf{X}_i\\ \boldsymbol{\beta}_2^T\mathbf{X}_i\end{pmatrix}=\begin{pmatrix}\beta_{01}+\beta_{11}\ln(Ring)+\beta_{21}[\ln(Ring)]^2+\beta_{31}\ln(Height)+\beta_{41}[\ln(Height)]^2\\ \beta_{02}+\beta_{12}\ln(Ring)+\beta_{22}[\ln(Ring)]^2\end{pmatrix} \quad (1)$$

A likelihood ratio test indicated that the reparameterization of Model 6 was justified. The *P*-Value of the LRT was found to be 0.223, and the final model has parameters, -2log-likelihood and AIC of 8, -3174, and -3158 respectively.

With the fixed-effects model completely specified, we now proceed by fitting different mixed-effects specifications to Equation (1) where $(\boldsymbol{\mu}_1 \quad \boldsymbol{\mu}_2)^T = (\boldsymbol{\beta}_1^T\mathbf{X}_i+\mathbf{b}_{i1}^* \quad \boldsymbol{\beta}_2^T\mathbf{X}_i+\mathbf{b}_{i2}^*)^T$. Table 3 contains comparisons of the mixed-effects SPML model with different variance-covariance structures. In Models 6.1 and 6.2, we allow for single random effects in both $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$. It can be seen in Table 3 that inclusion of single random effects in both Models 6.1 and 6.2 improved upon the -2log-likelihood, and AIC criterion. We then fit the mixed-effects SPML model with independent (Model 6.3) and correlated (Model 6.4) random effects in both components. In both instances, improvement was found in the -2log-likelihood and AIC. LRT's comparing the different variance-covariance structures are presented in Table 4. Comparing models with nested random effects via LRT's seems reasonable. However, the usual χ^2 reference distribution is no longer appropriate, as the conditions of Wilk's theorem are violated, and its application results in an overestimated *P*-Value (Pinheiro and Bates 2000). The *P*-Values for the LRT's in Table 4 were all found to be less than 0.05. According to the LRT, and AIC, Tables 4 and 3 indicate that the variance-covariance structure with correlated random effects best fits the data.

Parameter estimates for Model 6.4 are given in Table 5. All parameter estimates were found to be significant at the 0.0001 level. The model was found to account for 82 and 85 percent of the variation of MFA at the population and tree levels respectively (Table 6). From

Table 6, it can be seen that the mixed-effects SPML improved all of the statistical indices compared to the population level statistics. Mean residual values were found to be 0.12 and 0.02 at the population and tree levels respectively, suggesting that overall the models are underestimating the true value of MFA.

Plots of the residuals versus fitted values, the natural logarithm of ring and height, are given in Figure 5. Figure 5 does not indicate any general trends or patterns, and overall the residuals are small, suggesting that the SPML model was successful in explaining the variation of MFA. The majority of the residuals versus fitted values are centered on zero, and overall prediction appears to be unbiased because the mean residual values are small. The model does appear to be under-predicting the value of MFA above 25 degrees. The scatterplot of residuals versus the natural logarithm of ring will reveal possible bias at different ring numbers. In the same vein, the scatterplot of residuals versus disk height can be used to assess bias in the longitudinal direction. The residuals are almost symmetric around zero on the scatterplot of residuals versus ring number, suggesting the model's mean specification is overall unbiased across ring number (Figure 5). The plot of residuals versus the natural logarithm of height appears to be mildly quadratic, but overall prediction appears to be unbiased.

It is possible to obtain population or "typical prediction values of MFA by setting the random effects estimates to 0, and substituting corresponding fixed-effects values into Equation (16). We constructed a plot of the population, or "typical" MFA value at heights of 1.4, 4.6, 7.6, 10.7, and 13.7 meters (Figure 6). From Figure 6, it can be seen that the typical responses mimic the observed patterns of MFA found in Figure 1. MFA is larger at 1.4 meters across all rings compared to the other height levels. Above 1.4 meters, MFA is initially larger at 4.6 and 13.7

meters thru the first 10 rings compared to 7.6 and 10.7 meters in height. Beyond 10 rings, MFA is indistinguishable and all values are similar.

2. THE JORDAN MODEL

Parameter estimates from the Jordan model are given in Table 7. All parameters were found to be significant at the 0.001 level. From Table 7, one can evaluate the fixed-effects model parameters at differing heights. At a height of 1 meter, which sets the value of height in the model to 0, and ring 1, the initial value of MFA was found to be 30.7, 24.8, 24.0, 24.1, and 24.3 degrees at heights of 1.4, 4.6, 7.6, 10.7, and 13.7 meters respectively. The lower asymptotic bounds for MFA were found to be 8.92, 10.0, 11.0, 11.7, and 12.4 degrees at heights of 1.4, 4.6, 7.6, 10.7, and 13.7 meters respectively. The lower asymptotic bounds are similar at all heights excluding 1.4 meters to those reported by Jordan et al. (2005) for loblolly pine in the Gulf and Hilly regions of Texas. Jordan found that MFA eventually stabilizes at lower bounds of 12.45, 11.2, 11.15, 11.9, and 14.4 degrees at 1.4, 4.6, 7.6, 10.7, and 13.7 meters in height respectively. The rate parameter increases with increasing disk height and decreases with increasing ring number, indicating that MFA at larger heights will achieve the lower bound at an earlier age.

The model was found to account for 83 and 91 percent of the variation of MFA at the population and tree levels respectively (Table 8). From Table 8, it can be seen that the mixed-effects Jordan model improved all of the statistical indices compared to the population level statistics. Mean residual values were found to be 0.17 and -5.97×10^{-7} at the population and tree levels respectively, suggesting that overall the models are mildly underestimating the true value of MFA. Plots of the residuals versus fitted values, ring and the natural logarithm of height, are given in Figure 7. Figure 7 does not indicate any general trends or patterns, and overall the

residuals are small, suggesting that the Jordan model was successful in explaining the variation of MFA. The majority of the residuals versus fitted values are centered on zero, and overall prediction is unbiased because the mean residual values are small. The model does appear to be over predicting the value of MFA above 25 degrees. Plots of the residuals versus ring and the natural logarithm of height are almost symmetric around zero. The plot of residuals versus the natural logarithm of height appears to be mildly quadratic, but overall prediction appears to be unbiased.

From Figure 8 it can be seen that the Jordan model accurately describes the trends of MFA. MFA is initially larger from 1.4 to 4.6 meters in height, then increases with increasing disk height at ~ ring 5. MFA at 1.4 meters in Figure 8 is similar to that shown in Figure 1 up until ring ~10, where the model continues to decrease at a much more rapid rate than the observed values shown in Figure 1. This indicates that MFA will be highly under predicted beyond ring ~10 at 1.4 meters in height.

Even though it have been shown that population predictions from the final SPML and Jordan models accurately describe the trends of MFA, inclusion of the random-effects estimates will allow for more precise prediction. As a comparison, we graphed the predictive curves with and without random-effects for a single tree in the data set. From this, it can be seen that more precise predictions are obtained with the inclusion of the estimated random-effects in the model (Figure 9).

Table 9 contains the population mean and absolute mean residuals for both the SPML and Jordan mixed-effects models by height level. It can be seen that the assumption of mild quadratic trends observed in both the SPML and Jordan models is plausible. MFA was found to be under predicted at 1.4, 4.6, and 13.7 meters, and over predicted at 7.6 and 10.7 meters in

height for both the SPML and Jordan models. It can also be seen that the mean residual values are smaller for the SPML model at all heights compared to the Jordan model. However, the mean absolute residual values, for the Jordan model are small at all but 13.7 meters in height. Comparison of the population fit statistics of the SPML (Table 6) and Jordan models (Table 8) suggest that both models are suitable for population MFA predictions. Even though the Jordan model tends to have "overall" better qualities of fit, the differences between the two models are trivial.

