

SIMULATION MODELING TO TEST DIFFERENT HARVEST REGIMES FOR STAND  
CONVERSION IN SOUTHERN PINE

By

NILESH TIMILSINA

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2010

© 2010 NileshTimilsina

To my family

## ACKNOWLEDGMENTS

I thank my parents for believing in me, and bestowing their support and encouragement throughout my career. Many thanks go to my wife also for her support during my life as a graduate student. My advisor Dr. Christina Staudhammer is an integral part of this work, and I thank her for her suggestions, help, and continuous support. I thank my committee members, Dr. Wendell Cropper Jr., Dr. Shibu Jose, Dr. Michael Andreu, and Dr. Gregory Kiker for their suggestions, help and critical evaluation of my work. Special thanks go to Dr. Cropper for teaching me simulation modeling, and providing his continuous help and feedback to this work. I also thank Dr. Mary Christman for statistical help.

I have to thank the Suwanee River Water Management district, Charlie Houder and Bob Heeke for the data and access to the property. This work would not be possible without their help and support. I thank Louise Loudermilk for sharing her knowledge on simulation modeling, and helping me think about my work. I also thank Sparkle Malone for her time spent while discussing my work and helping me find a proper solution. Thanks go to Todd Bush for reviewing some of my writings, and thank you Helen. I also thank the College of Agriculture and Life Sciences, School of Forest Resources and Conservation, faculty, and staff for giving me this opportunity to work and learn, and for making my stay memorable. I also thank James Colee, and Dr. Salvador Gezan for statistical consultation and help.

# TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS.....	4
LIST OF TABLES.....	8
LIST OF FIGURES.....	11
LIST OF ABBREVIATIONS.....	13
ABSTRACT.....	15
CHAPTER	
1 INTRODUCTION.....	17
Modeling Forest Growth.....	21
Significance of the Study.....	23
Specific Objectives.....	24
2 INDIVIDUAL TREE-BASED DIAMETER GROWTH MODEL OF SLASH PINE IN FLORIDA USING NON-LINEAR MIXED MODELING.....	25
Introduction.....	25
Methods.....	27
Data.....	28
Mixed Modeling of Growth.....	34
Model Fitting.....	36
Model Evaluation.....	39
Growth among Natural Communities.....	41
Results.....	42
Models.....	42
Mixed Modeling of Growth.....	45
Cross-validation of Mixed Effects Model.....	46
Difference in Growth between Natural Communities.....	48
Discussion and Conclusion.....	50
3 INDIVIDUAL TREE HEIGHT GROWTH MODEL FOR SLASH PINE IN FLORIDA.....	64
Introduction.....	64
Methods.....	67
Data.....	67
Mixed Modeling of Growth.....	71
Model Fitting.....	75
Model Evaluation.....	79

Results.....	81
Model Selection.....	81
Comparison of Fixed Versus Mixed Effects Models .....	84
Base and Expanded Non-linear Model.....	85
Fitting Covariance Structure to Non-linear Mixed Models.....	86
Comparison between Linear and Non-linear Model .....	86
Cross-validation of Models .....	88
Comparing Growth among Different Natural Communities.....	90
Interpretation of the Parameters of the Best Model .....	91
Discussion and Conclusions .....	92
4 INDIVIDUAL TREE MORTALITY MODEL FOR SLASH PINE IN FLORIDA: A MIXED MODELING APPROACH .....	106
Introduction .....	106
Methods .....	109
Data.....	109
Selection of Variables.....	113
Model Development .....	116
Model Evaluation and Validation .....	119
Results.....	122
Discussion .....	127
5 RECRUITMENT MODEL FOR SLASH PINE IN FLORIDA .....	138
Introduction .....	138
Methods .....	142
Data.....	142
Model Development and Evaluation.....	144
Results.....	149
Model for Predicting the Probability of Recruitment.....	149
Model for Conditional Number of Recruits.....	151
Discussion .....	152
6 SIMULATION OF DIFFERENT HARVEST REGIMES FOR STAND CONVERSION OF SLASH PINE IN FLORIDA.....	163
Introduction .....	163
Methods .....	168
Description of the Model and Variables .....	168
Model Evaluation and Sensitivity Analysis.....	170
Harvest Simulation .....	171
Thinning .....	171
Single-tree selection (VGDL) .....	172
Group selection.....	172
Comparison of Harvests .....	173
Economic Analysis .....	174

