MODELING AND ALLOCATING FORESTRY SURVIVAL: A LOBLOLLY PINE CASE STUDY

by

CHARLES E. ROSE Jr.

(Under the direction of Michael L. Clutter and Barry D. Shiver)

ABSTRACT

The forest industry has a major economic impact on the Southeastern United States and loblolly pine (*Pinus taeda* L.) is the primary commercial species. Consequently, many management tools have been developed to aid in the management of loblolly pine. These tools include growth and yield models that predict stand growth and corresponding wood yield. Hence, growth and yield systems, which generally include survival, basal area, height, and volume models, have garnered considerable research attention in recent years. Prediction of surviving stems per unit area, which is critical in forecasting wood yield, is an important component of these growth and yield systems. The importance of survival prediction can be demonstrated using whole stand and individual tree or *dbh* class survival predictions for projecting stand tables, which are used for generating stock tables. Our study, which uses permanent plot loblolly pine data, builds upon the existing foundation of forestry survival models: both whole stand and individual tree, and assesses the impact of mortality allocation in stand table projection algorithms. We develop a generalized methodology for deriving flexible whole stand survival models for the continuum of a stand's development by merging traditional survival analysis and existing whole stand survival methods. In addition, we demonstrate a methodology for modeling interval-censored individual tree survival data and show that the model derivation naturally leads to the complementary log-log survival function. Our individual tree survival model accounts for heterogeneity that occurs within and among plots by using multilevel modeling techniques. Moreover, since logistic regression is the most common technique used for modeling individual tree survival, we document the utility of using a multilevel individual tree logit model. Lastly, the multilevel logit individual tree survival model is used in projecting stand tables and assessed with a commonly used stand table projection algorithm.

INDEX WORDS: Loblolly pine, Survival modeling, Multilevel models, Whole stand, Individual tree, Stand table projection

MODELING AND ALLOCATING FORESTRY SURVIVAL: A LOBLOLLY PINE CASE STUDY

by

Charles E. Rose Jr.

B.S., Northern Arizona University, 1995

M.S., Oklahoma State University, 1998

M.S., The University of Georgia, 2000

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2002

MODELING AND ALLOCATING FORESTRY SURVIVAL: A LOBLOLLY PINE CASE STUDY

by

Charles E. Rose Jr.

Approved

Major Professors: Barry D. Shiver
Michael Clutter

Committee: Bruce Borders
Chris Cieszewski
Daniel Hall
Glenn Ware

Electronic Version Approved:

Gordhan L. Patel
Dean of the Graduate School
The University of Georgia
August 2002

DEDICATION

As I pursued an education, from Northern Arizona University to Oklahoma State and finally here at The University of Georgia, I would never have completed the journey and finished this dissertation without the support and encouragement from family and friends. I often wished that my mother Vivian, who died on August 9, 1981, was alive to see me complete my dissertation, but her presence was always felt. During the 10 years of my educational journey, there have been numerous factors that enabled me to complete this dissertation, however, there was always one constant, my dog Dusty. Dusty was there to greet me and to relieve my stress at the end of some long days as I completed this dissertation. She was there as I pursued my education from Arizona to Oklahoma to Georgia and never wavered in her undying devotion. Dusty was almost fourteen when she died during the afternoon on the day before I defended this dissertation, which ironically focuses on survival analysis, long may she run. I dedicate this dissertation in the memory of my mother Vivian (February 18, 1931 – August 9, 1981) and dog Dusty (September, 1988 – July 9, 2002), both will forever live in my memories.

ACKNOWLEDGEMENTS

It has been a long and arduous journey from Northern Arizona University to the completion of my Ph.D. here at the University of Georgia. I wish to thank Dr. Robert L. Bailey who encouraged me to pursue a Ph.D. and served as my advisor before retiring in April 2001. I want to express my gratitude to my two major professors whose patience, encouragement, support, and advice were invaluable: Dr. Michael L. Clutter and Dr. Barry D. Shiver. Thanks to Dr. Daniel Hall, whose idea led me to pursue survival analysis in a forestry context, and who always had the time and patience to explain the statistical concepts. My sincere gratitude to committee members Dr. Chris Cieszewski, Dr. Bruce Borders, and Dr. Glenn Ware for their support and encouragement.

Thanks to all the people who believed in me and provided encouragement while I completed my dissertation at the University of Georgia. In addition, I would like to thank our Graduate Coordinator Rosemary Woods, and also Mike Harrison and Dr. Yujia Zhang for their support with the data. My sincere gratitude to the PMRC and CAPPS organizations, without their support this dissertation would not have been possible.

Thanks to my family, my father who was always willing to help, and my sister Liz and her husband Mike who were extremely supportive, and to Angie, who in pursuing her own educational goals has tolerated my long absences and supported my educational decisions. I could never have completed this dissertation without the above people and the hundreds of others who contributed in some fashion.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

The forest industry has a major economic impact on the southeastern United States and loblolly pine (*Pinus taeda* L.) is the primary commercial species. Loblolly pine is found naturally on habitats from the poorly drained floodplains of the lower coastal plains to well drained hilly slopes, e.g., Piedmont region. Its range extends from southern New Jersey to central Florida and west to eastern Texas and southeastern Oklahoma (Little 1996). Loblolly pine is highly adaptable and is one of the fastest growing southern pines; hence, its attraction as a commercial species.

Since the 1960's, the primary source of loblolly pine fiber has shifted from natural stands to plantations (Cost 1989). In Georgia, there are more than 24 million acres of forestland out of a land base of approximately 37 million acres. Currently, the Georgia forest product industry accounts for more than 177,000 jobs and 19.5 billion in income. Therefore, many management tools have been developed to effectively manage the forest resource. These tools include growth and yield models that predict the expected fiber flow, within some degree of certainty, from managed plantations and natural stands.

Growth and yield systems, which generally include survival, basal area, height, and volume models, have garnered considerable research attention in recent years. This increased research attention is due, in part, because it has become increasingly common to manage loblolly pine plantations using intensive management, e.g., fertilization, herbicide, and mechanical site preparation. Consequently, there has been considerable

focus on developing growth and yield models that account for intensive management practices. An important component of a growth and yield system is the prediction of surviving stems per unit area, which is critical in forecasting fiber flow.

Forestry survival modeling is generally done at two resolutions, whole stand and individual tree. Whole stand survival models predict the future stems per unit area given an initial number of stems and corresponding age, and have proven reliable within the data range. There has however, been difficulty predicting survival for a stand's early development phase. In addition, whole stand survival models developed for early stand survival have typically been modeled separately from the rest of the stand's continuum of survival (e.g., Hitch et al. 1996, Matney and Farrar 1992). Moreover, most whole stand survival models developed for plantation forestry in the southern United States have been developed using age 5-25 (yr) data. Therefore, it has become common to use a system of equations, tables, or combination thereof to predict survival for a stand's continuum (e.g., Amateis et al. 1997). In addition, use of multiple equations and/or tables to model whole stand pine plantations is often due to the difficulty in adequately modeling whole stand survival for the continuum of a stand's development. Here we propose and develop the methodology for modeling the lifespan of whole stand plantation survival using a single equation.

Most forestry survival modeling has focused on the individual tree level, probably because of the prolific use of individual tree simulators. It is common knowledge that forest survival tends to be highly variable from stand to stand and within a stand, i.e., two identical stands growing under identical conditions at age_1 can have vastly different survival rates at age_2. This heterogeneity may be accounted for by using a multilevel

modeling approach for individual tree survival. A multilevel approach, which accounts for the groupings of trees within a plot, should improve our understanding of processes that affect survival. For example, modeling heterogeneity at different levels will allow for quantification of survival variability when using different silvicultural treatments. In addition, methods developed for multilevel individual tree survival models can be easily used in other circumstances where the attribute of interest is a binary or proportional response. For example, it has been suggested that the gain in leaf biomass when using some intensive management techniques (e.g., nitrogen fertilization) can result in a decrease in the fine root mass (Landsberg and Gower 1997). Some current studies involve photographing fine roots and death of the root during an interval is recorded. Hence, developing and demonstrating methods for developing multilevel survival models is directly applicable to current and future binary and proportional response studies in forestry. Additionally, survival models that account for the heterogeneity at different levels can further our understanding of the underlying biological processes. Survival estimation for individual trees can be viewed as a two-step process: developing a model for predicting each tree's mortality probability and mortality allocation based upon the predicted probabilities. In individual tree simulators, mortality allocation is usually thought of as a classification problem, i.e., trees are classified using a threshold as dead or alive based upon their respective predicted mortality probability. Here, for individual tree survival, we focus on developing the foundation and demonstrating the viability of a multilevel individual tree survival modeling approach. Therefore, our purpose is not to develop the best classification threshold, however we will demonstrate threshold sensitivity using several alternative thresholds for classifying trees as dead or alive.

Regardless of forestry survival resolution, i.e., whole stand or individual tree, lack of survival research, relative to other growth and yield models, is probably due to insufficient and inadequate long-term survival data. It's common for permanent plot forestry survival data to have more trees recorded at the second measurement occasion relative to the first measurement occasion, assuming no ingrowth. This recorded measurement discrepancy impacts survival estimation more profoundly than other growth and yield models because the survival curve would be increasing from t_1 (time 1) to t_2 (time 2), where $t_2 > t_1$. Conversely, a measurement discrepancy impact on an attribute such as basal area per unit area is to over estimate the amount of basal area, which is still logical. In fact, growth and yield attributes affected by an illogical increase in trees per unit area due to a measurement discrepancy can logically increase or decrease from t_1 to t_2. For example, volume per unit area can increase or decrease, although without excessive mortality or large trees succumbing it is unlikely. Hence, modeling forestry survival is usually hindered by insufficient long-term and deficient data. In the southeastern United States, with respect to plantation forestry, survival is the major plantation growth and yield component where there is a lack of adequate models (Dr. Bruce Borders FORS8480 notes), which is probably due to a lack of adequate data for predicting survival under a wide range of stand conditions.

Prediction of whole stand and individual tree or *dbh* class survival is important for projecting stand tables. Usually whole stand survival models are used to predict stems per unit area and the implied mortality is then allocated to the *dbh* classes based upon some method. Obviously, our whole stand survival model can be used to predict stems per unit area, however our interest is in using the individual tree survival model, by *dbh*

class and individual tree, to predict mortality probabilities for use in projecting stand tables. This study proposes to build upon the existing foundation of forestry survival models: both whole stand and individual tree, and to assess the impact of mortality allocation in stand table projection algorithms. Our study objectives, presented in the order of development, are:

- Develop a generalized methodology for deriving flexible whole stand survival models for the continuum of a stand's development. Model flexibility is defined as the ability to model more than one inflection point for a survival curve, if the data warrants. It has been noted in previous studies (e.g., Siler 1979) that for many biological populations it is reasonable to assume the "true" survival curve has more than one inflection point. Our model development will merge traditional survival analysis with existing whole stand survival methods.
- Demonstrate methodology for modeling interval-censored forestry survival data. The developed methodology will show that model derivation naturally leads to the complementary log-log survival function. Moreover, we account for heterogeneity that occurs within and among plots by using multilevel modeling techniques. The random effect prediction procedure is demonstrated and mortality probability predictions are compared, with and without, the inclusion of random effects.
- Logistic regression is the most common technique used for modeling individual tree survival, therefore we document the utility of using a multilevel individual tree logit survival model. Our purpose is twofold: demonstrate and develop a

multilevel logit model and use the model as a component in the proceeding Chapter for projecting stand tables.

- Stand table projection: compare and assess the liability of two alternative methods with Pienaar and Harrison's (1988) method and with each other. Both alternatives use the multilevel individual tree logit survival model as a component. One alternative will use predicted mortality probabilities by *dbh* class and the other by individual tree.

Our whole stand and individual tree survival models will provide a basis for developing future models. In addition, modeling heterogeneity in the individual tree model is directly applicable to future development of whole stand survival models.

CHAPTER 2

SYNOPSIS OF SURVIVAL ANALYSIS AND FORESTRY SURVIVAL METHODS

Survival Analysis Synopsis

Survival analysis focuses on a defined event for a group(s) of individuals and there are three basic requirements for analyzing survival data. Time of origin must be precisely specified, a scale determined for time, and the exact meaning of failure specified (Lawless 1982). Generally, we are interested in how the hazard and/or survival functions change over time for groups and strata. Standard statistical techniques are generally not applicable to survival data because survival data are typically positively skewed. Therefore, it is unreasonable to assume that survival data are normally distributed but data non-normality is sometimes remedied using a transformation (Collet 1994). In addition, all subjects are not generally followed through time until the event of interest has occurred.

Since not all subjects are followed through time until failure and/or the exact time of failure is unknown, censoring often occurs in survival analysis (Cox and Oakes 1984). There are three general censoring classifications: left, right, and interval censoring. Right censoring occurs when not all failures are observed, e.g., suppose that trees enter a study at t_0 and some die at $t_0 + t$, where t is the last observation, then those trees that are still alive at time t are right censored. A consequence of not accounting for right censoring is that average time until failure is usually underestimated (Collet 1994). Left and

interval censoring occur when the actual origin time is unknown and when it is only known that a failure occurred within two adjacent time intervals, respectively. Interval censoring occurs frequently in forestry, because the time of death is not generally observed for an individual tree, often it is only known that the tree died within an interval (t_0, $t_0 + \delta$), where δ generally ranges from one to five-years. Irrespective of the censoring classification, survival analysis usually focuses on the survival and hazard functions.

Survival can be defined as the proportion of the population still alive at time t, hence, survival functions are defined as the probability that an individual survives longer than t units of time, i.e., $S(t) = \Pr(T > t)$. Here T represents a random variable for the distribution of survival time from the initialization to the event of interest. The hazard function is defined as the instantaneous mortality rate assuming the individual has survived to time t (Allison 1995). Hazards are typically thought of as the instantaneous probability of failure at time t, but technically, it is not a probability (Lawless 1982) since it is bounded below by zero but has an unbounded upper asymptote. According to Cox and Oakes (1984), there are several reasons why the hazard function is important. First, the immediate risk can be determined for an individual known to be alive at age t. Second, comparing groups of individuals may be more effectively conducted using the hazard function. Lastly, hazard function models can be convenient when there is censoring or there are several types of failure.

Generally for survival analysis, the interest lies in the hazard of failure at anytime after the origin of the study, therefore the hazard function is usually modeled directly (Lawless 1982). Cumulative hazard functions ($H(t)$) are obtained for a time interval by integrating over the hazard function from t_0 to t_1, which can be thought of as the

cumulative risk, i.e., the sum of all risks faced by an individual going from t_0 to t_1. Survival analysis modeling usually begins by assuming a hazard function form arising from the three general families of survival models: the proportional hazards, accelerated failure time, and proportional odds models (Collet 1994).

Proportional Hazards Model

Cox (1972) developed the proportional hazards model, which assumes no probability density function for survival time. It is a semi-parametric method and has the property that different groups have proportional hazards and the ratio of the hazards for different groups don't depend on time. For example, suppose the hazard function for the i^{th} individual is $h_i(t)$ and the baseline hazard function is $h_0(t)$. Let x be a vector of covariates. If they are proportional, we can write the hazard function as $h_i(t) = h_0(t)g(x_i)$. Where both $h_0(t)$ and $g(x_i)$ may involve parameters, and $g(x_i)$ is a function of the predictor variables for the i^{th} individual (Collet 1994). The function $g(x_i)$ is interpreted as the hazard at time t for the i^{th} individual relative to the hazard for an individual for whom $\boldsymbol{x}_i$ equals zero. Since the relative hazard $g(x_i)$ cannot be negative, it is typical to express it as $\exp(\eta_i)$. Where η_i is a linear combination of p covariates, which is known as the risk score or prognostic index for the i^{th} individual in medicine (Collet 1994). There is no assumption about the baseline hazard function form and hence, the coefficients in the proportional hazards model can be estimated without making any assumptions about $h_0(t)$ (Cox and Oakes 1984). Proportional hazard models commonly arise for heterogeneous populations and a common model is the two-parameter Weibull *cdf* that has a common rate parameter but a group specific scale parameter.

Accelerated Failure Time Model

Accelerated failure time models assume the predictor variables measured for an individual act multiplicatively on the time scale. Given a baseline survival function $S_0(t)$, the survival for the i^{th} individual is $S_i(t) = S_0(\phi t)$. Hence, the effects of the covariates are modeled by the acceleration parameter ϕ (Collet 1994), which means that the model can be interpreted in terms of the progression speed. Assuming the endpoint is mortality, if $\phi < 1$ then there is acceleration in the time to death, conversely if $\phi > 1$ there is a deceleration in the time to death. The hazard function for the i^{th} individual can be expressed as $h_i(t) = \phi^{x_i} h_0(\phi^{x_i} t)$. Since ϕ has to be non-negative, it is convenient to set ϕ to $\exp(\beta)$ and express the hazard function as $h_i(t) = e^{\beta x_i} h_0(e^{\beta x_i} t)$. The popularity of the Weibull distribution for survival analysis is, in part, because it possesses the properties of both the proportional hazard and accelerated failure time models. It is the only distribution possessing the properties of both of these models, which are the two most common models for survival analysis (Lawless 1982).

Proportional Odds Model

Proportional odds models assume that the covariates act multiplicatively on the odds of survival beyond *t* (Collet 1994) and the log of the cumulative odds ratio is proportional to the distance between the values of the explanatory variables (Agresti 1990). The hazard function for the i^{th} individual to the baseline hazard for the proportional odds model converges from $\exp(-\eta_i)$ at time zero to one as time approaches infinity. For most disciplines, the proportional odds model is not extensively utilized because the Cox proportional hazards model can incorporate time dependent covariates and/or interactions to produce non-proportional hazards, which will likely give similar results to the

proportional odds model (Collet 1994). However, in individual tree forestry survival models it is widely used since this is a property of the logistic distribution.

Forestry Survival Synopsis

Forest survival models were generally an overlooked component in the early days of forest growth and yield models, largely due to the difficulty in estimating mortality given insufficient long-term data. Therefore, mortality was rarely modeled in the early days of professional forestry; instead, it was implied by yield and stand tables. Miscellaneous Publication 50 (USDA 1929) presented implied mortality using stand tables for normal stands of southern pines. Given average *dbh* of a normal stand, the cumulative percentage of trees in each diameter class was predicted. Therefore, if the same stand was measured in the future and average *dbh* was computed, a new stand table could be calculated that contained the implied mortality. Other early examples are mortality by dbh class (e.g., Thomson 1932, Krauch 1930) and life tables (Deen 1933).

Reineke (1933) put forth the idea of a limiting stand density by developing a stand density index (SDI) for pure even-aged stands. Stand density index, which is species dependent but independent of site and age, implies self-thinning for fully stocked even-aged stands. Interestingly, SDI can be thought of as a survival function, i.e., $S_i(R) = R^{1.605}$ where R is the ratio of the quadratic mean diameters at t_1 to t_2. This survival function implies a constant rate of survival for a given ratio that is independent of age, species, and site. The 3/2 power law developed by Yoda et al. (1963) has the same implied survival functional form as Reineke (1933) except the exponent is 1.5. These forestry mortality estimation methods were common until the 1960's and 70's

when a period of rapid development began for whole stand and individual tree mortality models.

Whole Stand Mortality Models

Early whole stand survival models depended on functions that could be easily estimated using existing analytical tools. These models include linear (Lee 1971), logarithmic transformations (Schumacher and Coile 1960, Smalley and Bailey 1974), and probit (Lenhart and Clutter 1971, Lenhart 1972). Early whole stand survival models laid the foundation for a period of rapid development of whole stand survival models based on sound mathematical reasoning and empirical evidence.

Modeling of whole stand survival should consider the empirical and theoretical behavior for the species and conditions, which includes reasonable survival extrapolation properties. Although recognized in earlier studies, Clutter et al. (1983) formalized desirable properties for whole stand survival projection models, which includes:

1) if $A_2 = A_1$ then $N_2 = N_1$,

2) for even-aged stands: if $A_2 > A_1$ then $N_1 \geq N_2$,

3) for even-aged stands: $A_2 \rightarrow \infty$ then $N_2 \rightarrow 0$, and

4) path invariance, i.e., if we use N_1 to predict N_2 which is then used to predict N_3 it should be the same as if we used N_1 to predict N_3.

Here N_i represents the number of trees per unit area at age i (A_i). Property 3) has been sometimes argued as unreasonable since for a given species there is likely to be a limiting capacity higher than zero for a site. However, it can be argued that for a given stand, as age increases the cumulative mortality for the trees in the initial stand approaches one. Both arguments have merit and oftentimes the choice of whether to

include a carrying capacity limit is dependent on the empirical evidence for a particular species and locale. Most whole stand survival models developed since the late 1970's have these desirable properties and these latter whole stand models have usually been developed using a cumulative distribution function (*cdf*) or the differential equation approach, which are not necessarily distinctly different.

Cumulative Distribution Function Approach

Cumulative distribution functions most commonly used for modeling whole stand survival are derivatives of the generalized Gamma distribution (GGD), which has a probability density function (*pdf*) of

$$f(t) = \frac{\lambda\beta}{\Gamma(\kappa)}(\lambda t)^{\kappa\beta-1}\exp\left[-(\lambda t)^{\beta}\right], \quad t > 0\,. \tag{2.1}$$

Where the β, λ, and κ parameters are positive. The GGD was introduced by Stacy (1962) and has the following properties. If $\kappa = 1$ or $\kappa = 1$ and $t = t - a$ then we have the two and three-parameter Weibull *pdf* models, respectively. Defining $\beta = \kappa = 1$ results in the exponential *pdf* and defining $\beta = 1$ results in the Gamma *pdf*. Lastly, letting $\kappa \to \infty$, the limiting distribution is the log-normal. Thus, the GGD is a flexible function for modeling forestry survival.

An early example of using the GGD is the use of the Weibull *cdf* to model survival (Pinder et al. 1978), in which he modeled wildlife survival. In the context of forestry, the Weibull *cdf* has been effectively used to model whole stand survival (e.g., Pienaar and Shiver 1981, Somers et al. 1980, Amateis et al. 1997, Belli and Ek 1988). In addition, a study by Buford and Hafley (1985) compared the Weibull, Gamma, and Negative Binomial distributions, and the Chapman-Richards function for fitting mortality data. The Chapman-Richards function enables the underlying hazard function to have an

inflection point. It therefore has added flexibility when compared to the Weibull distribution hazard function, which has no inflection point. Some other examples of studies that have successfully used the GGD to model whole stand survival are the exponential (e.g., Matney and Farrar 1992) and Gamma (Kobe and Coates 1997). Other types of distributions for modeling whole stand survival include the inverse logistic Rennolls and Peace (1986) and logistic Hitch et al. (1996).

Difference Equation Approach

The Schumacher and Coile (1960) whole stand survival model is an early example of using the forestry modeling techniques that have become known as algebraic difference and difference equation approaches. It is interesting to note that, although not referred to as such by the authors, this is probably the earliest whole stand survival model developed using a site-specific parameter. This is illustrated by noting that their original prediction equation, for some species, is given by

$$\log(N) = \beta_{00} + \frac{\beta_0}{A} + \beta_1 \log(H) + \beta_2 (B). \tag{2.2}$$

Here N is trees per unit area, A is age, H equals dominant height, B is the basal area per unit area, and β_{00}, β_0, β_1, and β_2 are parameters. Let β_{00} be site-specific. Then using the algebraic difference approach, their resulting whole stand survival projection mode1 is

$$\log\left(\frac{N_1}{N_0}\right) = \beta_0\left(\frac{1}{A_1} - \frac{1}{A_0}\right) + \beta_1 \log\left(\frac{H_1}{H_0}\right) + \beta_2 \log\left(\frac{B_1}{B_0}\right), \tag{2.3}$$

which has the desirable whole stand survival model properties, assuming $\beta_0 < 0$. Model (2.3) is flexible and has subsequently been used in several other studies (e.g., Dell et al. 1979, Feduccia et al. 1979). Furthermore, an interesting note with the Schumacher and

Coile (1960) model (2.3) is that their whole stand survival model can be derived from the following differential equation.

$$\frac{1}{N}\frac{\partial N}{\partial A} = \frac{\beta_1}{H}\frac{\partial H}{\partial A} + \frac{\beta_2}{B}\frac{\partial B}{\partial A} + \beta_0 \tag{2.4}$$

This illustrates that the difference equation and algebraic difference approaches were in use many years before the ideas were formalized. Since the late 1970's, many of the whole stand survival models have been derived using the difference equation approach. The simplest model assumes that the instantaneous relative rate of mortality is constant, i.e.,

$$\frac{1}{N}\frac{dN}{dA} = \beta . \tag{2.5}$$

Using the initial conditions that when $A_1 = A_2$ then $N_1 = N_2$, and after integration yields

$$\frac{N_2}{N_1} = e^{-\beta(A_2 - A_1)} . \tag{2.6}$$

This exponential difference equation implies that the proportional mortality rate is constant for all ages, densities, and site indices (Clutter et al. 1983), but has been successfully used in some studies (e.g., Martin et al. 1999, Devine and Clutter 1985). Hence, this model form implies both underlying proportional hazard and accelerated failure time models, i.e., since the exponential is a special case of the Weibull, which possesses these properties. A flexible whole stand survival model was developed by Clutter and Jones (1980) that has subsequently been used in several other studies (e.g., Clutter et al. 1983, Pienaar and Rheney 1993). Brister (see Varner 1981) modified the lower asymptote of the Clutter and Jones (1980) model, and Martin and Brister (1999) used this same function but added a modifier to incorporate the effect of hardwood

competition. There are numerous additional examples where the differential equation approach has been used to model whole stand survival (e.g., Bailey et al. 1985, McTague and Stansfield 1994). The generalized Gamma distribution and difference equation approaches are the most widely used methods for modeling whole stand survival.

Individual Tree Survival Models

Modern individual tree survival modeling likely began with a study by Hamilton (1974) in which he used the logistic equation for modeling survival data. Since his pioneering study, the logistic equation has become the individual tree survival model of choice (e.g., Lowell and Mitchell 1987, Vanclay 1991, Krumland et al. 1977, Wykoff et al. 1982, Teck and Hilt 1990, Hitch et al. 1996). This is likely because of the ease of parameter interpretation and function flexibility. It has become common since the studies by Monserud (1976) and Hamilton and Edwards (1976) to annualize mortality, so that the response is yearly-predicted mortality. As discussed, the logistic possesses the properties of both the accelerated failure time and proportional odds models, which is expected since the logistic is a special case of the log-logistic distribution, which has these properties. Monserud (1976) developed an annualized generalized logistic model that uses data from any measured time interval in the parameter estimation.

The logistic equation is the most widely used model for individual tree mortality, likely being used in greater than 90 percent of all individual tree survival models. Some other functional forms used for modeling individual tree survival are exponential type models (e.g., Burkhart et al. 1987, Amateis et al. 1989). In addition, several studies have modified the predicted individual tree survival to achieve compatibility with predicted whole stand survival (e.g., McTague and Stansfield 1994). Several recent novel

approaches to modeling individual tree survival have been developed. Guan and Gertner (1991) modeled individual tree survival using an artificial neural network and a binary classification tree was used by Dobbertin and Biging (1998). However, these and other recent methods have not proven to be more effective in modeling individual tree survival than traditional statistical models. Hence, the logistic has remained the common approach to modeling individual tree survival as evident by its use in recent studies (e.g., Monserud and Sterba 1999, Huebschmann et al. 2000, Yao et al. 2001, Shen et al. 2001). Therefore, the flexibility, ease, and interpretability will likely continue to fuel the use of the logistic for modeling individual tree survival.

Individual tree or *dbh* class and whole stand survival models are typically used in stand table projection methods. Whole stand survival models are used to predict the total plot mortality and predicted mortality is usually allocated using an individual tree or *dbh* class survival model. Hence, stand table projection methods delve into the question of mortality allocation once the total mortality for a plot is predicted for a desired projection period. Our study interest is in developing whole stand and individual tree survival models, and to use an individual tree survival model to predict the mortality probabilities in projecting stand tables. The study proceeds in the following manner: Chapter 3 focuses on modeling whole stand survival, Chapters 4 and 5 concentrate on modeling individual tree survival, and Chapter 6 uses the individual tree survival model developed in Chapter 5 in projecting stand tables. All ensuing Chapters expand on the survival and forestry background where pertinent.

CHAPTER 3

A GENERALIZED METHODOLOGY FOR DEVELOPING FLEXIBLE WHOLE STAND SURVIVAL MODELS

Introduction

Whole stand survival models are critical in accurately reflecting growth and yield for plantations because of the sensitivity of the basal area growth model to the underlying mortality. However, development of whole stand survival models has received relatively little attention compared to individual tree survival models. Forest mortality has been usually classified as either natural or irregular (Staebler 1953) and our focus is on natural mortality, which generally occurs because of competition for light, nutrients, and water. Hence, our study focuses on plantation whole stand natural forest survival.

Forest survival, or complementary, mortality, can be analyzed using traditional survival analysis. The distribution of the random variable *T* from initialization to the event of interest can be represented by the survival and hazard functions. Let the distribution of *T* be $F(t) = \Pr[T \le t]$, then $f(t)$ is the corresponding density function and the survival function is defined as $S(t) = \Pr(T > t) = 1 - F(t)$. The hazard function, which is the instantaneous rate of mortality assuming the individual has survived to time *t* (Collet 1994), is defined as $h(t) = f(t)/[1 - F(t)]$. Analogous to the continuous time hazard function, the discrete time hazard function is defined as $q(t) = [F(t+1) - F(t)]/1 - F(t)$ (Wilson 1972). A nonparametric estimation of the

survival function is the Kaplan-Meier product limit estimator (KM) (Kaplan and Meier 1958), which is defined as

$$\hat{S}(t) = \prod_{i:t_i \le t} \left[1 - \frac{d_i}{n_i} \right]$$

for $t_1 \le t \le t_j$. Where n_i and d_i are the subjects at risk (n_i) and that die (d_i) at time t_i. An empirical estimate of the hazard function is simply the number of trees that died in the interval divided by the total length of time all the trees were observed. Survival curves are non-increasing over time, whereas the hazard function can increase, decrease, remain constant, or assume a combination of these shapes.

A common cumulative distribution function (*cdf*) used in survival analysis is the Weibull because it is capable of describing the three most common types of hazard curves, which are monotonically increasing or decreasing, and constant. However, the Weibull distribution imposes strong restrictions on the data and is unable to model complex hazard shapes such as a bathtub shape (Hjorth 1980). Increasing monotonic hazard functions are the most common because many studies focus on a snapshot of the subject's lifespan in which gradual aging takes place (Lawless 1984). Because of the snapshot focus, there is usually no empirical motivation to find distributions that are capable of producing bathtub shaped hazard functions. However, a bathtub shaped hazard function is a reasonable assumption when viewing the entire lifespan of many, if not most, biological organisms. Consider the human lifespan. There are typically three distinct phases: Infant (mortality decreases), juvenile to adulthood (mortality generally stable), and mature adults (mortality increases). Taking a snapshot from any of these phases will likely result in a specific type of survival and hazard curve, but when viewed as an entity it is reasonable to assume a bathtub shaped curve. In forestry, a study by

Lorimer and Frelich (1984) illustrated that the diameter-specific mortality rates for a given stand could be bathtub shaped. Moreover, mortality has been modeled as bathtub shaped with respect to diameter in individual tree models (e.g., Buchman et al. 1983, Monserud and Sterba 1999), but mortality has not been explicitly modeled as bathtub shaped for whole stand models.

Whole stand survival models have commonly been developed using a derivative of the generalized Gamma *cdf* (e.g., Weibull and exponential *cdf's* are special cases) or the difference equation approach. Both of these approaches, either implied or explicitly stated, use presuppositions about the relative rate of instantaneous mortality that is based upon empirical evidence. The Weibull *cdf* is a flexible distribution that has been widely used for whole stand survival models (e.g., Pinder et al. 1978, Glover and Hool 1979, Somers et al. 1980, Pienaar and Shiver 1981, Belli and Ek 1988, Amateis et al. 1997). As discussed, the Weibull *cdf* hazard function is capable of describing the three most common hazard shapes, hence, its popularity. Oftentimes the hazard function shapes that are capable when using the Weibull distribution are applicable since the study data, in which plots are often established after the initial seedling mortality, only contain a snapshot of a stand's lifespan. The hazard function for the two-parameter Weibull *cdf* survival model is $h(t) = \frac{c}{b}\left(\frac{t}{b}\right)^c$, where t is time and b, c are parameters. Cumulative distribution based whole stand survival models use suppositions with respect to the distribution's ability to model the empirical survival trends. Conversely, difference equation models use suppositions about the relative rate of instantaneous mortality change. Although it appears these two approaches are distinct, oftentimes the difference equation supposition leads, after integration, to a *cdf* based whole stand survival model.

This is illustrated by using a simple difference equation for the relative rate of mortality, which assumes that the instantaneous mortality rate is constant, i.e.,

$$\frac{1}{N}\frac{dN}{dA} = \beta .$$

Where N is the number of trees per unit area, A is age, and β is a parameter. After integration and using the initial conditions that when $A_2 = A_1$ then $N_2 = N_1$, the result is

$$\frac{N_2}{N_1} = S(A_2) = e^{\beta(A_2 - A_1)} .$$

Thus, this difference equation results in the exponential distribution and implies that the instantaneous mortality rate is constant for all ages, densities, and site indices (Clutter et al. 1983). Exponential *cdf's* have been used in several whole stand survival studies (e.g., Martin's et al. 1999, Devine and Clutter 1985), however it imposes a strong assumption of a constant hazard rate. Nonetheless, a constant hazard rate is oftentimes reasonable because of the age range for the study data. For example, Devine and Clutter (1985) used survival data from 161 plots, of which only two had measurement data less than five-years of age. Hence, we wouldn't expect to detect early stand survival trends. Clutter and Jones (1980) presented a more flexible difference equation in which they assumed that the relative rate of instantaneous mortality is proportional to age and initial trees per acre, which are raised to a power, i.e.,

$$\frac{1}{N}\frac{dN}{dA} = \alpha A^{\delta} N^{\phi} .$$

Integrating over the initial conditions of when $A_2 = A_1$, then $N_2 = N_1$ yields

$N_2 = \left[N_1^{\beta} + \eta\left(A_2^{\phi} - A_1^{\phi}\right)\right]^{\frac{1}{\beta}}$. This flexible whole stand survival model has subsequently been used with slight modifications in several other studies (e.g., Martin and Brister 1999

Pienaar and Rheney 1993). A study by Tait (1988) related the relative rate of instantaneous mortality to the rate of stand development. This differential equation for whole stand mortality was unique because it related the differential equation to density dependent and independent components. This whole stand survival model has been applied successfully in several subsequent studies (e.g., Tait et al. 1988, Tait and Jahraus 1988).

It has become common to model whole stand mortality for a stand's lifespan using a system of equations (e.g., Matney and Farrar 1992, Amateis et al. 1997). These systems disaggregate the lifespan of a stand's survival into distinct phases, typically some combination of the seedling, juvenile, adult, and mature phases. Disaggregating whole stand survival into phases is primarily conducted because of the difficulty in developing a flexible biologically reasonable function that can model survival throughout stand development and data limitations.

We demonstrate a method for deriving flexible biologically reasonable whole stand survival models, which are capable of modeling complex underlying hazard functions. Furthermore, it is hypothesized, for our data, that the continuum of whole stand forestry survival has an underlying bathtub shaped hazard function.

Data

Data were obtained from the Consortium for Accelerated Pine Production Studies (CAPPS), which is overseen by the Warnell School of Forestry at the University of Georgia. CAPPS purpose is to investigate the effects of intensive forest management on the productivity of loblolly pine plantations in the Southeastern United States and plots

were established throughout Georgia. Study protocol called for two complete blocks to be established at each location with each block containing four 0.15 ha treatment plots. A 0.15 ha treatment plot was established at each location using bare-root seedlings on a 2.44 m by 2.44 m spacing. A 0.05 ha measurement plot was centered within each of the treatment plots. Each of the following four cultural treatments were randomly assigned to the blocks at each location.

1) Herbicide: plot sprayed with non-soil active herbicide as needed to maintain complete control of woody and herbaceous vegetation,
2) Fertilization: apply recommended rates of fertilizer annually, if necessary, to ensure that nutrients are not the limiting factor,
3) Herbicide – Fertilization: apply both herbicide and fertilization treatments, and
4) Control treatment: no cultural treatment other than mechanical site preparation.

The original study protocol called for replicating all treatment plots every two years for the first ten years of the study, and each location has two treatments with two levels (herbicide versus no herbicide, fertilization versus no fertilization). The actual study varies from the protocol because of funding limitations and the replications have been repeated at different intervals for different locations. Plots have been measured annually with available survival data beginning at age two and data are summarized by plot age and treatment (Table 3.1).

Model Development

To develop whole stand survival models, which are capable of reflecting complex underlying hazard functions, we first computed the KM survival estimates and the

corresponding discrete empirical hazard function (Table 3.2). The KM survival estimates for the spectrum of plots illustrates that the underlying discrete hazard decreases from age 2-5 and then increases from ages 5-14. Kaplan-Meier survival and mortality estimates, and the corresponding discrete hazard function were computed by treatment (Figure 3.1). Survival curves by treatment illustrate that the herbicide and fertilizer treatments result in the most favorable and unfavorable survival, respectively. In addition, the herbicide/fertilizer treatment has favorable early survival but mortality increases rapidly after about age eight. The hazard function trend for the spectrum of plots appears to be bathtub shaped, but the oldest plots are 14 years and it is difficult to infer the future trend of the hazard function. Nevertheless, it is reasonable to assume that the hazard function will continue to increase with time, which is consistent with most biological organisms (Pinder 1978). We assumed that whole stand plantation survival could be modeled using a generalized differential equation to describe the relative rate of mortality, i.e.,

$$\frac{1}{N}\frac{dN}{dt} = f(t)\, X$$

Where N is the number of trees per unit area, $f(t)$ is a function of time, and X can be a function of any whole stand attribute. We narrowed the scope for viable survival models by assuming X equals one or N^d. Our search for a viable $f(t)$ function began, from empirical evidence (Figure 3.1), assuming that the function should be flexible enough to model a bathtub shaped hazard function. This criterion led us to the functional form of

$$f(t) = \frac{1}{1+a\,t} + \frac{b}{1+t} + c\,t\,.$$

Where t is time, and a, b, c are parameters. This function has the flexibility to model both monotonically increasing or decreasing hazards as well as a bathtub shape hazard

function (Figure 3.2). Substituting $f(t)$ into the differential equation, and integrating for both values of X using the initial condition that when $N_2 = N_1$ then $t_2 = t_1$ results in model (3.1a)

$$N_2 = N_1\left(\frac{1+a\,t_2}{1+a\,t_1}\right)^{\frac{1}{a}}\left(\frac{1+t_2}{1+t_1}\right)^{b} e^{\frac{c}{2}\left(t_2^2-t_1^2\right)},\ X=1 \qquad (3.1a)$$

and after re-parameterization, model (3.2a)

$$N_2 = \left[N_1^d + \frac{d}{a}\ln\left(\frac{1+a\,t_2}{1+a\,t_1}\right) + b'\ln\left(\frac{1+t_2}{1+t_1}\right) + c'\left(t_2^2 - t_1^2\right)\right]^{\frac{1}{d}},\ X = N^d. \qquad (3.2a)$$

Here $b' = -db$ and $c' = -dc/2$. Both models possess the desirable properties of path invariance and when $t_2 \to t_1$ then $N_2 \to N_1$. In addition, their respective lower asymptotes are zero, if c for (3.1a) and d for (3.2a) are negative. Model (3.2a) also has the ability to distinguish among the density classes for survival if the data exhibit this trend.