Comparison of the two models at the tree level indicates overwhelmingly that the Jordan model is preferred. The statistics of fit at the tree level all favor the Jordan model and thus will yield more precise predictions of MFA with the inclusion of random effects. It must be noted that the SPML model was fit with only random intercept terms, whereas the Jordan model included three random-effects parameters allowing for individual initial value, rate, and asymptotic parameters. A logical extension of the mixed-effects SPML model presented in this work would be the addition of random-effects in the slope terms. It can only be assumed, that inclusion of slope random-effects in the SPML model would benefit the model fit, thus improving prediction.

TABLES AND FIGURES FOR CHAPTER 5

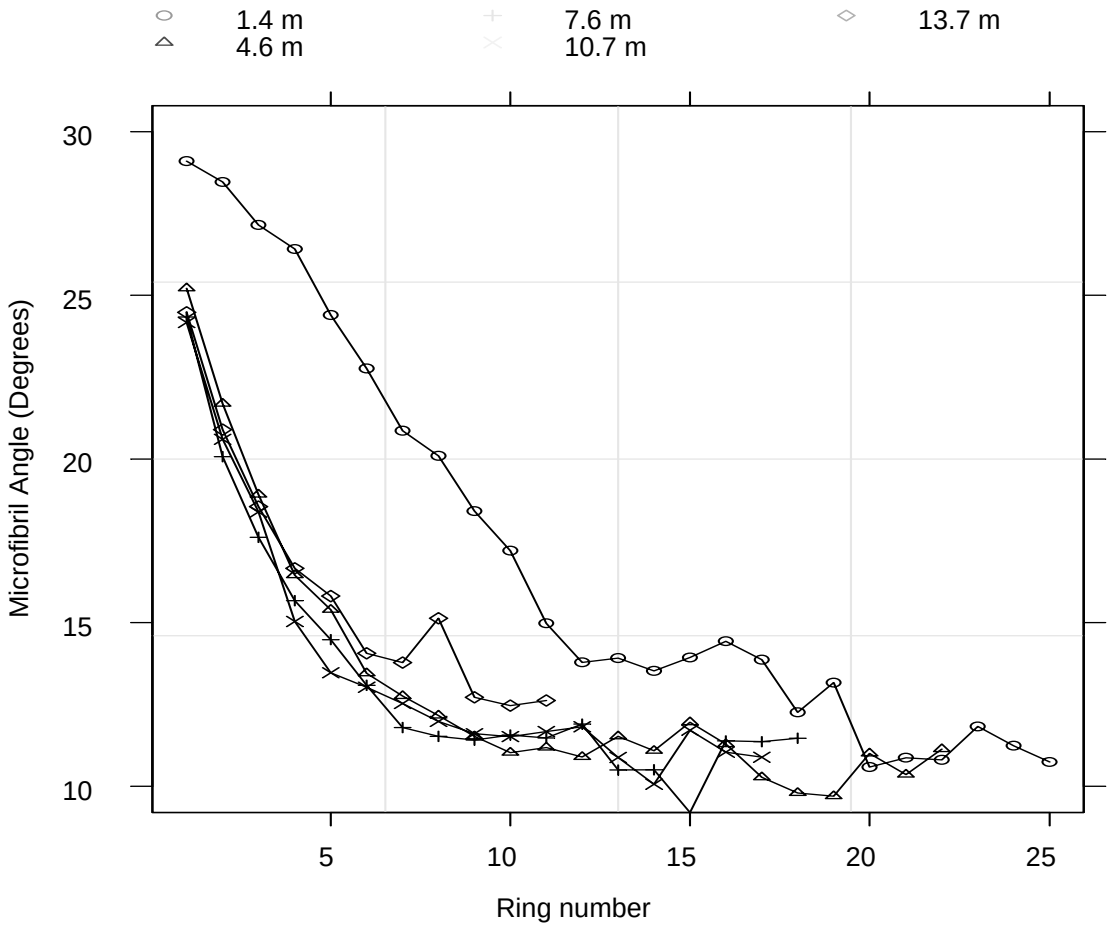

Figure 1. Plot of mean microfibril angle versus ring number by height level (m).

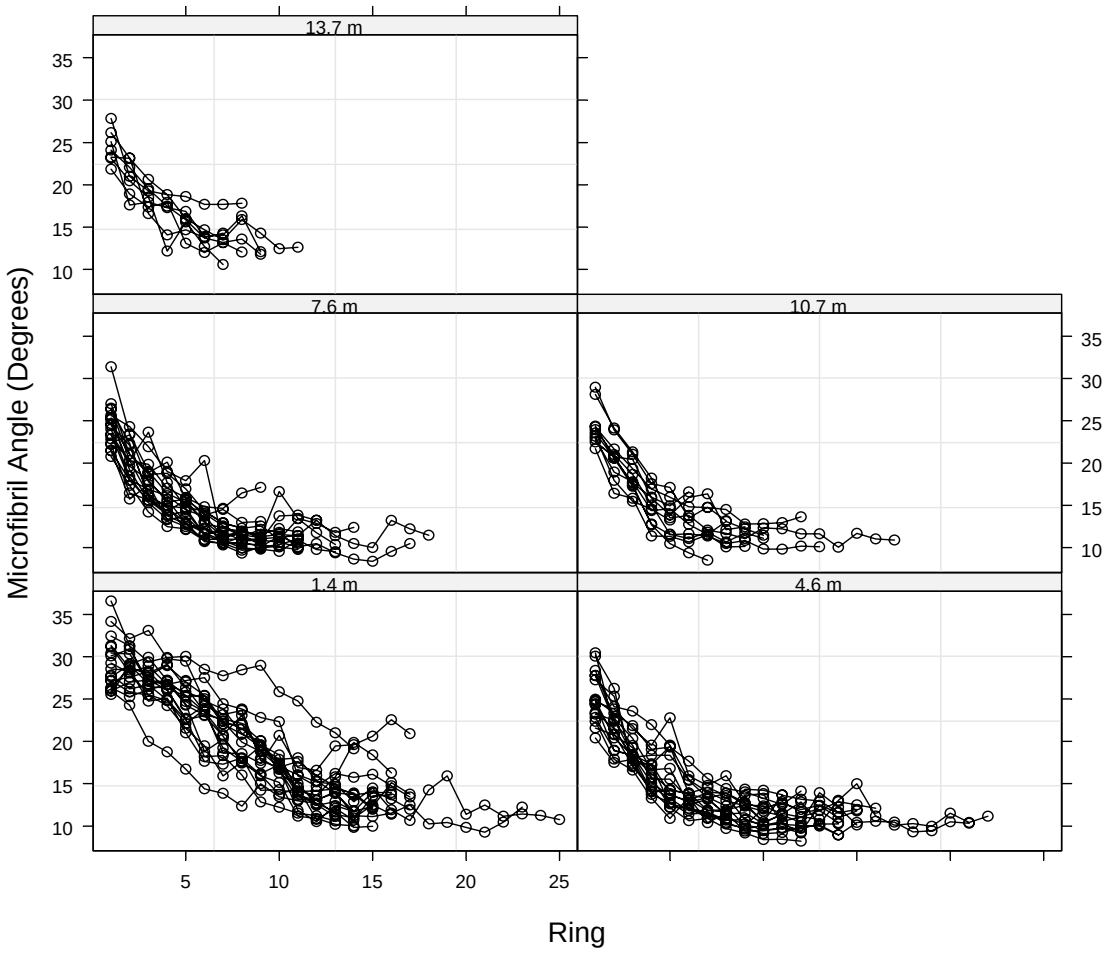

Figure 2. Plot of microfibril angle versus ring number by height level (m) for the 18 trees selected for microfibril analysis