Results.....	175
Evaluation of Model Prediction .....	175
Sensitivity Analysis.....	176
Diameter Distribution of Trees.....	177
Height Distribution of Trees .....	178
Harvest Volume and Economic Analysis of Harvest.....	178
Discussion and Conclusion.....	179
7 CONCLUSION.....	193
LIST OF REFERENCES .....	198
BIOGRAPHICAL SKETCH.....	216

## LIST OF TABLES

<u>Table</u>	<u>page</u>
2-1	Descriptive statistics for slash pine dataset (70 plots). ..... 56
2-2	Parameter estimates and their respective standard errors (SE) and P-values for diameter growth Model 1 with fixed effects only. .... 56
2-3	Parameter estimates and their respective standard errors (SE) and P-values for diameter growth Model 2 with fixed effects only. .... 56
2-4	Parameter estimates and their respective standard errors (SE) and P-values for Model 3 fit with fixed effects only. .... 56
2-5	Fit statistics used to compare the three annual diameter growth models. Models are given in the text. .... 57
2-6	Fit statistics used to compare the prediction ability of the three annual diameter growth models using the cross validation data. Models are given in the text. .... 57
2-7	Parameter estimates and their respective standard errors (SE) and P-values for selected mixed effects annual diameter growth model (Model 3; equation 2-12). .... 57
2-8	Comparison of Model 3 with various fit statistics using both the entire dataset and cross validation approach. .... 58
2-9	Likelihood ratio test results for comparison of a combined slash pine dbh growth model (a single model for all community types) with a separate model for UMF and a full model. See text for the description of community types. .... 58
2-10	Parameter estimates and their respective standard errors (SE) and P-values for selected mixed effects annual diameter growth model for Upland Mixed Forest (equation 2-17). .... 58
3-1	Parameter estimates with standard errors for linear fixed and mixed effects height growth model for slash pine in north central Florida ..... 98
3-2	AIC and BIC values of the non-linear model (equation 3-8e), treating various parameters as having both fixed and random effects. .... 98
3-3	Goodness of fit statistics for non-linear model (equation 3-8e), treating various parameters as having both fixed and random effects. .... 98
3-4	Comparison of AIC and BIC values when including fixed versus mixed effects, for linear and non-linear models (Models 1 and 2, respectively). .... 99

3-5	Comparison of various fit statistics when including fixed versus mixed effects, for linear and non-linear (Models 1 and 2, respectively).....	99
3-6	Fit statistics for fixed effects linear model, non-linear base model, and non-linear model expanded with different methods without intercept term. ....	99
3-7	Fit statistics for mixed effects linear model, non-linear base model, and non-linear model expanded with different methods without intercept. ....	100
3-8	AIC and BIC for expanded non-linear mixed model (equation 3-14) fit with residual covariance structures to account for spatial correlation between trees within a plot.....	100
3-9	Fit statistics for linear and non linear mixed base model, and non linear mixed model expanded with different methods using the cross validation approach. ....	101
3-10	AIC and BIC values, and likelihood ratio test results of constrained height growth model with full model, and a separate model for the UMF.....	101
3-11	Parameter estimates and their respective standard errors and p-values for the selected annual height growth model (equation 3-14) with fixed effects only.....	101
3-12	Parameter estimates and their respective standard errors and p-values for the selected annual height growth model (equation 3-14) with random effect..	102
3-13	Parameter estimates and their respective standard errors and p-values for the selected annual height growth model for UMF (equation 3-17) with random effect.....	102
3-14	Parameter estimates and their respective standard errors and p-values for the selected annual height growth model for UMF (equation 3-17) with fixed effects only. ....	102
4-1	Parameter estimates, their standard errors, and statistical significance of the linear logistic regression model for slash pine mortality in north Florida.....	133
4-2	Parameter estimates, their standard errors, and statistical significance of the generalized linear logistic regression model for slash pine mortality in north Florida. ....	133
5-1	Parameter estimates and their associated standard errors and p-values of the model predicting the probability of recruitment (equation 5-4) for slash pine in north Florida.....	158
5-2	Parameter estimates of the model predicting conditional number of recruits (equation 5-5) for slash pine in north Florida.....	158