Our motive for choosing $f(t)$ is its ability to model bathtub shapes. We can establish if the resulting survival function is capable of modeling an underlying bathtub shaped hazard function by noting the concavity and number of inflection points in the survival curve. To obtain a bathtub shaped hazard curve, it is obvious that the hazard curve must first decrease over time, level off, and then increase. A bathtub shaped hazard curve corresponds to a survival curve with at least two inflection points. Furthermore, the concavity of the survival curve between the two inflection points must be concave up to the first inflection point, then concave down from the first inflection to the second inflection point, and then concave up after the second inflection point. Models (3.1a) and (3.2a) have the flexibility to model multiple inflection points.

Models (3.1a) and (3.2a) were fitted to the CAPPS study data and evaluated by examining the residuals, fit index (defined as one minus the error sum of squares divided by corrected total sum of squares), mean square error (MSE), root mean square error (RMSE), and error sum of squares (SSE). In addition, the behavior of the fitted functions was examined, both within the range of the data and extrapolating to a reasonable age.

Results

Models (3.1a) and (3.2a) asymptotes were modified to allow for a biologically reasonable lower asymptote. This has been demonstrated to be a reasonable assumption for plantation loblolly pine of the Southeastern United States (Harrison and Borders 1996, Martin and Brister 1999). In addition, models (3.1a) and (3.2a) were initially fit separately by treatment. Then the estimated parameters by treatment for models (3.1a) and (3.2a) were plotted and linear trends were detected for each parameter. However, some treatment parameters were not substantially different. Therefore, we refitted models (3.1a) and (3.2a), now referred to as models (3.1b) and (3.2b), and allowed the fertilizer, herbicide, and herbicide/fertilizer cultural treatments to vary systematically from the baseline parameters. The cultural treatment effects were coded as *fert* = 1 if fertilized and zero otherwise, similarly for the herbicide (*herb*) and herbicide/fertilizer (*hf*) treatments. This resulted in a re-parameterization of the a, b, c, and d parameters, here for model (3.2b) we re-define $b=b'$ and $c = c'$, where applicable for models (3.1b) and (3.2b) as

$$\left.\begin{aligned} a &= a_0 + a_1(fert) + a_2(herb) + a_3(hf) \\ b &= b_0 + b_1(fert) + b_2(herb) + b_3(hf) \\ c &= c_0 + c_1(fert) + c_2(herb) + c_3(hf) \\ d &= d_0 + d_1(fert) + d_2(herb) + d_3(hf) \end{aligned}\right\}$$

Cultural treatment parameters were removed from models (3.1b) and (3.2b) using a stepwise procedure ($\alpha = 0.05$). Models (3.1b) and (3.2b) achieved convergence easily during the cultural treatment parameter elimination process. Residual plots for the models were examined and since there was no evidence of heteroscedasticity, no weighting or transformations were necessary. The fitted models (3.1b) and (3.2b) that allow for systematic cultural treatment effects and a modified lower asymptote are

$$N_2 = N_{\min} + (N_1 - N_{\min})\left(\frac{1 + a\,t_2}{1 + a\,t_1}\right)^{\frac{1}{a}}\left(\frac{1 + t_2}{1 + t_1}\right)^{b} e^{\frac{c}{2}\left(t_2^2 - t_1^2\right)}, \qquad (3.1b)$$

Where the parameters that allow for cultural treatment are defined as

$$\left.\begin{aligned} a &= a_0 + a_1(fert) + a_2(herb) + a_3(hf) \\ b &= b_0 + b_1(fert) + b_2(herb) \\ c &= c_0 + c_1(fert) + c_3(hf) \end{aligned}\right\}$$

Here $N_i = TPH/100$ (TPH = trees per hectare), $N_{\min}$ equals 2.5, which is the lower asymptote for $TPH/100$, t_i is plot age at time i, and a, b, c are parameters. The $N_{\min}$ of 2.5 corresponds to approximately 100 trees per acre, which has been deemed a reasonable lower limit for loblolly pine of this region (Harrison and Borders 1996). Model (3.2b) is

$$N_2 = N_{\min} + \left[(N_1 - N_{\min})^d + \frac{d}{a}\ln\left(\frac{1 + a\,t_2}{1 + a\,t_1}\right) + b\ln\left(\frac{1 + t_2}{1 + t_1}\right) + c\left(t_2^2 - t_1^2\right)\right]^{\frac{1}{d}}, \qquad (3.2b)$$

Where the final parameters that allow for systematic cultural treatment effects are

$$\left.\begin{aligned} a &= a_0 + a_2(herb) + a_3(hf) \\ b &= b_0 + b_2(herb) + b_3(hf) \\ c &= c_0 + c_1(fert) \\ d &= d_0 + d_2(herb) + d_3(hf) \end{aligned}\right\}$$

Models (3.1b) and (3.2b) and their respective parameter estimates, stand errors, and p-values are presented in Table 3.3. Summary fit statistics reveal that model (3.2b) explains more of the variation in survival. Models (3.1b) and (3.2b) RMSE are 0.4590 and 0.4570, respectively. This means that for the average *TPH* (approximately 1600), there is less than a three-percent error. The fit index for models (3.1b) and (3.2b) are 0.9507 and 0.9511, respectively. There is no substantial difference between models (3.1b) and (3.2b) for these criteria. Hence, to further assess model performance, the mean survival and corresponding hazard functions were computed for the spectrum of plots and stratified by treatment.

Models (3.1b) and (3.2b) predicted mean survival and corresponding discrete hazard functions for the spectrum of plots illustrate that both models adequately mirror the empirical survival and hazard functions trends (Figure 3.3). Note that small differences in the survival curves can have a profound impact on the shape of the hazard function. Model (3.1b) more closely mirrors the empirical hazard function for the early ages but model (3.2b) exhibits more overall flexibility. Mean survival and their corresponding hazard function were computed and stratified by treatment for both models (Figures 3.4 and 3.5). Both models adequately reflect the underlying hazard function associated with the survival curves.

Further assessment was conducted by examining the behavior of the fitted models to predict survival by treatment for ages 1-30 (Figure 3.6). Both models behave similarly

within the data range (age 2-14). However, model extrapolation properties are substantially different for these models when fitted to our data. Model (3.1b) provides reasonable survival curves by treatment and has a lower asymptote of 250 *TPH*. In contrast, model (3.2b) has reasonable extrapolation properties only for the herbicide treatment and the herbicide/fertilizer treatment declines rapidly beyond the data range. Additionally, the control and fertilizer treatments are only able to predict to age 17 using the estimated parameters because the term in the bracket for model (3.2b) when predicting survival for these treatments is negative after age 16, which is raised to a negative fractional power. Model (3.1b) provides more reasonable extrapolation predictions when fitted to our data. However, both models behave adequately and provide flexible solutions within our data range.

Discussion

Flexible whole stand survival models were developed and demonstrated to provide biologically reasonable solutions for complex underlying hazard functions. In addition, our empirical hazard curves illustrated that non-proportionality exists among the treatments, which was easily modeled by including, where necessary, treatment parameters. It is easily demonstrated that functional forms such as $f(t) = at^b$, that are simple linear or nonlinear functions of time are incapable of modeling complex hazard functions. These types of functions can be expanded to include other covariates, such as site index, that impact survival. However, the additional covariates are not necessarily enough in and of themselves to produce bathtub shaped hazard functions. In contrast, the addition of two (or more) linear/nonlinear functions can be developed that will model

bathtub shaped hazard curves. Furthermore, the functional form does not need to have numerous parameters in order to model complex hazard functions, e.g., model (3.1a) has only three baseline parameters. Model (3.2b) demonstrated that different assumptions placed on the generalized differential equation form using the same $f(t)$ function as model (3.1b) can produce more flexible hazard curves. Nevertheless, model (3.2b) provided only slightly more additional flexible solutions. Moreover, this increased flexibility has an extrapolation cost when model (3.2b) is fitted to our data. Additional covariates in the general difference equation form may provide more flexibility in the solutions. However, the flexibility from using additional covariates is unlikely to approach the flexible whole stand survival solutions gained from a flexible $f(t)$ function.

Our demonstrated method is relatively straightforward for developing a whole stand survival model that is capable of modeling a complex underlying hazard function. The Weibull distribution and the Clutter and Jones (1980), hereafter referred to as C&J, models are probably the most widely used whole stand survival models; therefore, these models were fitted to our data for comparison purposes. Neither model was able to achieve convergence for our data, probably because of the large decrease in early stand survival (Figure 3.1). Since we could not achieve convergence and because it is common to have whole stand plantation survival data that begins at age four or five, we fitted these models and our model (3.1a) after eliminating all data prior to age four, i.e., our first measurements for survival began at age five. All three models easily achieved convergence using the age 5-14 survival data. The results revealed that all three models fit well, but the C&J model has a more favorable fit index of 0.9842 than the Weibull (0.9834) and model (3.1a) (0.9834). Our motive for fitting these models to the age 5-14

data was not to determine which model fit the data better, but to establish which model would more accurately predict the empirical hazards for ages 2-4. The predicted survival curves and corresponding hazard curves for these models, using $N_1 = TPH = 1600$ (approximately the average) for the ages 2-14 reveal that model (3.1a) is able to extrapolate extremely well for these data (Figure 3.7). All three models adequately mirror the empirical hazard for the range of the data (5-14 years). However, the Weibull and C&J models extrapolate poorly when predicting the age two and three hazards. Model (3.1a) performs excellent when its predicted hazards are compared to the ages 2, 3, and 4 empirical hazards. This demonstrates that although the C&J model fits marginally better than model (3.1a) when fitted to the age 5-14 data, the extrapolation predictions for the age 2-4 hazards are substantially improved using model (3.1a). Furthermore, the age one hazard for the C&J and model (3.1a) are 0.0005 and 0.1387, respectively. Hence, model (3.1a) is behaving more biologically reasonable with respect to early whole stand loblolly pine survival of this region. This demonstrates that a model may provide adequate future extrapolation predictions, but may not behave reasonable for early survival extrapolation predictions because its hazard function is more restrictive. Hence, extrapolating both forward and backwards (if the data set warrants) are equally important in establishing if a model has biologically reasonable behavior. Oftentimes the model developer is only interested in predicting survival within the range of the data. Regardless of the data time frame for a given study, plotting the empirical hazard function can aid in establishing the complexity of the function that may be necessary to adequately model whole stand survival.

Conclusion

Forestry survival is a complex and difficult process to model; but the difficulty can be reduced by using the empirical hazard function behavior to aid in selecting an appropriate survival function. Use of the hazard function to aid in forestry survival model selection is not novel. Preisler and Slaughter (1997) demonstrated that they could limit their individual tree survival model selection to a model that was capable of reflecting the empirical hazard function behavior. However, their study did not consider any survival functions capable of having bathtub shaped behavior. Our study established that improved whole stand survival models could result by considering the underlying hazard function. We demonstrated this by using the empirical hazard function to limit our selection to an appropriate function that could model the bathtub shape. Note that although our fitted models (3.1b) and (3.2b) consist of 10 and 11 parameters, respectively, it is not the number of parameters that allows our model to reflect an underlying bathtub shaped hazard function behavior. Our models when fitted to the data set without cultural treatment parameters still exhibit the underlying bathtub hazard function behavior. The cultural treatment parameters allow additional model flexibility, but it is the $f(t)$ function that allows our model to exhibit bathtub shaped trends if the data warrants. Whole stand lifespan survival is commonly modeled using a system of equations, which oftentimes creates a cumbersome and difficult system to implement. It was demonstrated that one equation might provide the desired flexibility when derived from knowledge about the underlying hazard function. Furthermore, we demonstrated the ability of a simple yet flexible function to be integrated to obtain an initial condition equation. Survival is generally the least understood and hardest to model within a whole

stand forestry growth and yield system, however our method is relatively easy to implement and can model a whole stand survival curve that exhibits a complex underlying hazard function behavior.

Literature Cited

Amateis, R.L., H.E. Burkhart, and L. Jiping. 1997. Modeling survival in juvenile and mature loblolly pine plantations. For. Ecol. Manage. 90:51-58.

Belli, K.L., and A.R. Ek. 1988. Growth and survival modeling for planted conifers in the great lakes region. For. Sci. 34(2):458-473.

Buchnam, R.G., S.P. Pederson, and N.R. Walters. 1983. A tree survival model with applications to species of the Great Lakes region. Can. J. For. Res. 13: 601-608.

Buford, M.A., and W.L. Hafley. 1985. Probability distributions as models for mortality. For. Sci. 31(2):331-341.

Collet, D. 1994. Modelling survival data in medical research. Chapman & Hall. London. 347 p.

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

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

Devine, O.J., and J.L. Clutter. 1985. Predictions of survival in slash pine plantations infected with fusiform rust. For. Sci. 31(1):88-94.

Glover, G.R., and J.N. Hool. 1979. A basal area ratio predictor of loblolly pine plantation mortality. For. Sci. 25(2):275-282.

Harrison, W.M., and B.E. Borders. 1996. Yield prediction and growth projection for site-prepared loblolly pine plantations in the Carolinas, Georgia, Alabama, and Florida. The University of Georgia, PMRC Tech. Rep. 1996-1.

Hjorth, U. 1980. A realiability distribution with increasing, decreasing, constant and bathtub-shaped failure rates. Technometrics. 22(1):99-107.

Kaplan, E.L., and P. Meier. 1958. Nonparametric estimation from incomplete observations. JASA. 53:457-481.

Lawless, S.W. 1982. Statistical models and methods for lifetime data. New York: John Wiley & Sons, Inc. 580 p.

Lorimer, C.G., and L.E. Frelich. 1984. A simulation of equilibrium diameter distribution of sugar maple (*Acer saccarum*). Bull. Torrey Bot. Club 111:193-1999.

Martin, S.W., R.L. Bailey, and E.J. Jokela. 1999. Growth and yield predictions for lower coastal plain slash pine plantations fertilized at mid-rotation. SJAF. 23(1):1-7.

Martin, S.W., and G.H. Brister. 1999. A growth and yield model incorporating hardwood competition for natural loblolly pine stands in the Georgia piedmont. SJAF. 23(3):179-185.

Matney, T.G., and R.M. Farrar Jr. 1992. A thinned/unthinned loblolly pine growth and yield simulator for planted cutover site-prepared land in the mid-gulf south. SJAF. 16(2):70-75.

Monserud, R.A., and H. Sterba. 1999. Modeling individual tree mortality for Austrian forest species. For. Ecol. Manage. 113:109-123.

Pienaar, L.V., and J.W. Rheney. 1993. Yield prediction for mechanically site-prepared slash pine plantations in the southeastern coastal plain. SJAF. 17(4):163-173.

Pienaar, L.V., and B.D. Shiver. 1981. Survival functions for site-prepared slash pine plantations in the flatwoods of Georgia and northern Florida. SJAF. 5:59-62.

Pinder III, J.E., J.G. Wiener, and M.H. Smith. 1978. The Weibull distribution: a new method of summarizing survivorship data. Ecology. 59(1):175-179.

Preisler, H.K., and G.W. Slaughter. 1997. A stochastic model for tree survival in stands affected by annosum root disease. For. Sci. 43(1):78-85.

Somers, G.L., R.G. Oderwald, W.H. Harms, and O.G. Langdon. 1980. Predicting mortality with a Weibull distribution. For. Sci. 26(2):291-300.

Staebler, G.R., 1953. Mortality estimation in fully stocked stands of young growth Douglas-fir. USDA For. Serv. Pac Northwest Forest and Range Exp. Stn. Res. Pap. 4. 8 p.

Tait, D.E. 1988. The dynamics of stand development: a general stand model applied to Douglas-fir. Can. J. For. Res. 18: 696-702.

Tait, D.E., C.J. Cieszewski, and I.E. Bella. 1988. The stand dynamics of lodgepole pine. Can. J. For. Res. 18: 1255-1260.

Tait, D.E., and K.V. Jahraus. 1988. The contrast in stand dynamics as revealed by comparing parameter estimates for a general stand growth model. Can. J. For. Res. 18: 1479-1485.

Table 3.1. The CAPPS study plot distribution by year planted for the spectrum of plots and by treatment.

Year Planted	Plot Age	Plots	Plots by Treatment			
			Control	Fertilized	Herbicide	HF*
1986	14	26	8	4	8	6
1987	13	28	8	6	8	6
1988	12	36	10	8	10	8
1989	11	24	8	4	8	4
1992	8	20	6	4	6	4
1994	6	12	4	2	4	2
Total		146	44	28	44	30

* HF is the fertilizer and herbicide treatment.

Table 3.2. The Kaplan-Meier product limit estimator for the interval survival, cumulative survival and the hazard for the CAPPS study data.

Age	N	Alive	Dead	Censored	Survival	Hazard
2	11956	11425	532	0	0.9556	0.0444
3	11424	11269	156	4	0.9426	0.0136
4	11264	11193	72	1	0.9367	0.0063
5	11191	11134	62	3	0.9319	0.0051
6	11126	11061	70	950	0.9265	0.0058
7	10106	10034	76	15	0.9199	0.0071
8	10015	9940	79	1594	0.9130	0.0075
9	8342	8274	72	73	0.9055	0.0082
10	8197	8115	86	868	0.8965	0.0100
11	7243	7171	76	1341	0.8876	0.0099
12	5826	5745	81	1899	0.8752	0.0139
13	3846	3772	74	1943	0.8584	0.0192
14	1829	1788	41		0.8391	0.0224

Note: N = total number of trees.

Table 3.3. The estimated parameters and their associated standard errors and p-values for models (3.1b) and (3.2b) when fitted to the CAPPS study data.

	Model 3.1b			Model 3.2b		
Parameter	Estimate	Standard Error	Pr > | t |	Estimate	Standard Error	Pr > | t |
a	0.6843	0.0258	<0.0001	-0.03438	0.0045	<0.0001
a_F	-0.1200	0.0433	0.0056			
a_H	0.0920	0.0386	0.0173	0.2054	0.0737	0.0054
a_{HF}	-0.0500	0.0117	<0.0001	0.1194	0.0540	0.0273
b	-1.3539	0.0281	<0.0001	-2.8565	0.3884	<0.0001
b_F	-0.1759	0.0533	0.0010	-0.3898	0.1308	0.0029
b_H	0.1548	0.0414	0.0002			
c	-0.00118	0.000271	<0.0001	-0.05453	0.00328	<0.0001
c_F	-0.00172	0.000666	0.0098			
c_H				0.04294	0.00640	<0.0001
c_{HF}	-0.00219	0.000449	<0.0001	0.03039	0.00942	0.0013
d				0.8316	0.0619	<0.0001
d_H				0.2170	0.0519	<0.0001
d_{HF}				0.1104	0.0350	0.0016

Cultural treatments are F = fertilizer, H = herbicide, and HF = herbicide and fertilizer.

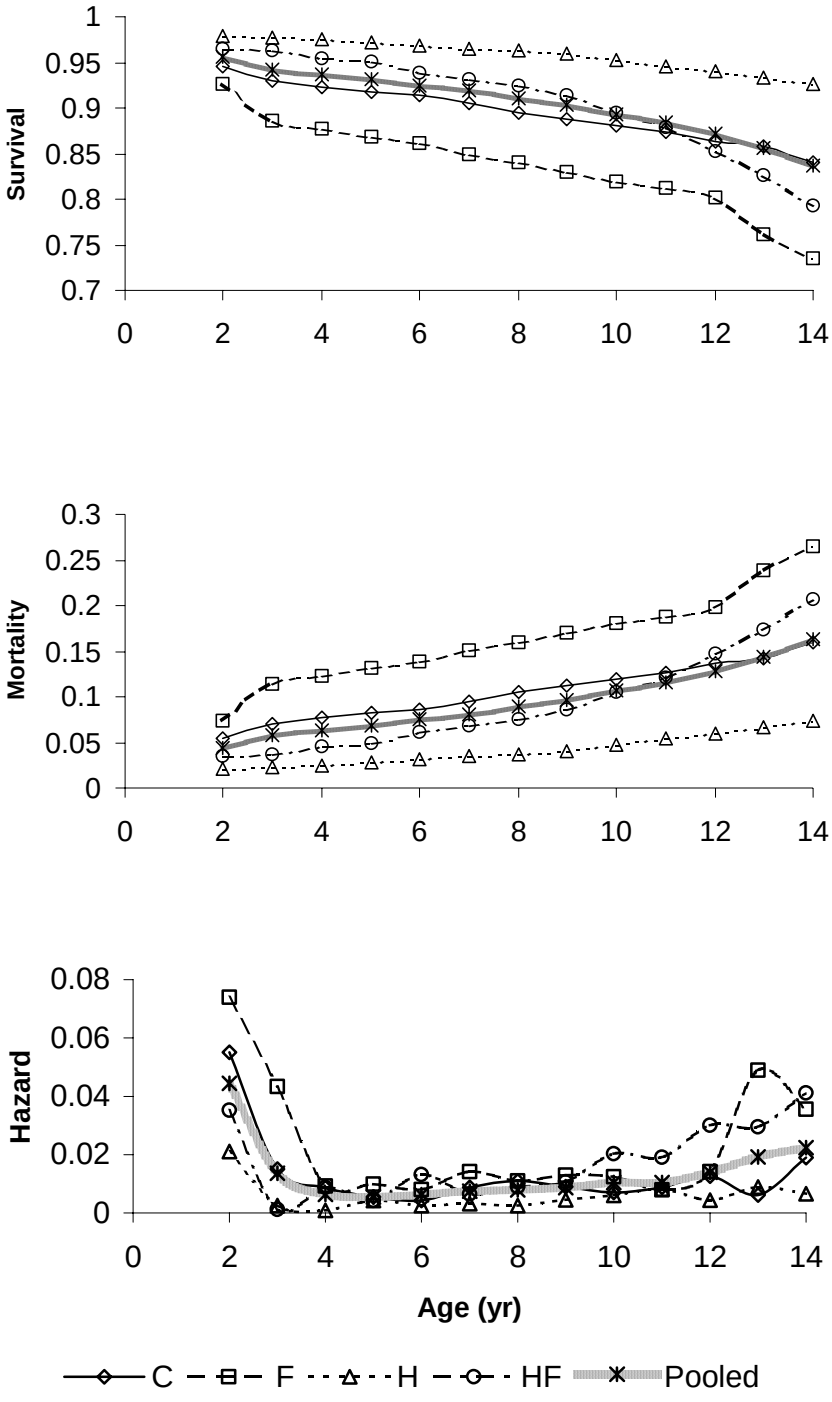


Figure 3.1. The CAPPS study empirical survival, mortality, and hazard functions by treatment (C = control, F = fertilized, H = herbicide, and HF = herbicide and fertilized) and for the spectrum of plots (pooled).

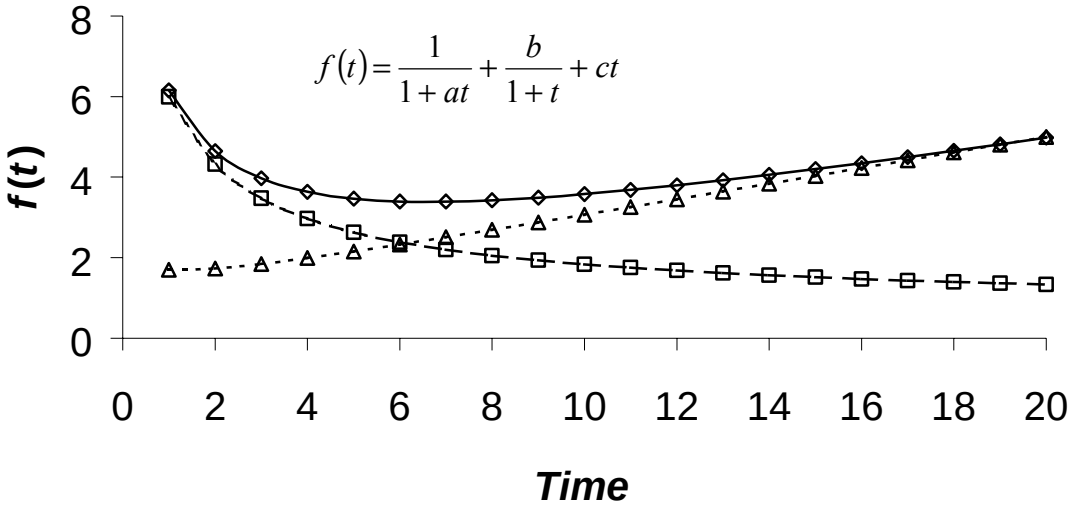


Figure 3.2. Examples of viable shapes for the chosen function $f(t)$ used in the differential equation for the whole stand survival model.

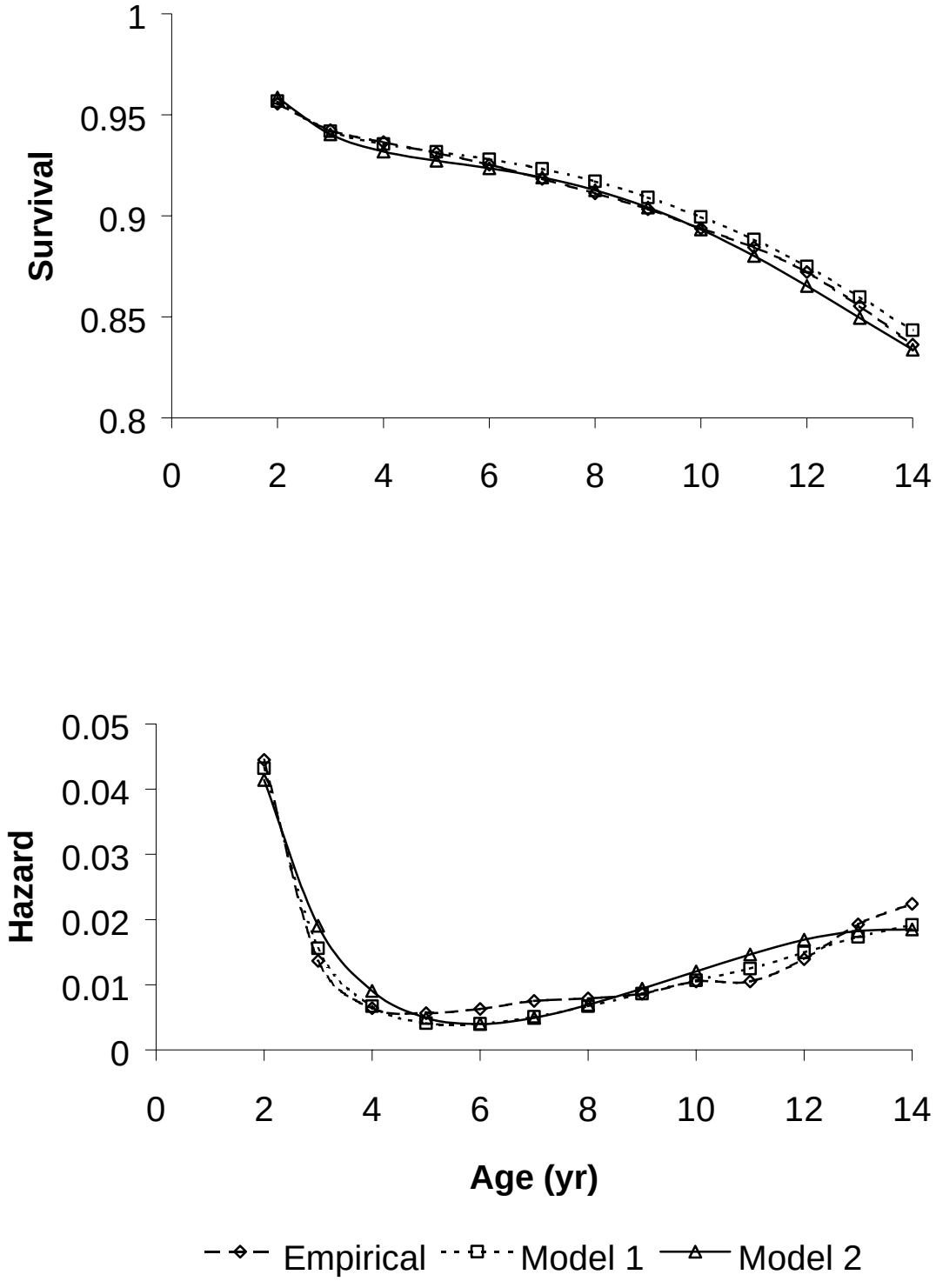


Figure 3.3. Models 1 (equation 3.1b) and 2 (equation 3.2b) fitted, and the empirical hazard and survival functions using the CAPPS study data.

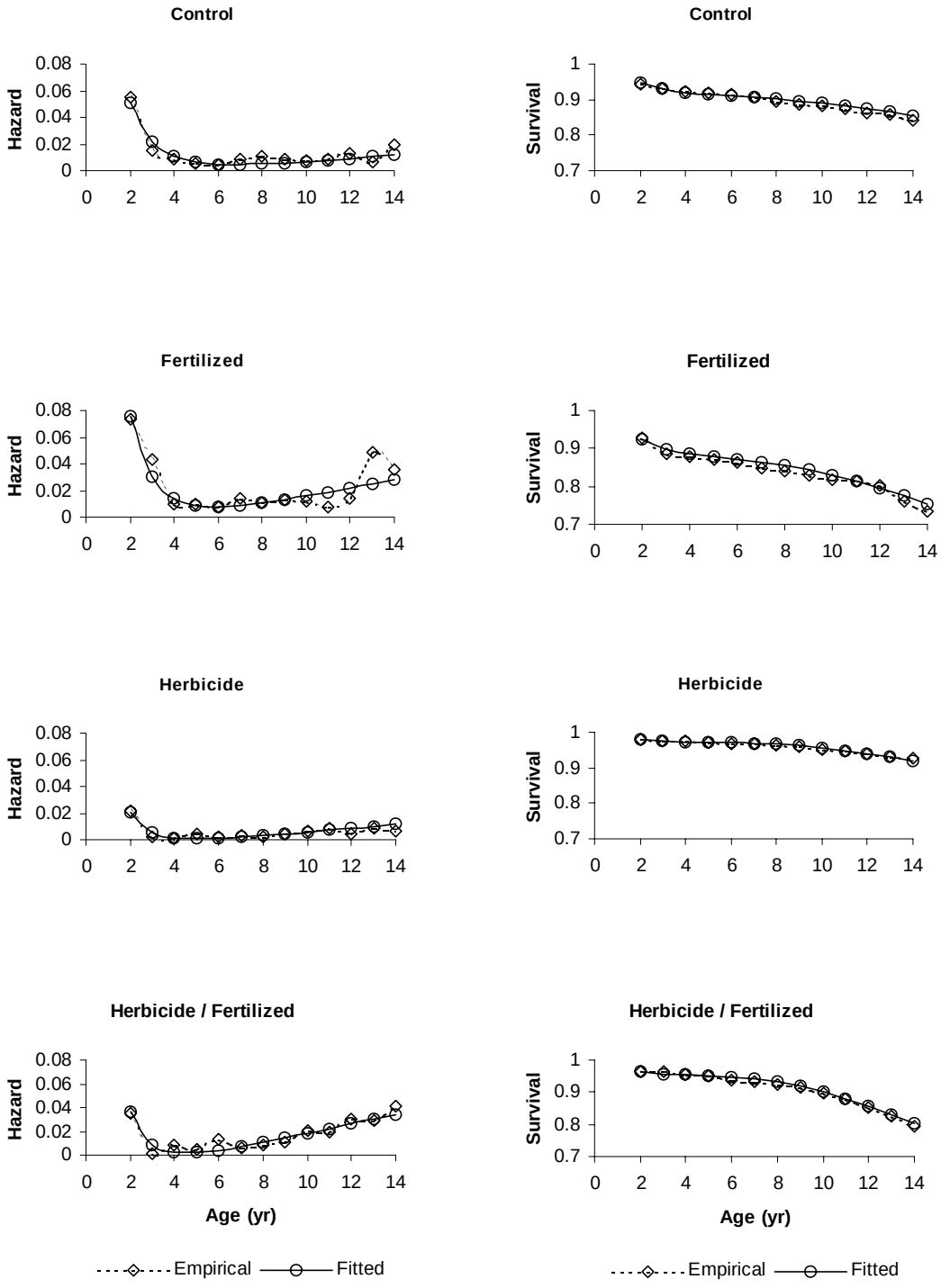


Figure 3.4. Model (3.1b) fitted and the empirical survival and hazard functions by treatment and age.

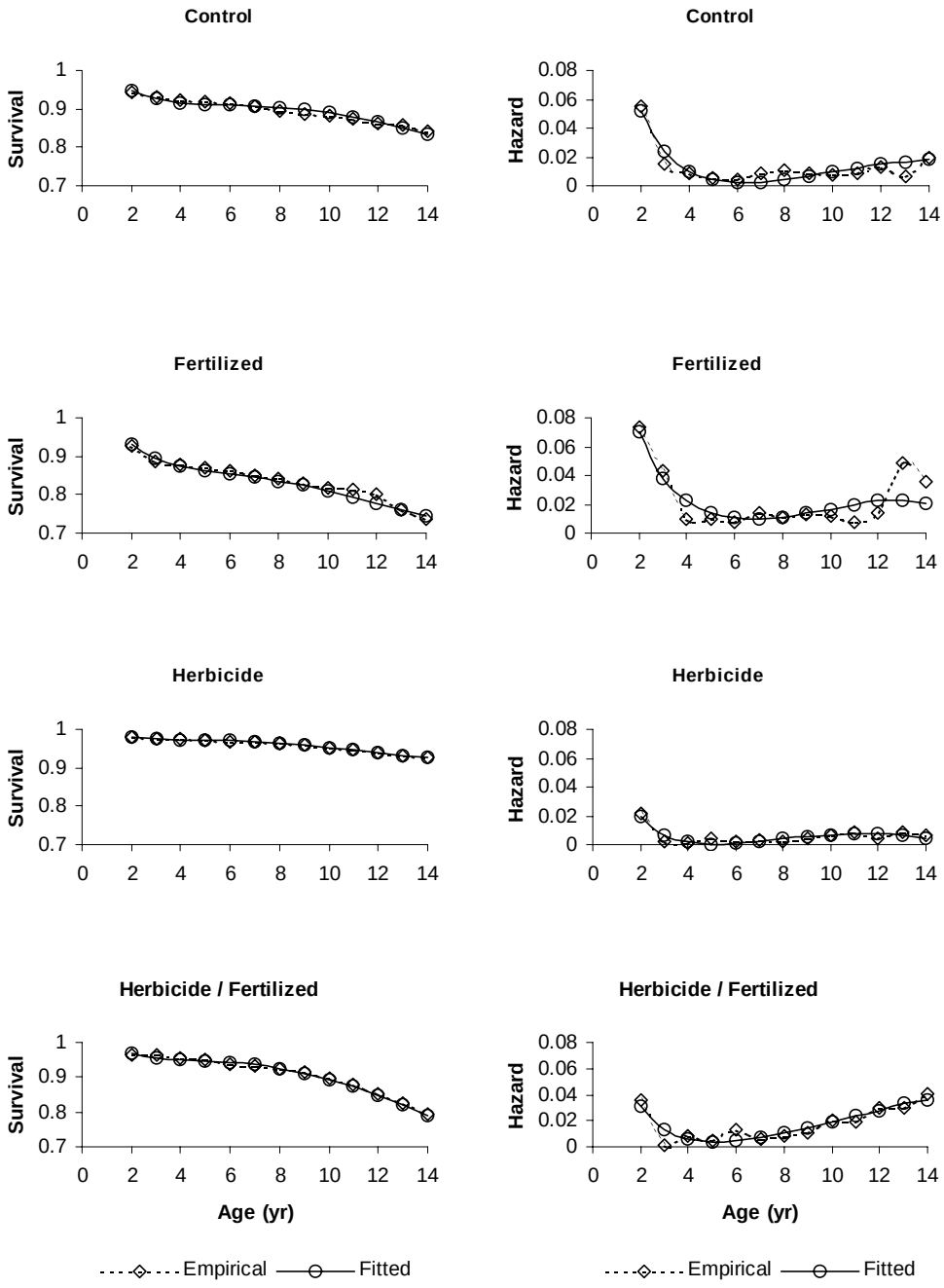


Figure 3.5. The CAPPS study model (3.2b) fitted and the empirical survival and hazard functions by treatment and age.

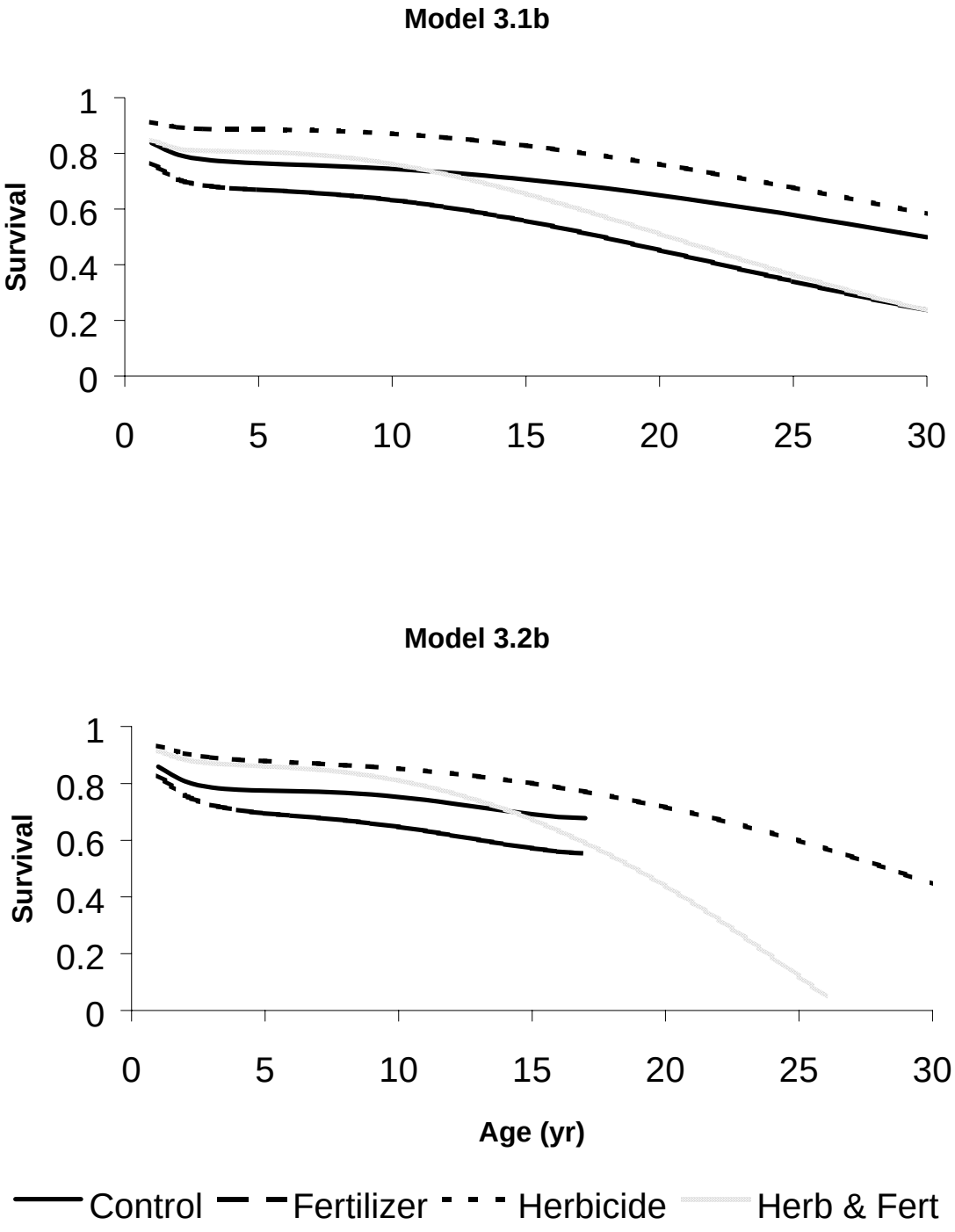


Figure 3.6. Predicted survival for ages 1-30 by treatment using models 3.1b and 3.2b.