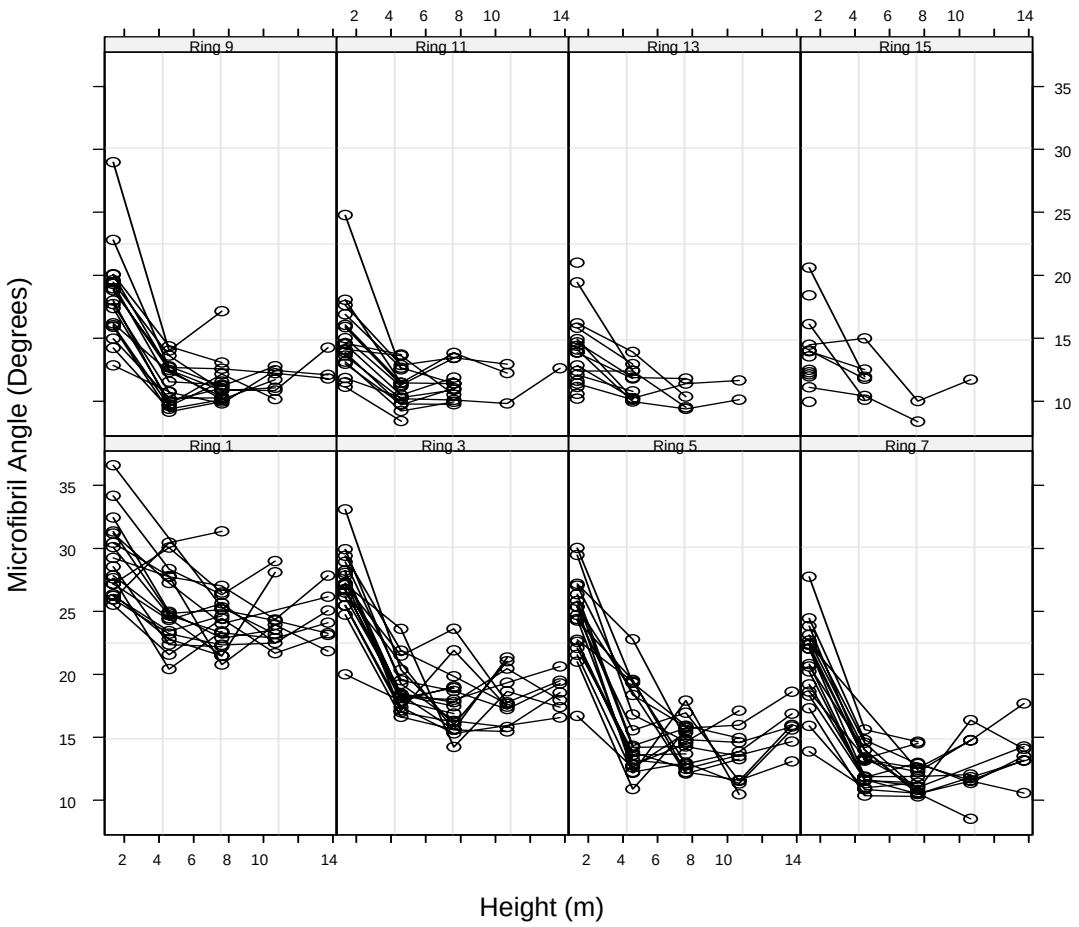

Figure 3. Plot of microfibril angle versus height level (m) by selected ring numbers for the 18 trees selected for microfibril analysis

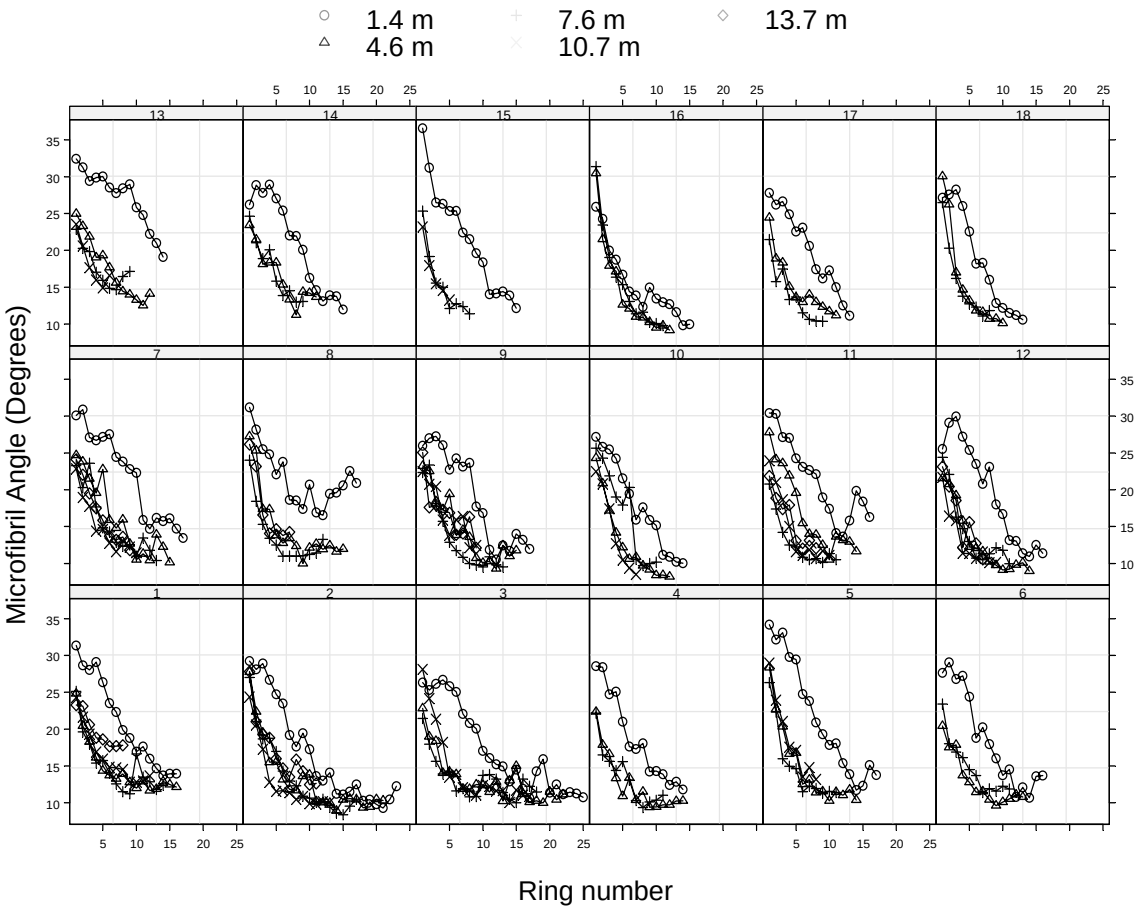

Figure 4. Plot of microfibril angle versus ring by height level (m) for each of the 18 trees selected for microfibril analysis.