5-3	Fit statistics for the cross validation of model for predicting number of recruits (Model 2) of slash pine in north Florida. ....	158
6-1	Variables and parameters used for sensitivity analysis of simulation model. ...	187
6-2	Percentage change in model output values relative to the base output value after changing the input variable and parameters.....	187
6-3	Various statistics describing the final diameter distributions of the simulated stand at age 68 under various harvest regimes with base site index 24.....	187
6-4	Various statistics describing the final height distributions of the simulated stands at age 68 under various harvest regimes.....	188
6-5	Volume of trees harvested (cubic meter outside bark) during each harvesting scenario simulated.....	188

## LIST OF FIGURES

<u>Figure</u>	<u>page</u>
2-1	Map showing location of Suwanee River Water Mangement District (SRWMD) in Florida. The black rectangular box in the bottom map shows plot locations within SRWMD..... 59
2-2	Plot of residual versus predicted values for Model 1. .... 60
2-3	Plot of residual versus predicted values for Model 2. .... 60
2-4	Plot of residual versus predicted values for Model 3. .... 61
2-5	Leave-one-out cross validation predictions of annual diameter growth versus tree size (dbh) from Model 3 for two example plots. .... 62
2-6	Predicted individual tree diameter growth for typical range of slash pine initial dbh-height data using Model 3(fixed effects only). .... 63
3-1	Height growth per year for slash pine in Florida using linear fixed effects model (equation 3-9a). .... 103
3-2	Annual height growth of slash pine in Florida against total tree height for different basal areas per hectare ..... 104
3-3	Annual height growth of slash pine in Florida against dbh..... 104
3-4	Annual height growth of slash pine in Florida against total tree height and dbh. .... 105
4-1	Receiver operating characteristic (ROC). .... 134
4-2	Probability of mortality of slash pine trees versus dbh (cm) as predicted by the mortality model. .... 135
4-3	Probability of mortality of slash pine trees versus height (m) for different values of basal area per hectare (BA) as predicted by the mortality model. .... 135
4-4	Probability of mortality of slash pine trees versus dbh (cm) for different values of site index (SPI in m) as predicted by the mortality model. .... 136
4-5	Probability of mortality of slash pine trees versus dbh (cm) for different values of competition index (RBAgd) as predicted by the mortality model..... 136
4-6	Subject-specific predicted (Pred mort) and observed (Obs mort) probability of mortality for different threshold values of assigning dichotomous survival status: a) 0.460 and b) 0.303..... 137