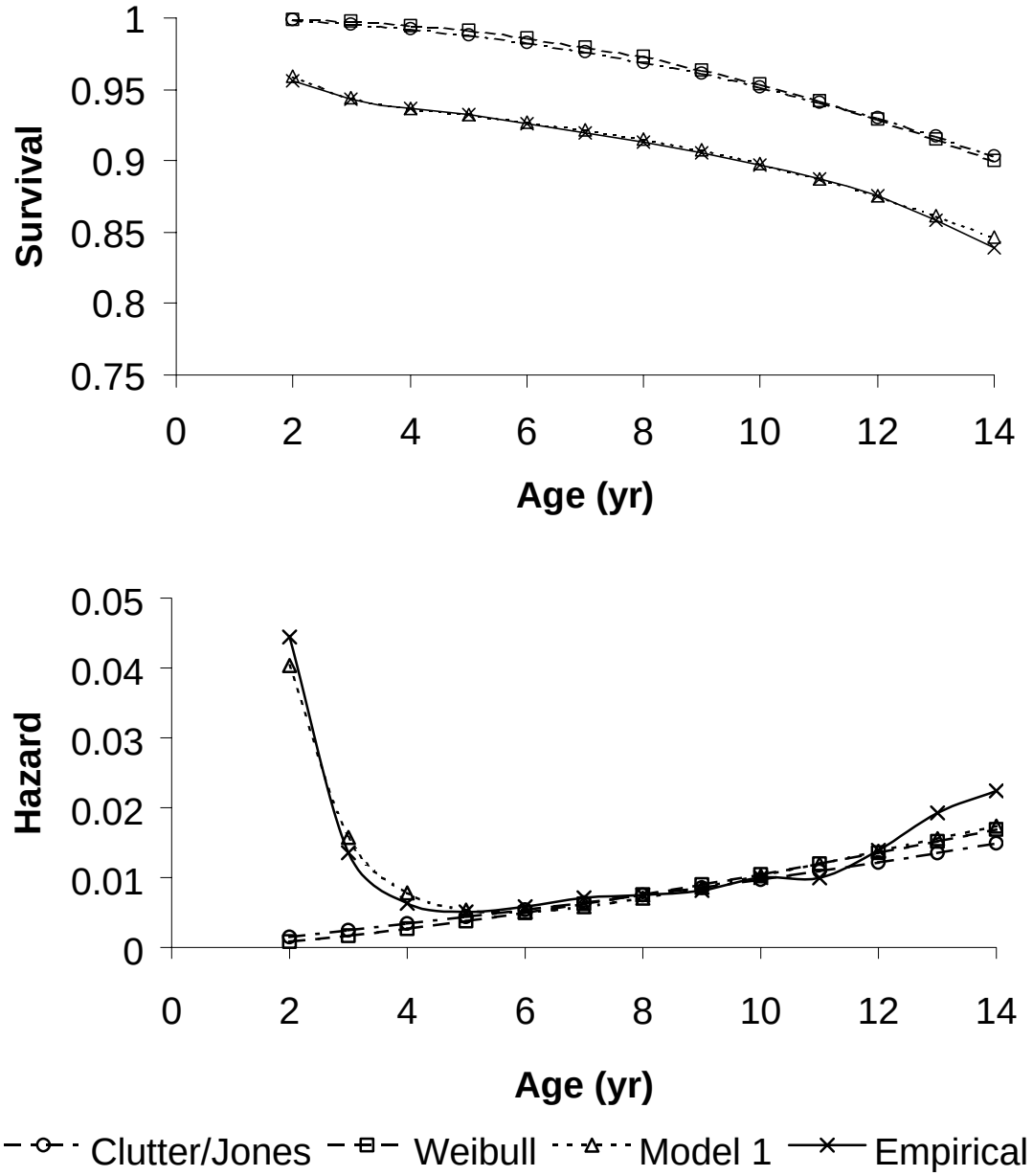


Figure 3.7. Predicted survival and their respective hazard curve for empirical, and the Weibull, Clutter and Jones (1980), and Model 1 (equation 3.1a) models. The models were fitted excluding the data from ages 2-4.

CHAPTER 4

ANALYSIS OF PERMANENT PLOT SURVIVAL DATA: A MULTILEVEL APPROACH

Introduction

Forest growth and yield systems rely on mortality or alternatively survival prediction, yet mortality is generally the least understood and has the most variability of all other components of forest growth and yield models (Hamilton 1986). It would usually be expected that two plots having identical site indices and density, and growing under the same conditions, that the basal area and volume per unit area would be similar at say age 15. However, the mortality experienced by these plots during the next 5-years can vary immensely, hence there can be high among plots survival variability. The mortality model is critical in accurately predicting growth and yield because of the sensitivity of most basal area growth models to the underlying mortality (Monserud and Sterba 1999, Lorimer and Frelich 1984). This sensitivity means that the contribution to total variability due to the mortality component increases as the projection period increases. Mortality can be classified as either natural or irregular (Staebler 1953). Natural mortality occurs because of competition for light, nutrients, and water and irregular mortality is caused by insects, diseases, and other catastrophic events. Individual tree mortality varies both among and within plots, i.e., these are sources of heterogeneity. Here we focus on natural mortality in developing an individual tree mortality model that accounts for different sources of heterogeneity.

Sources of heterogeneity occur naturally for permanent plot forestry studies because they typically have a multilevel data structure, i.e., measurement occasions are nested within trees, i.e., repeated measurements, and trees are nested within plots. In addition, the data are usually interval censored, i.e., data is collected for groups of trees at regular measurement occasions and it is only known that a tree died between adjacent measurement occasions. Fang and Bailey (2001) noted that the concept of multilevel forestry models likely began with Dr. J.L. Clutter's Duke University Ph.D. dissertation (1961) in which he recognized that forestry studies typically use repeated measurements, which require unique parameter estimation techniques. Additionally, since the pioneering study by Bailey and Clutter (1974), it has become widely recognized that incorporating site-specific parameters and using prior measurement data generally results in more precise predictions. Lappi and Bailey (1988) presented an innovative nonlinear mixed-effects height growth model that is acknowledged as an early example of a multilevel model in forestry. In recent years forestry multilevel models have become more common, both linear and nonlinear (e.g., Gregoire et al. 1995, Tasissa and Burkhart 1998, Hall and Bailey 2001, Fang and Bailey 2001). However, these multilevel models have a continuous response and the lowest level variation is assumed normally distributed. This assumption is not valid when modeling binary or proportion data, since the data are discrete or limited in range. We focus on modeling the multiple levels for individual tree survival in which the data are interval censored and responses are assumed binomially distributed.

Use of binary response multilevel models has gained popularity in recent years (e.g., Barbosa and Goldstein 2001, Biggeri et al. 2001, Yang et al. 2001). Multilevel

binary response models assume the responses are from a binomial distribution and the link function chosen is usually the probit, logit, or complementary log-log function (e.g., Biggeri et al. 2001, Hedeker et al. 2000). Recent efforts have focused on the parameter estimation techniques that are unique to binary or proportional response models (e.g., Rodriguez and Goldman 2001). These studies have demonstrated the importance of considering heterogeneity sources for a binary response model. In addition, multilevel or subject specific (SS) models have distinguished themselves from population averaged (PA) models.

Zeger et al. (1988) developed the framework for distinguishing between PA and SS models. Population-averaged models focus on the marginal expectation by modeling the mean response as a function of fixed covariates. Conversely, SS models focus on effects of covariates at a subject-specific level by relating the conditional mean response given latent subject-specific variables to covariates. Population averaged models do not specify a unique SS model but the SS model does specify a unique PA model, which can be inferred by integrating over the random effects. Subject-specific and PA models with identical linear predictors are usually incompatible for nonlinear models. In addition, unlike linear models that have the same parameter interpretation for SS and PA models, SS and PA parameters for a nonlinear model do not usually have the same interpretation. For PA models, the parameters describe how the average response changes across clusters given the covariates and only the link function needs to be specified correctly to make consistent inferences about the PA coefficients (Neuhaus et al. 1991). Subject-specific models pertain to expected changes for an individual or group of similar individuals (Vonesh and Chinchilli 1997). If fixed effects hypothesis testing is the study

focus, then either modeling approach will usually result in similar inferences. However, the PA approach will not provide any information about the heterogeneity that may exist at the different levels (Ten Have and Uttal 1991). If prediction is our primary purpose, then the SS approach is usually preferred (Ten Have and Uttal 1991). Individual tree forestry survival models naturally focus on prediction; therefore, the SS approach is the more appropriate choice. Our purpose is to demonstrate the methodology for developing an individual tree survival model that uses interval-censored data and accounts for the different sources of heterogeneity.

Survival Analysis Synopsis

Survival analysis concentrates on a group or groups of individuals that have a defined event. There are three basic requirements for analyzing survival data: time of origin must be precisely specified, a scale determined for time, and the exact meaning of failure specified (Lawless 1984). The main interest in survival analysis is usually on how failure times change across groups and strata. Usually, this is explored through estimation and/or modeling of the survival function or of the hazard function.

A survival function can be defined by letting the random variable T represent the time until the event of interest. Let the distribution of T be $F(t) = \Pr[T \le t]$, where $f(t)$ is the corresponding density function. Then the survival function is defined as $S(t) = \Pr(T > t) = 1 - F(t)$, which is the proportion of the population still alive at time t. The hazard function is the instantaneous rate of mortality assuming the individual has survived to time t, and is often referred to as a conditional density (Collet 1994). To define the hazard function, we first define the probability of death in the interval t_j to t_{j+1} given the individual survives to t_j, which is

$$\Pr\left[t_{j+1} \geq t \geq t_j \mid t \geq t_j\right] = \frac{\Pr\left[t_{j+1} \geq t \geq t_j\right]}{1 - F(t_j)}.$$

This defines the age-specific mortality rate for the interval t_j to t_{j+1}. If we divide this function by $(t_{j+1} - t_j)$ and take the limit as $(t_{j+1} - t_j) \rightarrow 0$, it yields the instantaneous mortality rate, i.e.,

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)}.$$

Here, $h(t)$ is known as the *hazard function* and its shape over time is generally monotonically increasing, decreasing, or constant. However, the hazard function can have an inflection point and can change directions over time. According to Cox and Oates (1984), there are several reasons why the hazard function is important. First, the immediate risk can be determined for an individual known to be alive at age t. Second, comparing groups of individuals may be more effectively conducted using the hazard function. Lastly, hazard function models can be convenient when there is censoring or there are several types of failure.

The cumulative hazard function, which is thought of as the cumulative risk, i.e., the sum of all risks faced by an individual going from zero to t, is defined as

$$H(t) = \int_0^t h(x)dx.$$

Using the fact that

$$h(t) = -\frac{\partial}{\partial t}\log[S(t)],$$

the survivor function is related to the cumulative hazard function as

$$S(t) = \exp\left(- \int_0^t h(x)dx \right),$$

which implies that the probability of surviving to *t* is a function of the hazard at all durations up to *t*. The expected lifespan is obtained by integrating the survival function from zero to infinity. In addition, $S(\infty) = 0$ and the limit as *t* approaches infinity for the cumulative hazard function is infinity. This implies that the event will occur with certainty only if the cumulative risk over the long duration is substantially high. These expressions illustrate that the survival and hazard functions provide alternative expressions for the distribution of *T*.

Modeling in survival analysis usually assumes a hazard function that arises from one of three general families of survival models: the proportional hazards, accelerated life failure, and proportional odds models (Collet 1994). The proportional hazards model, developed by Cox (1972), is a semi-parametric method that assumes only the functional form or mean and variance but not the entire distribution. Accelerated failure time models assume that the predictor variables measured for an individual act multiplicatively on the time scale. Proportional odds models assume that the covariates act multiplicatively on the odds of survival beyond *t* and the log of the cumulative odds ratio is proportional to the distance between the values of the explanatory variables (Agresti 1990). Here we adopt the Cox proportional hazard model, which is defined as $h_j(t) = \exp(\eta_j) h_0(t)$, where η_j is a function of explanatory variables for the j^{th} subject and $h_0(t)$ is the baseline hazard function. The corresponding survival function for the j^{th} subject is given by $S_j(t) = [S_0(t)]^{\phi}$, where $\phi = e^{\eta_j}$ and $S_0(t)$ is the baseline survivor function.

Individual Tree Survival Synopsis

Modeling individual tree mortality or alternatively survival, hereafter the modeling of survival or mortality is used interchangeably, began in earnest in the 1970's with a study by Hamilton (1974) in which he used the logistic equation. The logistic equation has become the individual tree survival model of choice (e.g., Hamilton 1974, Krumland et al. 1977, Wykoff et al. 1982), which is likely because of the ease of parameter interpretation and function flexibility. Moreover, beginning with Hamilton and Edwards (1976), it has become common to annualize mortality. Monserud (1976) modified Hamilton and Edwards (1976) method by developing an annualized generalized logistic equation that uses data from any measured time interval in the parameter estimation. The logistic equation has been used to model individual tree mortality for a variety of species and stand conditions (e.g., Hamilton 1986, Teck and Hilt 1990, Avila and Burkhart 1992). Recent individual tree mortality models have continued the trend of using the logistic equation (e.g., Monserud and Sterba 1999, Huebschmann et al. 2000, Yao et al. 2001, Shen et al. 2001). Although the logistic cumulative distribution is the most widely used for modeling individual tree survival, some other cumulative distributions used are the Richard's function (Buford and Hafley 1985) and gamma (Kobe and Coates 1997). In addition, non-traditional individual tree survival methods have been developed such as the binary classification tree (CART) (Dobbertin and Biging 1998) and an artificial neural network (Guan and Gertner 1991). Regardless of the chosen model or methodology, analysis of individual tree survival makes assumption about the survival and hazard functions.

Individual tree mortality models often assume one or a combination of proportional hazards, proportional odds, and accelerated failure time models. For example, the logistic equation has the accelerated failure time and proportional odds properties. Other individual tree survival analysis studies have used the Cox proportional hazards model (Volney 1998) and the log normal distribution (Preisler and Slaugther 1997) to allow for more flexibility in the hazard function. We adopt a Cox proportional hazards model for the modeling of interval censored individual tree survival data.

Individual Tree Survival Model Development

Our survival model formulation assumes the data are from a permanent plot plantation study in which measurement occasions are nested within a tree and trees are nested within a plot. We begin by defining the probability of a tree dying during the i^{th} time interval (t_{i-1}, t_i), $i = 1, 2, \ldots, n_{jk}$. A tree's mortality noted at time t_i, had an actual death time of t, where $t_{i-1} \le t < t_i$. All trees enter the study at $t_0 = 0$, however the calendar time corresponding to t_0 may vary by plot. All trees are followed to time t_l, where $l = n_{jk}$ and n_{jk} is the number of measurement occasions for the j^{th} tree on the k^{th} plot. In addition, $t_{l+1} = \infty$ for trees that are still alive at the last measurement occasion. Let p_{ijk} be the probability of mortality occurring for the j^{th} tree on plot k during the i^{th} time interval, i.e., $p_{ijk} = \Pr\left(t_{i-1} \le T_{jk} < t_i\right)$. Here T_{jk} is a random variable associated with the survival time for the j^{th} tree on plot k. The conditional probability of mortality during the i^{th} time interval given that the death occurs at or after t_{i-1} is given by $\pi_{ijk} = \Pr\left(t_{i-1} \le T_{jk} < t_i \mid T_{jk} \ge t_{i-1}\right)$, where $i = 1, 2, \ldots\ldots, l+1$. This means that $p_{ijk} = \left(1-\pi_{1jk}\right)\left(1-\pi_{2jk}\right)\ldots\ldots\ldots\ldots\left(1-\pi_{i-1,jk}\right)\pi_{ijk}$, where $i = 2, 3, \ldots\ldots, l+1$ and $p_{1jk} =$

π_{1jk}. Complementary probabilities are associated with a tree not dying during an interval and p_{1jk} is the probability it died during the first interval.

The likelihood function is developed using the conditional probabilities. Let δ_{ijk} = 1 if the j^{th} tree on the k^{th} plot succumbs in the interval from t_{i-1} to t_i and zero if it is alive. Let s_{ijk} = 1 if the j^{th} tree on the k^{th} plot succumbs after t_i and zero otherwise. Defining s_{ijk} = $\delta_{i+1,jk} + \delta_{i+2,,jk} + \ldots\ldots + \delta_{l+1,jk}$, the likelihood function for the total number of observations (s_{ijk}) for all j trees on all k plots, is

$L = \prod_{k=1}^{n} \prod_{j=1}^{n_k} \prod_{i=1}^{l+1} p_{ijk}^{\delta_{ijk}}$. Substituting for p_{ijk}, the sample likelihood function becomes

$L = \prod_{k=1}^{n} \prod_{j=1}^{n_k} \prod_{i=1}^{l+1} \left[\left(1-\pi_{1jk}\right)\left(1-\pi_{2jk}\right)\ldots\ldots\ldots\ldots\left(1-\pi_{i-1,jk}\right)\pi_{ijk} \right]^{\delta_{ijk}}$. Which can be expressed as

$L = \prod_{k=1}^{n} \prod_{j=1}^{n_k} \pi_{1jk}^{\delta_{1jk}} \left[\left(1-\pi_{1jk}\right)\pi_{2jk} \right]^{\delta_{2jk}} \ldots\ldots\left[\left(1-\pi_{1jk}\right)\left(1-\pi_{2jk}\right)\ldots\ldots\ldots\left(1-\pi_{i-1,jk}\right)\pi_{l+1,jk} \right]^{\delta_{l+1,jk}}$.

Hence, the sample likelihood function simplifies to $L = \prod_{k=1}^{n} \prod_{j=1}^{n_k} \pi_{l+1,jk}^{\delta_{l+1,jk}} \prod_{i=1}^{l} \pi_{ijk}^{\delta_{ijk}} \left(1-\pi_{ijk}\right)^{s_{ijk}}$.

Our interest is tree mortality and this implies that the probability of death, conditional on the tree being alive at time t_{l+1}, in the $t_{l+1,\infty}$ interval is one. This means that $\pi_{l+1,jk}$ equals one and therefore, $\pi_{l+1,jk}^{\delta_{l+1,jk}}$ equals one. Thus, the likelihood function can be reduced to

$$L = \prod_{k=1}^{n} \prod_{j=1}^{n_k} \prod_{i=1}^{n_{jk}} \pi_{ijk}^{\delta_{ijk}} \left(1 - \pi_{ijk}\right)^{s_{ijk}},$$

which be expressed in terms of δ_{ijk} by noting that s_{ijk} = 1-δ_{ijk} , hence we have

$$L = \prod_{k=1}^{n} \prod_{j=1}^{n_k} \prod_{i=1}^{n_{jk}} \pi_{ijk}^{\delta_{ijk}} \left(1 - \pi_{ijk}\right)^{1-\delta_{ijk}}.$$

This is the likelihood function for a binomial distribution with parameters 1 and π_{ijk} that corresponds to the total measurement occasions of all trees on k plots, which is a series of N Bernoulli trials. Here N equals the total measurement occasions (n_{jk}) for all n_k trees on all n plots. For example, suppose that there is data on two trees recorded for four measurement periods. Furthermore, suppose the first tree's mortality was detected at the second measurement period and the second tree's mortality was undetected for the four intervals. Using the likelihood function, for these two trees we have $(1-\pi_{11})\ \pi_{21}$ and $(1-\pi_{12})\ (1-\pi_{22})\ (1-\pi_{32})\ (1-\pi_{42})$, respectively. This means that for the first tree, the probability of the mortality occurring during the second interval is simply the probability that it survives the first interval and dies during the second interval. For the second tree, the probability that it succumbs in the l+1 interval is the product of the probabilities that it doesn't succumb during the first l intervals.

Model formulation proceeds using the conditional survivorship relationship. Given that the j^{th} tree on the k^{th} plot succumbs after time i is the survival function, then we can define the conditional survivor function for the j^{th} tree on the k^{th} plot as

$$1-\pi_{ijk} = \Pr\left(T_{jk} \geq t_i \mid T_{jk} \geq t_{i-1}\right) = \frac{S_{jk}\left(t_i\right)}{S_{jk}\left(t_{i-1}\right)} \ .$$

Adopting a Cox proportional hazards model, the hazard rate at time t_i for the j^{th} tree on plot k is $h_{jk}\left(t_i\right) = \exp\left(\eta_{jk}\right) h_0\left(t_i\right)$. Where $\eta_{jk} = \beta_1 x_{1jk} + \beta_2 x_{2jk} + .. + \beta_p x_{pjk}$ and $h_0(t_i)$ is the baseline hazard function. This model assumes the hazards are proportional only at the specified times, which is a relaxation of usual proportional hazards model that always assumes proportionality (Collet 1994). Expressing the Cox proportional hazard model in terms of the survivor function yields

$$S_{jk}(t_i) = \left[\frac{S_0(t_i)}{S_0(t_{i-1})}\right]^{\exp(\eta_{jk})}.$$

After taking logs, we have the grouped time version of the continuous time proportional hazards model (McCullagh 1980), i.e.,

$$\log\{-\log[1-\pi_{ijk}]\} = \eta_{jk} + \log\left[-\log\left(\frac{S_0(t_i)}{S_0(t_{i-1})}\right)\right] = \eta_{jk} + \kappa_i.$$

Hence, this is a linear model for the complementary log-log transformation of π_{ijk} in which the parameters κ_i are associated with the i^{th} time interval. The parameters κ_i yield an estimate of the baseline log-log survival function. These parameters can be incorporated into the model by fitting terms corresponding to the i^{th} time interval using indicator variables. The estimated coefficients can be interpreted just as in a proportional hazards model (Allison 1995). Hence, the model expresses the covariate effects on the log of the integrated hazard function, defined as $H(t) = -\log(1-\pi_{ijk}) = \exp(\eta_{jk} + \kappa_i)$. Thus for a covariate *x*, it expresses the $100(e^{\beta x} - 1)$ percent of increase or decrease in the hazard of death for a one unit increase in *x*. Failure probability for a given interval is $\pi_{ijk} = 1 - \exp[-\exp(\eta_{jk} + \kappa_i)]$, which assumes the hazards are proportional at the cut points κ_i.

Assuming proportional hazards at the cut points can be relaxed by incorporating time dependent covariates, which is referred to as the Cox regression model (Hedeker et al. 2000), and is expressed as $h_{jk}(t_i) = \exp(\eta_{jk}(t_i))h_0(t_i)$. The baseline hazard function, $h_0(t_i)$, can be interpreted as the hazard function for a subject for whom all the covariates are zero at the time origin and remains this value through time. This means that the relative hazard is time dependent and implies that the hazard of death at time *t* is no

longer proportional to the baseline hazard model. In addition, the survivor function for the j^{th} tree on the k^{th} plot depends on $h_0(t_i)$ and the time varying covariates. Therefore, $S_{jk}(t_i) \neq S_0(t_i)^{\varphi}$, which means the survivor function for the j^{th} tree on the k^{th} plot is difficult to obtain, however, we are usually only interested in obtaining the hazard of mortality for a given tree during a specified interval. The complementary log-log link function that includes time varying covariates would likely be sufficient for modeling forest mortality if the data structure was not hierarchical. However, it is typical for data from permanent plot forestry studies to have a multilevel structure.

Multilevel Individual Tree Survival Model

Multilevel models assume multiple sources of heterogeneity and recognize units at one level as grouped (nested) in the next higher level. Pinheiro and Bates (1999) number the levels by excluding the error term; here we adopt the nomenclature used by Goldstein (1995) in which the error term is recognized as the lowest level. Hence, we have measurement occasions nested within a tree (level 1), trees nested within a plot (level 2), and plots (level 3). According to Goldstein (1995), there are advantages in explicitly modeling the manner in which subjects are grouped. It allows the analyst to obtain statistically efficient estimates of the regression coefficients. The use of grouping information provides the correct standard errors, confidence intervals, and tests of significance. Lastly, it allows measuring covariates at any grouping level and then obtaining the corresponding predictions. In addition, if the relationship between the response and the covariates is nonlinear, e.g., logistic regression, then ignoring groups can result in large biases in the parameter estimates (Rodriguez and Goldman 2001).

Our multilevel model formulation accounts for the nesting of trees within a plot and plots by defining y_{ijk} as one if the j^{th} tree on the k^{th} plot dies during the i^{th} interval, zero otherwise. We assume given the random effects b_{jk} and b_k, that the y_{ijk}'s are independent Bernoulli random variables with conditional expectation π_{ijk}. Define π_{ijk} as the conditional probability that the j^{th} tree on the k^{th} plot succumbs during the i^{th} interval given that it has survived the previous intervals and given tree and plot specific random effects. Our multilevel individual tree survival model that uses the complementary log-log function, which accounts for interval censoring, is

$$\left.\begin{aligned} & y_{ijk} \mid b_k, b_{jk} \sim Binomial\left(1, \pi_{ijk}\right) \\ & \log\left\{-\log\left(1-\pi_{ijk}\right)\right\} = \kappa_i + \sum_l \beta_l x_{ijkl} + b_{jk} z_{(1)ijk} + b_k z_{(2)ijk} \\ & b_k \sim N\left(0, \sigma^2_{z_k}\right) \\ & b_{jk} \sim N\left(0, \sigma^2_{z_{jk}}\right) \end{aligned}\right\} \qquad (4.1)$$

Where $\boldsymbol{x}_{ijk}$ is the p x n_{ijk} covariate matrix associated with the fixed effects β, $\boldsymbol{z}_{(m)ijk}$ (m = 1, 2) are the q_m x n_{ijk} covariate matrices associated with the random effects b, and κ_i are the baseline hazards associated with the i^{th} interval. The random effects at the different levels are assumed independent of each other and are assumed to have a symmetric positive definite covariance matrix Σ_m. If the subject (tree) is the only source of heterogeneity, i.e., there is no need to incorporate multiple levels, then the above model reduces to the complementary log-log model of McCullagh (1980).

Our previous likelihood function is still valid for the multilevel model but the response is now conditional on the random effects. Let $\delta_{ijk} = y_{ijk}$, now define δ as the binary response vector pattern for the n plots having n_k trees observed at n_{jk} measurement occasions. Assuming independence of the response vector conditional on the random effects, the likelihood function can be expressed as (Gibbons and Hedeker 1997)

$$L(\delta \mid \boldsymbol{b}_k, \boldsymbol{b}_{jk}, \beta) = \prod_{k=1}^{n} \prod_{j=1}^{n_k} \prod_{i=1}^{n_{jk}} (\pi_{ijk})^{\delta_{ijk}} (1 - \pi_{ijk})^{1-\delta_{ijk}}$$

The unconditional distribution for the level 2 and level 3 random effects are assumed MN(0, $\sigma^2\Sigma$), where Σ is a block diagonal matrix with blocks Σ_2 and Σ_3, respectively. The marginal distribution of δ for the j^{th} tree on the k^{th} plot can be obtained from the likelihood by integrating over the distribution of the random effects ($f(\boldsymbol{\theta})$), i.e.,

$$L(\delta) = \prod_{k=1}^{n} \int_{\vartheta} \prod_{j=1}^{\boldsymbol{n_k}} \prod_{i=1}^{n_{jk}} (\pi_{ijk})^{\delta_{ijk}} (1 - \pi_{ijk})^{1-\delta_{ijk}} f(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

This is a difficult integral to evaluate because there is not a closed form solution when assuming a multivariate normal distribution for the random effects. Conceptually, the proposed multilevel survival model is straightforward. However, because of the difficulty in parameter estimation for multilevel binary response models, the use of multilevel models to analyze binary response variables is relatively recent (e.g., Goldstein 1991, Hedeker et al. 2001, Biggeri et al. 2001, Rodriguez and Goldman 2001).

Parameter estimation challenges are mainly related to the response having a binomial distribution rather than the more tractable normal distribution assumed in classical linear and nonlinear models. In our study, it was necessary to restrict attention to methods of estimation that were computationally feasible, i.e., methods that converged

for the models considered. A common multilevel parameter estimation technique for a binary response variable, and the one we adopt, is the marginal quasilikelihood (MQL-1) method. This method is motivated by using a linearization of the multilevel model. MQL-1 approximates $L(\delta)$ by using a 1st order Taylor series expansion of the function around fixed effects $\beta = \beta_0$ and random effects $b = 0$,

i.e., $f(H_{01}) = f(H_0) + X(\hat{\beta}_{01} - \hat{\beta}_0) f'(H_0)$. Here $f(\bullet)$ is some nonlinear function, e.g., Chapman-Richards function. Where β_0 and β_{01} are the current and updated estimates for the fixed effects from the iterative generalized least squares (IGLS) or restricted iterative generalized least squares (RIGLS) algorithm, respectively, and $H_0 = X\hat{\beta}_0$. The MLwiN (Rasbash et al. 2000) software was used to estimate our multilevel survival model parameters.

Data

Permanent plot loblolly pine data was obtained from the Consortium for Accelerated Pine Production Studies (CAPPS), which is overseen by the Warnell School of Forestry at the University of Georgia. Loblolly pine plantations were established throughout Georgia at Athens, Dawsonville, Eatonton, Thompson, Tifton, and Waycross. The study called for two complete blocks to be established at each location with each block containing four 0.15 ha treatment plots, which was established at each location using bare-root seedlings planted using a 2.44 m by 2.44 m spacing. A 0.05 ha measurement plot was centered within each of the treatment plots. Each of the four cultural treatments was randomly assigned to the blocks at each location. These cultural treatments are

1) Herbicide (H): spray plot with non-soil active herbicide as needed to maintain

complete control of woody and herbaceous vegetation,

2) Fertilization (F): apply recommended rates of fertilizer to achieve and maintain accelerated growth rates,

3) Herbicide – Fertilization (HF): apply both herbicide and fertilization treatments, and

4) Control (C): no cultural treatment.

The original study called for a replication of all treatment plots every two years for the first ten years of the study. This protocol would have resulted in five complete sets of experimental plots at all installations, where the plots have a staggered initiation time. Each location has two treatments with two levels: herbicide versus no herbicide and fertilization versus no fertilization. The actual study varies from the protocol because of limitations and the replications have been repeated at different intervals for different locations. The plots have been measured annually beginning at age one. The CAPPS data structure is: 112,365 total observations for measurement occasion within a tree, 11956 trees within a plot, and 146 plots. Study survival data are summarized by plot distribution (see Table 3.1, page 36) and by plot and tree attributes (Table 4.1).

Preliminary Analysis

Preliminary analyses were used for detecting survival and hazards trends and to investigate the trends of time varying covariates for the linear component of the CLL function. Survival and hazards by age were estimated using the nonparametric Kaplan-Meier product limit survival estimator (Kaplan and Meier 1958), which is defined as

$$\hat{S}(t) = \prod_{i:t_i \le t} \left[1 - \frac{d_i}{n_i} \right]$$

for $t_1 \le t \le t_j$, where n_i and d_i are the subjects at risk (n_i) and that die (d_i) at time t_i. A Kaplan-Meier type hazard function estimator is given by $h(t) = \frac{d_i}{n_i \tau_i}$, for $t_i \le t \le t_{i+1}$ and for our case $\tau_i = t_{i+1} - t_i = 1$ since all plots were measured annually.

Kaplan-Meier survival estimates for the plots reveal that its underlying estimated hazard function is bathtub shaped (Figure 4.1). In addition, the KM survival and hazard function estimates by treatment (Figure 4.2) reveal that, using the C treatment as the baseline, there is an acceleration in mortality for the F treatment in the early years and again at about age 12. The HF treatment hazards appear constant relative to the C treatment until about age 8, after which there is acceleration in mortality. The H treatment has a deceleration in mortality relative to the C treatment. Hazards variability is highest and lowest for the HF and H treatments, respectively. An age by treatment interaction is evident for the hazards, which suggests non-proportional hazards at the cut points. However, our Cox proportional hazard model is still valid, since we can allow for time varying covariates and/or a treatment by interval interaction, which relaxes the assumption of proportional hazards at the cut points.

It is important to determine if the dependence of the hazard function on the time varying covariates can be adequately modeled using linear terms (Collet 1994). If there is *a priori* information about the effect of a time varying covariate on the hazard function then this information can be used to develop the linear predictor. Lacking *a priori* information, covariates can be examined to determine if linearity is a reasonable assumption. Time dependent covariates considered are trees per hectare (*TPH*), basal area per hectare (*BA/ha*), quadratic mean diameter (D_q), individual-tee height (H_t),

diameter at breast height (*dbh*), and relative spacing (*RS*). All considered time dependent covariates were standardized. Plot level variables, which includes the plot mean H_t and *dbh*, are centered and standardized using the grand mean of the plots. Tree level covariates H_t and *dbh* are centered and standardized by plot. Let *A* be a factor that is formed from a time varying covariate, then linearity in the original values will correspond to linearity in the factor *A* (Collet 1994). We formed levels of factor *A* for each considered covariate by computing the 25^{th}, 50^{th}, and 75^{th} quartiles. These quartiles were used to create four levels of *A* by grouping the 0-25, 20-50, 50-75, and 75-100 quantiles. This results in each level of *A* having roughly the same number of observations. The empirical complementary-log-log was computed for each level of *A* for all time varying covariates using the given level empirical proportion of mortality (Figure 4.3). The *BA/ha*, D_q, and *RS* attributes exhibit quadratic trends, therefore, these attributes will be considered as both linear and quadratic terms in the model. The empirical CLL plots for *TPH*, H_t, and *dbh* illustrate linear trends and hence, these covariates are only considered as linear terms in the model.

Model Fitting Procedure

As Fang and Bailey (2001) noted, the determination of which parameters are purely fixed and which are both fixed and random is frequently data dependent. Exploratory analysis of potential random effects oftentimes fits the model by group, however this requires sufficient data for each group in order to obtain valid parameter estimates (Pinheiro and Bates 1999). These group parameter estimates and their respective confidence interval are plotted to determine if a specific covariate should be considered random, i.e., are the

confidence intervals by group disjoint? Level 1 data are typically assumed normally distributed and this method, given sufficient group observations, is often adequate. Conversely, the level 1 data for a binary response individual tree survival model are assumed binomially distributed. Therefore, sufficient observations by group may not be enough to obtain valid parameter estimates because some plots may experience little or no mortality.

There are numerous recommendations for determining random effects in a multilevel model. Pinheiro and Bates (1999) recommend, in the absence of prior information about the random effects variance-covariance matrix, to allow all effects to be fixed and random if convergence can be achieved. Oftentimes modelers who allow all parameters to be both random and fixed are using a specific functional form (e.g., Chapman-Richards function was used by Fang and Bailey 2001). Models such as the Chapman-Richards function have few parameters and it may be feasible to allow these parameters to vary at all levels. Conversely, a linear model without a specific form is often built from scratch and a baseline model is usually chosen somewhat arbitrary. Fang and Bailey (1999) fitted a multilevel linear model for plot basal area, however a specific model form was chosen. Oftentimes a variance component model is selected as the baseline model for a linear model.

Variance component models assume that the variation among the responses can be partitioned (Lindsey 1999), which for our data are among and within plot variations. Goldstein (1995) espouses fitting a variance component model and then adding covariates as fixed effects when developing a multilevel linear model. Hox (1995) suggests a more pragmatic view by recommending a model selection procedure that

begins with a random intercept model and then adding and removing fixed and random parameters until a suitable model is determined. Both the variance component and random intercept models are often referred to as hierarchical linear (HLM) and multilevel linear models (Goldstein 1995, Bryk and Raudenbush 1992). The basic structure for a HLM is an extension of the ANCOVA model in which some or all regression coefficients are assumed to arise from a random sample belonging to the population. Assuming a random cut point model, we can define a three level HLM model by letting log(-log(1-π_{ijk})) denote the CLL for the i^{th} interval of the j^{th} tree on the k^{th} plot, i.e., $\log\left(-\log\left(1-\pi_{ijk}\right)\right)=\kappa_i+b_{jk}+b_k$. Where κ_i is the cut point for the i^{th} interval, b_{jk} and b_k are random parameters for the trees within a plot and plots, $i = 1, 2, \ldots\ldots, n_{jk}$, $j = 1, 2, \ldots.., n_k$, $k = 1, 2, \ldots.., n$, and the distributions of the random effects are $b_{jk} \sim N\left(0,\sigma^2_{b_{jk}}\right)$ and $b_k \sim N\left(0,\sigma^2_k\right)$.