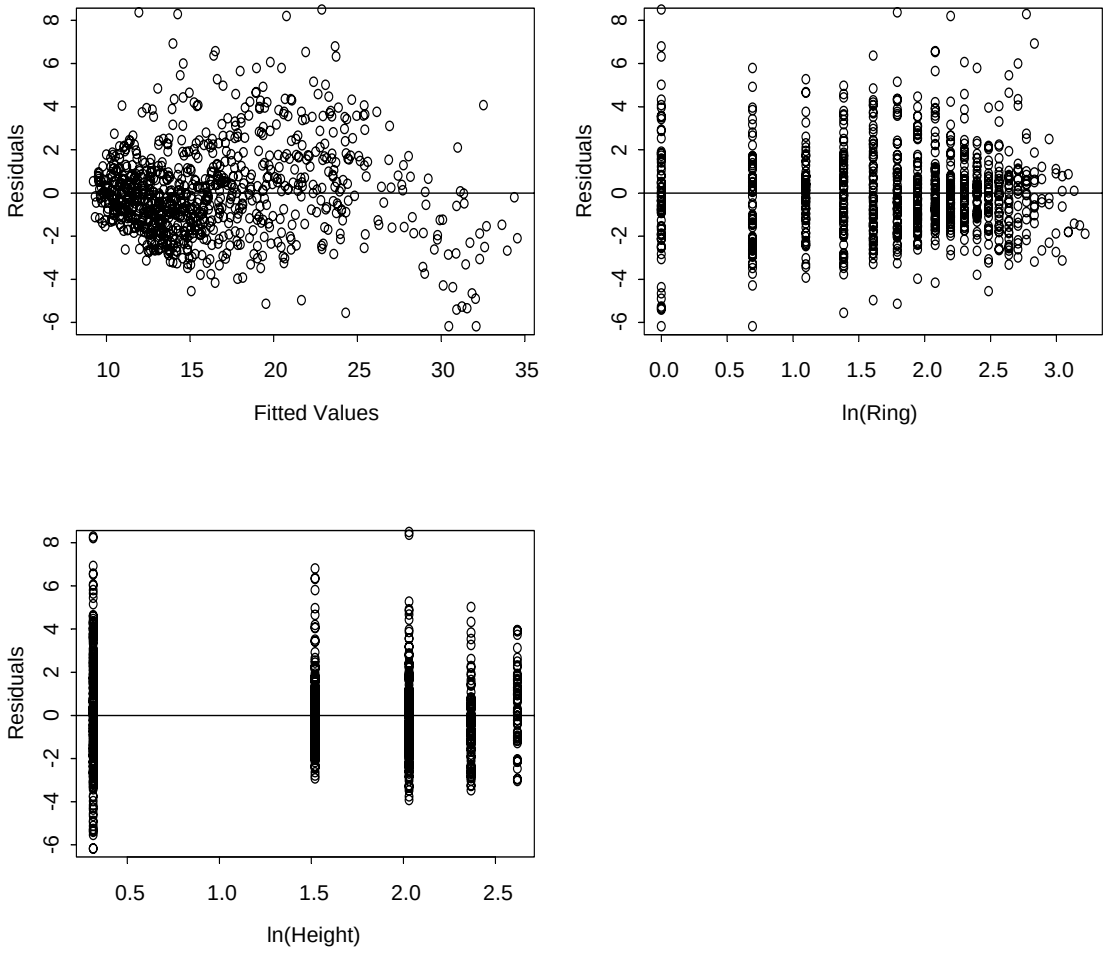

Figure 5. Plots of residuals versus fitted values, natural logarithm of ring, and the natural logarithm of height for the linear mixed-effects SPML model.

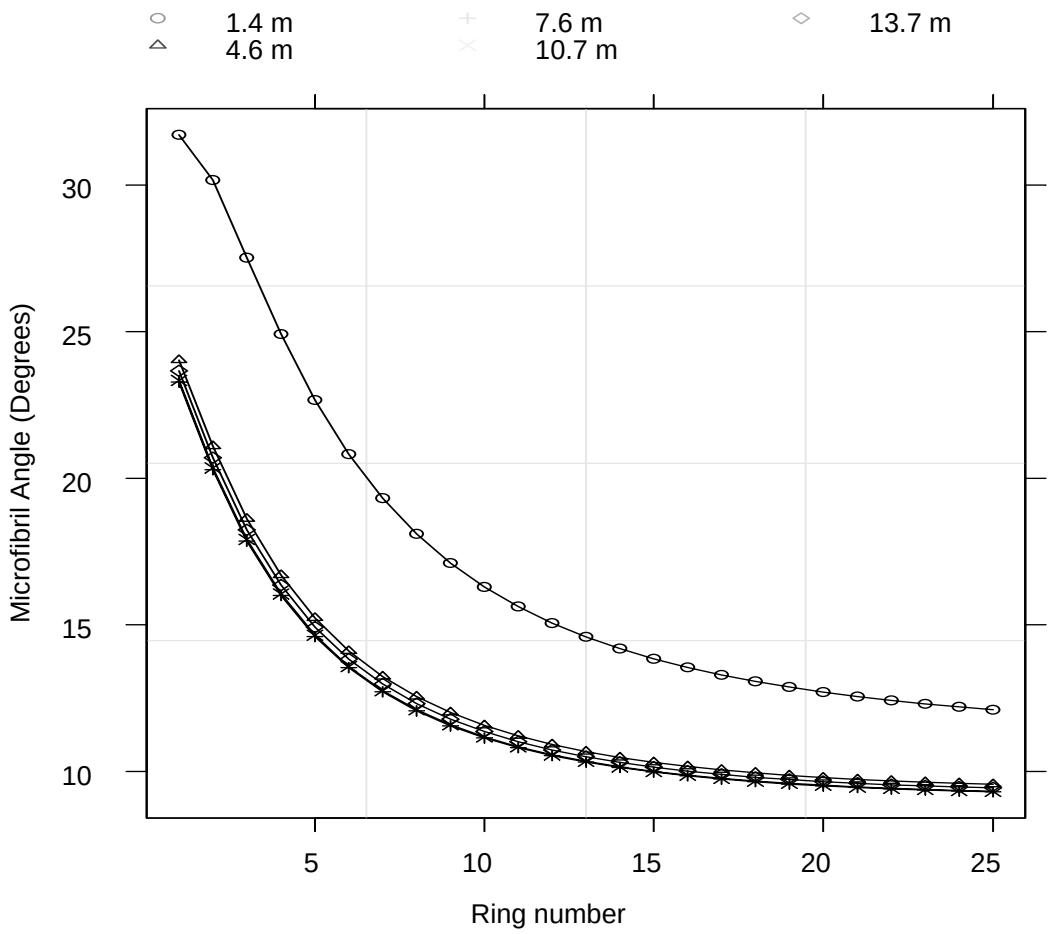

Figure 6. Plot of mean responses (typical responses) of microfibril angle versus ring by height level (m) for the SPML mixed-effects model.

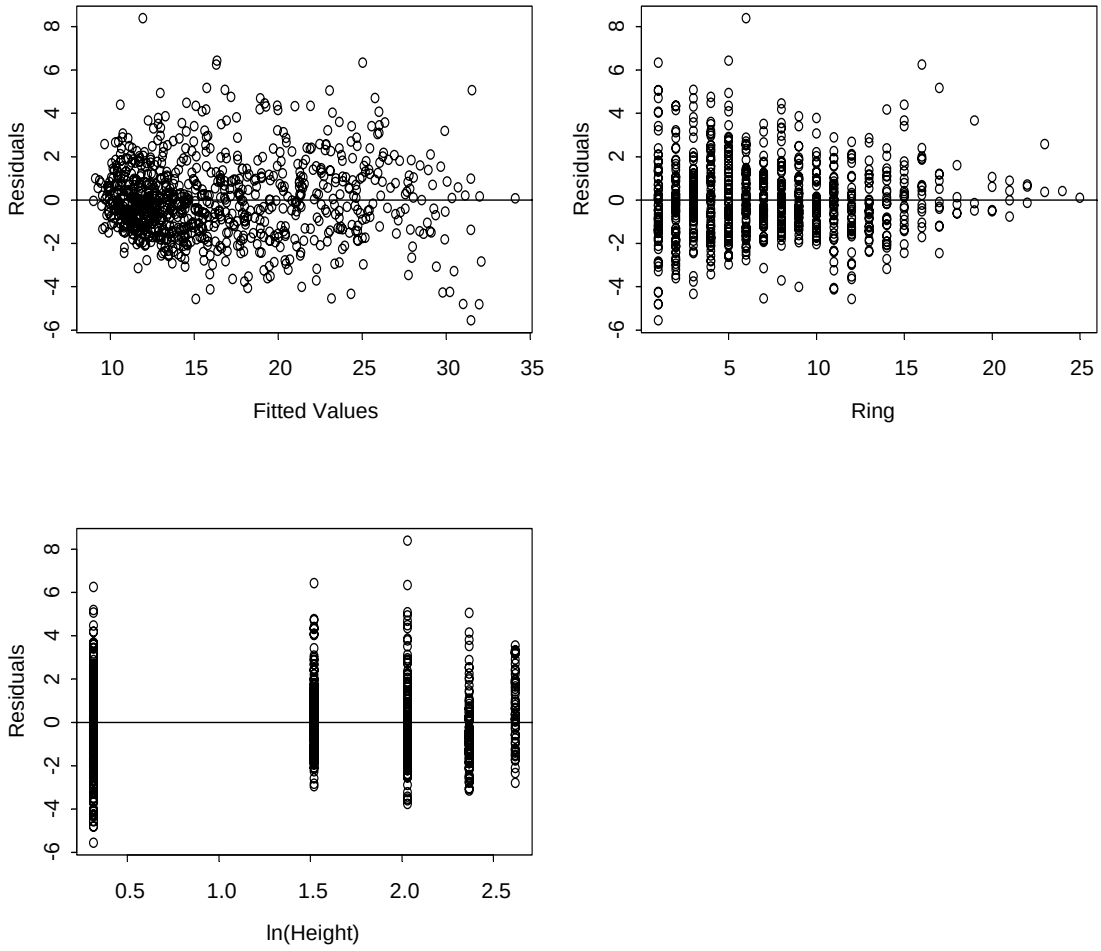

Figure 7. Plot of residuals versus fitted values, ring, and the natural logarithm of height for the Jordan model.