4-7	Population averaged predicted (Pred mort) and observed (Obs mort) probability of mortality for different threshold values of assigning dichotomous survival status: a) 0.590 and b) 0.292.....	137
5-1	Probability of recruitment at 12.7 cm (5 inches) dbh for slash pine in north Florida at different values of stand basal area.....	159
5-2	Probability of recruitment at 12.7 cm (5 inches) dbh for slash pine in north Florida at different values of site index.....	159
5-3	The relationship between probability of recruitment at 12.7 cm (5 inches) dbh and basal area for different values of stand quadratic mean diameter (qmd) of slash pine in north Florida.....	160
5-4	Receiver Operating Characteristics curve of the recruitment model using the predicted values from the cross validation approach.....	160
5-5	Relationship between the conditional number of recruits and stand basal area at different values of quadratic mean diameter for slash pine in north Florida.....	161
5-6	Relationship between the conditional number of recruits and number of trees in a stand for slash pine in north Florida.....	161
5-7	Observed and predicted number of recruits for site index classes from cross validation of Model 2 for slash pine in north Florida.....	162
5-8	Observed and predicted number of recruits for basal area classes from cross validation of Model 2 for slash pine in north Florida.....	162
6-1	Simulated basal area development over time using the model for control stand (no harvest) for (a) base site index 24 (b) and site index 18.....	189
6-2	Simulated net stand volume over time for control stand (no harvest) with base site index 24.....	190
6-3	Simulated mean annual volume increment over time for the control stand (no harvest) with base site index 24.....	190
6-4	Diameter distribution (number of trees/ha) at the end of the simulation (stand age 68 years) under different harvest techniques.....	191
6-5	Height distribution (number of trees/ha) at the end of the simulation (stand age 68 years) under different harvest techniques.....	192

## LIST OF ABBREVIATIONS

AIC	Akaike's Information Criteria
BIC	Bayesian Information Criteria
BS	Bay Swamp
CI	Competition Index
dbh	diameter at breast height
ECA	Exposed Crown Area
GLM	General Linear Model
GLMM	Generalized Linear Mixed Model
GPS	Global Positioning System
LRS	Likelihood Ratio Statistics
MF	Mesic Flatwood
MSE	Mean Square Error
NC	Natural Communities
PBIAS	Percentage Bias
QMD	Quadratic Mean Diameter
RD	Relative Diameter
RMSE	Root Mean Square Error
ROC	Receiver Operating Characteristics
SA	Sand Hill
SC	Schwartz Criterion
SD	Standard Deviation
SDI	Stand Density Index
SRWMD	Suwanee River Water Management District
TCA	Total Crown Area

UMF	Upland Mixed Forest
USA	United States of America
USDA	United States Department of Agriculture
VDT	Variable Density Thinning
VGDL	Volume Guided Diameter Limit

Abstract of Dissertation Presented to the Graduate School  
of the University of Florida in Partial Fulfillment of the  
Requirements for the Degree of Doctor of Philosophy

SIMULATION MODELING TO TEST DIFFERENT HARVEST REGIMES FOR STAND  
CONVERSION IN SOUTHERN PINE

By

Nilesh Timilsina

December 2010

Chair: Christina Staudhammer  
Major: Forest Resources and Conservation

Managing forests with multiple ages and size classes can promote both economic and ecological services. Interest among landowners in managing for multiple uses under structurally diverse uneven-aged management has increased recently in Florida (USA), partly due to a trend toward public landownership. Increasingly, forest managers and land owners are interested in converting structurally homogenous plantations into more diverse conditions commonly found in uneven-aged stands. The objective of the present study is to assist in this process by developing growth, mortality and recruitment models for slash pine (*Pinus elliotti*. var. *elliotti*) in Florida, and also to use these models to demonstrate the use of a simulation model to test harvesting regimes such as thinnings, single tree selection and group selection. The regimes were evaluated based on their utility in achieving a set of desirable structural characteristics. Using 70 permanent plots in natural and plantation forests, we developed diameter and height growth, recruitment and mortality models. We used a mixed modeling approach, comparing linear and non-linear mixed models. Models were evaluated with information criteria, bias, and precision measures. Due to the unavailability of an independent dataset, we validated the models using a leave-one-out cross validation. Our results

indicated that the mixed models produced better estimates of growth versus fixed effects models. A Mixed model also gave better probability estimates of mortality than a fixed effects model. Probability of recruitment was best estimated with logistic regression with number recruited conditionally with a generalized Poisson regression. Simulation results demonstrated that light thinning, heavy thinning, and group selection could be useful for the conversion process.