This random cut point model would oftentimes be an adequate baseline model. However, there is no evidence to suggest that the variability, both among and within plots, can be adequately modeled using one variance component for all cultural treatments. Moreover, it is of interest to analyze the variance by treatment. In addition, preliminary analysis suggests a treatment by age interaction with respect to the hazards (Figure 4.2) and the necessity of time varying covariates (Figure 4.3). A baseline model was developed by first defining π_{ijk} as the probability that the j^{th} tree on the k^{th} plot succumbs during the i^{th} interval. Our baseline three-level mixed effects survival model that uses the CLL link function and includes fixed effects for the: intervals, treatment by interval interaction, and time varying covariates, and allows the random parameters to vary by treatment is

$$
\left.
\begin{aligned}
& y_{ijk} \mid b_k, b_{jk} \sim Binomial\left(1, \pi_{ijk}\right) \\
& \log\left(-\log\left(1-\pi_{ijk}\right)\right) = \kappa_{00,i} + \kappa_{01,i} x_{F,i} + \kappa_{02,i} x_{H,i} + \kappa_{03,i} x_{HF,i} + \beta_l x + b_{jk}^{(2)} + b_k^{(3)} \\
& \beta_l x = \beta_1 BA/ha_{ik} + \beta_2 TPH_{ik} + \beta_3 RS_{ik} + \beta_4 H_{ik}^{Plot} + \beta_5 dbh_k^{Plot} + \beta_6 dbh_{ijk} + \beta_7 H_{ijk} \\
& b_{jk}^{(2)} = b_{20} z_C + b_{21} z_F + b_{22} z_H + b_{23} z_{HF} \\
& b_k^{(3)} = b_{30} z_C + b_{31} z_F + b_{32} z_H + b_{33} z_{HF} \\
& \boldsymbol{b}_k \sim N\left(0, \Sigma_{(3)}\right) \\
& \boldsymbol{b}_{jk} \sim N\left(0, \Sigma_{(2)}\right) \\
& \Sigma_{(2)} = \Sigma_{(3)} = \begin{bmatrix} \sigma_C^2 & 0 & 0 & 0 \\ 0 & \sigma_F^2 & 0 & 0 \\ 0 & 0 & \sigma_H^2 & 0 \\ 0 & 0 & 0 & \sigma_{HF}^2 \end{bmatrix}
\end{aligned}
\right\} \quad (4.2)
$$

Where y_{ijk} is a binary response that equals one if the j^{th} tree on the k^{th} plot dies during the i^{th} interval and zero otherwise, x and z are indicator variables that equal one if the response belongs to the cultural treatment and zero otherwise (C = control, F = fertilizer, H = herbicide, and HF herbicide and fertilizer). Interval, fixed and random effects parameters are denoted by κ, β, and b, respectively. Intervals are defined as 1, 2, 3, …, 13, which correspond to the ages 1-2, 2-3, 3-4, ….., 13-14. Random effects are normally distributed with mean zero and level 2 (trees within a plot) and 3 (plots) covariance matrices of $\Sigma_{(2)}$ and $\Sigma_{(3)}$, respectively. Off-diagonal elements of the random effects covariance matrices are assumed to equal zero, i.e., it is assumed that the plot-specific treatment effects are independent. Standardized time varying covariates are *BA/ha*, *TPH*,

RS, D_q, H^{Plot}, dbh^{Plot}, *dbh*, and *H*. Our baseline model (equation 1) includes 13 intervals, 39 treatment by interval interactions (C, H, and HF), 8 time varying covariates, and 8 variance and covariance components; a total of 68 parameters.

A likelihood ratio test is typically used to determine the necessity of fixed and random effects for a mixed effects model that assumes the lowest level variance is normally distributed (Pinheiro and Bates 1999). However, for multilevel binary response models that use quasi-likelihood to estimate the parameters, the likelihood ratio test is a crude approximation and the preference for testing the fixed and random effects is the Wald chi-square test (Goldstein 1995). Therefore, the Wald chi-square test was used to test for parameter significance using $\alpha = 0.05$. Our first hypothesis tested the necessity of the age by treatment interaction. This was accomplished using a joint Wald χ^2 test and the joint treatment by age interaction test statistic is 231.08 with 39 degrees of freedom (p-value<0.0001). Since our joint test for the interactions is significant, we did not remove insignificant individual interaction terms. However, time varying covariates were removed from the model using a stepwise procedure, which included adding the quadratic terms for *BA/ha*, D_q, and *RS*. The time varying quadratic covariates (BA/ha^2, D_q^2, and RS^2) and the linear D_q covariate were not significant; therefore, they were removed from the model. Interval, treatment by age interactions for intervals 1 and 2 (for the sake of brevity), time varying covariates, and random parameters are presented in Table 4.2. Random effects, i.e., residuals for the trees within a plot and plots, are not presented because there are 146 plot and 11956 tree level random effects.

Model Diagnostics, Fit, and Parameter Interpretation

Several model diagnostic tools were used to assess model adequacy. Residual and Q-Q plots were inspected for levels 2 and 3 and there is no evidence of any serious heterogeneity or outliers. For example, plot level rank residuals and Q-Q plots by treatment illustrate no serious outliers or departures from the assumption of normality (Figure 4.4). Q-Q plots are useful for checking normality of the residuals and for our assumed normal distribution; it should be approximately a straight line. Hence, there is no evidence of non-normality for any of the treatments. Graphs of the plots by treatment and rank, where plots for each treatment are ranked from smallest to largest residuals, illustrate that H and C have the least and most variability, respectively. Standardized residual for the plot level residuals are presented for the mean predicted values and *BA/ha* (Figure 4.5), the other attributes behave similarly with no evidence of heterogeneity. Standardized residuals for five of the 146 plots are greater than 2.0 in absolute value. However, this is within our expectation and there are no abnormally large standardized residuals. The *BA/ha* plot illustrates that most plots for the HF treatment are above average, whereas the C and F plots tend to be below average for this attribute.

Additional model fit was assessed by estimating the marginal proportions by age for our multilevel model, often referred to as a subject specific (SS) model. We also compared several fit statistics with a population averaged (PA) model. The PA model was obtained by re-fitting equation (4.2) and excluding the random effects. Integration over the random effects was necessary to compare our SS model with the marginal mortality probabilities. This is a difficult integral to evaluate and computer intensive; therefore, we used Monte Carlo integration to estimate the marginal means (Figure 4.6).

The comparison of the marginal means to the actual raw proportions by treatment and for the plots is virtually identical. To compare the PA and SS models we computed the Pearson chi-square statistic for survival by plot. Plot survival predictions were grouped by age and the degrees of freedom is the number of measurement occasions for the given plot minus one. The SS and PA models have 6 and 10 plots, respectively, which are significantly different from the actual stand tables for the 146 plots ($\alpha = 0.05$). Here stand table refers to the number of stems by age for a given plot. In addition, the SS model fits the stand tables better relative to the PA model for 115 of the 146 plots. Hence, it provides a substantial improvement in fit for over ¾ of the plots. For example, two stand tables were chosen at random to represent "normal" (plot 40) and "extreme" (plot 59) cases with respect to plot survival (Table 4.3). Plot 40 results illustrate that the SS model offer a substantial improvement for predicting the actual surviving trees at the last plot measurement than the PA model. In addition, the SS model for plot 59 predicts closer to the actual remaining trees at age 12 relative to the PA model. However, it is substantially over predicting the mortality for this plot. In general, our SS model substantially improves the model fit. Moreover, it provides information about the plot to plot and within plot variability.

Parameters are interpreted as for a proportional hazards model. A covariate effect on the hazard can be expressed as $-\log(1-\pi_{ijk}) = e^{\eta_{ijk}} e^{\beta x_{ijk}}$, where the LHS is the cumulative hazard function for a given interval and the effect of the covariate on the hazard as a percentage is expressed as $100\left(e^{\beta x_{ijk}} - 1\right)$. For example, the estimated parameter for the F and interval 1 interaction is 0.2921. This means, holding all other covariates constant, that there is a $100\left(e^{0.2921} - 1\right) \cong 33.92$ percent increase in the hazard

of mortality for a fertilized tree during interval 1 relative to the C treatment. Conversely, for H there is a $100\left(e^{-1.4053}-1\right)\cong 75.47$ percent reduction in the hazard for a herbicide tree during the first interval relative to the C treatment. The HF treatment decreases the hazard for an individual tree during the first interval by 69.75 percent relative to the C treatment. Time varying covariates are interpreted similarly. For example, if a plot is one standard deviation below the mean for *BA/ha*, then there is a $100\left(e^{-2.6893}-1\right)\cong 93.21$ percent decrease in the hazard, holding all other covariates constant. Conversely, if a plot is one standard deviation above the mean for *BA/ha* it means there is an approximate increase in the hazard of 1372.1 percent. Parameter estimates reveal that *BA/ha* and *dbh* have the largest and *RS* has the smallest impact on survival, respectively. For *dbh*, if a tree is one standard deviation below and above the mean, it results in an approximate increase and decrease of the tree's hazard of 92.50 and 1233.38 percent, respectively. Individual tree *dbh* has a greater impact on survival than the average plot *dbh*. Note that if a plot is one standard deviation above the mean for the average plot *dbh* then there is an increase in an individual tree's hazard of 283.74 percent. These *dbh* results imply that plots that are above the mean have a larger hazard and trees within a plot that are below the mean for *dbh* have an increase in their hazards. This means that smaller trees on older plots are more likely to die. In addition, note that the signs for the plot average and individual tree parameter estimates for H_t are negative. This means that as a tree or plot increases in size relative to the mean, its hazard decreases. However, the plot average height has a greater effect on the hazard than the individual tree height. The estimated parameter for *TPH* is negative, which means that as *TPH* increases the probability of mortality decreases. This seems counterintuitive, however, our study plots were all

planted at the same density and plots that have more *TPH* as time increases are those plots with a lower rate of mortality. Therefore, for our data it is logical for the *TPH* estimated parameter to be negative.

Random Parameter Interpretation and Contrasts

Random parameter variance component estimates are indicative of the survival variability for the trees within a plot and among plots for a given treatment. Note that the C treatment is the only treatment where the estimated variance is larger at the plot level than for trees within a plot. Hence, there is more plot-to-plot survival variability for the C treatment. Herbicide has the least variability at the plot level and F has the least variability for trees within a plot. Note that the H treatment has the largest level 2 to level 3 ratio of the variability, i.e., there is 15.4 times more variability at the tree level relative to the plot level. However, the results for a Wald type test reveal that the H plot level variability is not significantly difference from zero and its tree level variability is borderline significant. Control and HF treatments have the highest variability for the plot and trees within plot levels, respectively. In addition, the Wald type test for the C, F, and HF treatments are significant at both levels. A Wald type test can used to construct contrasts among the treatment random parameters, i.e., to test for differences in the variability among the treatments.

We demonstrate the procedure for testing orthogonal contrasts of the random parameters using the plot random parameters. Define a r x q contrast matrix C, where r is the number of contrasts being tested and q is the number of random parameters. Then we can form q-1 linearly independent functions of the q random parameters. For our

model, we can test a set of three linear independent plot level random parameter functions by defining

$$C = \begin{bmatrix} z_C & z_F & z_H & z_{HF} \\ z_C & z_F & z_H & z_{HF} \\ z_C & z_F & z_H & z_{HF} \end{bmatrix}$$

Now define *g* to be the vector of plot level random parameters, i.e.,

$$g = \begin{bmatrix} \sigma_C^2 & \sigma_F^2 & \sigma_H^2 & \sigma_{HF}^2 \end{bmatrix}^T$$

Then define a general hypothesis as $H_0 : f = k$, where $f = Cg$ and *k* is a *q* x 1 vector that is usually assumed to be zero. The Wald type test statistic for this general hypothesis is

$$\mathrm{T} = (f-k)^T \left[C\left(Z^T V^{-1} Z\right)^{-1} C^T \right]^{-1} (f-k)$$

If the null hypothesis is true then this is distributed approximately χ^2 with *r* degrees of freedom (Goldstein 1995). Note $\left(Z^T V^{-1} Z\right)^{-1}$ is the covariance matrix of the random parameters and since this is unknown, the estimated covariance matrix of the random parameters is substituted. Our set of plot level orthogonal contrasts will test the mean of the F, H, and HF treatments versus the C, the mean of the H and HF treatments versus the F treatment, and the H versus HF. The *C* matrix corresponding to this set of orthogonal contrasts is

$$C = \begin{bmatrix} 3 & -1 & -1 & -1 \\ 0 & 2 & -1 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

and *g*, defined as the vector of plot level random parameters, is

$$g = \begin{bmatrix} 0.5872 & 0.3903 & 0.0321 & 0.4063 \end{bmatrix}^T$$

Our null hypothesis is

$$H_0 : Cg = k \Rightarrow \begin{bmatrix} 3 & -1 & -1 & -1 \\ 0 & 2 & -1 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \sigma_C^2 \\ \sigma_F^2 \\ \sigma_H^2 \\ \sigma_{HF}^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 3\sigma_{\mathrm{C}}^2 - \sigma_F^2 - \sigma_H^2 - \sigma_{HF}^2 \\ 2\sigma_{\mathrm{F}}^2 - \sigma_{\mathrm{H}}^2 - \sigma_{\mathrm{HF}}^2 \\ \sigma_H^2 - \sigma_{HF}^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The Wald type test statistic is

$$\mathrm{T} = \begin{bmatrix} 0.9329 & 0.3422 & -0.3742 \end{bmatrix} \left[C\left(Z^T \hat{V}^{-1} Z \right)^{-1} C^T \right]^{-1} \begin{bmatrix} 0.9329 \\ 0.3422 \\ -0.3742 \end{bmatrix}$$

Where the covariance matrix $\left(Z^T \hat{V}^{-1} Z \right)^{-1}$ for the random effects is

$$\begin{bmatrix} 0.025148 & -4.7894E-15 & 5.9826E-16 & -2.3596E-15 \\ -4.7894E-15 & 0.017356 & -1.2763E-15 & -7.7346E-16 \\ 5.9826E-16 & -1.2763E-15 & 0.0026325 & -1.0299E-16 \\ -2.3596E-15 & -7.7346E-16 & -1.0299E-16 & 0.022108 \end{bmatrix}$$

After pre and post multiplying the random effects covariance matrix by *C* and estimating this inverse we have

$$\mathrm{T} = \begin{bmatrix} 0.9329 & 0.3422 & -0.3742 \end{bmatrix} \begin{bmatrix} 4.0820 & 1.3102 & -4.2447 \\ 1.3102 & 13.1055 & -11.3479 \\ -4.2447 & -11.3479 & 52.6939 \end{bmatrix} \begin{bmatrix} 0.9329 \\ 0.3422 \\ -0.3742 \end{bmatrix} = 19.172$$

Where T is approximately χ^2 with 3 degrees of freedom and the p-value is 0.0002519. Therefore, we would reject the null hypothesis for any reasonable α. Individual Wald type χ^2 test statistics (1 degree of freedom) for these three contrasts are 3.241 (p-value = 0.07182), 1.241 (p-value = 0.26528), and 5.658 (p-value = 0.01738). Hence, there is some evidence of a significant difference between the C and the mean of the F, H, and HF treatments. There is no significant difference between the F and mean of the H and HF treatments. However, there is a significant difference between the H and HF

treatments. Since this set of orthogonal contrasts informs us that there are some significant differences among the plot level random parameters, we can construct hypothesis tests to determine which random parameters are significantly different. We have already determined that $\sigma_H^2 \neq \sigma_{HF}^2$. In addition, the $\sigma_H^2 = \sigma_F^2$ contrast results in a χ^2 test statistic of 6.415 (p-value = 0.011316), which is indicative of a significant difference between these treatments for the random parameters at the plot level. Furthermore, contrasts of the HF treatment versus the C and F treatments results in χ^2 test statistics of 0.691 and 0.007, and their respective p-values are 0.40582 and 0.93332. Hence, there is no evidence that the plot level variability is different for the C, F, and HF treatments, but there is evidence that the H treatment is significantly different from the other treatments.

Contrasts for level 2 reveal that the HF treatment is significantly different from the C and F treatments, and has borderline significance (p-value = 0.05477) with the H treatment. The C, F, and H treatments are not significantly different. Therefore, it is reasonable to assume that the variability among trees within a plot for the C, F, and H treatments can be modeled adequately using one variance parameter.

Individual Tree Mortality Predictions

Individual tree mortality predictions can be obtained at all levels and we can obtain predictions for new plots assuming different resolutions of information are available. Here we demonstrate predicting individual tree mortality for our study plots and then illustrate mortality predictions for a new plot.

Our first scenario assumes we would like a prediction of the typical mortality, setting the random effects to zero, for the C and H treatments at interval 1, which corresponds to ages 1-2, given the following plot and tree information. Plot and tree standardized covariate values are *BA/ha* (-1.0), *TPH* (0.0), *RS* (2.0), H^{Plot} (-1.5), dbh^{Plot} (-1.5), *dbh* (tree 1 = -1.5 and tree 2 = -1.0), and H_t (tree 1 = -1.5 and tree 2 = -1.0). Using equation (4.2) for tree 1, the predicted mortality for the C and H treatments are $\log\left[-\log\left(1-\pi_{1j}\right)\right] = -1.0038 - 1.4053x_H$, which results in tree 1 predictions of 0.3068 and 0.08597, respectively. Similarly for the second tree, the C and H mortality predictions are 0.0669 and 0.01685, respectively. Hence, as expected the H treatment reduces the probability of mortality substantially for interval 1 relative to the C treatment.

Our next scenario demonstrates how predictions are obtained at the different levels for a tree on the HF treatment during interval 8, which corresponds to ages 8-9. The first tree is from plot 1 (HF) and has the following attributes for this interval. Standardized values of the covariates are *BA/ha* (1.430172), *TPH* (0.1378), *RS* (-.5683655), H^{Plot} (0.229449), dbh^{Plot} (0.3832622), *dbh* (0.7104519), and H_t (0.702079). Hence, this plot is above the mean for *BA/ha*, *TPH*, H^{Plot}, and dbh^{Plot}, which would generally be expected since the HF treatment accelerates tree growth. In addition, this particular tree is above the plot average for *dbh* and H_t. A typical response for a tree with these attributes and treatment is $\log\left[-\log\left(1-\pi_{1j}\right)\right] = -5.164107$, which results in a 0.005702 mortality probability. However, by including the random effects we can obtain a more precise estimate for this tree's mortality. The plot level random effect is –0.08526806. Hence, using the plot level random effect the CLL prediction is $\log\left[-\log\left(1-\pi_{1j}\right)\right] = -5.164107 + b_k = -5.164107 - 0.08526806 = -5.24937506$ and after

transforming, the tree's mortality prediction is 0.005237. Including this tree's random effect of –0.1000968 in the CLL function gives a mortality prediction of 0.004739. This tree is still alive at the last measurement period and hence; the probability is more precise as expected by including the random effects. The plot level random effect reduces the hazard by $100\left(e^{-0.08526806}-1\right)\cong 8.17$ percent. Whereas, the tree level random effect reduces the hazard by $100\left(e^{-0.1000968}-1\right)\cong 9.53$ percent. Moreover, inclusion of the plot and tree level random effects results in a $100\left(e^{-0.0852688-0.1000968}-1\right)\cong 16.92$ percent decrease in the hazard. However, for this tree, the tree level random effect has a more profound impact upon survival than the plot random effect. Conversely, another tree was chosen from this plot that died during this interval. The fixed, plot, and tree specific effects are –2.023651, -0.08526806, and 0.6395384, respectively, which following the previous procedure correspond to mortality predictions of 0.1238, 0.1143, and 0.2055. Improvement in this tree's mortality prediction is evident and also by noting that the tree level random parameter increases the hazard by $100\left(e^{0..6395384}-1\right)\cong 89.56$ percent relative to the fixed effects. Mortality predictions at all levels for a given tree are easily obtainable using the MLwiN software. It is typical in forestry to desire predictions for a plot that is not included in the original study, we present several scenarios for mortality predictions given a new plot.

Case I: Mortality Predictions for trees with no prior records and the plot is not associated with the study treatments

Suppose mortality predictions are desired for trees on a new plot that has no previous records and is different from our study cultural treatments. Hence, the plot or tree level random effects can't be estimated; however we can estimate the typical response. Let the

age 3 standard deviations for *BA/ha*, *TPH*, *RS*, H^{Plot}, and dbh^{Plot} be –0.25, -0.50, -2.0, -0.25, and –0.25, respectively. For simplicity, we will consider four trees on the plot that have standardized *dbh's* and H_t*'s* of –1.0, -0.5, 0.5, and 1.0. The estimated mortality probabilities for the age interval 3-4 are obtained using model (4.2) and setting the random effects to zero, i.e.,

$$\log\left(-\log\left(1-\pi_{ijk}\right)\right)=\kappa_{00,3}+\beta_1 BA/ha_{ik}+\beta_2 TPH_{ik}+\beta_3 RS_{ik}+\beta_4 H_{ik}^{Plot}+\beta_5 dbh_k^{Plot}+\beta_6 dbh_{ijk}+\beta_7 H_{ijk}$$

Here all terms of the complementary-log-log model are as defined previously. The substitution of the plot attributes and their respective parameter estimates yields

$$\log\left(-\log\left(1-\hat{\pi}_{ijk}\right)\right)=-5.2296+1.3448\,dbh_{ijk}-0.5418\,H_{ijk}$$

Mortality predictions are obtained by substituting the standardized *dbh's* and H_t*'s* for the four trees. The matrix of *dbh's* and H_t*'s* for these four trees is given by

$$\begin{bmatrix}-1.0 & -0.5 & 0.5 & 1.0\\ -1.0 & -0.5 & 0.5 & 1.0\end{bmatrix}^T,$$

where the first row represents *dbh* and the second row represents H_t.

The resulting mortality predictions for the four trees are 0.115527, 0.025315, 0.001118, and 0.000234. Hence, as expected, the probability of mortality decreases as the standardized variables increases. These mortality predictions are typical responses for a new plot that has a different treatment from the study plots. However, the plot random effect could be estimated if this new plot was associated with one of the study cultural treatments.

Case II: Prediction for a new Plot Associated with the HF Treatment but with no Prior Measurements

Suppose the *Case I* plot is associated with the HF cultural treatment but has no prior measurements, hence we can estimate the plot level random effect. However, because model (4.2) is nonlinear, there is some background given to motivate the method used to estimate the random effects. Goldstein's (1991) MQL-1 parameter estimation method is used to estimate the random effects, which is motivated by a generalized linear mixed model. A linear mixed effects model can be expressed as

$$\boldsymbol{y} = \boldsymbol{x\beta} + \boldsymbol{zb} + \boldsymbol{\varepsilon},$$

where $\boldsymbol{Z}$ and $\boldsymbol{X}$ are the design matrices corresponding to the random and fixed effects, $\boldsymbol{b}$ and $\boldsymbol{\beta}$ are the respective parameters, and $\boldsymbol{\varepsilon}$ is usually assumed to be normally distributed with mean zero. Random effects can be estimated using the best linear unbiased predictor (BLUP) (Vonesh and Chinchilli 1997), which is often referred to as an empirical Bayes (EB) or shrinkage type estimator (Goldstein 1995), and is defined as

$$\hat{\boldsymbol{b}} = \boldsymbol{DZ}^{T}\left(\boldsymbol{ZDZ}^{T} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\right),$$

where $\boldsymbol{D}$ and $\boldsymbol{R}$ are the variance-covariance matrices of the random effects and errors, respectively. Using Goldstein's (1991) method, we can write our multilevel binary response model as (Rodriguez and Goldman, 1995)

$$\boldsymbol{y} = \boldsymbol{g}\left(\boldsymbol{X\beta} + \boldsymbol{Z}_2\boldsymbol{b}_2 + \boldsymbol{Z}_3\boldsymbol{b}_3\right) + \boldsymbol{\varepsilon},$$

where $\boldsymbol{\varepsilon}$ has a mean of zero and a variance which depends upon the mean $\mathbf{g}$. The inverse link $\mathbf{g}(\bullet)$ is approximated using a first order Taylor series expansion, around $\boldsymbol{\beta} = \boldsymbol{\beta}_0$ and $\boldsymbol{b_2} = \boldsymbol{b_3} = 0$, as

$$\boldsymbol{y} \approx \boldsymbol{g}(\boldsymbol{\eta}_0) + \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{\eta}_0} \boldsymbol{X}(\boldsymbol{\beta} - \boldsymbol{\beta}_0) + \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{\eta}_0} \boldsymbol{Z}_2 \boldsymbol{b}_2 + \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{\eta}_0} \boldsymbol{Z}_3 \boldsymbol{b}_3 + \boldsymbol{\varepsilon},$$

where $\partial g / \partial \eta_0$ is a diagonal matrix of derivatives of the mean with respect to the conditional linear predictor evaluated at $\boldsymbol{\eta} = \boldsymbol{\eta}_0$. For our CLL model, the derivative is

$$\frac{\partial g}{\partial \eta} = e^{\eta} e^{-e^{\eta}}$$

Furthermore, call this derivative $\boldsymbol{\Lambda}$, our model which approximates the nonlinear model can be expressed as

$$\boldsymbol{y}^* \approx \boldsymbol{X}^* \boldsymbol{\beta} + \boldsymbol{Z}_2^* \boldsymbol{b}_2 + \boldsymbol{Z}_3^* \boldsymbol{b}_3 + \boldsymbol{\varepsilon},$$

which has the same form as a multilevel linear model. The dependent variable is $\boldsymbol{y}^* = \boldsymbol{y} - \boldsymbol{g}_0 + \boldsymbol{X}^* \boldsymbol{\beta}_0$, $\boldsymbol{X}^* = \boldsymbol{\Lambda X}$, and $\boldsymbol{Z}^* = \boldsymbol{\Lambda Z}$. Notice that $\boldsymbol{E}(\boldsymbol{y}^*) = \boldsymbol{X}^* \boldsymbol{\beta}$ and $\boldsymbol{\Lambda}$ is an estimate of the variance. The BLUP of the random effects for a given level is

$$\hat{\boldsymbol{b}}_i = \boldsymbol{D}_i \boldsymbol{Z}_i^{*T} \left(\boldsymbol{Z}_2^* \boldsymbol{D}_2 \boldsymbol{Z}_2^{*T} + \boldsymbol{Z}_3^* \boldsymbol{D}_3 \boldsymbol{Z}_3^{*T} + \boldsymbol{\Lambda} \right)^{-1} \left(\hat{\boldsymbol{y}}^* - \hat{\boldsymbol{X}}^* \boldsymbol{\beta} \right),$$

where *i*= 2 (tree level) or 3 (plot level) for our model. Using the fact that at convergence $\boldsymbol{\eta} = \boldsymbol{\eta}_0$, we can write the BLUP of the plot level random effects as

$$\hat{\boldsymbol{b}}_3 = \boldsymbol{D}_3 \boldsymbol{Z}_3^* \left(\boldsymbol{Z}_2^* \boldsymbol{D}_2 \boldsymbol{Z}_2^{*T} + \boldsymbol{Z}_3^* \boldsymbol{D}_3 \boldsymbol{Z}_3^{*T} + \boldsymbol{\Lambda} \right)^{-1} (\boldsymbol{y} - \hat{\boldsymbol{g}})$$

Continuing with our example, using the CLL's from *Case I*, our matrix of evaluated differentials is

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0.108571 & 0 & 0 & 0 \\ 0 & 0.024991 & 0 & 0 \\ 0 & 0 & 0.001117 & 0 \\ 0 & 0 & 0 & 0.000234 \end{bmatrix}$$

To simplify the computations for our example and since the covariances of our plot level random parameters are zero we can let $\boldsymbol{D}$ = 0.4063, which is the estimated variance component parameter for the HF treatment at the plot level. In addition, $\boldsymbol{Z}$ is a vector of ones corresponding to the four trees and let $\boldsymbol{V} = \boldsymbol{Z}_2^* \boldsymbol{D}_2 \boldsymbol{Z}_2^{*T} + \boldsymbol{Z}_3^* \boldsymbol{D}_3 \boldsymbol{Z}_3^{*T} + \boldsymbol{\Lambda}$. The level 2 and 3 design matrices are identical for our study, hence

$$\boldsymbol{Z}_2^* = \boldsymbol{Z}_3^* = \boldsymbol{\Lambda}\boldsymbol{Z}_3 = \begin{bmatrix} 0.108571 & 0 & 0 & 0 \\ 0 & 0.024991 & 0 & 0 \\ 0 & 0 & 0.001117 & 0 \\ 0 & 0 & 0 & 0.000234 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.108571 \\ 0.024991 \\ 0.001117 \\ 0.000234 \end{bmatrix}$$

Then estimate $\boldsymbol{V}$ using

$$\boldsymbol{Z}_2^* \boldsymbol{D}_2 \boldsymbol{Z}_2^{*T} = \begin{bmatrix} 0.0135016 & 0.0031078 & 0.0001389 & 0.0000291 \\ 0.0031078 & 0.0007154 & 0.0000320 & 6.6982E-6 \\ 0.0001389 & 0.0000320 & 1.4291E-6 & 2.9938E-7 \\ 0.0000291 & 6.6982E-6 & 2.9938E-6 & 6.2718E-8 \end{bmatrix}$$

and

$$\boldsymbol{Z}_3^* \boldsymbol{D}_3 \boldsymbol{Z}_3^{*T} = \begin{bmatrix} 0.0047893 & 0.0011024 & 0.0000493 & 0.0000103 \\ 0.0011024 & 0.0002538 & 0.0000113 & 2.3760E-6 \\ 0.0000493 & 0.0000113 & 5.0694E-7 & 1.0620E-7 \\ 0.0000103 & 2.376E-6 & 1.0620E-7 & 2.2247E-8 \end{bmatrix}$$

Hence,

$$\boldsymbol{V} = \boldsymbol{Z}_2^* \boldsymbol{D}_2 \boldsymbol{Z}_2^{*T} + \boldsymbol{Z}_3^* \boldsymbol{D}_3 \boldsymbol{Z}_3^{*T} + \boldsymbol{\Lambda} = \begin{bmatrix} 0.1268619 & 0.0042102 & 0.0001882 & 0.0000394 \\ 0.0042102 & 0.0259601 & 0.0000433 & 9.0742E-6 \\ 0.0001882 & 0.0000433 & 0.0011189 & 4.0558E-7 \\ 0.0000394 & 9.0742E-6 & 4.0558E-7 & 0.0002341 \end{bmatrix}$$

Suppose that tree 2 had recently died, our corresponding vector of residuals is then defined as

$$\boldsymbol{y}-\hat{\boldsymbol{g}}=\begin{bmatrix}0\\1\\0\\0\end{bmatrix}-\begin{bmatrix}0.115527\\0.025315\\0.001118\\0.000234\end{bmatrix}=\begin{bmatrix}-0.115527\\0.974685\\-0.001118\\-0.000234\end{bmatrix}$$

Thus, the random effect for this plot is estimated as

$$\hat{\boldsymbol{b}}_3=[0.4063]\begin{bmatrix}0.108571\\0.024991\\0.001117\\0.000234\end{bmatrix}^{\boldsymbol{T}}\boldsymbol{V}^{-1}\begin{bmatrix}-0.115527\\0.974685\\-0.001118\\-0.000234\end{bmatrix}=0.2881946$$

Using the predicted random effect for this plot and adding it to the previous obtained CLL's for *Case I*, the estimated age 3-4 mortality probabilities for the three living trees can be estimated using

$$\log\left(-\log\left(1-\pi_{ijk}\right)\right)=-5.2296+1.3348\,dbh_{ijk}-0.5418\,H_{ijk}+0.2881946$$

Thus, the predicted mortality probabilities for the three living trees are 0.151061, 0.001491, and 0.000312. Inclusion of the plot level random effect increases the predicted probability of mortality, which is expected since in our example we are using only four trees and assuming one tree died. It may be of interest to assess the plot effect in terms of the proportional hazards. This plot has about a $100\left(e^{0.2881946}-1\right)=33.40$ percent increase in the hazard of mortality relative to the typical HF treatment plot in which the random effect is zero. Our purpose is to illustrate the random effect prediction process and it is obvious that given prior information and a larger sample of trees we can get a better estimate of the plot random effect.

Case III: Mortality Predictions for a new Plot with Prior Measurements and Associated with the HF Treatment

Suppose that for *Case II* we have prior tree measurements, for simplicity we will ignore the prior measurements for the estimation of the plot random effects, i.e., we should use all information to re-estimate the plot random effect. Assume that in addition to the present measurements that there is also plot information for the ages 1 and 2. We will illustrate the tree random effect estimation process assuming the predicted age 1-2, 2-3, and 3-4 probabilities are 0.22, 0.19, and 0.11 for a given tree. Our tree level BLUP is given by

$$\hat{\boldsymbol{b}}_2 = \boldsymbol{D}_2 \boldsymbol{Z}_2^* \left(\boldsymbol{Z}_2^* \boldsymbol{D}_2 \boldsymbol{Z}_2^{*T} + \boldsymbol{Z}_3^* \boldsymbol{D}_3 \boldsymbol{Z}_3^{*T} + \boldsymbol{\Lambda} \right)^{-1} \left(\boldsymbol{y} - \hat{\boldsymbol{g}} \right)$$

The diagonal matrix of derivative evaluated at $\mathbf{g}^{-1}$ is

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0.193795 & 0 & 0 \\ 0 & 0.170687 & 0 \\ 0 & 0 & 0.103713 \end{bmatrix}$$

The level 2 and 3 design matrices are identical, therefore

$$\boldsymbol{Z}_3^* = \boldsymbol{Z}_2^* = \boldsymbol{\Lambda} \boldsymbol{Z}_2^* = \begin{bmatrix} 0.193795 & 0 & 0 \\ 0 & 0.170687 & 0 \\ 0 & 0 & 0.103713 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.193795 \\ 0.170687 \\ 0.103713 \end{bmatrix},$$

and the vector of residuals for this tree is

$$\boldsymbol{y} - \hat{\boldsymbol{g}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0.22 \\ 0.19 \\ 0.11 \end{bmatrix} = \begin{bmatrix} -0.22 \\ -0.19 \\ -0.11 \end{bmatrix}$$

The $\boldsymbol{V_2}$ estimated matrix is

$$\boldsymbol{V}_2 = \boldsymbol{Z}_2^* \boldsymbol{D}_2 \boldsymbol{Z}_2^{*T} + \boldsymbol{Z}_3^* \boldsymbol{D}_3 \boldsymbol{Z}_3^{*T} + \boldsymbol{\Lambda} = \begin{bmatrix} 0.2520714 & 0.0513276 & 0.0311877 \\ 0.0513276 & 0.2158943 & 0.0274689 \\ 0.0311877 & 0.0274689 & 0.1204037 \end{bmatrix}$$

Hence, our BLUP for this tree's random effect is

$$\hat{\boldsymbol{b}}_2 = [1.14541] \begin{bmatrix} 0.193795 \\ 0.170687 \\ 0.103713 \end{bmatrix}^{\boldsymbol{T}} \boldsymbol{V}^{-1} \begin{bmatrix} -0.22 \\ -0.19 \\ -0.11 \end{bmatrix} = -0.34498$$

This predicted random effect results in about a $100\left(1 - \boldsymbol{e}^{-0.34498}\right) = 29.18$ percent decrease in the hazard of mortality relative to a typical tree on this HF plot that has a random effect of zero, holding all other variables constant. Our purpose here was to demonstrate the prediction of the random effects. However, note that if for a new plot the information is contained in a stand table, the plot level and *dbh* class random effects could be predicted.

Discussion

Survival analysis concentrates on the hazard and survival functions and how these functions change over time and across strata. Individual tree mortality models oftentimes assume one or a combination of proportional hazards, proportional odds, and accelerated failure time models. For example, the logistic equation has the accelerated failure time and proportional odds properties. Other forestry survival analysis methods include adopting a Cox proportional hazards model (Volney 1998) and the log normal distribution was used by Preisler and Slaugther (1997) to allow for more flexibility in the hazard function. However, the logistic equation has become the most widely used individual tree survival model because of its flexibility, ease of parameter estimation, and interpretability.

Although the logistic equation is flexible, it does impose restrictions on the data, i.e., its estimated parameters are not invariant to the interval length. This means that switching from an interval of five years to one year changes the model and hence, the coefficients are not directly comparable when using these different interval lengths (Allison 1991). Conversely, the CLL function is invariant to the length of the measurement interval and hence the CLL the parameter estimates are comparable using the same data that is fitted for different interval lengths. One reason for the logistic equation popularity is the ease of parameter interpretation. Parameters in the logit model act multiplicatively on the odds of survival. As demonstrated, the CLL coefficients have a relative risk interpretation just as in the Cox proportional hazards model. Thus for a covariate x, it expresses the $100\left(e^{\beta x}-1\right)$ percent of increase or decrease in the hazard of death for a one unit increase in x.

Adopting a Cox proportional hazards model using the likelihood function for interval censored data leads to the grouped time version of the continuous time proportional hazards model (McCullagh 1980) of

$$\log\left\{-\log\left[1-\pi_{ijk}\right]\right\}=\eta_{jk}+\log\left[-\log\left(\frac{S_0\left(t_i\right)}{S_0\left(t_{i-1}\right)}\right)\right]=\eta_{jk}+\kappa_i$$

As discussed, this is a linear model for the CLL transformation of π_{ijk} in which the parameters κ_i are associated with the i^{th} time interval. Therefore, for interval censored permanent plot forestry data the CLL model is a more natural model choice. We demonstrated that the assumption of proportional hazards at the cut points can be relaxed by incorporating time dependent covariates and/or interaction terms. In addition, sources

of heterogeneity that are inherent in permanent plot repeated measurement studies are easily included in the model.

Permanent plot forest inventory data naturally have a hierarchical structure, i.e., measurement occasions (repeated measurements) are nested within trees, trees are nested within plots, and plots. Consideration of the multilevel structure allows the analyst to obtain statistically efficient estimates of the regression coefficients. In addition, incorporating the heterogeneity from the groups provides the correct standard errors, confidence intervals, and tests of significance. Lastly, it allows measuring covariates at any grouping level and then obtaining the corresponding predictions. Moreover, if the relationship between the response and the covariates is nonlinear, e.g., logistic, then ignoring groups can result in large biases in the parameter estimates (Rodriguez and Goldman 2001).

Conceptually, our multilevel survival model is straightforward. However, multilevel binary response models have presented some difficulties in parameter estimation. This is mainly due to the response having a binomial distribution rather than the more tractable normal distribution assumed in classical linear and nonlinear models. There have been several methods proposed for multilevel binary response model parameter estimation. These parameter estimation methods include maximum marginal likelihood (MML), 1st and 2nd order marginal quasi-likelihood (MQL-1 and MQL-2, respectively), 1st and 2nd order penalized quasi-likelihood (PQL-1 and PQL-2, respectively), and Markov Chain Monte Carlo (MCMC). However, there is no current consensus on which parameter estimation method is preferable under a given set of conditions.

There is a consensus that some of these methods are deficient in certain respects and that some methods are better than others, however the best methods are not feasible computationally. Moreover, in some instances the choice of parameter estimation method may be limited to those methods that converge. For our example, only the MQL-1 method converged. It has been suggested by Goldstein and Rasbash (1996) that a rule of thumb is to compare the MQL-1 and PQL-1 estimates and accept them if they are similar. However, a simulation study by Rodriguez and Goldman (2001) demonstrated that even if the MQL-1 and PQL-1 estimates are similar, there can be substantial bias in the estimates, i.e., similarity does not guarantee the estimates are close to the true values. They suggest using all four of the approximation methods (MQL-1, MQL-2, PQL-1, and PQL-2) to estimate the parameters and to accept the results if there is similarity among the methods, since this will likely indicate small biases. In addition, they recommend computing five iterations of the bootstrap for the PQL-1 and if the trajectories of the estimates remain flat then the results are likely to be adequate. If there is disagreement among the approximation methods or the bootstrap estimates exhibit variability, then the Bayesian modeling methods should be considered.