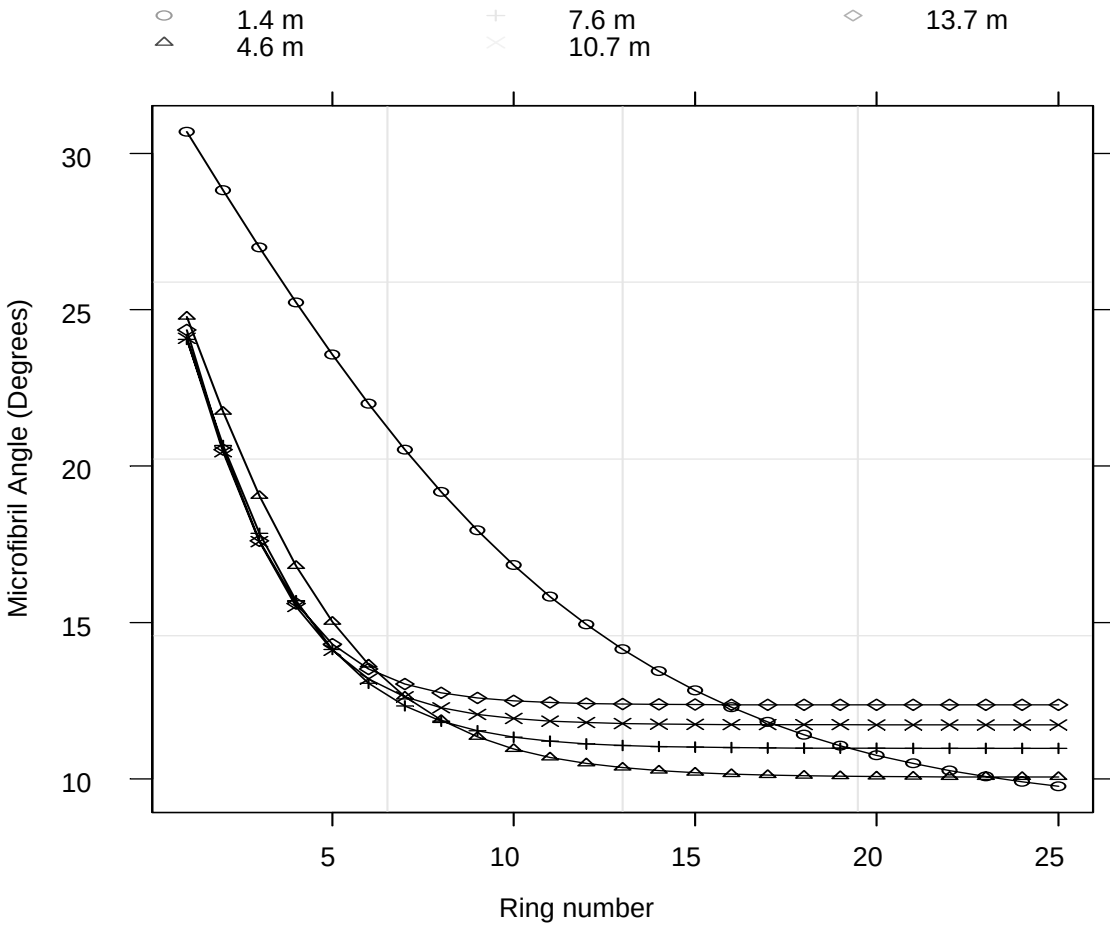

Figure 8. Plot of mean responses (typical responses) of microfibril angle by height level (m) for the Jordan model.

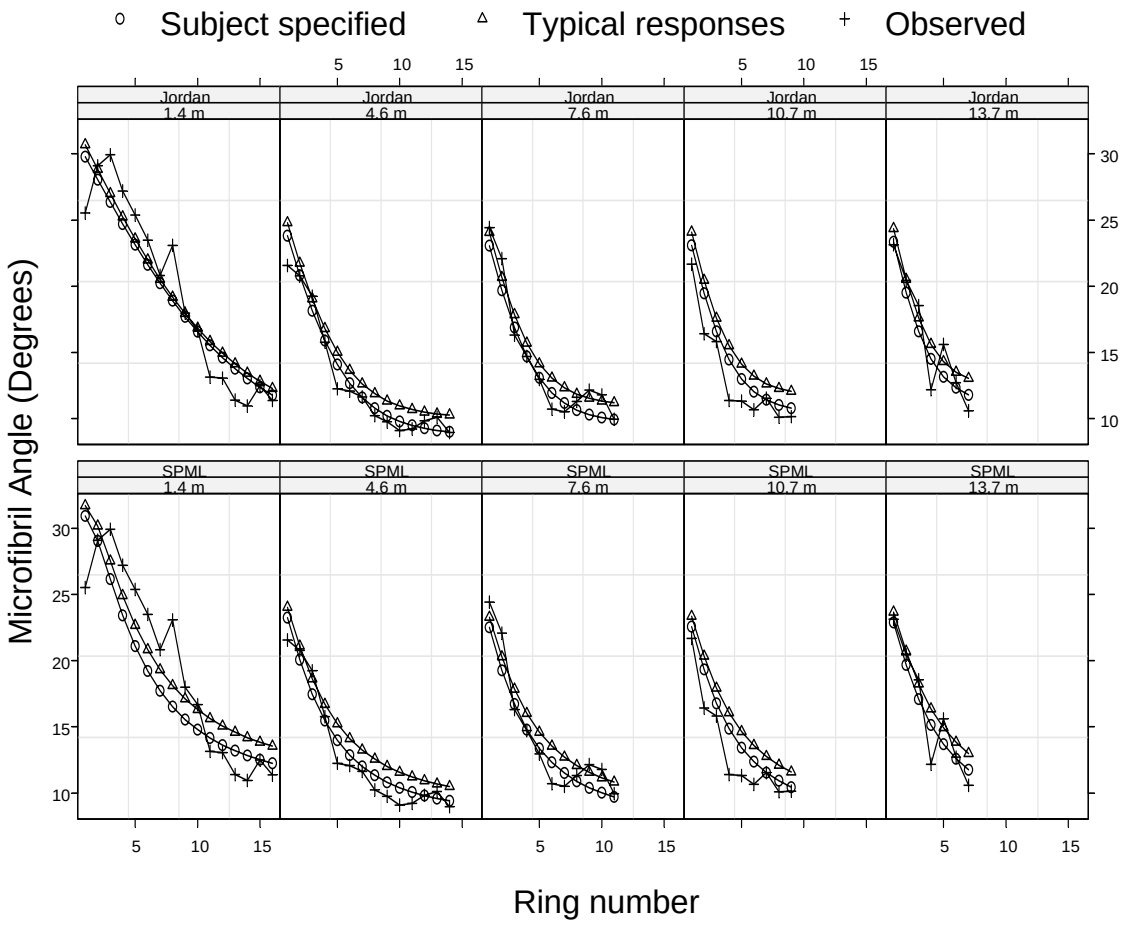

Figure 9. A comparison of individual microfibril angle predictions by typical response and subject-specified response with the Jordan and SPML mixed-effects models.