## CHAPTER 1 INTRODUCTION

Forest management has evolved according to the demands of society. Custodial forestry, sustained yield forestry, multiple uses, and ecological forestry are all different management approaches taken based on societal needs and concerns (Plochman 1992, Sedjo 1996, Seymour and Hunter 1999). The former two approaches are more timber oriented and focused more on consumptive uses, whereas the latter two approaches are focused on balancing between consumption and other ecosystem services and wildlife habitat provided by a forest. In addition to these, intensive forestry, which started in United States during the 1960s, aimed at the objective of efficient timber production and led to the creation of fertilized and bedded plantations which were managed on a short-rotation clearcut harvest schedule

There is growing evidence that traditional forest management such as sustained yield forestry and intensive forestry do not retain diverse forest structure as natural forests do, because the aim of management is to optimize a few species of high value. These practices simplified forest structure and affected wildlife habitat and other ecological processes. Hence the importance of these practices in conservation of biodiversity is minimal (Seymour and Hunter 1999). Clearcutting and plantation establishment – although believed to follow the pattern of stand development after stand replacing disturbances – do not retain enough structural components, such as large trees, standing dead trees, and woody debris that are found after stand replacing fires. Usually plantations have trees of similar ages and are homogenous in terms of vertical and horizontal structure.

To manage forests for both ecological and economic objectives, silvicultural methods that mimic natural stand development – termed ecological forestry – are becoming popular (Seymour and Hunter 1999, Will et al. 2005). In ecological forestry, harvesting methods will leave enough old trees in a stand so that the natural pattern of forest regeneration and succession is continued (Franklin et al. 2007); enough snags and coarse woody debris are left after harvest, and priority is given to maintaining diverse vertical and horizontal structure in the future. An uneven-aged stand will have horizontal and vertical structure that are compatible with biodiversity conservation objectives. Because uneven-aged stands offer more natural and diverse structure, there is increasing interest in uneven-aged management, and land managers in both Europe and United States are eager to convert former even-aged plantations (especially in public lands) into uneven-aged structure (Nyland 2003).

In an uneven-aged stand there is a continuous forest cover and different size classes; seedlings, saplings and large trees occur at one time. For this to happen, there should be constant regeneration, and this is achieved naturally in gaps created by disturbances. In managed forests such as former plantations, natural processes have been altered, and they do not contain gaps and adequate environment for regeneration to occur. Therefore, regeneration environments should be created artificially by forest management. This could be achieved by using harvest methods such as single tree selection and group selection (Smith et al. 1997). Both methods open up gaps in the canopy to allow regeneration (Smith et al. 1997).

In single tree selection, individual trees are selected and cut based on their form, vigor, maturity, quality, and growth rate. Cutting is spread over the stand uniformly to

achieve regeneration at different locations. As opposed to single tree, group selection selects trees for cutting in groups at various locations within a stand, thereby creating larger openings. Other methods such as irregular shelterwood and diameter limit cutting can also be used to achieve an uneven-aged structure (Smith et al. 1997, Brockway et al. 2005). All these methods are also useful in converting even aged stands to uneven-aged. There is no “one size fits all” type of prescription, and each method can be modified to fit the site, management objective and species of interest. In some situations, rather than a single method, a combination of different methods along with thinning has been used for the conversion process (Guldin and Farrar 2002, Nyland 2003). Although, the methods discussed above have been recommended and tested for loblolly pine (Guldin and Farrar 2002), their effectiveness has not been tested for the conversion prescriptions in slash pine (*Pinus elliotii*. var. *elliotii*) in north Florida.