Our model considered and estimated the random parameters by treatment at the plot and trees within a plot levels. We did not consider time varying covariates as random effects because our interest was on assessing the variability among the treatments. There may be some questions about our failure to consider or model the correlation over time within a tree. It is well known that ignoring the correlation among repeated measurements may result in biased estimates of the estimated parameter standard errors and test statistics can be inflated. However, we modeled a binary

response in which the event is non-repeated, hence, there is no theoretical reason to consider the correlation over time for our repeated measurements. For our model, the use of repeated measurements is not a result necessarily of the data but of factoring the likelihood function. Our previously discussed likelihood function is

$$L = \prod_{k=1}^{n} \prod_{j=1}^{n_k} \prod_{i=1}^{n_{jk}+1} \pi_{\mathrm{ijk}}^{\delta_{\mathrm{ijk}}} \left(1 - \pi_{\mathrm{ijk}}\right)^{1-\delta_{ijk}}$$

This likelihood function was developed using the conditional probabilities, i.e., the probability of mortality in the i^{th} interval is conditional on mortality not occurring during the previous intervals. For example, if the j^{th} tree on the k^{th} plot succumbs during the second interval we can factor the likelihood function for this tree as $(1-\pi_{1\mathrm{jk}})\ \pi_{2\mathrm{jk}}$. This means that the probability of the mortality occurring during the second interval is simply the probability that it survives the first interval and dies during the second interval. Therefore, each of these terms for this tree may be treated as though it came from a distinct independent observation (Allison 1995).

Our example demonstrated the importance of considering the heterogeneity that may occur at different levels. Most treatments have significant survival variability at the plot and trees within plot levels. In addition, including the sources of heterogeneity in the model generally increases the precision of the mortality prediction. The differences among the random parameters by treatment can be used to assess the impact of the treatment on survival. For example, it is known that fertilization and herbicide usually accelerate the development of the stand and the impact of these treatments on stand development and survival can be obtained using our model. For example, we can determine at what age the treatments obtain the mean *BA/ha*, which are: C = age 8-9, F = age 6-7, H = age 6-7, and HF = age 5-6 (Figure 4.7). Hence, the HF treatment obtains the

mean *BA/ha* first and it occurs at about 5 years 7 months. Comparing the other treatments to the HF at this age, we can determine roughly how many standard deviations from the mean the other treatments are at this time. Calculating the standard deviations for these treatments and then obtaining the corresponding relative risks, we find that the C, F, and H treatments result in a 52.95, 41.96, and 22.46 percent decrease in the hazard relative to the HF treatment, holding all other variables constant. Corresponding mean *BA/ha* for the C, F, and H treatments are 71.4, 51.7, and 24.4 percent less than the *BA/ha* of the HF treatment (13.1 m^2) (Figure 4.7). In addition, the HF gain in H_t, *dbh*, and D_q are reflected when it achieves the mean *BA/ha*. For D_q, the C, F, and H treatments are 48.9, 33.5, and 13.4 percent less than the HF treatment, which is 10.1 cm. For *dbh*, the C, F, and H treatments are 103.4, 71.6, and 28.6 percent less than the average *dbh* of 10.0 cm for the HF treatment. For H_t, the C, F, and H treatments are 39.9, 26.5, and 9.7 percent less than the HF treatment, which is 6.7 m. It is important to note that when the HF treatment achieved the mean *BA/ha*, its survival was substantially greater than the C and F treatments. Hence, looking at one attribute in isolation is difficult for making inferences about the overall effect on a treatment since survival is dependent upon many factors. However, our random parameters results suggest that the F, H, and HF treatments will reduce the plot level survival variability relative to the C treatment. Moreover, the estimated random parameters at the tree level reveal that the survival variability for the C, F, and H treatments are not significantly different. Estimated random parameters for the HF treatment reveal that it has substantially more survival variability at the tree level. Estimated random parameters are larger for the F, H, and HF treatments at the tree level, and the converse is true for the C treatment. This higher

survival variability at the tree level for the F, H, and HF treatments is likely due to the inherent variability when applying the treatment.

Conclusion

Our model development demonstrates that the CLL link function is the natural choice for permanent plot binary response data because it is derived directly from the likelihood function that accounts for interval censoring. Moreover, the parameters can be interpreted as for a Cox proportional hazards model. In addition, it was demonstrated that our model easily relaxes the assumption of proportional hazards at the cut points and can include random effects at the different levels.

Multilevel models have become increasingly common in forestry, which is likely to continue as better software becomes available for estimation of large complex data sets. It is common for multilevel binary response models to use first and second order MQL and PQL to estimate the parameters. It has been suggested that MQL-1 tends to under estimate the variance components and that PQL-2 is the preferred method. However, in many instances, especially for large complex data sets such as our, the only method that may converge is MQL-1. Moreover, it is usually better allow for a multilevel structure and use MQL-1 than to ignore the multilevel structure. Additional studies are necessary to assess the effect of the different parameter estimation techniques on the fixed and random effects.

It was demonstrated that using the plot and tree level random effects generally results in more precise predictions. However, if our goal is to obtain the marginal probabilities, then these can be easily computed using Monte Carlo integration. Hence,

the multilevel model is empirically and theoretically correct for obtaining the marginal probabilities, but the converse is not true for the PA model. If fixed effects hypothesis testing is the study focus, then the SS and PA approaches will result in similar inferences. However, the PA approach will not provide any information about the heterogeneity that may exist at the different levels. If predictions are our primary purpose, then the SS approach is usually preferred (Ten Have and Uttal 1991). Individual tree forestry survival models naturally focus on prediction; therefore, the SS approach is the more natural method.

Literature Cited

Agresti, A. 1990. Categorical data analysis. John Wiley & Sons, Inc. New York. 558 p.

Allison, P.D. 1982. Discrete time methods for the analysis of event histories. in Sociological Methodology. ed. S. Leinhardt, San Francisco, CA: Jossey-Bass 61-98.

Allison, P.D. 1995. Survival analysis using the SAS® system: A practical guide. SAS Institute Inc. Cary, NC. 292 p.

Avila, O.B., and H.E. Burkhart. 1992. Modeling survival of loblolly pine trees in thinned and unthinned plantations. Can. J. For. Res. 22:1878-1882.

Bailey, R.L., and J.L. Clutter. 1974. Base-age invariant polymorphic site curves. For. Sci. 20:155-159.

Barbosa, M.F., and H. Goldstein. 2001. Discrete response multilevel models for repeated measures; an application to voting intentions data. In review.

Biggeri, L., M. Bini, and L. Grilli. 2001. The transition from university to work: a multilevel approach to the analysis of the time to obtain the first job. J. R. Statist. Soc. A 293-305.

Bryk, A.S., and S.W. Raudenbush. 1992. Hierarchical linear models: applications and data analysis methods. SAGE Publications, Inc. London.

Buford, M.A., and W.L. Hafley. 1985. Modeling the probability of individual tree mortality. For. Sci. 31:331-341.

Collet, D. 1994. Modelling survival data in medical research. Chapman & Hall. London. 347 p.

Cox, D.R. 1972. The analysis of multivariate binary data. Appl. Statist. 21:113-120.

Cox, D.R., and D. Oakes. 1984. Analysis of survival data. Chapman and Hall, New York 201 p.

Clutter, J.L. 1961. The development of compatible analytic models for growth and yield of Loblolly pine. Ph.D. dissertation. Duke University.

Dobbertin, M., and G.S. Biging. 1998. Using the non-parametric classifier CART to model forest tree mortality. For. Sci. 44(4):507-516.

Fang, Z., and R.L. Bailey. 1999. A simultaneous system of linear and nonlinear mixed effects models for forest growth and yield prediction. Ph.D. dissertation. The University of Georgia.

Fang, Z., and R.L. Bailey. 2001. Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. For. Sci. 47(3):287-300.

Gibbons, R.D., and D. Hedeker. 1997. Random effects probit and logistic regression models for three-level data. Biometrics. 53:1527-1537.

Goldstein, H. 1991. Nonlinear multilevel models, with an application to discrete response data. Biometrika 78: 45-51.

Goldstein, H. 1995. Multilevel statistical models. Ed. 2. Halstead Press, New York. 178 p.

Goldstein, H., and M. Rasbash. 1996. Improved approximations for multilevel models with binary responses. J. R. Statist. Soc. A 159(3): 505-513.

Greogoire, T.G., O. Schabenberger, and J.P. Barret. 1995. Linear modeling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. Can. J. For. Res. 25:137-156.

Guan, B.T., and G.Z. Gertner. 1991. Modeling red pine tree survival with an artificial neural network. For. Sci. 37(5): 1429-1440.

Hall, D.B., and R.L. Bailey. 2001. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. For. Sci. 47(3):311-321.

Hamilton, D.A. 1974. Event probabilities estimated by regression. USDA For. Serv. Intermt. Res. Stn. Res. Pap. INT-152. 18 pp.

Hamilton, D.A., 1986. A logistic model of mortality in thinned and unthinned mixed conifer stands of northern Idaho. For. Sci., 32:989-1000.

Hamilton, D.A. and B.M. Edwards. 1976. Modeling the probability of individual tree mortality. USDA For. Serv. Intermt. Res. Stn. Res. Pap. INT-185. 22 pp.

Hedeker, D., O. Siddiqui, and F.B. Hu. 2000. Random-effects regression analysis of correlated grouped-time survival data. Statistical Methods in Medical Research. 9:161-179.

Hox, J.J. 1995. Applied multilevel analysis. TT-Publikaties, Amsterdam. 114 p.

Huebschmann, M.M., L.R. Gering, T.B. Lynch, O. Bitoki, and P.A. Murphy. 2000. An individual-tree growth and yield prediction system for uneven-aged shortleaf pine stands. SJAF. 24(2): 112-120.

Kaplan, E.L., and P. Meier. 1958. Nonparametric estimation from incomplete observations. JASA. 53:457-481.

Kobe, R.K., and K.D. Coates. 1997. Models of sapling mortality as a function of growth to characterize interspecific variation in shade tolerance of eight tree species of northwestern British Columbia. Can. J. For. Res. 27:227-236.

Krumland, B., J. Dye, and L.C. Wensel. 1977. Individual tree mortality models for the north coast region of California. University of California, Co-op Redwood Yield Research Project, Berkeley, California. Research Note 6. 7 pp.

Lappi, J., and R.L. Bailey. 1988. A height prediction model with random stand and tree parameters: An alternative to traditional site index methods. For. Sci. 38(2):409-429.

Lawless, J.F. 1982. Statistical models and methods for lifetime data. Wiley, New York. 580 p.

Lindsey, J.K. 1999. Models for repeated measurements. Oxford University Press. Oxford. 515 p.

Monserud, R.A. 1976. Simulation of forest tree mortality. For. Sci. 22:438-444.

Monserud, R.A., and H. Sterba. 1999. Modeling individual tree mortality for Austrian forest species. For. Ecol. Manage. 113:109-123.

McCullagh, P. 1980. Regression models for ordinal data. JRSS Series B 42:109-142.

Neuhaus, J.M., J.D. Kalbfleisch, and W.W. Hauck. 1991. A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data. International Statistical Review. 59(1):25-35.

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

Preisler, H.K., and G.W. Slaughter. 1997. A stochastic model for tree survival in stands affected by annosum root disease. For. Sci. 43(1):78-85.

Rasbash, J., W. Browne, H. Goldstein, M. Yang, I. Plewis, M. Healy, G. Woodhouse, and D. Draper. 1999. A user's guide to MLwiN. 2nd ed. London: Institute of Education.

Rodriguez, G., and N. Goldman. 1995. An assessment of estimation procedures for multilevel models with binary responses. J. R. Statist. Soc. A: 158(1):73-89.

Rodriguez, G., and N. Goldman. 2001. Improved estimation procedures for multilevel models with binary response: a case study. J. R. Statist. Soc. A 339-355.

Shen, G., J.A. Moore, and C.R. Hatch. 2001. The effect of nitrogen fertilization, rock type, and habitat type on individual tree mortality. For. Sci. 47(2): 203-213.

Staebler, G.R., 1953. Mortality estimation in fully stocked stands of young growth Douglas-fir. USDA For. Serv. Pac Northwest Forest and Range Exp. Stn. Res. Pap. 4. 8 p.

Tasissa, G., and H.E. Burkhart. 1998. An application of mixed effects analysis to modeling thinning effects on stem profile of loblolly pine. For. Ecolog. Manag. 103(2): 119-127.

Teck, R.M., and D.E. Hilt. 1990. Individual-tree probability of survival model for the Northeastern United States. USDA For. Serv. Northeastern For. Exp. Stn. Res. Pap. NE-642. 10 pp.

Ten Have, T.R., and D.H. Uttal. 1994. Subject-specific and population-averaged continuation ratio logit models for multiple discrete time survival profiles. Appl. Statist. 43(2):371-384.

Volney, W.J.A. 1998. Ten-year tree mortality following a jack pine budworm outbreak in Saskatchewan. Can. J. For. Res. 28(12): 1784-1793.

Vonesh, E.F., and V.M. Chinchilli. 1997. Linear and nonlinear models for the analysis of repeated measurements. Marcel Dekker, New York.

Wykoff, W.R., N.L. Crookston, and A.R. Stage. 1982. User's guide for the stand prognosis model – version 5.0. USDA For. Serv. Region One, Missoula, Montana. For. Serv. Handbook 2409.21 Chapter 300. 113 pp.

Yang, M., and H. Goldstein. 2000. Multilevel models for repeated binary outcomes: attitudes and voting over the electoral cycle. J. R. Statist. Soc. A 163(1): 49-62.

Yao, X., S.J. Titus, and S.E. MacDonald. 2001. A generalized logistic model of individual tree mortality for aspen, white spruce, and lodgepole pine in Alberta mixedwood forests. Can. J. For. Res. 31: 283-291.

Zeger, S.L., K. Liang, and P.S. Albert. 1988. Models for longitudinal data: a generalized estimating equation approach. Biometrics. 44:1049-1060.

Table 4.1. The CAPPS study summary statistics (N = 112365) for *dbh*, tree height (H_t), trees per hectare (*TPH*), basal area per hectare (*BA/ha*), quadratic mean diameter (D_q), and relative spacing (*RS*) across the age range (1-13).

Attribute	Mean	Minimum	Maximum	Std. Error
dbh (cm)	8.45	0.00	33.78	6.49
H_t (m)	6.95	0.03	22.86	4.93
TPH	14.98	5.53	18.78	1.75
BA/ha (m^2)	13.06	0.00	46.41	12.35
D_q (cm)	8.64	0.00	24.08	6.29
RS	0.92	0.12	6.24	1.23

Table 4.2. Estimated fixed parameters and variance components for the multilevel complementary log-log individual tree survival model. Only the first two of the thirteen estimated parameters for treatment by interval interaction are presented.

Parameter	Estimate	Standard Error	p-value
Interval 1	-3.6298	0.3195	<0.0001
Interval 2	-4.9038	0.2692	<0.0001
Interval 3	-5.4412	0.2653	<0.0001
Interval 4	-5.3547	0.2880	<0.0001
Interval 5	-5.2564	0.3082	<0.0001
Interval 6	-4.6412	0.2527	<0.0001
Interval 7	-4.8289	0.2570	<0.0001
Interval 8	-5.0276	0.2763	<0.0001
Interval 9	-5.4391	0.3091	<0.0001
Interval 10	-5.3685	0.3062	<0.0001
Interval 11	-5.2006	0.3419	<0.0001
Interval 12	-5.9219	0.4515	<0.0001
Interval 13	-5.6449	0.4779	<0.0001
Fertilizer*I1	0.2921	0.2084	0.1610
Fertilizer*I2	0.1813	0.2416	0.4530
Herbicide*I1	-1.4053	0.2198	<0.0001
Herbicide*I2	-2.3563	0.3981	<0.0001
H&F*I1	-1.1958	0.2712	<0.0001
H&F*I2	-3.2856	0.6465	<0.0001
BA/ha	2.6893	0.1601	<0.0001
TPH	-1.0020	0.0558	<0.0001
dbh	-2.5903	0.1514	<0.0001
H(plot)	-1.7569	0.5005	0.0004
RS	-0.1505	0.0768	0.0500
dbh (*plot*)	1.3448	0.5954	0.0239
Height	-0.5418	0.1736	<0.0001
Variance Components			
Level 2			
Control	0.5219	0.1294	<0.0001
Fertilizer	0.4588	0.0985	<0.0001
Herbicide	0.4930	0.2686	0.0664
H&F	1.1454	0.2083	<0.0001
Level 3			
Control	0.5872	0.1586	0.0002
Fertilizer	0.3903	0.1317	0.0030
Herbicide	0.0321	0.0513	0.5315
H&F	0.4063	0.1487	0.0063

Note: Intervals 1, 2, 3, … , 13 correspond to ages 1-2, 2-3, 3-4, … , 13-14.

Table 4.3. The predicted stand progression for plots 40 and 59 using the multilevel complementary log-log survival model with random treatment effects for the trees within a plot and plots (Random). The fixed effect model (Fixed) uses the same covariates as the random effect model but with no random effects.

	Plot 40			Plot 59		
Age	Actual	Fixed	Random	Actual	Fixed	Random
2	73	74.4	70.4	76	79.0	69.5
3	72	71.5	66.8	73	78.3	64.9
4	70	70.8	65.7	70	77.7	61.3
5	69	70.4	65.0	68	77.2	58.8
6	69	70.2	64.5	64	76.4	55.9
7	68	69.7	63.8	64	75.4	52.9
8	67	69.3	63.2	60	74.0	49.5
9	66	68.9	62.6	54	71.5	44.3
10	65	68.6	62.2	52	69.6	40.8
11	65	68.4	61.9	50	67.3	37.2
12	63	68.1	61.5	45	65.5	33.8
13	61	67.2	60.8			

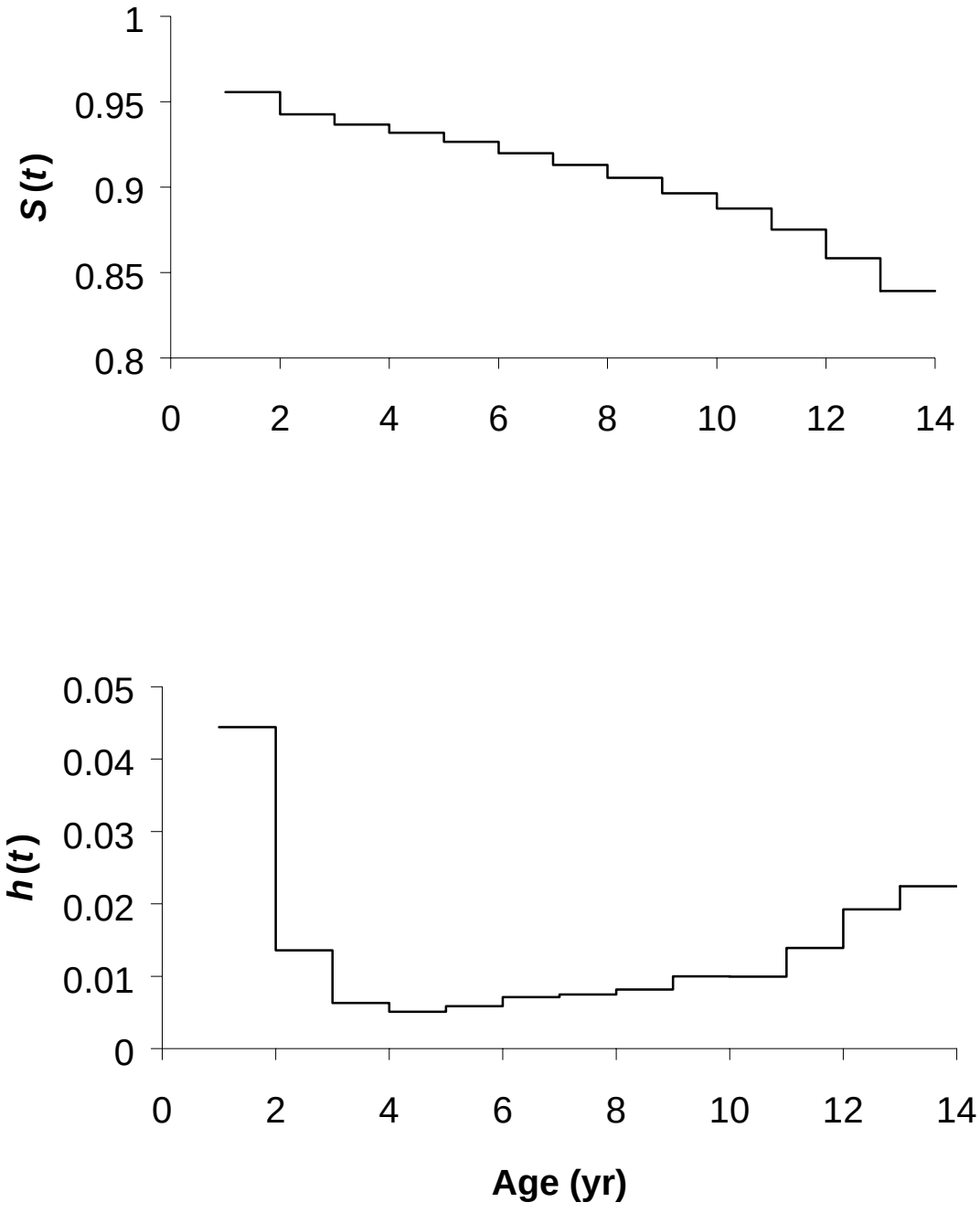


Figure 4.1. The Kaplan-Meier product limit survival ($S(t)$) and hazard functions ($h(t)$) estimates for the CAPPS study.

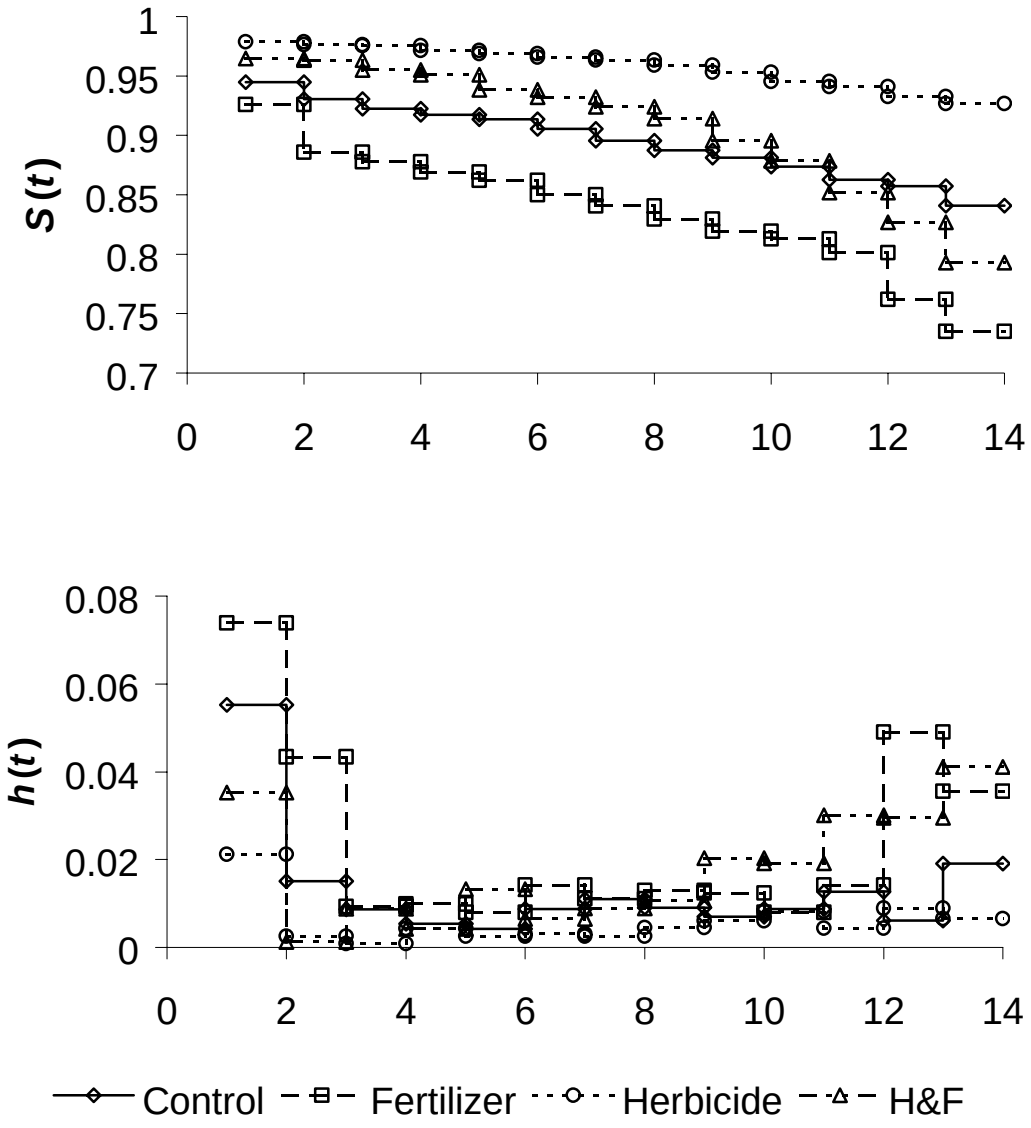


Figure 4.2. CAPPS study Kaplan-Meier (1958) product limit estimates for the survival ($S(t)$) and hazard functions ($h(t)$) by treatment.

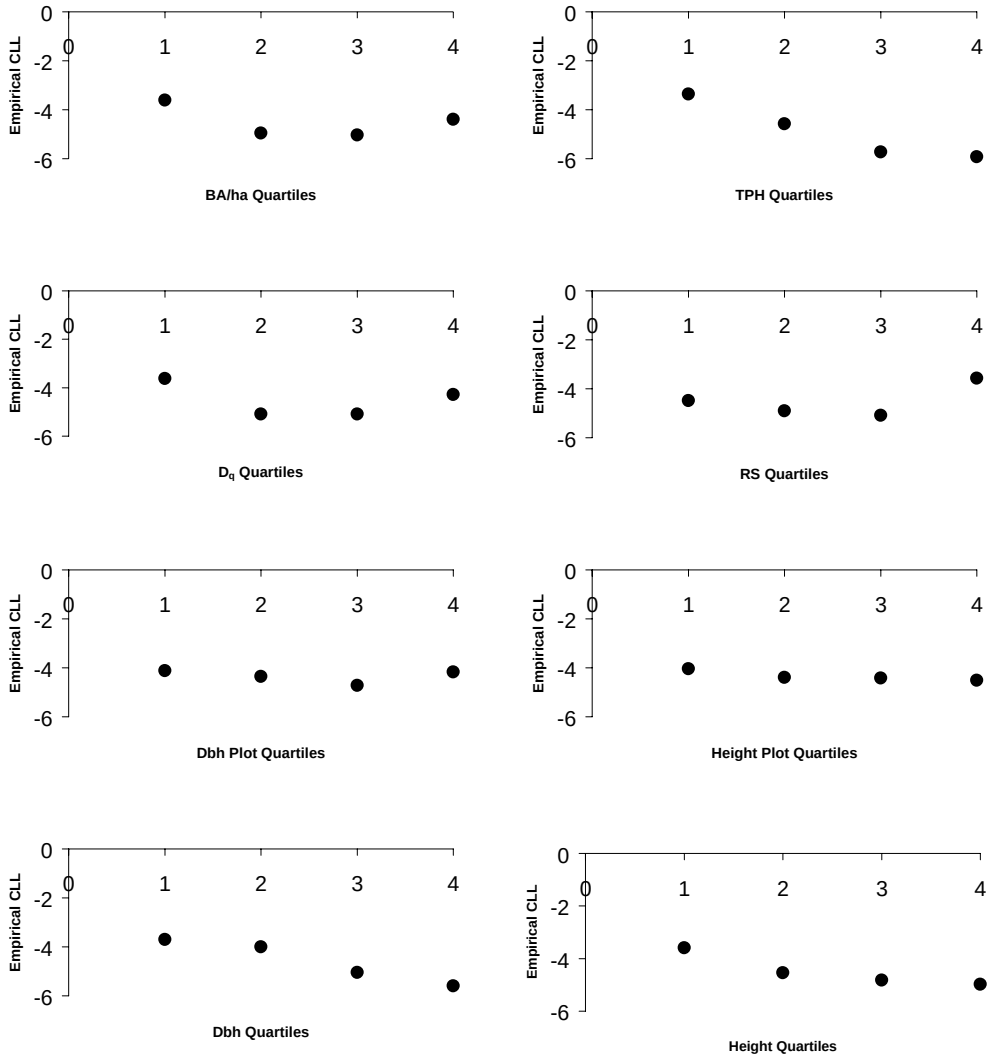


Figure 4.3. The computed complementary log-log for the considered covariates. The empirical CLL were computed by computing the 25th, 50th, and 75th quartiles, and then grouping these quartiles into the 0-25 (1), 25-50 (2), 50-75 (3), and 75-100 (4) quantiles.

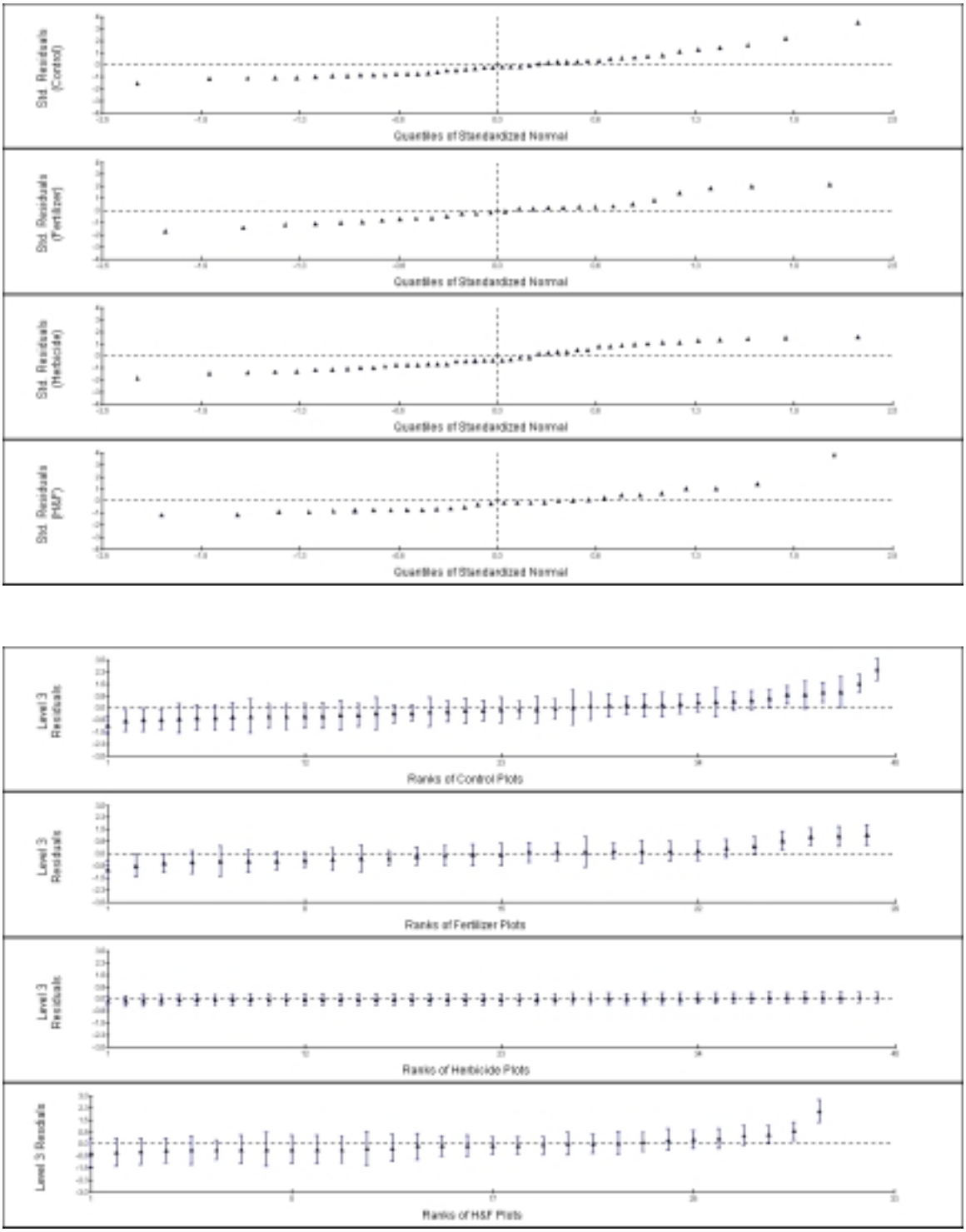
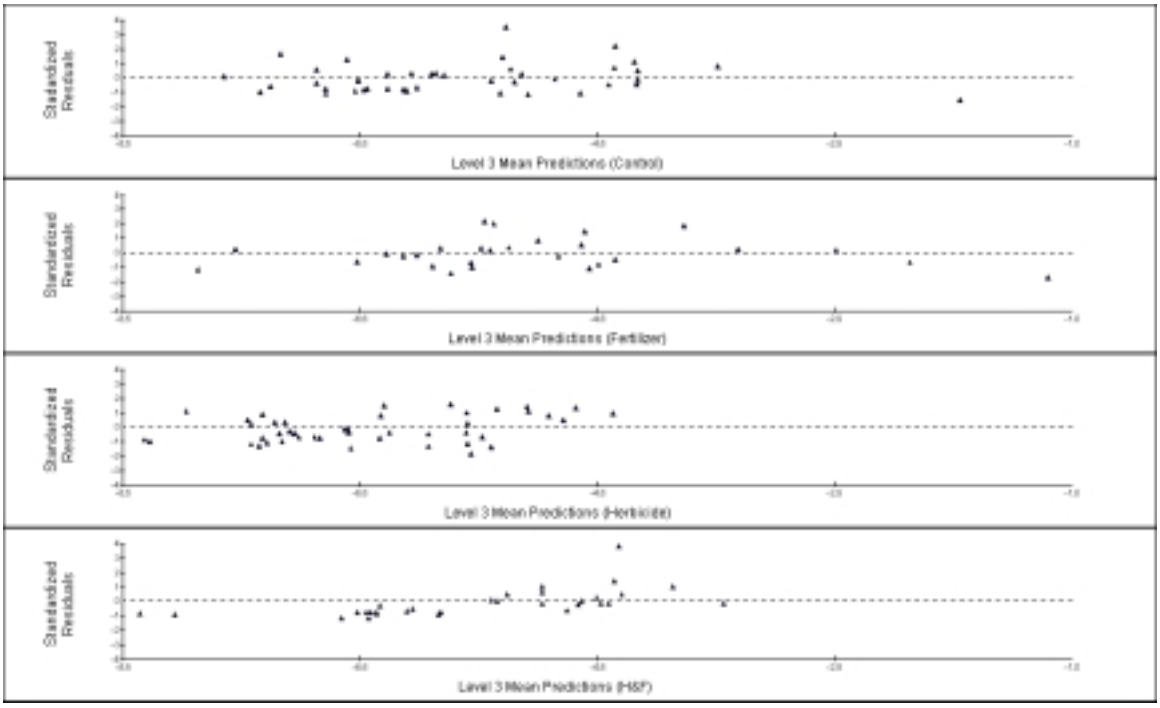


Figure 4.4. Standardized residuals by plot rank and quantiles of the standardized normal distribution for the plot level random effects.



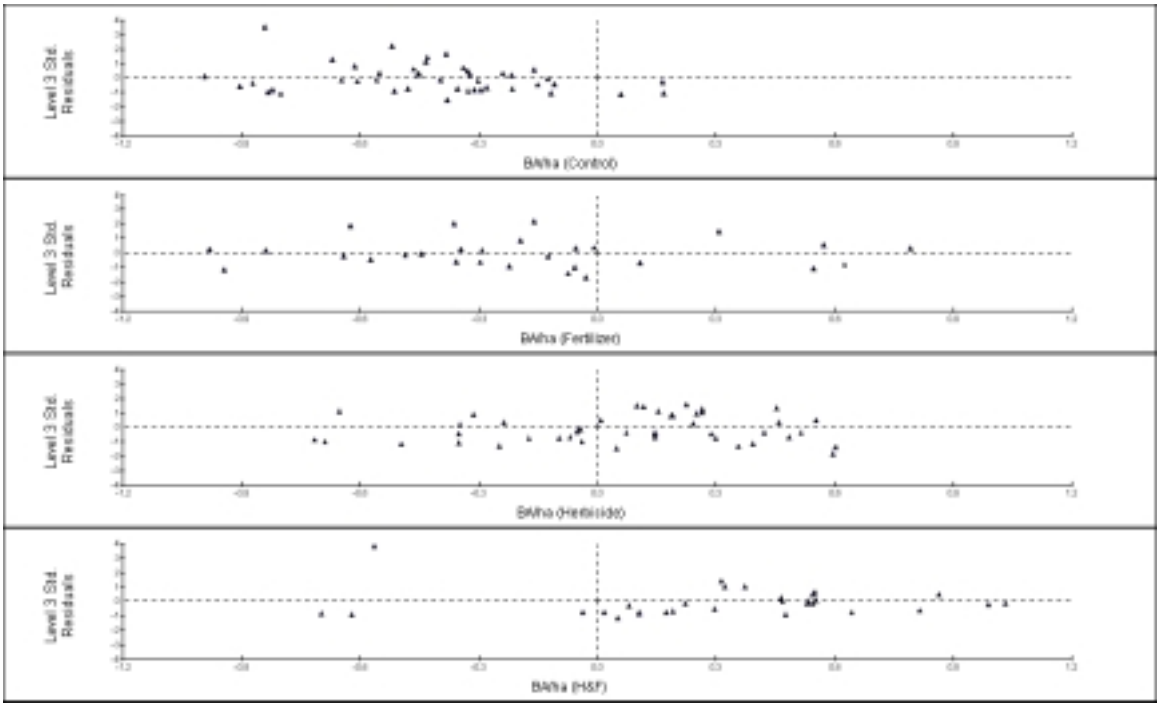


Figure 4.5. Plot level standardized residuals for the mean plot level predictions and average *BA/ha*.