Table 1. Preliminary fits of the SPML fixed-effects model.

Model	Parameters	Model Formula	-2log(L)	AIC
1	6	Ring + Height	-2620	-2608
2	6	Ring + ln(Height)	-2832	-2820
3	6	ln(Ring) + ln(Height)	-2975	-2963
4	8	$\ln(\text{Ring}) + [\ln(\text{Ring})]^2 + \ln(\text{Height})$	-3052	-3036
5	8	$\ln(\text{Ring}) + \ln(\text{Height}) + [\ln(\text{Height})]^2$	-3101	-3085
6	10	$\ln(\text{Ring}) + [\ln(\text{Ring})]^2 + \ln(\text{Height}) + [\ln(\text{Height})]^2$	-3177	-3157

Table 2. Comparison of preliminary model fits from the SPML fixed-effects model.

Model	DF	χ^2	Δ AIC	*P*-Value
3 vs. 4	2	77	73	0.0000
3 vs. 6	4	202	194	0.0000
4 vs. 6	2	125	121	0.0000
5 vs. 6	2	76	72	0.0000

Table 3. Comparisons of the mixed-effects SPML model with different variance-covariance structures (Indep = Independent, Sym = General positive-definite).

Model	Parameters	Variance-Covariance Structure	-2log(L)	AIC
6	8	None	-3174	-3158
6.1	9	Indep: $\mathbf{b}_{i1}^{*}$	-3317	-3299
6.2	9	Indep: $\mathbf{b}_{i2}^{*}$	-3346	-3328
6.3	10	Indep: $\mathbf{b}_{i1}^{*}, \mathbf{b}_{i2}^{*}$	-3356	-3338
6.4	11	Sym: $\mathbf{b}_{i1}^{*}, \mathbf{b}_{i2}^{*}$	-3368	-3346

Table 4. Comparison of preliminary model fits from the SPML mixed-effects model.

Model	DF	χ^2	Δ AIC	Δ BIC	*P*-Value
6 vs. 6.1	1	143	141	171	0.0000
6.2 vs. 6.3	1	10	10	10	0.0015
6.3 vs. 6.4	1	12	8	6	0.0005

Table 5. Estimated parameters for population prediction of microfibril angle (random effects input as 0) with the SPML model.

Parameter	Estimate	Std. Error	*P*-value
β_{01}	19.7528	1.4474	0.0001
β_{11}	-11.1588	1.5306	0.0001
β_{21}	4.4091	0.5045	0.0001
β_{31}	13.1957	0.7235	0.0001
β_{41}	-3.0631	0.2610	0.0001
β_{02}	14.5923	0.7180	0.0001
β_{12}	-6.9602	0.6641	0.0001
β_{22}	1.4452	0.1825	0.0001

Table 6. Fit statistics from the SPML model including R^2, root mean square error (*RMSE*) mean residual (*MR*), mean absolute residual (*MAR*) at both the population and tree levels.

Fit Statistic	Level	
	Population	Tree
R^2	0.82	0.85
RMSE	2.52	2.13
MR	0.12	0.02
MAR	1.89	1.59

Table 7. Estimated parameters for population prediction of microfibril angle (random effects input as 0) with the modified Logistic function.

Parameter	Estimate	Std. Error	*P*-value
β_{00}	51.6845	1.3905	0.0001
β_{01}	-14.6684	2.1837	0.0001
β_{02}	2.8536	0.8657	0.001
β_{10}	0.1068	0.0133	0.0001
β_{11}	0.1706	0.0095	0.0001
β_{20}	8.8734	0.3479	0.0001
β_{21}	0.5089	0.0664	0.0001

Table 8. Fit statistics from the Jordan model including R^2, root mean square error (*RMSE*) mean residual (*MR*), mean absolute residual (*MAR*) at both the population and tree levels.

Fit Statistic	Level	
	Population	Tree
R^2	0.83	0.91
RMSE	2.42	1.76
MR	0.17	-5.97×10^{-7}
MAR	1.82	1.34

Table 9. Mean residual (*MR*) and mean absolute residual (*MAR*) at the population for the SPML and Jordan mixed-effects models by height level.

	SPML		Jordan	
Height (m)	*MR*	*MAR*	*MR*	*MAR*
1.4	0.18	2.56	0.27	2.37
4.6	0.09	1.63	0.27	1.60
7.6	-0.02	1.53	-0.08	1.49
10.7	-0.03	1.56	-0.15	1.51
13.7	0.75	1.60	0.84	1.64

CHAPTER 6

DISCUSSION AND CONCLUSIONS

Mixed-effects models are useful tools for analyzing repeated measures data. Their inherent flexibility allows for development of unique variance-covariance structures, which is limiting in traditional linear and nonlinear regression. Mixed-effects extensions of Presnell et al's. (1998) SPML model were developed, treating directional observations as projections onto the unit circle, and we applied these models to characterize the patterns of MFA. The natural logarithm of ring number and disk height were found to be good predictor variables and were incorporated into the SPML model. Statistical indices including LRT's, AIC, BIC, mean residual, and R^2 values show that the mixed-effects SPML provides more accurate predictions and an overall better fit, compared to fixed-effects SPML models.

For comparison against the mixed-effects SPML model, we used a variant of the Logistic function to characterize the patterns of MFA with ring and height. Population level predictions were found to be similar amongst the two models. Overall, both models were found to be unbiased across ring number, with a mild quadratic trend observed in the plot of residuals versus the natural logarithm of height. It should be noted that of the 18 trees used in this analysis, only 7 trees were tall enough to have disks removed at 13.7 meters. From these 7 trees there were a total of 118 ring observations. The bias of the models at upper heights may possibly be overcome by fitting to a more complete data set, properly calibrating the models.

The population level SPML and Logistic models are suitable for prediction of MFA for a population of trees. One way to attain better MFA predictions is to estimate random-effects by determining ring-by-ring MFA values from incremental cores taken from several sample trees. These estimated random effects will allow better prediction precision. This approach can be useful for predicting MFA values of a unique stand or plantation. Several trees from such a stand can be selected for MFA analysis. Each of these trees will yield unique tree level random-effects estimates, and the mean of these random effects can be used to adjust the MFA prediction for that stand or plantation.

The mixed-effects SPML model could be improved by incorporating random-effects in the slope terms, thus allowing for both random intercept and slope values, potentially increasing tree level prediction and the global population parameters. The mixed-effects SPML model could be generalized to higher dimensions and can potentially be used in analysis of directional data from designed experiments. In the analysis of variance setting, it may be possible to apply the mixed-effects SPML model to do a means only analysis (Presnell et al. 1998). The mixed-effects SPML models developed in this paper provide a prototype from which block designs, multilevel or repeated measures from directional data sets could be analyzed.

LITERATURE CITED

Abt, R., F. Cubbage, and G. Pacheco. 1999. Southern forest resource assessment using subregional timber supply (SRTS). Model Forest Prod. J. 50(4): 25-33 pp.

Addis, T., A.H. Buchanan, and J.C.F. Walker. 1995. A comparison of density and stiffness for predicting wood quality. Or Density: The lazy man's guide to wood quality. Journal of the Institute of Wood Science 13(6): 539-543 pp.

Bates, D.M. and D.G. Watts. 1988. Nonlinear regression analysis and its applications. Wiley, New York.

Batschelet, E. 1981. Circular statistics in biology. Academic Press, London.

Bendtsen, B.A., and J. Senft. 1986. Mechanical and anatomical properties in individual growth rings of plantation-grown eastern cottonwood and loblolly pine. Wood and Fiber Science 18(1): 23-38 pp.

Borders, B.E. and R.L. Bailey. 1997. Loblolly pine-pushing the limits of growth. CAPPS Tech. Rep. 1997-1. WSFR, University of Georgia, Athens, Georgia.