The concepts of landscape ecology emphasize the maintenance of spatial heterogeneity at the stand and landscape level. This sustains biodiversity, because heterogeneous forest patches provide diverse habitats, and also act as insurance against disturbances (Wiens 2008). Due to the increase in intensive forestry in the past few decades, there has been an increase in size and number of plantation in Florida (Frost 2006). Landscapes have become more homogenized due to the higher number of younger plantations. The size and number of old growth patches have significantly declined along with the structures related to those patches, such as large trees, big gaps, standing snags, and large coarse woody debris. This will affect other ecological processes (fire behavior) and wildlife habitats. To maintain spatial heterogeneity with patches of old growth and young forests, forest managers are interested in accelerating

the conversion of structure of young stands into old growth. This can be achieved by using silvicultural prescriptions such as thinning (Choi et al. 2007). Because thinning removes trees and allows growing space for the remaining trees in the stand, tree growth is faster in thinned than unthinned stand. This is also believed to speed up the process of old growth restoration. Thinning, such as Variable Density Thinning (VDT) have been recommended for stand conversion as well (Franklin et al. 2007). In Florida, land managers are also interested in restoring patches of old growth slash pine forests, especially in public lands, but there is no information about how methods of thinning will be useful in this process.

Testing different harvesting regimes for stand conversion and old-growth restoration in the field will take much time and resources. Another approach is to construct growth and yield models and simulate stand growth overtime after the application of various treatments (Vanclay 1994). The overall objective of this study is to create a simulation model to test several silvicultural prescriptions for converting even-aged slash pine (*Pinus elliottii*. var *elliottii*) stands into uneven-aged stands, and also to demonstrate the utility of the model by testing thinning, single tree selection, and group selection. The model can also be used to test thinning regimes for old growth restoration. This study will answer some of these questions: What variables are important for explaining height growth, diameter growth, mortality and recruitment in slash pine? Is it important to use a mixed effect model to incorporate the hierarchical nature of forestry data and achieve better prediction of growth and mortality? Are silvicultural regimes such as thinning, single tree selection and group selection useful to achieve stand conversion?

## **Modeling Forest Growth**

Growth and yield models can predict growth, mortality, reproduction, and other associated changes in forest trees and predict forest development over time; therefore, they have been widely used in forecasting forest production, and to test management alternatives and silvicultural prescriptions (Vanclay 1994). Three approaches have been used in modeling forest growth and development: empirical models, process models, and mechanistic models (Vanclay 1994, Choi et al. 2001). In empirical models, a best fit is found of the available growth, mortality, and recruitment data, to independent variables using regression techniques. The process based approach involves finding a theoretical model of forest growth based on earlier biological studies and uses physiological processes such as photosynthesis and respiration as inputs. Mechanistic models combine the approaches of empirical and process based models to simulate growth and development. Models created using any of these approaches can consist of a single equation or a number of submodels (equations) connected to simulate stand growth. Since the empirical models give the most precise estimates of forest growth (Vanclay 1994), they are the most popular; however, they are not good for extrapolation or predicting under different future conditions, e.g., estimating growth under the assumption of climate change or other environmental uncertainties. In this study, I will use an empirical approach to modeling forest growth.

Based on the variables used as inputs, growth models have been classified into whole stand models or individual tree models (Burkhart 1990, Vanclay 1994). Whole stand models use basal area, volume, and the size distribution of trees to simulate stand growth and yield (Peng 2000). Since they do not provide information about individual trees and only include stand level variables, they are less for modeling

uneven-aged stands. Individual tree models simulate establishment, growth, and mortality of individual trees and aggregate those estimates to produce stand level estimates; they are more appropriate for modeling uneven-aged stands.

Both distance independent and dependent individual tree models are used in forestry. Distance independent models such as the JABOWA model (Botkin 1993), and FORET model (Shugart 1984) do not require the location of individual trees in their input. On the other hand, distance dependent models (e.g., FOREST [Ek and Monserud 1974], and CANOPY [Choi et al. 2001]) do require locations for individual trees. Tree locations are useful for calculating competition indices – an important component of distance dependent models. Competition indices are significant in explaining the variation in stand structure and are useful tools to predict development of an individual tree (Daniels et al. 1986, Pukkala and Kolström 1991). For example, FOREST is a distance dependent individual tree based stand model that can simulate growth and development of both even-aged and uneven-aged stands of mixed northern hardwood forests. It uses submodels for overstory tree diameter growth, height growth, and survival; it also uses models for reproduction, establishment, understory tree height growth, and survival (Ek and Monserud 1974). All the sub-models in FOREST are affected by the spatial location of trees, as quantified by a competition index calculated on the basis of crown overlap between a subject tree and its neighbors, weighted by the relative size of competitors.