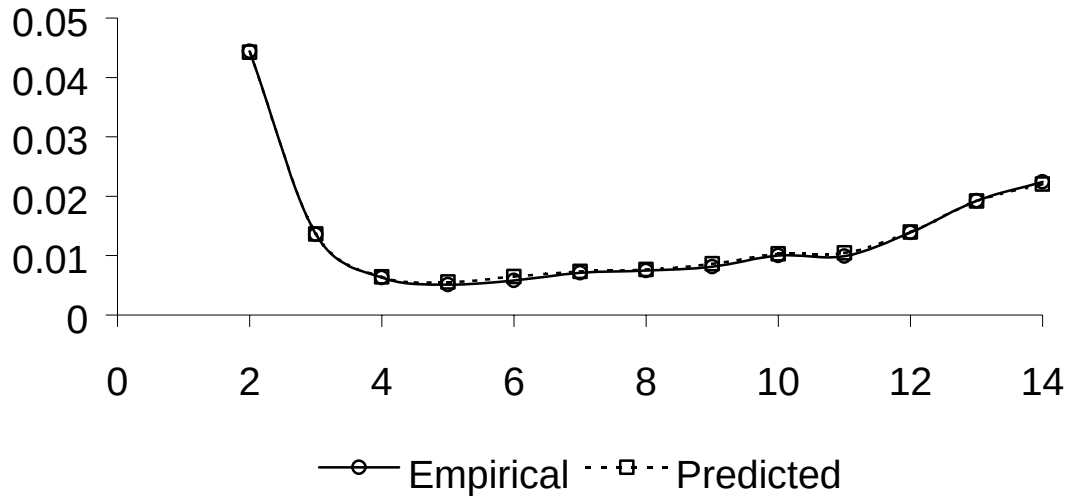


Figure 4.6. Predicted and empirical hazards by age for the CAPPS study data.

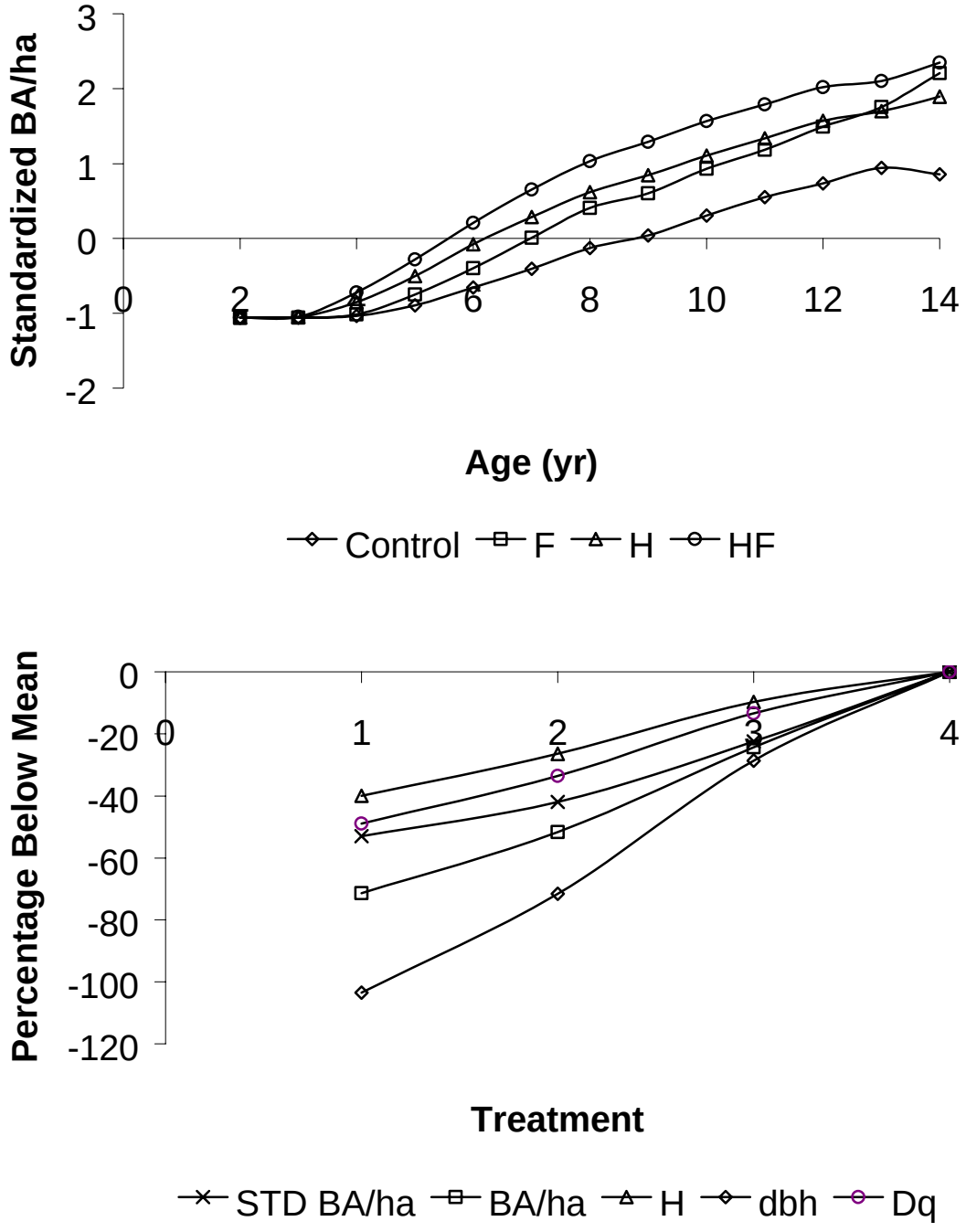


Figure 4.7. The average standardized *BA/ha* by treatment (C = 1 = control, F = 2 = fertilizer, H = 3 = herbicide, and HF = 4 = herbicide and fertilizer) and age. Mean *BA/ha*, H_t, *dbh*, and D_q at the age when the HF treatment achieves the mean *BA/ha* (approximately 5 years 7 months).

CHAPTER 5

A MULTILEVEL INDIVIDUAL TREE MORTALITY MODEL DEVELOPED FROM PERMANENT PLOTS IN LOBLOLLY PINE PLANTATIONS

Introduction

Our focus here is to develop a parsimonious multilevel individual tree logit mortality model for use in projecting stand tables. Since this Chapter is an extension of the previous Chapter, see Chapter 4 for further details on survival analysis and forestry survival background, and model development. However, inevitably there will be some overlap in the details of this Chapter with Chapter 4.

Modeling individual tree mortality or alternatively survival, hereafter the modeling of survival or mortality is used interchangeably, began in earnest in the 1970's with a study by Hamilton (1974) in which he used the logistic. Alternatives to the logistic for modeling individual tree survival have been developed but have not been demonstrated to be more effective. Therefore, the flexibility, ease, and interpretability of the logistic will likely enable its popularity to be maintained for modeling individual tree survival. Individual tree survival models typically focus on natural mortality, which occurs due to factors such as competition for light, nutrients, and water. Here we use permanent plot plantation data to model natural individual tree survival.

Plantation forestry studies frequently use repeated measurement data from permanent plots in which the data structure generally form three levels or groupings. Measurement occasions are nested within a tree (level 1), trees are nested within a plot

(level 2), and plots (level 3). Here we adopt the nomenclature used by Goldstein (1995) in which the error term is recognized as the lowest level. Forestry individual tree mortality models have typically been fitted assuming the trees within a plot are independent and it was acknowledged by Hamilton (1974) that because there is clustering, i.e., trees are sampled by plots, that the independence assumption is violated. Recent forestry studies have demonstrated the importance of accounting for the multiple levels of heterogeneity using a subject specific (SS) model and have distinguished themselves from population averaged (PA) models. Here, we develop a multilevel individual tree logit mortality model that accounts for the variation among and within plots.

Survival Model Formulation

Our survival model assumes the data are from a permanent plot plantation study in which measurement occasions are nested within a tree (level 1) and trees are nested within a plot (level 2). We began the modeling process by defining the probability of a tree dying during the i^{th} time interval (t_{i-1}, t_i), $i = 1, 2, \ldots, l$. A tree's mortality noted at time t_i, had an actual death time of t, where $t_{i-1} \le t < t_i$. All trees enter the study at $t_0 = 0$, however the calendar time corresponding to t_0 may vary by plot. All trees are followed to time t_l, the last measurement occasion or the time when the tree succumbs. Let p_{ijk} be the probability of mortality occurring for the j^{th} tree on plot k during the i^{th} time interval, i.e., $p_{ijk} = \Pr\left(t_{i-1} \le T_{jk} < t_i\right)$. Then T_{jk} is a random variable associated with the j^{th} tree in plot k. Conditional probability of mortality in the i^{th} time interval given that the death occurs after t_{i-1} is given by $\pi_{ijk} = \Pr\left(t_{i-1} \le T_{jk} < t_i \mid T_{jk} \ge t_{i-1}\right)$, where $i = 1, 2, \ldots, l+1$. This

means that $p_{ijk} = (1-\pi_{1jk})(1-\pi_{2jk})\ldots(1-\pi_{i-1,jk})\pi_{ijk}$, where $i = 2, 3, \ldots, l+1$ and $p_{1jk} = \pi_{1jk}$. The complementary probabilities are associated with a tree not dying during an interval and p_{1jk} is the probability it was planted and died during the first interval. In addition, $\delta_i = 1$if the j^{th} tree on plot k dies during the i^{th} interval, otherwise $\delta_i = 0$. The likelihood function for the j^{th} tree on plot k can be expressed as

$$L = \prod_{i=1}^{l} \pi_{ijk}^{\delta} (1-\pi_{ijk})^{1-\delta_i} .$$

It is common to analyze discrete time survival data by modeling the hazard function as a linear function of covariates using a transformation g(•) (e.g., Hedeker et al. 2000, Biggeri et al. 2001). The linear function typically includes an intercept, time (t) and a suite of explanatory variables $\boldsymbol{x}$, i.e.,

$$g(\pi_{ijk}) = \sum_{s=0}^{r} \alpha_s t^s + \sum_{u=0}^{v} \beta_u x_{ijk} .$$

Where α and β are parameters. Commonly, the link function chosen is either the logit (Cox 1972) or complementary log-log (Prentice and Gloeckler 1978). However, since the logistic regression model is the most common individual tree survival model, only the logit link is considered for our example. We can express the logit link function for the i^{th} interval of the j^{th} tree on the k^{th} plot as

$$\ln\left[\frac{\pi_{ijk}}{1-\pi_{ijk}}\right] = \eta_{ijk} = \sum_{s=0}^{r} \alpha_s t^s + \sum_{u=0}^{v} \beta_u x_{ijk} .$$

Here t is time, x is a suite of tree and/or plot level covariates, and α, β are parameters. Conceptually, this model can be easily extended to include random effects for the different levels. In summary, defining π_{ijk} as the probability that the j^{th} tree on the k^{th} plot

succumbs during the i^{th} interval, the three-level mixed effects mortality model that uses the logit link function and allows for time varying covariates is

$$\left.\begin{aligned} & y_{ijk} \mid b_k, b_{jk} \sim Binomial\left(1, \pi_{ijk}\right) \\ & g\left(\pi_{ijk} \mid x_{ijk}; \alpha, \beta\right) = \sum_{s=0}^{r} \alpha t^{s}_{\ i} + \sum_{l} \beta_l x_{ijkl} + \boldsymbol{b}_{jk} \boldsymbol{z}_{(1)ijk} + \boldsymbol{b}_{k} \boldsymbol{z}_{(2)ijk} \\ & \boldsymbol{b}_k \sim N\left(0, \Sigma_{(3)}\right) \\ & \boldsymbol{b}_{jk} \sim N\left(0, \Sigma_{(2)}\right) \end{aligned}\right\} \qquad (5.1)$$

Where $\boldsymbol{x}_{ijk}$ is the p x n_{ijk} covariate matrix associated with the fixed effects $\boldsymbol{\beta}$, $\boldsymbol{z}_{(m)ijk}$ are the q_m x n_{ijk} covariate matrices associated with the random effects, and t^s is the s x n_{ijk} time covariate matrix associated with the fixed effects $\boldsymbol{\alpha}$. The random effects at the different levels are assumed independent and the level 3 (plot) and level 2 (trees within a plot) random effects b_k and b_{jk} are distributed as N(0, ,$\Sigma_{(3)}$) and N(0,$\Sigma_{(2)}$), respectively. There are n_k trees within each plot and N_k total observations within plot k (the sum of all the measurement occasions for all trees within a plot). The random effects design matrix generally is a subset of the fixed effects design matrix and hence, the random effects represent deviations from the means.

Our previous likelihood function is valid for the multilevel model but the response is now conditional on the random effects. Define $g^{-1}{}_k(\pi_{ijk})$ as the inverse link function binary response vector pattern for plot k having n_k trees observed at n_{jk} measurement occasions. Assuming independence of the response vector conditional on the random effects, the likelihood for the k^{th} plot can be expressed as (Gibbons and Hedeker 1997)

$$L(\delta_k \mid \boldsymbol{b}_k, \boldsymbol{b}_{jk}, \beta) = \prod_{j=1}^{n_k} \prod_{i=1}^{n_{jk}} \left[g_k^{-1}(\pi_{ijk})^{-1} \right]^{\delta_{ijk}} \left[1 - g_k^{-1}(\pi_{ijk}) \right]^{1-\delta_{ijk}}$$

The marginal distribution of δ_k can be obtained from the likelihood by integrating over the distribution of the random effects ($f(\boldsymbol{\theta})$), i.e.,

$$L(\delta_k) = \int_{\theta} L(\delta_k \mid \boldsymbol{b}_k, \boldsymbol{b}_{jk}, \boldsymbol{\beta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

Conceptually, the multilevel survival model is straightforward. However, because of parameter estimation challenges for binary response multilevel models, use of multilevel models to analyze binary response variables is relatively recent, e.g., Goldstein 1991, Hedeker et al. 2001, Biggeri et al. 2001, Rodriguez and Goldman 2001. Most of the parameter estimation challenges are related to the response having a binomial distribution rather than the more tractable normal distribution assumed in classical linear and nonlinear models.

A common multilevel parameter estimation technique for a binary response variable, and the one we adopt, is the marginal quasilikelihood (MQL-1) method. This method is motivated by using the linear form of the multilevel model. MQL-1 approximates π by using a 1st order Taylor series expansion of the function around fixed effects $\beta = \beta_0$ and random effects $b = 0$,

i.e., $g^{-1}(\pi_{ijk,01}) = g^{-1}(\pi_{ijk,0}) + X(\hat{\beta}_{01} - \hat{\beta}_0) g^{-1'}(\pi_{ijk,0})$. Where β_0 and β_{01} are the current and updated estimates for the fixed effects from the iterative generalized least squares (IGLS) or restricted iterative generalized least squares (RIGLS) algorithm, respectively. The MLwiN (Rasbash et al. 2000) software was used to implement this multilevel binary response parameter estimation method.

Example

We used permanent plot data from plantation loblolly pine that were established throughout Georgia. The study called for two complete blocks to be established at each location with each block containing four 0.15 ha treatment plots, which were established at each location using bare-root seedlings planted on a 2.44 m by 2.44 m spacing. A 0.05 ha measurement plot was centered within each of the treatment plots. Four cultural treatments, herbicide, fertilization, herbicide and fertilization, and control were randomly assigned to the blocks at each location. See Chapter 4 (pages 60-61) for further study protocol and data details. Plots have been measured annually beginning at age one and the data structure is: 112,365 total observations for measurement occasions within a tree, 11957 trees within a plot, and 146 plots. Plot survival data are summarized by plot distribution (see Table 3.1, page 36) and by plot and tree attributes (Table 5.1).

Preliminary Analysis

The Kaplan-Meier (1958) product limit estimator for survival and the discrete hazard function were computed to detect trends among and within the cultural treatments (see Figures 4.1 and 4.2 on pages 96 and 97). Results illustrate that the H treatment has the most favorable survival and that the HF treatment crosses and declines relative to the C treatment for survival at about age 11. The discrete hazard functions by treatment reveal a bathtub shaped trend over time (Figure 4.2, page 97) and that the variability is highest for the HF treatment and lowest for the H treatment.

Empirical logits can aid in selecting an appropriate form of the systematic component (Agresti 1990). Empirical logits for trees within a plot can be computed,

however this is uninformative because the trees have a 0-1 binary response. Therefore, we computed the empirical logits by plot. The empirical logit is defined as

$$\text{logit}_E = \log\left(\frac{y_{ik} + 0.5}{n_{ik} - y_{ik} + 0.5}\right)$$

Where y_{ik} is the number of trees that succumbed on plot k and n_{ik} is the total trees for plot k at time i. The empirical logits were grouped and plotted by treatment (Figure 5.1). There is no substantial evidence of heterogeneity over time but there is moderate evidence that the variability of the intercepts differs by treatment.

To assess the intercept and slope variability by treatments, the logit model of $\eta_i = \alpha + \beta t_i$ was fitted by treatment (Figure 5.2). There is evidence that the intercepts can be stratified by F and C, and H and HF treatments. The HF treatment has a positive slope, whereas the other treatments have negative slopes. This suggests there is an acceleration of mortality over time for the HF treatment relative to the other treatments. Hence, the HF treatment is further along in stand development. The H logit eventually crosses both the fertilizer and control treatments, which suggests an acceleration of mortality over time versus these treatments. The estimated parameters and their respective 95 percent confidence intervals reveal disjoint intercepts and slopes (Figure 5.2), which suggests a need, at a minimum, for an adjustment in the fixed effects. The F and C have similar estimated slopes and intercepts. In addition, H and HF treatments have similar slopes. The variability among plots by treatment is visually negligible, but this does not imply equality of variability or that a random effect for all treatments rather than by treatment is adequate.

Model Fitting Procedure

Preliminary evidence does not strongly suggest equality of variability among plots for survival by treatment but does suggest that the intercepts and slopes are not equal for all treatments. It is reasonable to assume the H, F, and HF treatments will accelerate the growth rates relative to the C treatment, hereafter equivalently referred to as the baseline. This growth acceleration impacts survival, which may be explained using treatment effects or growth related covariates such as *dbh* and *BA/ha*. There is a tradeoff in the multilevel modeling process (Jones 1990), i.e., inclusion of growth related covariates could negate the necessity of some random effects.

Numerous recommendations for developing multilevel models abound, e.g., Pinheiro and Bates (2000) recommend, in the absence of prior information about the random effects variance-covariance matrix, to allow all effects to be fixed and random if convergence can be achieved. Other authors espouse fitting a variance component model and then adding covariates as fixed effects (Goldstein 1995). A variance component model would usually be an adequate starting point when fitting a multilevel forestry survival model. However, there is no overwhelming empirical evidence to suggest that the variability, both among and within plots, is equal for the cultural treatments. Moreover, there is evidence that the intercepts and slopes differ by treatment. Depending upon assumptions, there are numerous possible baseline models. For example, a variance component model could be assumed as our baseline model. This assumes that the slopes and intercept are adequately modeled, for the trees within a plot and the plots, using a typical response. Then the slopes and intercepts for the trees within a plot and plots are allowed to deviate from the typical response through the random effects, which assumes

the intercept and slope variations for the treatments can be adequately modeled using single parameters and a covariance. There is not an overwhelming motive from our preliminary analysis for choosing one baseline model. Therefore, to be somewhat pragmatic, we chose an intercept variance component model as suggested by Hox (1995) with age as a fixed covariate by treatment. The logit model intercepts were treated as random and fixed, for trees within a plot and plots, by treatment. In addition, the slopes (age) were treated as fixed effects by treatment. The baseline model is

$$\text{logit}(\pi_{ijk}) = \alpha_0 + b_0^{(2)} + b_0^{(3)} + \alpha_1 A_{ijk}$$

Where

$$\begin{aligned}
\alpha_0 &= \alpha_{00} z_C + \alpha_{01} z_F + \alpha_{02} z_H + \alpha_{03} z_{HF} \\
\alpha_1 &= \alpha_{10} A_{C(ijk)} + \alpha_{11} A_{F(ijk)} + \alpha_{12} A_{H(ijk)} + \alpha_{13} A_{HF(ijk)} \\
b_0^{(2)} &= b_{20} z_C + b_{21} z_F + b_{22} z_H + b_{23} z_{HF} \\
b_0^{(3)} &= b_{30} z_C + b_{31} z_F + b_{32} z_H + b_{33} z_{HF}
\end{aligned} \tag{5.1a}$$

$$\sum_{(2)} = \sum_{(3)} = \begin{bmatrix} \sigma_C^2 & 0 & 0 & 0 \\ 0 & \sigma_F^2 & 0 & 0 \\ 0 & 0 & \sigma_H^2 & 0 \\ 0 & 0 & 0 & \sigma_{HF}^2 \end{bmatrix}$$

Where z_l's are indicator variables for the l treatment, A is age, and α, b are parameters, all other terms are as previously defined. The off-diagonal elements are assumed to equal zero, i.e., it is assumed the plot-specific treatment effects are independent.

A likelihood ratio test is typically used to determine the necessity of fixed and random effects for a mixed effects model that assumes the lowest level variance is normally distributed (Pinheiro and Bates 2000). However, for multilevel binary response models that use quasi-likelihood to estimate the parameters, the likelihood ratio test is a crude approximation and the preference for testing the fixed and random effects is the

Wald chi-square test (Goldstein 1995). Therefore, the Wald chi-square test was used to test for fixed and random parameter significance using $\alpha = 0.05$. For example, the H treatment plot level variance component was dropped because it is not significantly different from zero. Inclusion of time dependent covariates negated the necessity of some baseline covariates. All considered covariates were included and removed from the model using a stepwise procedure. In addition, it was desirable to determine if the variance components could be combined for different treatments. A Wald type test was used to construct contrasts among the variance components, i.e., to test for differences in the variability among the treatments (see pages 71-74 for contrast details). Variance component contrasts revealed that the C, F, and HF treatment variance components could be pooled for the tree level, and C and F variance components could be pooled at the plot level. Hence, our final fitted model is

$$\text{logit}[\pi_{ijk}] = \alpha_0 + \alpha_1 A_k + \alpha_2 A_k^{\,2} + \alpha_3 A_k^{\,3} + \beta_0 BA_k + \beta_1 TPH_k + \beta_2 D_{q\,k} + \beta_3 dbh_{jk}$$

Where

$$\alpha_0 = \alpha_{00} + \alpha_{01} z_H + b_{jk} z_H + b_{jk} z_{C/F/HF} + b_k z_{C/F} + b_k z_{HF} \qquad (5.1b)$$

$$\sum_{(2)} = \begin{pmatrix} \sigma^2_H & 0 \\ 0 & \sigma^2_{C/F/HF} \end{pmatrix}$$

$$\sum_{(3)} = \begin{pmatrix} \sigma^2_{C/F} & 0 \\ 0 & \sigma^2_{HF} \end{pmatrix}$$

Where π_{ijk} is the estimated probability of mortality for the j^{th} tree on plot k at time i, z_H = 1 if H, otherwise $z_H = 0$, similarly for the other treatment indicator variables, A is age, BA is basal are per hectare (m), TPH = trees per hectare divided by 100, D_q is the quadratic

mean diameter (cm), *dbh* is the individual tree diameter at breast height (cm), and α, β are parameters. This model allows the intercepts to vary systematically by cultural treatment. In addition, the intercepts vary randomly for trees within a plot for the H treatment and pooled C/F/HF treatments, and by plot for the pooled C/F and HF treatments. Estimated parameters and their respective standard errors and p-values are presented in Table 5.2. This model has a large number of random effects (b_{jk} and b_k). For plots (b_k), there are 146 random effects, and trees within a plot (b_{jk}) have 11,956 estimated random effects, which are not presented here for brevity. However, these estimates are easily obtainable and can be used for mortality probability predictions.

Level 2 random parameters reveal that the variability for the pooled C/F/HF treatments is over 1.7 times larger than the H treatment for trees within a plot. In addition, the variability for the H treatment at level 3 was not significantly different from zero. The pooled C/F treatment variability is over 2.35 times larger than the HF treatment for level 3. These random parameters indicate that the H treatment in addition to having higher survival also has the lowest variability for both among and within plots. The fixed parameters are all behaving logically with respect to their signs for our data. Note that the parameters for D_q and *BA/ha* are positive, this means that as these attributes increase, the probability of mortality increases. This is consistent with trends in the limiting density relationships. The *TPH* parameter is negative, which implies that as *TPH* increases mortality will decrease. This seems counterintuitive, but the plots were all planted at the same density and the plots with higher *TPH* as time increases are those plots with higher survival, i.e., plots experiencing lower mortality. Therefore, for these data, it is reasonable that the sign for *TPH* parameter is negative.

Model Evaluation and Predictions

Model adequacy was determined using the residuals and predictions. Furthermore, model (5.1b) was compared with the same model fitted without the inclusion of random effects. Level 3 residuals versus ranks and their respective Q-Q plots are presented in Figure 5.3. The ideal Q-Q plot would be a straight line. The level 3 Q-Q plots reveal no evidence of non-normality, similarity for the level 2 residuals. Level 3 standardized residuals were plotted by predictions and attributes. Only the predictions and *dbh* plots are presented (Figure 5.4). There is no evidence of heteroscedasticity or any abnormally large standardized residuals. To compare the SS model to the marginal proportions, raw estimated mortality by age, we should integrate over the random effects. A more pragmatic approach is to obtain the SS averages by age and compare these estimates to the marginal probabilities (Figures 5.5 and 5.6). If the marginal probabilities were computed, i.e., integrate over the random effects; the predicted probabilities would be expected to more closely reflect the raw proportions. These model evaluation results reveal that the model fits adequately and there is no evidence of systematic bias. To further assess the model performance, the final model was fitted without the inclusion of random effects.

We compared the model performance with the same model fitted without any random effects, hereafter referred to as the subject-specific (SS) and population-averaged (PA) models. The SS and PA models were compared using several common individual tree evaluation criteria. One method is to determine a threshold for classifying trees as alive or dead based upon the predicted probabilities. According to Monserud and Sterba (1999), it is more important if a tree is correctly classified as dead or alive rather than if a

dead tree is predicted to be close to one and zero. In addition, it was inferred that the most logical classification threshold is to use the raw probabilities, i.e., for our case it would be the population proportion predicted to die by age class, which was used for the PA model. However, this was deemed an unacceptable threshold for the SS model because we are modeling the subjects and not the marginal average. Therefore, since our goal was to use an appropriate threshold, we used the predicted typical response by age for each plot. We conducted a sensitivity analysis for the thresholds by multiplying the thresholds by 1.25-2.00 (0.25 increments) to determine the effect of the chosen threshold (Table 5.3). An alternative to the discrete classification of trees is to treat the mortality as a continuous event and to sum over the probabilities to estimate the total predicted mortality (Stage 1973). Hence, we computed the total trees predicted to die for the SS and PA models and compared these with the actual number that died. The third criterion was to compare the ratio of the predicted probabilities for the dead and live trees of the SS model to the PA model. We expect the SS model to generally predict closer to zero for live trees and closer to one for dead trees relative to the PA model. Lastly, the mean predicted probabilities for the live and dead trees were compared for the SS and PA models.

The results for the continuous time mortality removal reveal that the SS model predicts 1667 and the PA model predicts 1741 dead trees, whereas the actual number that died is 1477 trees. The SS model reduces the number predicted to die by more than four percent relative to the PA model. The mean predicted probability for the live trees is 0.0122 and 0.0177 for the SS and PA models, respectively. For the dead trees, the mean predicted probabilities are 0.2141 and 0.1180 for the SS and PA models, respectively.

For the live trees, the SS model predicts almost 63 percent of them closer to 0 relative to the PA model. For the dead trees, the SS model predicts almost 98 percent of them closer to one relative to the PA model. Results for the percent correctly classified reveal that the SS model is substantially superior for classifying the dead trees. However, the PA performs more adequately for the live trees until the multiplier is 1.75, then the SS model performs more favorably for both classification categories. This is only meant as a guide to illustrate the live and dead tree classification tradeoff and further research is needed to determine a good classification threshold for the SS model. However, these classification results reveal the importance of predicting accurate probabilities for the trees that actually lived and died. This is clear because the larger the separation of the predicted probabilities for the dead and live trees, the greater the allowance in selecting a good threshold. The overall performance of SS model is a substantial improvement over the PA model and in addition, individual tree mortality probabilities can be predicted at all levels for these plots.

Individual tree mortality predictions can be obtained at all levels when using the multilevel mortality model (5.1b) for these data. Moreover, individual tree mortality can be predicted for a new plot at various levels, depending upon the information available, using (5.1b). Obviously if future individual tree mortality predictions are desired for a study plot, then the plot and tree random effects should be included in the prediction process. However, a typical forestry situation is prediction of mortality probabilities for trees within a plot that are not included in the original study. Hence, given a new plot, mortality probability predictions for individual trees can be obtained using the typical response, inclusion of tree and/or plot level random effects dependent upon available

information. Details for estimating plot and tree level random effects for a new plot are given in Chapter 4 (see pages 76-83).

Discussion and Conclusion

A multilevel forestry survival model is attractive because it allows us to increase the precision of the predicted probabilities by modeling the different levels of variation. Often in forestry, we want to predict a future attribute but the prior plot information is ignored, i.e., we make predictions based upon the current measurements. However, as demonstrated in this and other multilevel forestry modeling studies (e.g., Fang and Bailey 2001, Hall and Bailey 2001), there is generally an increase in the precision of the predictions when using prior information. Our SS individual tree mortality model demonstrated a clear gain in precision for the predicted probabilities relative to the PA model.

It has become common to use the mean probability as a threshold in PA models for classifying trees as dead or alive, which has been shown to be a reliable threshold. Conversely, for a SS model there are no previous studies to determine an adequate threshold. It has been inferred by Monserud and Sterba (1999) that it is not relevant for an individual tree survival model, which is to be used deterministically in an individual tree simulator how close the probabilities are to zero and one. For instance, if the trees that live and died are predicted to have probabilities of 0.49 and 0.51, respectively, we could simply use the threshold of 0.50 to correctly classify the trees. However, this is an unrealistic situation and the actual predicted probabilities will overlap, i.e., some trees that died will be in the larger diameter classes, which should realistically have a lower

probability of death than the smaller diameter classes. A higher degree of separation for the dead and alive trees should allow the modeler more flexibility in selecting a threshold that will increase the precision of the classifications.

Estimated parameters for PA and SS models can vary considerably, which is demonstrated by the parameter estimates for model (5.1b) fitted without random effects, which are quite different from model (5.1b) when fitted including the random effects. Although as noted by Ten Have and Uttal (1991), inferences about parameter estimates are usually similar, i.e., p-values are similar; the estimated parameters for our SS and PA models are dissimilar. This means that we could be over or under estimating the effect of a parameter when using the PA model (Rodriguez and Goldman 2001).

Subject-specific models have several advantages over PA models; one is that the population average predictions can be obtained by integrating over the random effects using numerical integration or Monte Carlo integration techniques. Conversely, the PA model is incapable of providing us with any information about the SS model. In addition, there are advantages in explicitly modeling the manner in which subjects are grouped (Goldstein 1995) and this shortcoming of PA models ignoring correlated individual tree groups was acknowledged by Hamilton (1974). Modeling the clusters allows for statistically efficient estimates of the regression coefficients. The use of grouping information provides the correct standard errors, confidence intervals, and tests of significance. Lastly, it allows measuring covariates at any grouping level and then obtaining the corresponding predictions. Furthermore, as revealed by our parameter estimates for the SS and PA models, if the relationship between the response and the

covariates is nonlinear, e.g., logistic regression, then ignoring the groups can result in large biases in the parameter estimates (Rodriguez and Goldman 2001).

Questions may arise about our failure to consider or model the correlation over time within a tree. Ignoring the correlation among repeated measurements can result in biased estimates of the estimated parameter standard errors and test statistics can be inflated. However, our use of repeated measurements is not a result necessarily of the data but of factoring the likelihood function, i.e., the probability of mortality during the i^{th} interval is conditional on mortality not occurring during the previous intervals. Therefore, each conditional probability for a tree may be treated as though it came from a distinct independent observation (Allison 1995). This study demonstrates the viability of using a multilevel model for modeling individual tree mortality. A study shortcoming is the lack of long-term data and further research is needed for multilevel individual tree survival models that are based on more encompassing data sets.

Literature Cited

Agresti, A. 1990. Categorical data analysis. John Wiley & Sons, Inc. New York. 558 p.

Allison, P.D. 1995. Survival analysis using the SAS® system: A practical guide. SAS Institute Inc. Cary, NC. 292 p.

Biggeri, L., M. Bini, and L. Grilli. 2001. The transition from university to work: a multilevel approach to the analysis of the time to obtain the first job. J. R. Statist. Soc. A 293-305.

Cox, D.R. 1972. The analysis of multivariate binary data. Appl. Statist. 21:113-120.

Fang, Z., and R.L. Bailey. 2001. Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. For. Sci. 47(3):287-300.

Gibbons, R.D., and D. Hedeker. 1997. Random effects probit and logistic regression models for three-level data. Biometrics. 53:1527-1537.

Goldstein, H. 1995. Multilevel statistical models. Ed. 2. Halstead Press, New York. 178 p.

Hall, D.B., and R.L. Bailey. 2001. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. For. Sci. 47(3):311-321.

Hamilton, D.A. 1974. Event probabilities estimated by regression. USDA For. Serv. Intermt. Res. Stn. Res. Pap. INT-152. 18 pp.

Hedeker, D., O. Siddiqui, and F.B. Hu. 2000. Random-effects regression analysis of correlated grouped-time survival data. Statistical Methods in Medical Research. 9:161-179.

Hox, J.J. 1995. Applied multilevel analysis. TT-Publikaties, Amsterdam. 114 p.

Jones, R.H. 1990. Serial correlation or random subject effects? Comm. Statist. Simulation Comput. 19: 1105-1123.

Kaplan, E.L., and P. Meier. 1958. Nonparametric estimation from incomplete observations. JASA. 53:457-481.

Monserud, R.A., and H. Sterba. 1999. Modeling individual tree mortality for Austrian forest species. For. Ecol. Manage. 113:109-123.

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

Prentice, R.L., and L.A. Gloeckler. 1978. Regression analysis of grouped survival data with application to breast cancer data. Biometrics. 34:57-67.

Rasbash, J., W. Browne, H. Goldstein, M. Yang, I. Plewis, M. Healy, G. Woodhouse, and D. Draper. 1999. A user's guide to MLwiN. 2nd ed. London: Institute of Education

Rodriguez, G., and N. Goldman. 2001. Improved estimation procedures for multilevel models with binary response: a case study. J. R. Statist. Soc. A 339-355.

Stage, A.R. 1973. PROGNOSIS model for stand development. USDA For. Ser. Res. Pap. INT-137.

Ten Have, T.R., and D.H. Uttal. 1994. Subject-specific and population-averaged continuation ratio logit models for multiple discrete time survival profiles. Appl. Statist. 43(2):371-384.

Table 5.1. The CAPPS study summary statistics (N = 112365) for *dbh*, trees per hectare (*TPH*), basal area per hectare (*BA/ha*), and quadratic mean diameter (D_q) across the age range (1-13).

Attribute	Mean	Minimum	Maximum	Std. Error
dbh	8.45	0.00	33.78	6.49
TPH	14.98	5.53	18.78	1.75
BA/ha	13.06	0.00	46.41	12.35
D_q	8.64	0.00	24.08	6.29

Table 5.2. The multilevel individual tree survival model estimated fixed parameters and variance components.

Parameter	Estimate	Standard Error	Chi-square	p-value
Constant	9.03777	0.43343	434.80	<0.0001
Herb	-0.38090	0.09956	14.64	0.00013
Age	-1.63021	0.12107	181.30	<0.0001
Age2	0.13748	0.02003	47.13	<0.0001
Age3	-0.00476	0.00090	28.04	<0.0001
Dbh	-0.51324	0.01481	1200.82	<0.0001
Dq	0.31824	0.02595	150.40	<0.0001
TPH	-0.62487	0.02347	708.56	<0.0001
BA/ha	0.16509	0.00960	295.54	<0.0001
Variance Components				
Level 2				
Herb	0.66794	0.30345	4.85	0.0276
C/F/HF	1.15271	0.10904	111.76	<0.0001
Level 3				
C/F	0.33316	0.08148	16.72	<0.0001
HF	0.14087	0.07204	3.82	0.0506

Table 5.3. The proportion of trees classified correctly using the population-averaged (PA) and subject-specific (SS) models. The PA model uses the mean raw proportions by age class as its threshold and the SS model uses the average typical response predicted by plot and age class as its threshold. The multiplier is a factor for adjusting the threshold, i.e., the 1.25 multiplier will adjust the threshold by multiplying by 1.25.

Multiplier	PA Model			SS Model		
	Alive	Dead	Overall	Alive	Dead	Overall
1.00	0.8234	0.7495	0.8225	0.7709	0.9370	0.7731
1.25	0.8662	0.6682	0.8636	0.8508	0.9188	0.8517
1.50	0.8941	0.6121	0.8904	0.8886	0.8395	0.8879
1.75	0.9098	0.5856	0.9055	0.9138	0.7204	0.9112
2.00	0.9220	0.5559	0.9171	0.9319	0.6053	0.9276

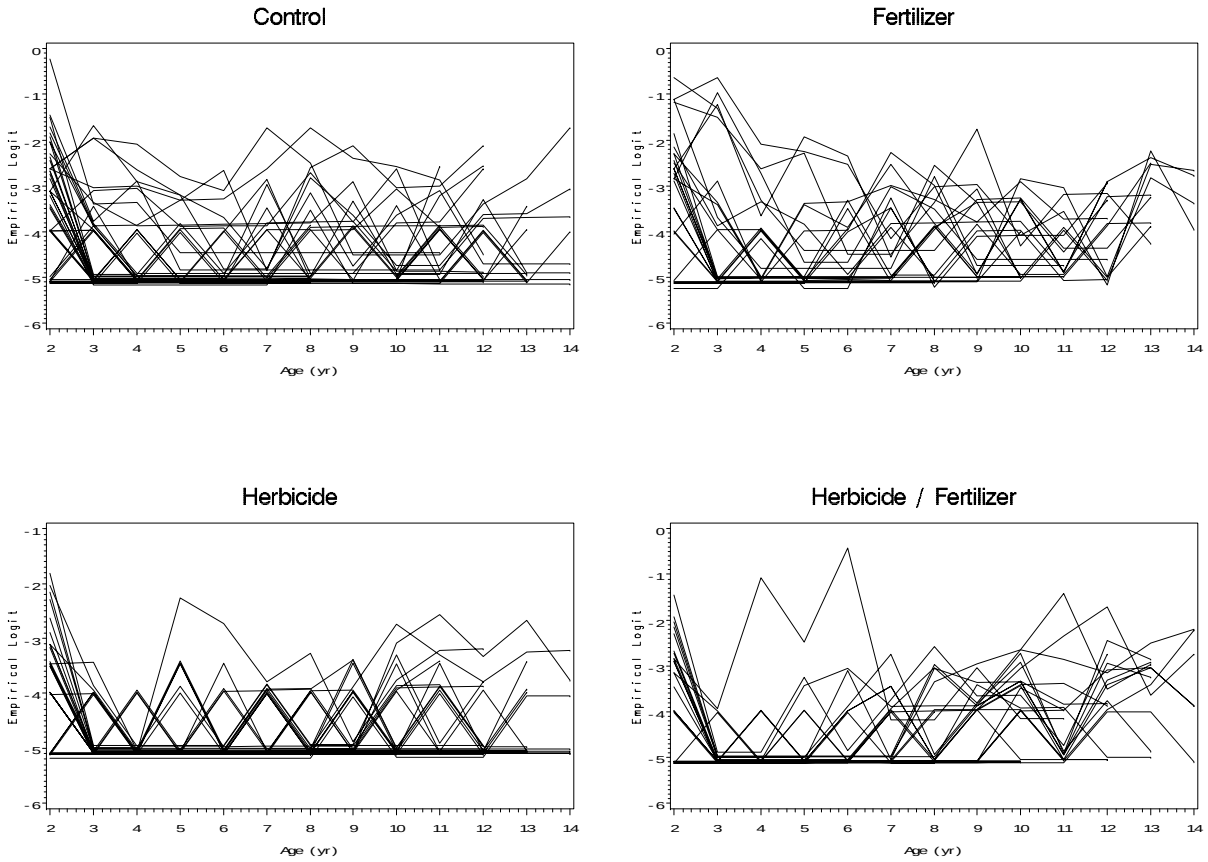


Figure 5.1. The empirical logit for each plot by treatment and age.