Briggs, D.G., and W.R. Smith. 1986. Effect of silvicultural practices on wood properties of conifers: a review. In Douglas Fir: Stand Management for the Future. C. Oliver, D. Hanley and J. Johnson (eds). University of Washington Press, Seattle, 108-117 pp.

Broerman, F.S. 1967. Nitrogen-phosphorous fertilization of slash pine. Woodlands Research Notes 18. Union Camp Corp. 4 pp.

Cave, I.D., and J.C.F Walker. 1994. Stiffness of wood in fast-grown plantation softwoods: the influence of microfibril angle. Forest Products Journal 44(5): 43-48 pp.

Clark III, A., and J.R. Saucier. 1991. Influence of planting density, intensive culture, geographic location, and species on juvenile wood formation in southern pine. Forest Products Journal 39(7/8): 42-48 pp.

Clark III, A., and M.B. Edwards. 1999. Effect of six site-preparation treatments on piedmont loblolly pine wood properties at age 15. A paper present at the tenth biennial southern silvicultural research conference, Shreveport, LA, February, 16-18 pp.

Clark III, A., J.R. Saucier, and T.I. Sarigumba. 1991. Effect of site preparation, planting density, and soil drainage on juvenile wood formation of slash pine. Gen. Tech. Rep. SE-70. Athens, GA: USDA Forest Service, Southern Research Station: 757-768 pp.

Cowdrey, D.R., and R.D. Preston. 1966. Elasticity and microfibrillar angle in the wood of sitka spruce. Proceedings of the Royal Society of London 166B(1004): 245-272 pp.

Dadswell, H.E., A.B. Waldrop, and A.J. Watson. 1958. The morphology , chemistry and pulping characteristics of reaction wood. In Fundamentals of Papermaking Fivers, ed. K Bolam. London: British Paper and Board Makers Association: 187-229 pp.

Daniels, R. F., R. He, A. Clark III, and R. Souter. 2003. Predicting wood properties of loblolly pine from stump to tip and pith to bark. In: Nepveu, G. (Ed). Proceedings of the Fourth Workshop: Connection Between Forest Resources and Wood Quality: Modelling Approaches and Simulation Software. IUFRO Working Party S5.01-04. Harrison Hot Springs, B.C. CA. Sept 8-14, 2002. INRA-Centre de Recherches de Nancy, France.

Daniels, R.F. 1999. Effects of tree, stand, silvicultural and environmental variables on wood properties of planted loblolly pine in Georgia. Georgia TIP3 Research Proposal. WSFR, University of Georgia, Athens, Georgia.

Davidian, M., and D.M. Giltinan. 1995. Nonlinear models for repeated measurement data. Chapman and Hall, London.

Dempster, A.P., N.M. Laird, and D.B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. J. Royal Stat. Soc, Ser. B. 39: 1-22 pp.

Dickson, R.L. and J.C.F. Walker. 1997. Selecting wood quality characteristics for pines. In Timber Management Toward Wood Quality and End-Product Value. S.Y. Zhang, R. Gosselin and G. Chauret (eds). Proceedings of the CTIA/IUFRO International Wood Quality Workshop, Quebec City. Part IV. 45-52 pp.

Elliott, G.K. 1970. Wood density in conifers. Tech. Comm. No. 8. Commonwealth Forestry Bureau, Oxford, England.

Fisher, N.I., and A.J. Lee. 1992. Regression models for an angular response. Biometrics 48: 665-677 pp.

Fisher, N.I. 1993. Statistical analysis of circular data. University Press, Cambridge.

Fosgate, H. 2001. Wood quality begins in the woods. The Forester's Log. Alumni Association Publication. WSFR, University of Georgia, Athens, Georgia. Fall/Winter: 4-5 pp.

Gregoire, T.G., D.R. Brillinger, P.J. Diggle, E. Russek-Cohen, W.G. Warren, and R.D. Wolfinger. 1997. Modeling longitudinal and spatially correlated data. Springer, New York. 402 pp.

Gould, A.L. 1969. A regression technique for angular variates. Biometrics 25: 683-700 pp.

Hall, D.B., and M. Clutter. 2004. Multivariate multilevel nonlinear mixed effects models for timber yield predictions. Biometrics. 60: 16-24 pp.

Hall, D.B., and K.S. Berenhaut. 2002. Score tests for heterogeneity and overdispersion in zero-inflated poisson and binomial regression models. Can. J. Stat. 30(3).

Harris, J.M. 1981. Effect of rapid growth on wood processing. In Proc, Div. 5, 17th IUFRO World Congress, Japan. 117-125 pp.

Harris, J.M., and B.A. Meylan. 1965. The influence of microfibril angle on longitudinal and tangential shrinkage in Pinus radiata. Holzforschung 19: 144-153 pp.

Haygreen, J.G., and J.L. Bowyer. 1996. Forest products and wood science. An introduction. Iowa State University Press, Ames, Iowa. 484 pp.

Haywood, J.D. and J.D. Burton. 1989/1990. Phosphorous fertilizer, soils, and site preparation influence loblolly pine productivity. New Forests 3:275-287 pp.

He, R., L. Jordan, R.F. Daniels, and A. Clark. 2005. Mixed effect smoothing spline analysis of variance models for loblolly pine microfibril angle. *Submitted for publication in Forest Science*

Hiller, C.H. 1964. Correlation of fibril angle with wall thickness of tracheids in summerwood of slash and loblolly pine. TAPPI Journal 47: 125-128 pp.

Jammalamadaka, S.R., and A. SenGupta. 2001. Topics in circular statistics. World Scientific. New Jersey.

Jokela, E.J., and L.A. Morris. 1998. Interactions between fertilizer additions and competition control on the growth of southern pine plantations. Proceedings, Southern Weed Science Society 51: 111-112 pp.

Jordan, L., R.F. Daniels, A. Clark, and R. He. 2005. Multilevel nonlinear mixed effects models for the modeling of earlywood and latewood microfibril angle. *Accepted for publication in Forest Science*.

Jordan, L., R.F. Daniels, and A. Clark. 2005. Variation in loblolly pine microfibril angle with physiological age, height, and physiographic region. University of Georgia Wood Quality Consortium. Tech. Rep. pp. 40.

Johnson, R.A., and T.E. Wehrly. 1978. Some angular-linear distributions and related regression models. JASA 73: 602-606 pp.

Kellogg, R.M., E. Thykeson and W.G. Warren. 1975. The influence of wood and fiber properties on kraft converting-paper quality. Tappi 58(12): 113-116 pp.

Kendall, D.G. 1974. Pole-seeking Brownian motion and bird navigation. Journal of the Royal Statistical Society, Ser. B, 36: 261-294 pp.

Kliger, I.R., G. Johansson, M. Perstorpher, and D. Engstrom. 1994. Formulation of requirements for the quality of wood properties used by the construction industry. Final Report EC Contract No. MA2B-0024. Chalmers University of Technology, Sweden.

Kvålseth, T.O. 1985. Cautionary note about R^2. Am. Stat. 39: 279-285.

Loague, K., and R.E. Green. 1991. Statistical and graphical methods for evaluating solute transport models: overview and application. J. Contam. Hydrol. 7:51-73.

Laird, N.M., and J.H. Ware. 1982. Random-effects models for longitudinal data. Biometrics. 38: 963-974 pp.

Laird, N.M., Lange, N., and D. Stram. 1987. Maximum likelihood computations with repeated measures: Application of the EM algorithm. JASA. 82: 97-102 pp.

Larson, P.R., D.E. Kretschmann, E. David, A. Clark, and J.G. Isebrands. 2001. Formation and properties of juvenile wood in southern pines: a synopsis. Gen. Tech. Rep. FPL-GTR-129. Madison, WI: USDA, Forest Service, Forest Products Lab. 42 pp.

Lichtenegger, H., A. Reiterer, S.E. Stanzl-Tschegg, and P. Fratzl. 1999. Variation of cellulose microfibril angles in softwoods and hardwoods-a possible strategy of mechanical optimization. Journal of Structural Biology 128: 257-269 pp.