The use of a competition index makes it easy to incorporate the effect on growth due to the presence of competitors in the vicinity. Since harvesting removes trees and reduces competition, a distance dependent competition index could be useful to

incorporate the effect of increased growing space on the growth of a subject tree. Therefore, I will use a distance dependent approach for modeling tree growth in the proposed study.

### **Significance of the Study**

Most of the models forecasting growth and development of slash pine are focused on plantation management. These models either focus on survival of planted trees (Bailey et al. 1985) or predicting the effect of silvicultural treatments such as weed control and fertilization on growth and yield of slash pine (Pienaar and Rheney 1995, Logan and Shiver 2006). Other models have been developed that synthesize physiological and ecological processes in slash pine plantations for understanding carbon dynamics and the effect of climate change on slash pine trees (Cropper and Gholz 1993, Cropper 1998). These existing models provide information on the ecology and management of the modeled species. In north Florida, models that can simulate harvest and growth of slash pine after harvest could fill a gap in helping forest managers with new planned silvicultural practices. The model should incorporate the effect of increased growing space and resources due to harvest on tree growth. This will allow simulation of thinnings as well as harvests to predict tree growth over time. The Forest Vegetation Simulator (Dixon 2002) is a popular model in the US that allows simulation of different thinning regimes. However, it is a distance independent tree model that does not take into account the location of trees and its neighbors. It also does not have a recruitment model; therefore, it is difficult to simulate uneven-aged management. Individual tree growth models allow for competitive effects of individual trees; therefore, they are suitable for forecasting tree growth regardless of species mixture, age distribution and applied silvicultural systems (Hasenauer 2006). This study will use data

from slash pine stands under different ages, site conditions, origins (natural and plantation), habitat types, densities, and stand structure with the goal of creating a model with wide applicability.

### **Specific Objectives**

The specific objectives of the proposed study are:

- To develop an individual tree diameter growth model for slash pine in Florida.
- To develop an individual tree height growth model for slash pine in Florida.
- To develop an individual tree mortality model for slash pine in Florida.
- To develop models for predicting probability of recruitment and number of recruits for slash pine in Florida.
- To create a simulation model by using all the developed submodels and demonstrate the utility of the model for testing regimes for stand conversion such as thinning, single tree selection and group selection.

CHAPTER 2  
INDIVIDUAL TREE-BASED DIAMETER GROWTH MODEL OF SLASH PINE IN  
FLORIDA USING NON-LINEAR MIXED MODELING

**Introduction**

There is growing evidence that traditional forest management regimes such as sustained yield forestry and intensive forestry do not retain the same diverse forest structures that natural forests do (Franklin et al. 2007) because the aim of such management is to optimize a few high value species (Seymour and Hunter 1999). More modern forest management regimes, however, encompass a broad range of services from timber production to carbon sequestration to biodiversity conservation (Franklin et al. 2000, Franklin et al. 2007). To manage forests for both ecological and economic objectives, silvicultural methods that mimic natural stand development – termed ecological forestry – are becoming popular (Seymour and Hunter 1999, Will et al. 2005, Franklin et al. 2007). In ecological forestry, harvesting methods aim to leave enough old trees and regeneration in a stand so that the natural pattern of forest regeneration and succession is continued (Franklin et al. 2007). The harvest is also planned so as to leave snags and coarse woody debris, with the goal of maintaining plant and animal habitat, as well as diverse vertical and horizontal structure for the future. This type of harvest promotes uneven-aged stands with horizontal and vertical structure that are compatible with biodiversity conservation objectives. Because uneven-aged stands offer more natural and diverse structure, there is increasing interest in uneven-aged management, and land managers in both Europe and United States are eager to convert even-aged plantations (especially on public lands) into uneven-aged structure (Nyland 2003).