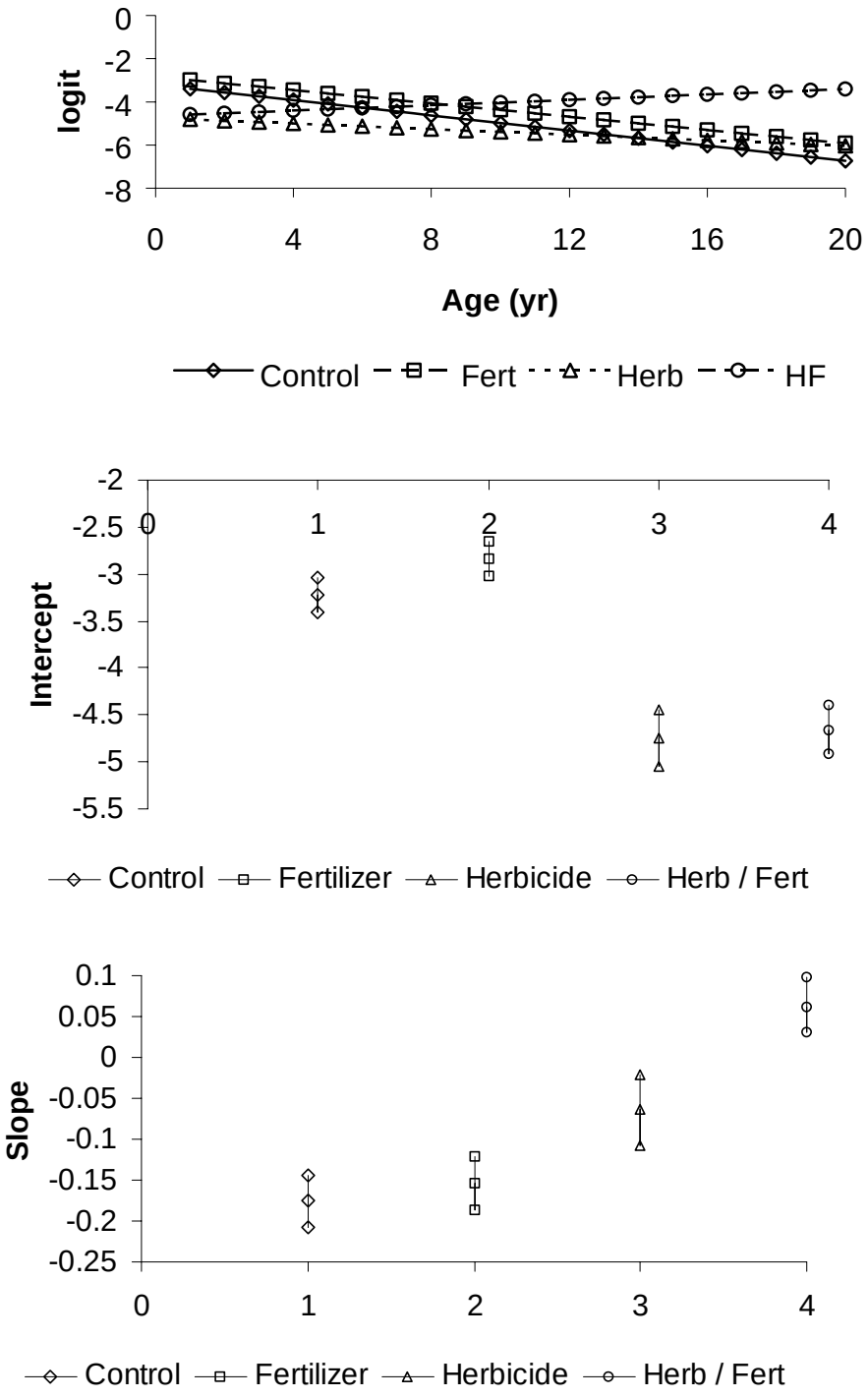


Figure 5.2. The fitted logit model by treatment and 95% confidence intervals for the intercept and slopes by treatment.

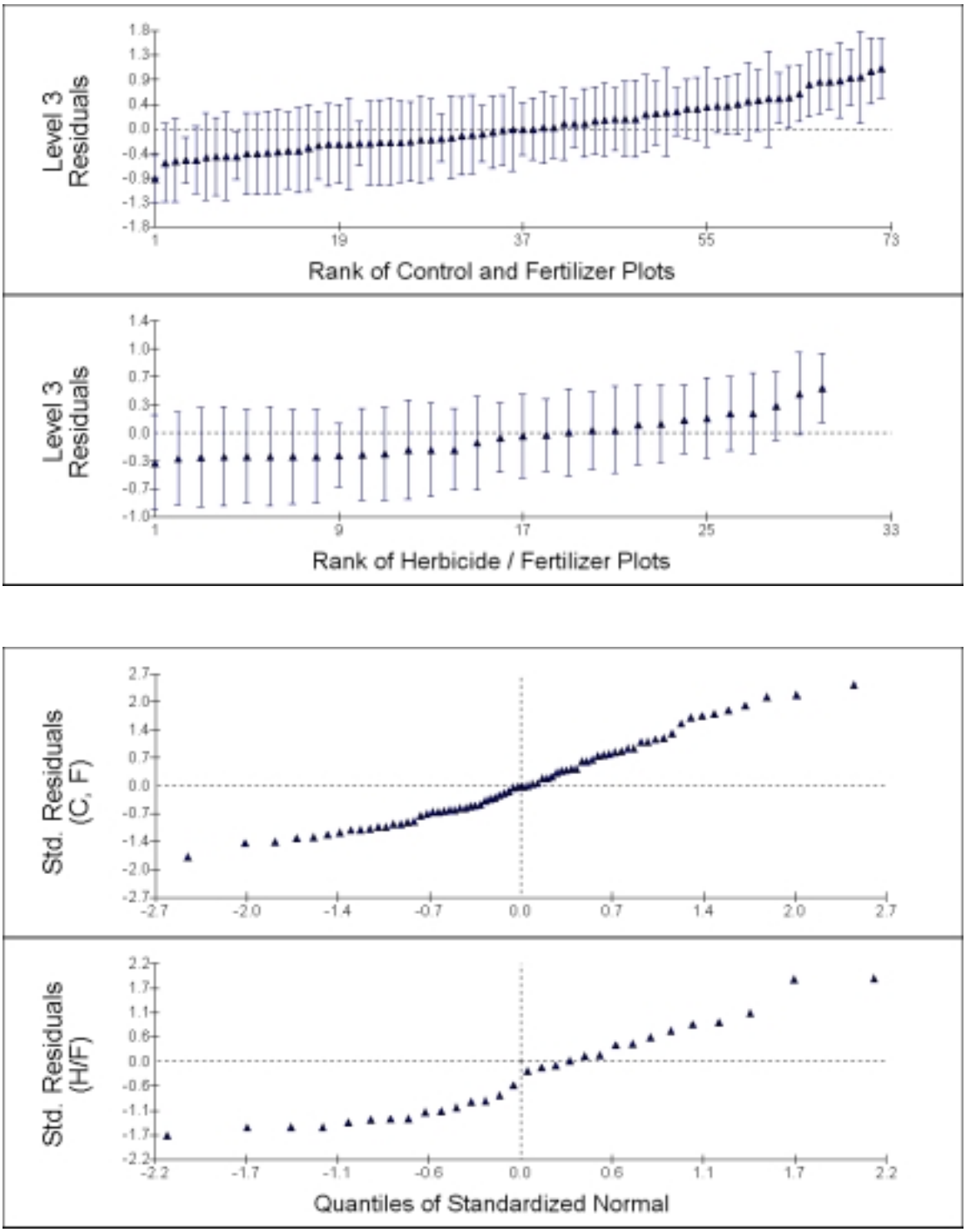


Figure 5.3. Level 3 residuals by rank and standardized residuals by quantiles of the standardized normal.

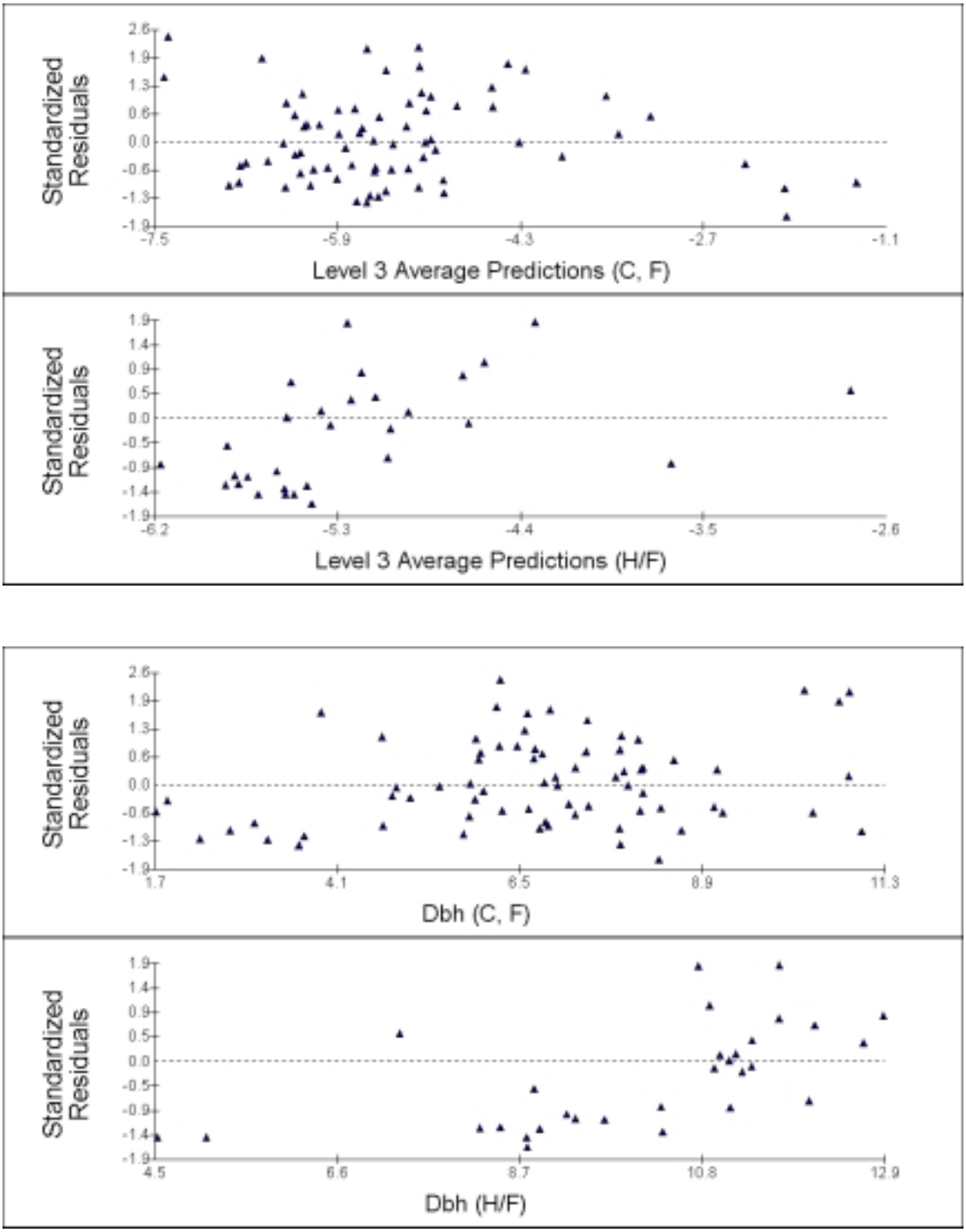


Figure 5.4. Level 3 standardized residuals by average level 3 predicted values and average *dbh*. The random effects correspond to the control and fertilizer (C, F), and herbicide/fertilizer (H/F) treatments.

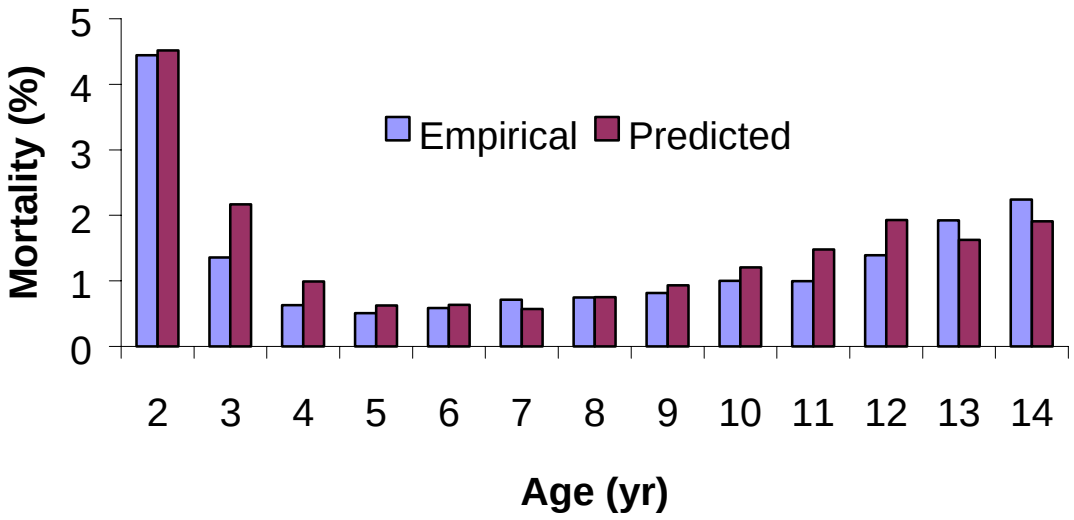


Figure 5.5. The CAPPS study empirical and multilevel logit model predicted probabilities for mortality by age.

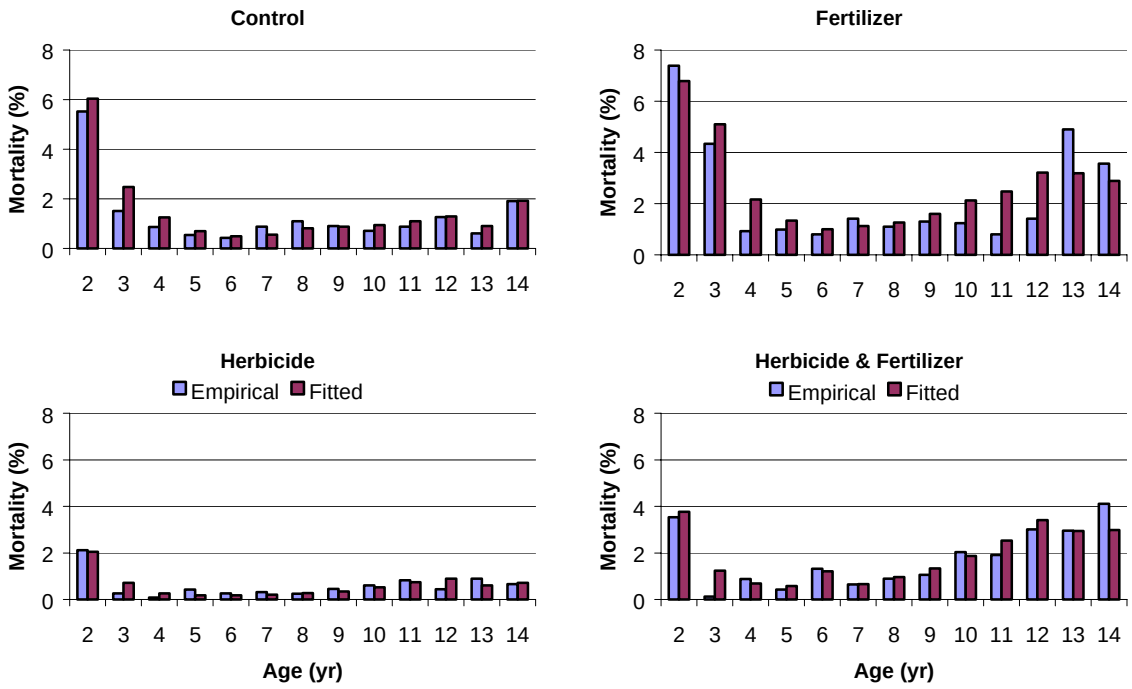


Figure 5.6. The CAPPS study empirical and multilevel logit model predicted probabilities by age and treatment.

CHAPTER 6

COMPARISON OF THREE STAND TABLE PROJECTION METHODS USING DATA FROM PERMANENT PLOT LOBLOLLY PINE PLANTATIONS

Introduction

Stand table projection methods predict the future number of stems by diameter class and the corresponding stock tables can be derived. Accurate and precise stand and stock tables are required for making sound forest management decisions (Matney and Sullivan 1982). Stand tables usually consist of trees per unit area and the corresponding average height by *dbh* class for a given species. Stock tables contain the information from the stand table and in addition, give the volume and/or weight on a tree and unit area by *dbh* class. Stand table projection methods require the availability of an initial stand table or tree list. If only stand level information is available, it is common to estimate initial stand tables using stand level attributes and the Weibull distribution. Parameters for the Weibull distribution can be estimated using a technique such as the parameter recovery method (Bailey et al. 1981), which relates stand level attributes to the Weibull distribution parameters. Here we focus on projecting stand tables for even-aged loblolly pine plantations given an initial stand table or individual tree list.

Clutter and Allison (1974) put forth the methodology that is commonly used in stand table projection. Their methodology is not the classic stand table projection method

based on increment cores as described by Meyer (1952) and discussed in Mensuration books (e.g., Avery and Burkhart 1994), but is a generalized stand table projection compatible with whole stand estimates of basal area and trees per unit area. Their stand table projection algorithm, assuming the availability of an initial stand table or tree list and future estimates of per unit area of basal area and trees, consists of three main components. These components are mortality allocation, growth prediction, and a constraint to ensure the basal area per unit implied by the stand table is consistent with observed or predicted basal area per unit area. They assume, as many subsequent studies have (e.g., Pienaar and Rheney 1993, Knowe et al. 1997), that the unconditional probability of a tree of a *dbh* class dying is inversely proportional to it relative size. Relative size is usually defined as the basal area per tree of *dbh* class *i* divided by the plot mean basal area per tree. The conditional probabilities of a tree belonging to *dbh* class *i* given a tree dies are obtained using Bayes' formula by assuming that the number of surviving trees is inversely proportional to relative size and directly proportional to the number of trees in diameter class *i*. Once they allocate mortality, the trees are grown such that the projected stand table is constrained to equal the predicted basal area per unit area. Clutter and Allison's (1974) methodology was subsequently modified by Clutter and Jones (1980), and Pienaar and Harrison (1988). Pienaar and Harrison simplified the algorithm by developing a parameter free method for estimating the *dbh* class mortality probabilities.

Pienaar and Harrison's (1988) stand table projection method, which is versatile and easy to implement, has subsequently been used in numerous studies (e.g., Pienaar and Rheney 1993, McTague and Stansfield 1994). In addition, some studies have

modified the Pienaar and Harrison (1988) method by defining relative size as the ratio of individual tree diameter to quadratic mean diameter (e.g., Knowe 1994, Knowe and Stein 1995, Knowe and Hibbs 1996, and Knowe et al. 1997). Nepal and Somers (1992) proposed a stand table projection algorithm that simultaneously adjusted for stand mortality and growth. With their method, rather than develop separate algorithms for mortality and growth, these attributes are adjusted simultaneously. However, unlike previous methods, they grow all initial trees, not just the survivors. Cao and Baldwin (1999) subsequently revised the Nepal and Somers (1992) method, primarily by projecting only the surviving trees.

Our purpose is to compare two alternative methods to the Pienaar and Harrison (1988) method for projecting stand tables. The first alternative assumes that the initial and all intermediate stand tables are available for the projection period; hence, the stand table projection algorithm is iterated annually until reaching the projection period. This alternative modifies their method by using a logit model to predict the *dbh* class mortality probabilities. Our second alternative assumes that the initial and intermediate individual tree lists are available. Thus, as for our first alternative, the stand table projection algorithm is an annual process. This method removes trees from the tree list annually using the logit model individual tree predicted mortality probabilities until the number of trees equals the observed number of trees. Then each surviving tree is grown and the number of trees per *dbh* class is computed.

Data

Permanent plot loblolly pine plantation data were obtained from the Consortium for Accelerated Pine Production Studies (CAPPS). Four cultural treatments were randomly assigned to blocks at each location: herbicide, fertilization, herbicide and fertilization, and control. Plots have been measured annually beginning at age one and only plots at least 11 years of age were considered for this study, which results in 101 plots that range in age from 11-14 years (Table 6.1). Further study protocol and data details were discussed in Chapter 4 (pages 60-61).

Methods

Our stand table projection methods require the current stand table or individual tree list, an individual tree *dbh* growth model, future survival and basal area per unit area. For analysis purposes, we use observed survival and basal area per hectare. For all plots, the projection period initiates at age 6, which was chosen because there was adequate separation of the diameter classes. Depending upon plot age the projection period ranges from 5-8 years (Table 6.1). We compare two alternative stand table projection methods with Pienaar and Harrison's (1988) method, hereafter referred to as P&H. All considered stand table projection methods assume that mortality occurs at the beginning of the projection period.

The P&H method assumes that the probability of a tree dying for a given *dbh* class is inversely proportional to its relative size. Relative size is defined as the ratio of per tree diameter class basal area to the average per tree plot basal area. After identifying the mortality by *dbh* class, the *dbh* class midpoints are grown for the projection period

subject to the *BA/ha* constraint that ensures compatibility with the observed basal area per hectare. Once the adjusted *dbh* class midpoints are obtained, they are placed back into 2.5 cm classes by assuming trees are uniformly distributed within a *dbh* class and the class limits extend halfway between adjacent class midpoints. In addition, for the first and last occupied *dbh* classes, their respective lower and upper class limits extend the same distance as their class limit for the adjacent *dbh* class.

Our first alternative modifies the P&H method (MP&H). Rather than assuming that the probability of a tree dying for a given *dbh* class is inversely proportional to its relative size, we compute the *dbh* class mortality probabilities using an individual tree mortality model that was fitted to the data. These *dbh* class mortality probabilities are annual predictions; therefore, we iterate the stand table projection algorithm annually until reaching the projection period. Observed stand tables are used for all intermediate computations. Predicted *dbh* class midpoints are not placed into 2.5 cm classes for the intermediate projection periods; only the final projection period *dbh* class midpoint, are placed back into the 2.5 cm *dbh* classes.

Our second alternative (IND) assumes that we have the individual tree lists; i.e., the initial and ending tree lists, and at all intermediate times for the projection period. Mortality probabilities are predicted annually for surviving trees. Then, annual plot mortality is allocated by removing the trees with the highest predicted mortality probabilities until the number of surviving trees equals the observed number of trees. Surviving individual trees are grown subject to the constraint that the sum of the basal area for all trees within a plot (multiplied by an expansion factor) equals the observed plot basal area per hectare. The stand table projection algorithm is iterated until reaching

the projection period, and then trees are placed into 2.5 cm diameter classes. All stand table methods considered consist of two main components: allocation of mortality and growing the basal area *dbh* class midpoints or individual trees subject to the *BA/ha* constraint.

Survival

Predicted P&H *dbh* class mortality probabilities assume that a *dbh* class that has a basal area less than the plot average basal area will have a higher probability of mortality. This general approach was used by Clutter and Allison (1974), and Clutter and Jones (1980). The P&H method assumes that the probability of mortality (p_i) for the i^{th} *dbh* class relative basal area is inversely proportional to its size, i.e.,

$$p_i = \frac{\left(\frac{\bar{b}}{b_i}\right)}{\sum_{i=1}^{j}\left(\frac{\bar{b}}{b_i}\right)} = \frac{b_i^{-1}}{\sum_{i=1}^{j} b_i^{-1}}. \qquad (6.1)$$

Where $\bar{b}$ is the average per tree basal area for plot *k*, and b_i is the *dbh* class *i* basal area. It is interesting to note that these predicted mortality probabilities (p_i's) only depend upon the number of *dbh* classes for a given plot.

The MP&H and IND methods predict annual mortality probabilities by *dbh* class and individual trees, respectively. These mortality probabilities are predicted using a previously developed multilevel logit mortality model for these data (Chapter 5). This multilevel model has three levels, measurement occasions are nested within a tree (level 1), trees are nested within a plot (level 2), and plots (level 3). The multilevel logit mortality model is:

$$\text{logit}[p_{ijk}] = \alpha_0 + \alpha_1 A_k + \alpha_2 A_k^{\,2} + \alpha_3 A_k^{\,3} + \beta_0 BA_{ik} + \beta_1 TPH_{ik} + \beta_2 D_{q\,ik} + \beta_3 dbh_{ijk}$$

where

$$\alpha_0 = \alpha_{00} + \alpha_{01} z_H + b_{jk} z_H + b_{jk} z_{C/F/HF} + b_k z_{C/F} + b_k z_{HF}$$

$$b_{jk} \sim N(0, \Sigma_{(2)}) \text{ and } b_k \sim N(0, \Sigma_{(3)}) \qquad (6.2)$$

$$\sum_{(2)} = \begin{pmatrix} \sigma^2_H & 0 \\ 0 & \sigma^2_{C/F/HF} \end{pmatrix}$$

$$\sum_{(3)} = \begin{pmatrix} \sigma^2_{C/F} & 0 \\ 0 & \sigma^2_{HF} \end{pmatrix}$$

Where p_{ijk} is the estimated annual probability of mortality for the j^{th} tree or *dbh* class on plot k at time i, $z_H = 1$ if treatment is herbicide, otherwise $z_H = 0$, similarly for the other treatment indicator variables, A is age, BA is basal are per hectare (m), TPH = trees per hectare divided by 100, D_q is the quadratic mean diameter (cm), *dbh* is the individual tree diameter at breast height (cm), α, β are parameters, and b_{jk} and b_k are the random effects. Intercepts vary systematically by cultural treatment and vary for trees within a plot for the H treatment and pooled C/F/HF treatments, and by plot for the pooled C/F and HF treatments. Estimated parameters and their respective standard errors and p-values were presented previously (see Table 5.2, page 126). MP&H *dbh* class mortality probabilities were computed using the *dbh* class midpoint, plot level attributes, and the plot random effect. The IND method also includes the tree level random effect in the prediction of the individual tree mortality probabilities.

For the P&H and MP&H methods, the p_i's are used in Bayes' formula to predict each *dbh* class conditional mortality probability given a tree dies. The p_i's are associated with the probability of a tree belonging to a *dbh* class succumbing. Whereas the total

mortality observed on a plot is allocated to the *dbh* classes using the conditional probabilities. The conditional probability of a dead tree belonging to a particular *dbh* class can be calculated using Bayes' formula (e.g., Clutter and Allison 1974, Clutter and Jones 1980, Pienaar and Harrison 1988), i.e.,

$$P(dbh_i \mid D) = \frac{P(dbh_i \cap D)}{P(D)} = \frac{P(D \mid dbh_i)P(dbh_i)}{\sum_{i=1}^{j} P(D \mid dbh_i)P(dbh_i)}.$$

Where dbh_i is the midpoint diameter for class *i,* and *D* is a tree that has died. Using Baye's formula, the *dbh* class conditional probabilities (π_i) are given by:

$$\pi_i = \left[\frac{\frac{n_{1i}}{\sum_{i=1}^{j} n_{1i}} p_i}{\sum_{i=1}^{j} \frac{n_{1i}}{\sum_{i=1}^{j} n_{1i}} p_i} \right]. \qquad (6.3)$$

Surviving trees per *dbh* class are obtained by multiplying the conditional probabilities by the mortality for plot *k* during the projection period, and subtracting the predicted number of dead trees from the number of trees in the initial stand table for the *dbh* class. Hence, the surviving number of trees for the i^{th} *dbh* class can be expressed as

$$n_{2i} = n_{1i} - \left[\frac{\frac{n_{1i}}{\sum_{i=1}^{j} n_{1i}} p_i}{\sum_{i=1}^{j} \frac{n_{1i}}{\sum_{i=1}^{j} n_{1i}} p_i} \right] \mathrm{M} = n_{1i} - \left[\frac{n_{1i} p_i}{\sum_{i=1}^{j} n_{1i} p_i} \right] \mathrm{M} = n_{1i} - \pi_i \mathrm{M}. \qquad (6.4)$$

Where n_{2i} and n_{1i} are the surviving and initial *TPH* for *dbh* class *i*, and M is the *TPH* mortality. P&H and MP&H methods compute the conditional mortality probabilities (equation 4) using the p_i's from equations (6.1) and (6.2), respectively.

The IND method allocates the mortality to individual trees based on the p_i's. Annual mortality using the IND method was allocated by removing the trees with the highest predicted mortality probabilities until the number of trees remaining equaled the observed number of trees. For example, suppose that during a one-year period three trees on a plot died, then the three trees with highest predicted mortality probabilities were removed from the tree list. Note that the P&H method allocates mortality for the 5-8 year projection period in one step. By comparison, the MP&H and IND methods allocate mortality annually at the beginning of each algorithm iteration until reaching the 5-8 year projection period. Once mortality is allocated, surviving trees are grown under the compatibility constraint that the *BA/ha* implied by the stand table (P&H and MP&H) or individual trees (IND) is consistent with the observed *BA/ha*.

Basal Area Growth

A basal area growth equation of the form developed by Clutter and Allison (1974) and subsequently used by P&H was fitted to the CAPPS age 3-14 data. Years 1 and 2 data were excluded because *BA/ha* was generally zero. We fitted the model to minimize the errors in predicted basal area rather than minimizing the errors with respect relative size. In addition, the basal area growth equation was modified to allow for systematic cultural treatment effects. Namely,

$$b_{2i} = \bar{b}_2 \left(\frac{b_{1i}}{\bar{b}_1} \right)^{A^{\beta}} \qquad (6.5)$$
$$\beta = \beta_0 + \beta_F z_F + \beta_H z_H + \beta_{HF} z_{HF}$$

Where b_{2i} and b_{1i} are the basal area at Age_2 and Age_1 for *dbh* class *i*, $\bar{b}_2$ and $\bar{b}_1$ denote the mean tree basal area for plot *k* at ages 2 and 1, respectively, *A* is the ratio of Age_2 to Age_1,

and z_j's are indicator variables for the j^{th} treatment. The model used 73779 individual tree observations (Table 6.2) and estimated parameters are β_0 = -0.48639, β_F = 0.048693, β_H = 0.191459, and β_{HF} = 0.36506. The RMSE and pseudo-R^2 are 0.000970 and 0.9907, respectively, and there was no evidence of heteroscedasticity. Since β_0 is negative and all treatment effect adjustments still result in a negative value, the relative contribution to basal area of trees smaller (larger) than average relative size will increase (decrease) over time. However, this will occur more quickly in the C plots than the H&F plots.

Using the observed future survival and *BA/ha*, the following equation ensures compatibility between the observed *BA/ha* and the *BA/ha* implied by the stand table or individual trees.

$$b_{2i} = b_2 \left[\frac{\left(\frac{b_{1i}}{\overline{b_1}} \right)^{(A)^{\beta}}}{\sum_{i=1}^{j} n_{2i} \left(\frac{b_{1i}}{\overline{b_1}} \right)^{(A)^{\beta}}} \right] \tag{6.6}$$

For the P&H and MP&H methods n_{2i} is the number of trees in a given diameter class. Whereas for the IND method n_{2i} equals one, i.e., each tree is treated as a unique diameter class. At the end of the projection period, the P&H and MP&H trees were allocated to the 2.5 cm diameter classes using the projected class midpoints and assuming that the class midpoints extend halfway between adjacent class midpoints and that the trees are uniformly distributed within these diameter classes. The IND method trees were placed into 2.5 cm classes based upon the predicted *dbh* for each tree.

Two criteria are used to compare the predicted stand tables using the three methods with the observed stand tables. The two-sample Kolmogorov-Smirnov (KS) and sum of the absolute deviations (SAD) statistics, which are measures of goodness of fit,

were used to compare the predicted and observed stand tables. The KS statistic measures the maximum departure of the cumulative proportion of predicted versus observed trees in each *dbh* class. The SAD statistic is similar to the error index proposed by Reynolds et al. (1988) and is computed by plot as the absolute difference between the observed and predicted proportion of trees in each diameter class. If the predicted stand table approximates the observed stand tables closely then the KS and SAD statistics values will be small.

Results

Total plots rejected using the KS statistic for the three methods were computed for α levels 0.01, 0.05, and 0.10. Using $\alpha = 0.01$, the IND, P&H, and MP&H methods rejected 23, 1, and 4 plot(s), respectively. For $\alpha = 0.05$, the IND, P&H, and MP&H rejected 21, 1, and 2 plot(s), respectively. Lastly, for $\alpha = 0.10$, the IND, P&H, and MP&H methods rejected 19, 0, and 1 plot(s), respectively. Thus, for $\alpha = 0.05$, the IND, P&H, and MP&H methods reject approximately 20.79, 0.99, and 1.98 percent of the plots.

Mean KS statistic values for the IND, P&H, and MP&H methods are 0.09465, 0.04179, and 0.05892, respectively. The P&H KS statistic is smaller than the IND for 84 of the 101 plots and the mean difference is 0.06850. Conversely, when the IND method KS statistic is smaller than the P&H method the mean difference is 0.02437. Moreover, the P&H method KS statistic is smaller than the MP&H method for 72 of the 101 plots and has a mean difference for these 72 plots of 0.02694. The MP&H method has a smaller KS statistic than the P&H method for 27 plots and the mean difference is 0.007773 (two of the plots had a mean difference of zero for the P&H and MP&H

methods). The MP&H method KS statistic is smaller for 61 of the 101 plots versus the IND method and the mean difference for these plots is 0.08208. Conversely, when the IND KS statistic is smaller than the MP&H method the mean difference is 0.03493.

Mean SAD statistics for the IND, P&H, and MP&H methods are 0.4697, 0.2990, and 0.3779, respectively. The P&H SAD statistic is smaller than the IND and MP&H for 70 and 77 of the 101 plots, respectively and the mean difference versus the IND method is 0.30274. When the IND SAD statistic is smaller relative to the P&H method the mean difference is 0.1218. In addition, if the P&H SAD statistic is smaller than the MP&H SAD statistic the mean difference is 0.1242. Conversely, when the MP&H SAD statistic is smaller the mean difference is 0.06767. The MP&H method SAD statistic is smaller than the IND method for 54 of the 101 plots. When the MP&H SAD statistic is smaller relative to the IND method the mean difference is 0.3466. When the IND SAD statistic is smaller, the mean difference is 0.2003.

The poorest predicted stand tables using the three methods are presented in Figure 6.1. Plot 23 (Figure 6.1a) is the poorest predicted stand table using the P&H and MP&H methods, the KS statistics are 0.11454 and 0.14627, respectively. Their respective KS statistic p-values are 0.003300 and 0.03710. The p-value for plot 23 using the IND method is 0.8002. Plot 87 is the poorest predicted stand table for the IND method and the KS statistic is 0.4900 with a p-value <0.0001. Plot 87 KS statistic for the P&H and MP&H methods is 0.065403, which has a p-value of 0.7959.

In addition to these assessment criteria, the results were evaluated by treatment to determine if one method was projecting better than the other methods for a given cultural treatment. Results revealed no systematic trends in the ability of the methods to predict

the stand tables by treatment. The P&H method performs substantially and slightly better relative to the IND and MP&H methods for all cultural treatments.

Discussion

Results reveal that the P&H and MP&H stand table projection methods substantially outperform the IND method for these plots and projection periods. In addition, the P&H stand table projection method performs better than the MP&H method. At the $\alpha = 0.05$ level, P&H increases the proportion of plots not rejected versus the IND method by 25 percent. The KS and SAD statistics revealed the plots where the P&H performs better relative to the MP&H and IND methods there was a much greater mean difference than when the MP&H and IND methods performed better than the P&H method. In addition, when the KS and SAD statistics were smaller for the MP&H and IND versus the P&H method, the difference is negligible. Moreover, the P&H method has an advantage over the IND and MP&H methods in the ease of use, i.e., it allocates mortality and projects *dbh* growth for the entire projection period, whereas the MP&H and IND methods invoke the algorithm annually until reaching the end of the projection period.

Using the KS statistic ($\alpha = 0.05$), the P&H and MP&H methods rejected 1 and 2 of the 101 plots, respectively. Plot 23, which has the control treatment and is 11 years old, was rejected using the P&H and MP&H methods. The initial (age 6) stand table for plot 23 has *dbh* classes of 2.5, 5.0, and 7.5 cm with 869.8, 632.6, and 19.8 TPH. So there is minor differentiation in the *dbh* classes. The plot 23 ending stand table (age 11) has *dbh* classes of 5.0, 7.5, 10.0, 12.5, 15.0, and 17.5 cm with 59.3, 79.1, 711.6, 533.7, 118.6, and 19.8 TPH, respectively. No mortality occurred for plot 23 during the 5-year

projection period; therefore, any differences in the ending stand table can be attributed to the basal area growth equation and assumption that trees are uniformly distributed within a *dbh* class. Since plot 23 received the control treatment, a population average (PA) estimate of the basal area growth equation parameter β_0 is –0.48639. To determine the effect of the basal area growth equation (6.5) on the ending plot 23 stand table when using the P&H method, we fitted the basal are growth equation (6.5) to plot 23 separately. The Plot 23 subject-specific (SS) estimated parameter β_0 is –0.85641, which implies that less separation of the *dbh* classes will occur using the SS parameter than when using the PA parameter. The P&H method was used to predict the plot 23 ending stand table using the SS estimated parameter for equation (6.5), then the stand table using the SS parameter was compared to the stand table using the PA parameter (Table 6.3). The PA parameter results in a greater separation of the trees among the *dbh* classes. A KS test was conducted for plot 23, as given previously, the P&H PA results are KS = 0.114536 and p-value = 0.03710, whereas the P&H SS results are KS = 0.082907 and p-value = 0.2504. Hence, using the subject specific estimated parameter for the basal area growth equation results in an improvement of the plot 23 predicted stand table. This result is not surprising since fitting the basal area growth equation parameter β_0 by plot revealed a high amount of variability among the plots (Figure 6.2). This would indicate that the precision of basal area growth predictions would likely improve by including a plot level random effect for this parameter, thereby increasing the precision in general of the stand table projections.