Lindstrom, M.J., and D.M. Bates. 1990. Nonlinear mixed-effects models for repeated measures data. Biometrics. 46: 673-687 pp.

Lindstrom, M.J., and D.M. Bates. 1988. Newton-Raphson and EM algorithms for linear mixed-effects models for repeated measures data. JASA. 83(404): 1014-1022 pp.

MacDonald, E., and J. Hubert. 2002. A review of the effects of silviculture on timber quality of sitka spruce. Forestry 75(2): 107-138 pp.

MacPeak, M.D., L.F. Burkhart, and D. Weldon. 1987. A mill study of the quality, yield, and mechanical properties of plywood produced from fast-grown loblolly pine. Forest Products Journal 37(2): 51-56 pp.

MacPeak, M.D., L.F. Burkhart, and D. Weldon. 1990. Comparison of grade, yield, and mechanical properties of lumber produced from young fast-grown and older slow-grown planted slash pine. Forest Product Journal 40(1): 11-14 pp.

Mardia, K.V. 1972. Statistics of directional data. Academic Press, London.

Mardia, K.V. 1975. Statistics of directional data. Journal of the Royal Statistical Society. Series B (Methodological). 37(3): 349-393 pp.

Mardia, K.V., and P.E. Jupp. 2000. Directional Statistics. Wiley, New York.

Mayer, D.G., and D.G. Butler. 1993. Statistical validation. Ecol. Model. 68:21-32.

McCalister, R.H. and A. Clark. 1992. Shrinkage of juvenile and mature wood of loblolly pine from three locations. Forest Products Journal 42(7/8): 25-28 pp.

McMillan, C.W. 1968. Morphological characteristics of loblolly pine wood as related to specific gravity, growth rate, and distance from pith. Wood Science and Technology 2: 166-176 pp.

Megraw, R.A. 1985. Wood quality factors in loblolly pine. The influence of tree age, position in tree, and cultural practice on wood specific gravity, fiber length, and fibril angle. Tappi Press. Atlanta, Georgia. 88 pp.

Megraw, R.A., D. Bremer, G. Leaf, and J. Roers. Stiffness in loblolly pine as a function of ring position and height, and it relationship to microfibril angle and specifc gravity. Proceeding of the Third Workshop-Connection between silviculture and wood quality through modeling approaches. IUFRO S5.01-04. Sept. 5-12, 1999. 341-349 pp.

Meylan, B.A. 1968. Cause of high longitudinal shrinkage in wood. Forest Products Journal 18(4): 75-78 pp.

Mitchell, H.L. and P.R. Wheeler. 1959. The search for wood quality. Forest Farmer 18(4): 4-6 pp.

Navidi, W. 1997. A graphical illustration of the EM algorithm. The American Statistician. 51(1): 29-31 pp.

Phillips, C. 2002. Modeling within-tree changes on wood specific gravity and moisture content for loblolly pine in Georgia. M.S. Thesis. University of Georgia. 83 pp.

Pinheiro, J.C. and D.M. Bates. 2000. Mixed-effects models in S- and S-PLUS, Springer, New York. 528 pp.

Pinheiro, J.C. and D.M. Bates. 1996. Unconstrained parameterizations for variance-covariance matrices. Statistics and Computing. 6: 289-296.

Pillow, M.Y., B.Z. Terrell, and C.H. Hiller. 1953. Patterns of variation in fibril angles in loblolly pine. Mimeo. D1953. Washington, D.C.: USDA, Forest Service. 11 pp.

Presnell, B., S.P. Morrison, and R.C. Littell. 1998. Projected multivariate linear models for directional data. JASA. 93(443): 1068-1077 pp.

Ratkowsky, D.A. Handbook of nonlinear regression models. 1990. Marcel Dekker Inc. New York. 231pp.

SAS Institute Inc. 1999. SAS OnlineDoc. http://v8doc.sas.com/sashtml/.

Scapini, F., Aloia, A., Bouslama, M.F., Chelazzi, L., Colombini, I., ElGtari, M., Fallaci, M., and Marchetti, G.M. 2002. Multiple regression analysis of the sources of variation in orientation of two sympatric sandhoppers, *Talitrus saltator* and *Talorchestia brito*, from an exposed Mediterranean beach. Behavioral Ecology and Sociobiology. 51: 403-414 pp.

Schlaegel, B.E. 1982. Boxelder (Acer negundo L.) biomass component regression analysis for the Mississippi Delta. Forest Science. 28: 355-358 pp.

Searle, S.R., G.Casella, and C.E. McCulloch. 1992. Variance Components. John Wiley and Sons. New York. 501 pp.

Senft, J.F., B.A. Bendtsen, and W.L. Galligan. 1985. Weak wood: Fast-grown trees make problem lumber. Journal of Forestry. August: 477-484 pp.

Schmidtling, R.C. 1973. Intensive culture increases growth without affecting wood quality of young southern pines. Canadian Journal of Forest Research. 3: 565-573 pp.

Sheiner, L.B., and S.L. Beal. 1980. Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model: Routine clinical pharmacokinetic data. Journal of Pharmacokinetics and Biopharmaceutics. 8(6): 553-571 pp.

Tamalong, F.N., F.F. Wangaard, and R.M. Kellogg. 1967. Strength and stiffness of hardwood fibers. Tappi 56(2): 68-72 pp.

Taylor, F.W. and J.S. Moore. 1981. A comparison of earlywood and latewood tracheid lengths in loblolly pine. Wood and Fiber 13: 159-165 pp.

Treacy, M., J. Evertsen, and A.I. Dhubhain. 2000. A comparison of mechanical and physical wood properties of a rang of sitka spruce provenances. COFORD, Helsinki, Finland. 35 pp.

Verbeke, G and G. Molenberghs. 2000. Linear mixed models for longitudinal data. Springer, New York. 568 pp.

Vonesh, E.F. and R.L. Carter. 1992. Mixed-effects nonlinear regression for unbalanced repeated measures. Biometrics. 48: 1-18 pp.

Vonesh, E.F. and V.M. Chinchilli. 1997. Linear and nonlinear models for the analysis of repeated measures. Marcel Deker, New York. 560 pp.

Walker, J.C.F., and B.G. Butterfield. 1996. The importance of the microfibril angle for the processing industries. New Zealand Forestry. Feb. 1996: 34-40 pp.

Walker, J.F.C. 1993. Primary wood Processing: Principles and Practice. Chapman and Hall, London.

Wardrop, A.B. 1951. Cell wall organization and properties of the xylem. Aust. J. Sci. Res. Series B, Biol. Sci. 4: 391-414 pp.

Ware, J.H. 1985. Linear models for the analysis of longitudinal studies. The American Statistician. 39: 95-101 pp.

Watson, A.J. and H.E. Dadswell. 1964. Influence of fibre morphology on paper properties. 4. Micellar spiral angle. Appita. 17: 151-156 pp.

Watson, G.S. 1983. Statistics on spheres. New York, Wiley.

Wear, D.N. and J.G. Greis. 2002. Southern forest resource assessment: Summary of findings. Journal of Forestry. 100(7):6-14 pp.

Wolfinger, R.D. 1999. Fitting nonlinear mixed models with the new NLMIXED procedure. Paper 287, Proceedings of the 99 Joint Statistical Meetings.

Ying, L., D.E. Kretschmann, and B.A. Bendtsen. 1994. Longitudnal shrinkage in fast-grown loblolly pine plantation wood. Forest Products Journal 44(1): 58-62 pp.

Zhang, S.Y. 1995. Effect of growth rate on wood specific gravity and selected mechanical properties from distinct wood categories. Wood Sci. Technol. 29: 451-465 pp.

Zobel, B.J. and J.B. Jett. 1995. Genetics of wood production. Springer-Verlag, Berlin. 337 pp.

Zobel, B.J. and R. Blair. 1976. Wood and pulp properties of juvenile wood and topwood of the southern pines. Journal of Applied Polymer Sciences 28: 421-433 pp.