Diameter class and individual tree mortality probabilities for the IND and MP&H methods are estimated using a multilevel individual tree mortality model that was

developed from these data. Whereas the P&H method predicts mortality using a parameter free method that is species independent. Moreover, the P&H p_i's are only dependent upon a given site in the sense that the probability of mortality for a *dbh* class depends upon the number of diameter classes. However, two stands with the exact same *dbh* classes that have a different number of trees in the diameter classes, growing under different conditions and for different species will have the same mortality probabilities by *dbh* class. This is an interesting development because the individual tree mortality prediction equation (6.2) developed for these data is site specific, i.e., it incorporates random effects for the trees within a plot and plots. It is reasonable to assume that the individual tree mortality model is predicting the *dbh* class mortality probabilities more precisely than the P&H method, which assumes that the mortality probabilities are inversely proportional to basal area. If the p_i's are more precise for the MP&H method versus the P&H method, then the question is why is the P&H method performing better for these data. The only differences between the P&H and MP&H methods are the prediction of the p_i's and the MP&H method is iterative annually for the projection period. To provide insight, we fitted equation (6.2) such that the mortality probabilities for plot 125 are based on the 8-year projection period by *dbh* class rather than the annual mortality probabilities. These p_i probabilities and the conditional probabilities π_i are compared to the P&H probabilities (Table 6.4). There is a substantial difference in the p_i probabilities by *dbh* class using the P&H and MP&H methods. However, there is no substantial difference for the P&H and MP&H conditional probabilities. For plot 125, approximately 296.5 TPH were observed to die between age 6 and 14. A KS test was conducted for the mortality allocation of plot 125 using the P&H and MP&H methods

and there is no significant difference. This would indicate that the conditional probabilities are robust if the p_i's are substantially different but the overall p_i trends for *dbh* classes are similar. In addition, the predicted plot 125 stand table using the MP&H method iterated annually versus the MP&H method when projected in one step are virtually identical.

Conclusion

Stand table projection can be an important part of the management decision process and there have been several recent methods proposed to increase the accuracy and precision of the predicted stand tables (Somers and Nepal 1992, Cao and Baldwin 1999). Our results illustrate that the P&H method is a simple yet effective method for projecting stand tables. The fact that the P&H and MP&H methods substantially outperformed the IND method may indicate that the IND method needs a growth equation that is more plot and tree specific to accurately project stand tables. In addition, the P&H method appears to be robust against departures from the unconditional probabilities (p_i's) by *dbh* class, given the overall mortality probabilities by *dbh* class are similar. In general, mortality allocation is unlikely to be as important as the basal area growth equation for accurate and precise stand table projections, which is not surprising since mortality is usually a rare event. Furthermore, our results for plot 23, which had no mortality, suggest that stand table projections may be improved by using a site-specific basal area growth equation. It is interesting to note that stand table projection method developed by Nepal and Somers (1992), which they inferred to be more precise versus the P&H method for

their data, uses plot specific parameters in the equation that adjusts the *dbh* class mortality and growth estimates.

The IND and MP&H methods require a higher degree of information, i.e., they require the development of a mortality model and a tree list for the IND method. In addition, these methods invoke the stand table projection algorithm annually until reaching the projection period. Hence, the IND and MP&H have a higher degree of complexity compared to the P&H method, but the P&H method outperformed these methods for these specific data and projection periods. Moreover, the P&H method is easy to implement and appears robust against departures of the *dbh* class mortality predictions (p_i's).

Literature Cited

Avery, T.E., and H.E. Burkhart. 1994. Forest measurements. Fourth edition. McGraw-Hill, Inc. New York. 408 p.

Bailey, R.L., N.C. Abernathy, and E.P. Jones. 1981. Diameter distribution models for repeatedly thinned slash pine plantations. P. 115-126 *in* Proc. 1st Bienn. South. Silvic. Res. Conf., J.P. Barnett (ed.). USDA For. Serv. Tech. Rep. SO-34. 375 p.

Cao, Q.V., and V.C. Baldwin Jr. 1999. A new algorithm for stand table projection models. For. Sci. 45(4): 506-511.

Clutter, J.L., and B.J. Allison. 1974. A growth and yield model for *Pinus radiata* in New Zealand. *In* Growth models for trees and stand simulation. R. Coll. For. Res. Note 30, p. 136-160. Stockholm, Sweden.

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

Knowe, S.A. 1994. Incorporating the effects of interspecific competition and vegetation management treatments into stand table projection models for Douglas-fir saplings. For. Ecol. Manage. 67: 87-99.

Knowe, S.A., G.R. Ahrens, and D.S. Bell. 1997. Comparison of diameter-distribution-prediction, stand-table-projection, and individual-tree-growth modeling approaches for young red alder plantations. For. Ecol. Manage. 98: 49-60.

Knowe, S.A., and D.E. Hibbs. 1996. Stand structure and dynamics of young red alder as affected by planting density. For. Ecol. Manage. 82: 69-85.

Knowe, S.A., and W.I. Stein. 1995. Predicting the effects of site preparation and protection on development of young Douglas-fir plantations. Can. J. For. Res. 25: 1538-1547.

Matney, T.G., and A.D. Sullivan. 1982. Compatible stand and stock tables for thinned and unthinned loblolly pine stands. For. Sci. 28(1): 161-171.

McTague, J.P., and W.F. Stansfield. 1994. Stand and tree dynamics of uneven-aged Ponderosa pine. For. Sci. 40(2): 289-302.

Meyer, H.A. 1952. Structure, growth, and drain in balanced, uneven-aged forests. J. Forestry 50: 85-92.

Nepal, S.K., and G.L. Somers. 1992. A generalized approach to stand table projection. For. Sci. 38(1): 120-133.

Pienaar, L.V., and W.M. Harrison. 1988. A stand table projection approach to yield prediction in unthinned even-age stands. For. Sci. 34(3): 804-808.

Pienaar, L.V., and J.W. Rheney. 1993. Yield prediction for mechanically site-prepared slash pine plantations in the southeastern coastal plain. South. J. Appl. For. 17(4): 163-173.

Reynolds, M.R., T.E. Burk, and Won-Chin Huang. 1988. Goodness-of-fit tests and model selection procedures for diameter distribution models. For. Sci. 34(2): 373-399.

Table 6.1. The CAPPS study plot distribution by age for the spectrum of plots and by treatment (C = control, F = fertilizer, H = herbicide, and H&F = herbicide and fertilization).

Plot Age	Plots	Plots by Treatment			
		C	F	H	H&F
14	26	8	4	8	6
13	28	8	6	8	6
12	29	10	6	8	5
11	18	8	2	6	2
Total	101	34	18	30	19

Table 6.2. The CAPPS study summary statistics for *dbh*, trees per hectare (*TPH*), basal area per hectare (*BA/ha*), and quadratic mean diameter (D_q) for the data age range 3-14.

Attribute	N	Mean	Minimum	Maximum	Std. Error
dbh (cm)	73779	12.05	0.254	35.05	0.01966
BA (m)	73779	0.01364	5.067E-6	.09649	0.00003826
TPH	1028	1418.7	533.7	1858.2	7.1610
BA/ha (m^2)	1028	19.35	0.03937	49.96	0.3825
D_q (cm)	1028	12.37	0.7770	24.59	0.1555

Table 6.3. The observed age 11 stand table for Plot 23 contrasted with the P&H stand table estimation method. The P&H (PA) and (SS) methods use the population average (PA) and subject (SS) estimated parameters for the basal area growth equation.

dbh class (cm)	Observed	P&H (PA)	P&H (SS)
5.0	59.3	107.0	
7.5	79.1	383.7	386.5
10.0	711.6	382.5	483.3
12.5	533.7	297.8	383.5
15.0	118.6	297.8	253.8
17.5	19.8	50.9	13.1
20.0		2.0	2.0

Table 6.4. The P&H and MP&H mortality probabilities and respective conditional probabilities given a tree dies for plot 125.

	P&H Method		MP&H Method	
Dbh	p_i	Conditional (π_i)	p_i	Conditional (π_i)
2	0.539245	0.193168	0.706043	0.153445
3	0.239664	0.357718	0.436103	0.394912
4	0.134811	0.346092	0.235141	0.366242
5	0.086279	0.103023	0.117884	0.085400

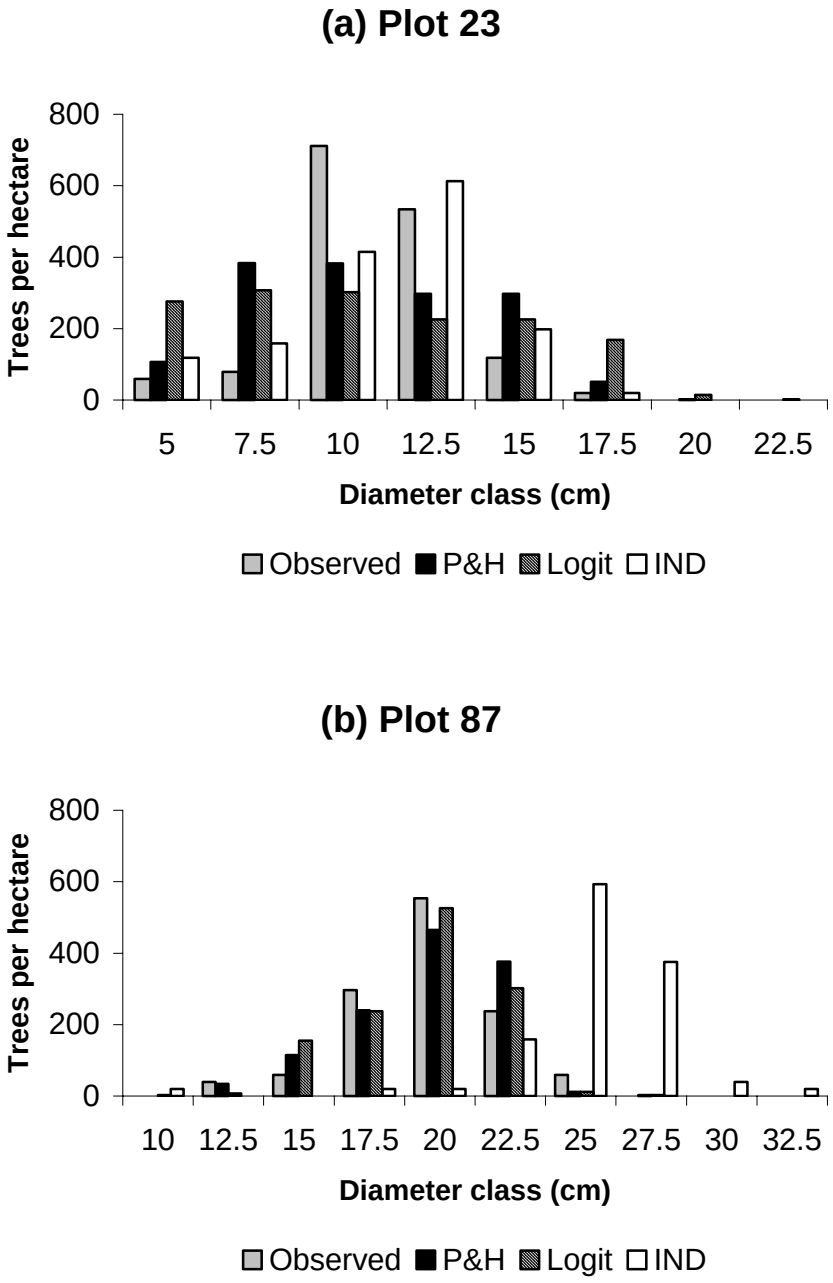


Figure 6.1. Examples of the worst predicted stand table projections using the P&H and Logit models (Plot 23), and the individual tree model (Plot 87) contrasted with the observed trees per hectare.

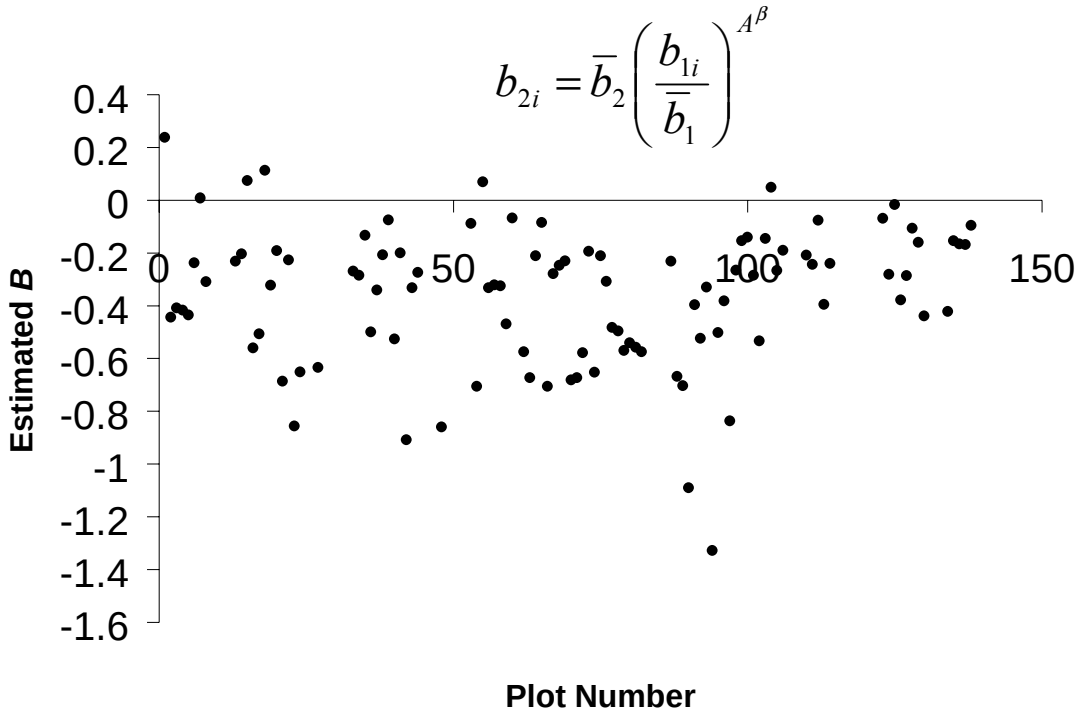


Figure 6.2. The estimated basal area growth equation (3) β_0 parameter when fitted by plot (N = 101 plots).

CHAPTER 7

CONCLUSION

Our study provides a foundation for developing biologically reasonable whole stand and multilevel individual tree survival models. Although forestry survival is difficult to model, the difficulty can be reduced by using the empirical hazard function to aid in the modeling process. We demonstrate the importance of the hazard function in modeling forestry survival data. Use of the hazard function to aid in forestry survival model selection is not novel. Preisler and Slaughter (1997) limited their individual tree survival model selection to those that were capable of reflecting the empirical hazard function behavior, but this is the first known time that the empirical hazard function has been used to aid in the development of a whole stand survival model. As mentioned, survival prediction can be important in the management decision process, e.g., prediction of whole stand survival can be critical in the volume predicted per unit area. However, our stand table projection study (Chapter 6) suggests that mortality allocation in the stand table projection method is unlikely to be the primary critical component.

Our whole stand survival model methodology demonstrates that using the empirical hazard function can aid in the development of biologically reasonable whole stand survival models that adequately reflect the continuum of a plantation's survival. Using the empirical hazard function, we limited our function selection in the differential equation to those functions that are capable of reflecting a "bathtub" shape. Moreover, ignoring cultural treatment effects, our whole stand survival models mirrored

a “bathtub” shaped hazard function using as few as three parameters. Cultural treatment parameters allowed additional model flexibility, but as shown, it is the $f(t)$ function that allowed our whole stand survival model to exhibit an underlying bathtub shaped hazard function. In addition, our methodology produced models that have excellent extrapolative properties, which was demonstrated by fitting model (1) to the age 5-14 data. Whole stand lifespan survival is commonly modeled using a system of equations, tables, or combination thereof, which increases the complexity. We demonstrated that one equation could model the continuum of a stand’s survival when derived from knowledge about the underlying hazard function.

Multilevel individual tree survival models are attractive for two main reasons. First, these models allow for site-specific predictions and secondly, they can account for sources of heterogeneity. Frequently in forestry survival predictions for a new plot are made by ignoring prior information, i.e., predictions are based upon current measurements. Using prior information generally increased the mortality prediction precision of our individual tree survival models, which has been demonstrated in other multilevel forestry models (e.g., Fang and Bailey 2001, Hall and Bailey 2001). Here, we developed two individual tree survival models: the logit and complementary log-log (CLL). Model development demonstrates the CLL function is the natural choice for interval-censored permanent plot binary response data. The logistic is the most commonly used model for individual tree survival therefore, a logit model was developed to illustrate the procedure and for use in the stand table projection. In addition, it is often difficult to distinguish between the fit of the CLL and logit models if the data are

somewhat symmetrical (Agresti 1990). Both individual tree survival models illustrate the advantages of using a subject (SS) versus population averaged (PA) modeling approach.

Subject-specific models have several advantages over PA models; one is that the marginal predictions can be obtained by integrating over the random effects using numerical integration or Monte Carlo integration techniques. Conversely, the PA modeling approach does not uniquely specify a SS model. As discussed, modeling the grouping structure allows for statistically efficient estimates of the regression coefficients and provides the correct standard errors, confidence intervals, and tests of significance. Ten Have and Uttal (1991) noted that if predictions are the primary modeling purpose, then a SS approach is preferable. Individual tree forestry survival models usually focus on prediction; therefore, the SS approach is the more natural method. If hypothesis testing about factors that influence tree survival is the main study purpose, (e.g., our CLL model) then it is imperative to account for the variability within and among plots. Failing to account for the sources of heterogeneity when modeling individual tree survival could result in an over or under estimation of a given parameter effect (Rodriguez and Goldman 2001).

A study limitation with respect to the multilevel individual tree survival models was the non-convergence of most parameter estimation techniques. For our individual tree survival models, the MQL-1 parameter estimation technique was the only method that converged. Recently Rodriguez and Goldman (2001) assessed the impact of the different parameter estimation methods when fitting a multilevel binary response model. Their study suggests that MQL-1 tends to under estimate the variance components and that PQL-2 is the usually the preferred method. However, they inferred that Bayesian

methods will probably result in the best parameter estimates but can be computationally intensive. Although our parameter estimation method was limited, as noted by several authors (e.g., Goldstein and Rasbash 1996), it is likely better to allow for a multilevel structure and use MQL-1 than to ignore the multilevel structure. Although beyond the scope of our study, an assessment of using the different parameter estimation methods should be conducted in a forestry context.

Individual tree survival models often employ a threshold to classify trees as dead or alive, based upon the predicted mortality probabilities, and the mean predicted mortality probability has been suggested as the most logical threshold for a PA model (Monserud and Sterba 1999). Our sensitivity analysis revealed that the mean mortality prediction by plot might not be the best threshold for a SS individual tree survival model. An adequate threshold for classifying trees as dead or alive needs to be addressed in future research of multilevel individual tree survival models.

Stand table projection is an important part of the management decision process and our results reveal, for our data, that the Pienaar and Harrison (1988) (P&H) method is a simple, yet effective, method for projecting stand tables. It substantially and slightly outperformed the IND and MP&H methods, respectively. The fact that the P&H and MP&H methods substantially outperformed the IND method may suggest that the individual tree basal area growth equation is the critical component for projecting stand tables. We demonstrated the impact of the basal area growth equation by comparing the projected stand tables using the P&H, MP&H, and IND methods for a plot that had no mortality during the projection period. Thus, our results suggest that mortality allocation is unlikely to be as important as the basal area growth equation for accurate and precise

stand table projections. This is not surprising since mortality is a relatively rare event. Furthermore, our results suggest that stand table projections may be improved using a mixed modeling approach for the basal area growth equation, which should be addressed in future research. Interestingly, the conditional probabilities used in the P&H method appear to be robust against departures from the unconditional probabilities (p_i's) by *dbh* class. The IND and MP&H methods require a higher degree of information, i.e., they require the development of a mortality prediction model and a tree list for the IND method. In addition, these methods invoke the stand table projection algorithm annually until reaching the projection age. Although, the IND and MP&H methods increased complexity, there was no gain in the precision of the projected stand tables relative to the P&H method.

In conclusion, we demonstrate a relatively easy methodology for developing flexible whole stand survival models based on the behavior of the underlying hazard function. Furthermore, our study emphasizes the advantages of using the multilevel modeling approach for modeling individual tree survival, for both hypothesis testing and prediction. Lastly, while mortality prediction and allocation is an important component of stand table projection, our study demonstrates that the basal area growth equation is a critical component in projecting stand tables.

LITERATURE CITED

Agresti, A. 1990. Categorical data analysis. John Wiley & Sons, Inc. New York. 558 p.

Allison, P.D. 1982. Discrete time methods for the analysis of event histories. in Sociological Methodology. ed. S. Leinhardt, San Francisco, CA: Jossey-Bass 61-98.

Allison, P.D. 1995. Survival analysis using the SAS® system: A practical guide. SAS Institute Inc. Cary, NC. 292 p.

Amateis, R.L., H.E. Burkhart, and T.A. Walsh. 1989. Diameter increment and survival equations for loblolly pine trees growing in thinned and unthinned plantations on cutover, site-prepared lands. South. J. Appl. For. 13(4): 170-174.

Amateis, R.L., H.E. Burkhart, and L. Jiping. 1997. Modeling survival in juvenile and mature loblolly pine plantations. For. Ecol. Manage. 90:51-58.

Avery, T.E., and H.E. Burkhart. 1994. Forest measurements. Fourth edition. McGraw-Hill, Inc. New York. 408 p.

Avila, O.B., and H.E. Burkhart. 1992. Modeling survival of loblolly pine trees in thinned and unthinned plantations. Can. J. For. Res. 22:1878-1882.

Bailey, R.L., and J.L. Clutter. 1974. Base-age invariant polymorphic site curves. For. Sci. 20:155-159.

Bailey, R.L., N.C. Abernathy, and E.P. Jones. 1981. Diameter distribution models for repeatedly thinned slash pine plantations. P. 115-126 *in* Proc. 1st Bienn. South. Silvic. Res. Conf., J.P. Barnett (ed.). USDA For. Serv. Tech. Rep. SO-34. 375 p.

Bailey, R.L., B.E. Borders, K.D. Ware, and E.P. Jones Jr. 1985. A compatible model relating slash pine plantation survival to density, age, site index, and type and intensity of thinning. For. Sci. 31(1): 180-189.

Barbosa, M.F., and H. Goldstein. 2001. Discrete response multilevel models for repeated measures; an application to voting intentions data. In review.

Belli, K.L., and A.R. Ek. 1988. Growth and survival modeling for planted conifers in the great lakes region. For. Sci. 34(2):458-473.

Biggeri, L., M. Bini, and L. Grilli. 2001. The transition from university to work: a multilevel approach to the analysis of the time to obtain the first job. J. R. Statist. Soc. A 293-305.

Bryk, A.S., and S.W. Raudenbush. 1992. Hierarchical linear models: applications and data analysis methods. SAGE Publications, Inc. London.

Buchnam, R.G., S.P. Pederson, and N.R. Walters. 1983. A tree survival model with applications to species of the Great Lakes region. Can. J. For. Res. 13: 601-608.

Buford, M.A., and W.L. Hafley. 1985. Modeling the probability of individual tree mortality. For. Sci. 31:331-341.

Burkhart, H.E., K.D. Farrar, R.L. Amateis, and R.F. Daniels. 1987. Simulation of individual tree growth and stand development in loblolly pine plantations on cutover, site-prepared areas. School of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University, Blacksburg, Virginia. Pub. No. FWS-1-87.

Cao, Q.V., and V.C. Baldwin Jr. 1999. A new algorithm for stand table projection models. For. Sci. 45(4): 506-511.

Clutter, J.L. 1961. The development of compatible analytic models for growth and yield of Loblolly pine. Ph.D. dissertation. Duke University.

Clutter, J.L., and B.J. Allison. 1974. A growth and yield model for *Pinus radiata* in New Zealand. *In* Growth models for trees and stand simulation. R. Coll. For. Res. Note 30, p. 136-160. Stockholm, Sweden.

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

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

Collet, D. 1994. Modelling survival data in medical research. Chapman & Hall. London. 347 p.

Cost, N. D. 1989. Multiresource inventories in the southeast. USDA For. Ser. Res. Pap. SE-269.

Cox, D.R. 1972. The analysis of multivariate binary data. Appl. Statist. 21:113-120.

Cox, D.R., and D. Oakes. 1984. Analysis of survival data. Chapman and Hall, New York 201 p.

Deen, J.L. 1933. A survival table for even-aged stands of northern white pine. J. Forestry. 31: 42-44.

Dell, T.R., D.P. Feduccia, T.E. Campbell, W.F. Mann Jr., B.H. Polmer. 1979. USDA Forest Service. Res. Pap. SO-147. 16 p.

Devine, O.J., and J.L. Clutter. 1985. Predictions of survival in slash pine plantations infected with fusiform rust. For. Sci. 31(1):88-94.

Dobbertin, M., and G.S. Biging. 1998. Using the non-parametric classifier CART to model forest tree mortality. For. Sci. 44(4):507-516.

Fang, Z., and R.L. Bailey. 1999. A simultaneous system of linear and nonlinear mixed effects models for forest growth and yield prediction. Ph.D. dissertation. The University of Georgia.

Fang, Z., and R.L. Bailey. 2001. Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. For. Sci. 47(3):287-300.

Feduccia, D.P., T.R. Dell, W.F. Mann Jr., T.E. Campbell, and B.H. Palmer. 1979. Yields of unthinned loblolly pine plantations on cutover sites in the west gulf region. USDA Forest Service. Res. Pap. SO-148.

Gibbons, R.D., and D. Hedeker. 1997. Random effects probit and logistic regression models for three-level data. Biometrics. 53:1527-1537.

Glover, G.R., and J.N. Hool. 1979. A basal area ratio predictor of loblolly pine plantation mortality. For. Sci. 25(2):275-282.

Goldstein, H. 1991. Nonlinear multilevel models, with an application to discrete response data. Biometrika 78: 45-51.

Goldstein, H. 1995. Multilevel statistical models. Ed. 2. Halstead Press, New York. 178 p.

Goldstein, H., and M. Rasbash. 1996. Improved approximations for multilevel models with binary responses. J. R. Statist. Soc. A 159(3): 505-513.

Greogoire, T.G., O. Schabenberger, and J.P. Barret. 1995. Linear modeling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. Can. J. For. Res. 25:137-156.

Guan, B.T., and G.Z. Gertner. 1991. Modeling red pine tree survival with an artificial neural network. For. Sci. 37(5): 1429-1440.

Hall, D.B., and R.L. Bailey. 2001. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. For. Sci. 47(3):311-321.

Hamilton, D.A. 1974. Event probabilities estimated by regression. USDA For. Serv. Intermt. Res. Stn. Res. Pap. INT-152. 18 pp.

Hamilton, D.A., 1986. A logistic model of mortality in thinned and unthinned mixed conifer stands of northern Idaho. For. Sci., 32:989-1000.

Hamilton, D.A. and B.M. Edwards. 1976. Modeling the probability of individual tree mortality. USDA For. Serv. Intermt. Res. Stn. Res. Pap. INT-185. 22 pp.

Harrison, W.M., and B.E. Borders. 1996. Yield prediction and growth projection for site-prepared loblolly pine plantations in the Carolinas, Georgia, Alabama, and Florida. The University of Georgia, PMRC Tech. Rep. 1996-1.

Hedeker, D., O. Siddiqui, and F.B. Hu. 2000. Random-effects regression analysis of correlated grouped-time survival data. Statistical Methods in Medical Research. 9:161-179.

Hitch, K.L., B.D. Shiver, and B.E. Borders. 1996. Mortality models for newly regenerated loblolly pine plantations in the Georgia piedmont. South. J. Appl. For. 20(4): 197-202.

Hjorth, U. 1980. A realiability distribution with increasing, decreasing, constant and bathtub-shaped failure rates. Technometrics. 22(1):99-107.

Hox, J.J. 1995. Applied multilevel analysis. TT-Publikaties, Amsterdam. 114 p.

Huebschmann, M.M., L.R. Gering, T.B. Lynch, O. Bitoki, and P.A. Murphy. 2000. An individual-tree growth and yield prediction system for uneven-aged shortleaf pine stands. SJAF. 24(2): 112-120.

Jones, R.H. 1990. Serial correlation or random subject effects? Comm. Statist. Simulation Comput. 19: 1105-1123.

Kaplan, E.L., and P. Meier. 1958. Nonparametric estimation from incomplete observations. JASA. 53:457-481.

Knowe, S.A. 1994. Incorporating the effects of interspecific competition and vegetation management treatments into stand table projection models for Douglas-fir saplings. For. Ecol. Manage. 67: 87-99.

Knowe, S.A., G.R. Ahrens, and D.S. Bell. 1997. Comparison of diameter-distribution-prediction, stand-table-projection, and individual-tree-growth modeling approaches for young red alder plantations. For. Ecol. Manage. 98: 49-60.

Knowe, S.A., and D.E. Hibbs. 1996. Stand structure and dynamics of young red alder as affected by planting density. For. Ecol. Manage. 82: 69-85.

Knowe, S.A., and W.I. Stein. 1995. Predicting the effects of site preparation and protection on development of young Douglas-fir plantations. Can. J. For. Res. 25: 1538-1547.

Kobe, R.K., and K.D. Coates. 1997. Models of sapling mortality as a function of growth to characterize interspecific variation in shade tolerance of eight tree species of northwestern British Columbia. Can. J. For. Res. 27:227-236.

Krauch, H. 1930. Mortality in cut-over stands of western yellow pine. J. Forestry. 28: 1085-1097.

Krumland, B., J. Dye, and L.C. Wensel. 1977. Individual tree mortality models for the north coast region of California. University of California, Co-op Redwood Yield Research Project, Berkeley, California. Research Note 6. 7 pp.

Landsberg, J.J., and S.T. Gower. 1997. Applications of physiological ecology to forest management. Academic Press. San Deigo, CA. 354 p.

Lappi, J., and R.L. Bailey. 1988. A height prediction model with random stand and tree parameters: An alternative to traditional site index methods. For. Sci. 38(2):409-429.

Lawless, J.F. 1982. Statistical models and methods for lifetime data. Wiley, New York. 580 p.

Lee, Y. 1971. Predicting mortality for even-aged stands of lodgepole pine. For. Chron. 29-32.

Lenhart, J.D. 1972. Predicting survival of unthinned, old-field loblolly pine plantations. J. For. 754-755.

Lenhart, J.D., and J.L. Clutter. 1971. Cubic-foot yield tables for old-field loblolly pine plantations in the Georgia Piedmont. Georgia Forest Res. Con. Rep. No. 22, Series 3. 12 p.

Lindsey, J.K. 1999. Models for repeated measurements. Oxford University Press. Oxford. 515 p.

Little, E.L. 1996. National Audubon Society field guide to North American trees: Eastern region. Alfred A. Knopf, New York. 714 p.

Lorimer, C.G., and L.E. Frelich. 1984. A simulation of equilibrium diameter distributions of sugar maple (*Acer saccarum*). Bull. Torrey Bot. Club 111:193-199.

Lowell, K.E., and R.J. Mitchell. 1987. Modeling growth and mortality probabilistically using logistic regression. IUFRO For. Growth Modeling and Prediction Con. Minneapolis, MN, August 24-28, 1987. p. 708-715.

Martin, S.W., R.L. Bailey, and E.J. Jokela. 1999. Growth and yield predictions for lower coastal plain slash pine plantations fertilized at mid-rotation. SJAF. 23(1):1-7.

Martin, S.W., and G.H. Brister. 1999. A growth and yield model incorporating hardwood competition for natural loblolly pine stands in the Georgia piedmont. SJAF. 23(3):179-185.

Matney, T.G., and R.M. Farrar Jr. 1992. A thinned/unthinned loblolly pine growth and yield simulator for planted cutover site-prepared land in the mid-gulf south. SJAF. 16(2):70-75.

Matney, T.G., and A.D. Sullivan. 1982. Compatible stand and stock tables for thinned and unthinned loblolly pine stands. For. Sci. 28(1): 161-171.

McCullagh, P. 1980. Regression models for ordinal data. JRSS Series B 42:109-142.

McTague, J.P., and W.F. Stansfield. 1994. Stand and tree dynamics of uneven-aged Ponderosa pine. For. Sci. 40(2): 289-302.

Meyer, H.A. 1952. Structure, growth, and drain in balanced, uneven-aged forests. J. Forestry 50: 85-92.

Monserud, R.A. 1976. Simulation of forest tree mortality. For. Sci. 22:438-444.

Monserud, R.A., and H. Sterba. 1999. Modeling individual tree mortality for Austrian forest species. For. Ecol. Manage. 113:109-123.

Nepal, S.K., and G.L. Somers. 1992. A generalized approach to stand table projection. For. Sci. 38(1): 120-133.

Neuhaus, J.M., J.D. Kalbfleisch, and W.W. Hauck. 1991. A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data. International Statistical Review. 59(1):25-35.

Pienaar, L.V., and W.M. Harrison. 1988. A stand table projection approach to yield prediction in unthinned even-age stands. For. Sci. 34(3): 804-808.

Pienaar, L.V., and J.W. Rheney. 1993. Yield prediction for mechanically site-prepared slash pine plantations in the southeastern coastal plain. SJAF. 17(4):163-173.

Pienaar, L.V., and B.D. Shiver. 1981. Survival functions for site-prepared slash pine plantations in the flatwoods of Georgia and northern Florida. SJAF. 5:59-62.

Pinder III, J.E., J.G. Wiener, and M.H. Smith. 1978. The Weibull distribution: a new method of summarizing survivorship data. Ecology. 59(1):175-179.

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

Preisler, H.K., and G.W. Slaughter. 1997. A stochastic model for tree survival in stands affected by annosum root disease. For. Sci. 43(1):78-85.

Prentice, R.L., and L.A. Gloeckler. 1978. Regression analysis of grouped survival data with application to breast cancer data. Biometrics. 34:57-67.

Rasbash, J., W. Browne, H. Goldstein, M. Yang, I. Plewis, M. Healy, G. Woodhouse, and D. Draper. 1999. A user's guide to MLwiN. 2nd ed. London: Institute of Education.

Reineke, L.H. 1933. Perfecting a stand-density index for even-aged forests. Jour. Ag. Res. 46: 627-638.

Rennolls, K., and A. Peace. 1986. Flow models of mortality and yield for unthinned forest stands. For. 59(1): 47-58.

Reynolds, M.R., T.E. Burk, and Won-Chin Huang. 1988. Goodness-of-fit tests and model selection procedures for diameter distribution models. For. Sci. 34(2): 373-399.

Rodriguez, G., and N. Goldman. 1995. An assessment of estimation procedures for multilevel models with binary responses. J. R. Statist. Soc. A: 73-89.

Rodriguez, G., and N. Goldman. 2001. Improved estimation procedures for multilevel models with binary response: a case study. J. R. Statist. Soc. A 339-355.

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

Shen, G., J.A. Moore, and C.R. Hatch. 2001. The effect of nitrogen fertilization, rock type, and habitat type on individual tree mortality. For. Sci. 47(2): 203-213.

Siler, W. 1979. A competing-risk model for animal mortality. Ecology 60(4): 750-757.

Smalley, G.W., and R.L. Bailey. 1974. Yield tables and stand structure for shortleaf pine plantations in Tennessee, Alabama, and Georgia Highlands. USDA For. Ser. Res. Pap. SO-96, 80 p. South. For. Exp. Stn., New Orleans, La.

Somers, G.L., R.G. Oderwald, W.H. Harms, and O.G. Langdon. 1980. Predicting mortality with a Weibull distribution. For. Sci. 26(2):291-300.

Stacy, E.W. 1962. Generalization of Gamma distribution. Ann. Math. Stat. 33(3): 1187-.

Staebler, G.R., 1953. Mortality estimation in fully stocked stands of young growth Douglas-fir. USDA For. Serv. Pac Northwest Forest and Range Exp. Stn. Res. Pap. 4. 8 p.

Stage, A.R. 1973. PROGNOSIS model for stand development. USDA For. Ser. Res. Pap. INT-137.

Tait, D.E. 1988. The dynamics of stand development: a general stand model applied to Douglas-fir. Can. J. For. Res. 18: 696-702.

Tait, D.E., C.J. Cieszewski, and I.E. Bella. 1988. The stand dynamics of lodgepole pine. Can. J. For. Res. 18: 1255-1260.

Tait, D.E., and K.V. Jahraus. 1988. The contrast in stand dynamics as revealed by comparing parameter estimates for a general stand growth model. Can. J. For. Res. 18: 1479-1485.

Tasissa, G., and H.E. Burkhart. 1998. An application of mixed effects analysis to modeling thinning effects on stem profile of loblolly pine. For. Ecolog. Manag. 103(2): 119-127.

Teck, R.M., and D.E. Hilt. 1990. Individual-tree probability of survival model for the Northeastern United States. USDA For. Serv. Northeastern For. Exp. Stn. Res. Pap. NE-642. 10 pp.

Ten Have, T.R., and D.H. Uttal. 1994. Subject-specific and population-averaged continuation ratio logit models for multiple discrete time survival profiles. Appl. Statist. 43(2):371-384.

Thomson, R.B. 1932. Mthods of determining increment over large areas. Jour. For. 30: 47-57.

USDA. 1929. Volume, yield, and stand tables for second-growth southern pines. Misc. Publ. U.S. Dept. Agric. No. 50.

Vanclay, J.K. 1991. Mortality functions for north Queensland rain forests. J. Tropical For. Sci. 4(1):15-36.

Varner, A.E. 1981. Yield prediction in seeded stands of loblolly pine in the Piedmont of Georgia and Alabama. M.S. Thesis, The University of Georgia, Athens, GA. 95 p.

Volney, W.J.A. 1998. Ten-year tree mortality following a jack pine budworm outbreak in Saskatchewan. Can. J. For. Res. 28(12): 1784-1793.

Vonesh, E.F., and V.M. Chinchilli. 1997. Linear and nonlinear models for the analysis of repeated measurements. Marcel Dekker, New York.

Wykoff, W.R., N.L. Crookston, and A.R. Stage. 1982. User's guide for the stand prognosis model – version 5.0. USDA For. Serv. Region One, Missoula, Montana. For. Serv. Handbook 2409.21 Chapter 300. 113 pp.

Yang, M., and H. Goldstein. 2000. Multilevel models for repeated binary outcomes: attitudes and voting over the electoral cycle. J. R. Statist. Soc. A 163(1): 49-62.

Yao, X., S.J. Titus, and S.E. MacDonald. 2001. A generalized logistic model of individual tree mortality for aspen, white spruce, and lodgepole pine in Alberta mixedwood forests. Can. J. For. Res. 31: 283-291.

Yoda, K., T. Kira, H. Ogawa, and K. Hozumi. 1963. Self-thinning in overcrowded pure stands under cultivated and natural conditions. J. Biol. Osaka City Univ. 14: 107-129

Zeger, S.L., K. Liang, and P.S. Albert. 1988. Models for longitudinal data: a generalized estimating equation approach. Biometrics. 44:1049-1060.