Field trials have been installed in the Southeast (Shibu Jose, pers. comm., University of Florida, September 10, 2009) in an effort to better understand the best harvesting techniques for the conversion process. Another approach is to use simulation modeling of forest growth and harvest, which will provide projected estimates on the best techniques. To simulate forest growth, growth models are needed, and for simulating harvest, spatially explicit growth models are needed (Peng 2000).

Various growth models have been developed for a range of purposes. Based on the variables used as inputs, growth models have been classified into whole stand models or individual tree models (Burkhart 1990, Vanclay 1994). Whole stand models use basal area, volume, and the size distribution of trees to simulate stand growth and yield (Peng 2000). Since they do not provide information about individual trees and only include stand level variables, their utility is limited in modeling uneven-aged stands (Peng 2000). Individual tree models simulate establishment, growth and mortality of individual trees, and aggregate those estimates to produce stand level estimates; they are therefore more appropriate for modeling uneven-aged stands (Vanclay 1994, Peng 2000, Lee et al. 2004).

Various stand level models have been developed in the past for slash pine in Florida. However, most growth studies of natural slash pine stands have been concerned with yield only (e.g., Benett and Clutter 1968, Benett 1970, Benett 1980). Other models have been developed for plantations (e.g., Pienaar et al. 1985; Pienaar and Shiver 1984), using age as an independent variable in the growth model, and some studies have investigated silvicultural treatments, such as thinning (e.g., Bailey et al. 1985). While the main purpose of the latter model was to predict the change in number

of trees per acre due to timing and intensity of thinning, Schreuder et al. (1979) developed models for predicting diameter distributions, and average height and volume by diameter for unthinned natural stands of slash pine. Although stand level models such as that of Schreuder, et al. (1979) are available for slash pine, they have limited utility for the simulation of uneven-aged forest growth. While stand level models have been adequate for simulating growth of even-aged pure stands, individual tree growth models are recommended for simulating uneven-aged forest growth (Vanclay 1994, Peng 2000).

The objective of this study is to develop a spatially explicit diameter (dbh) growth model that can be used in a forest growth simulator for testing different harvest regimes for stand conversion of slash pine. We used a non-linear mixed modeling approach (Vonesh and Chinchilli 1997, Pinheiro and Bates 2000) to account for the hierarchical nature of the data. While mixed models have been in widespread use in many disciplines over the last decade, the use of mixed model is a fairly recent approach in forestry (e.g., see Fang and Bailey 2001, Calama and Montero 2004, Adame et al. 2008), and in particular as applied to slash pine (except see Fang and Bailey 2001). This study will add to the mixed model literature in forestry, as well as further the slash pine growth modeling literature, where we know of no evidence of publications describing diameter growth with mixed models.

### **Methods**

Initial tree size, competition, and site characteristics are variables frequently used to model growth of individual trees (Stage 1973, Wykoff et al. 1982, Sterba et al. 2002). These variables have been used in vegetation simulators built in Europe and America including (but not limited to) STEMS (Belcher et al 1979), PROGNOISIS (Stage 1973),

and PROGNAUS (Monserud and Sterba 1996). In many models tree age is not used, often because it is unavailable or impractical and time-consuming to collect (Wykoff 1990, Monserud and Sterba 1996, Gourlet-Fleury and Houllier 2000). While Lee et al. (2004) preferred a practical model without age, models including age as a variable showed superior statistical performance.

Individual tree growth has been modeled in terms of basal area increment (Wykoff 1990, Quicke et al. 1994, Choi et al. 2001) and diameter increment (Lee et al. 1999, Gourlet-Fleury and Houllier 2000). Since the choice of dependent variable is largely personal, we chose to model diameter increment.

There are two main approaches that have been used to model diameter increment. The first approach is to calculate the potential diameter increment from competition-free trees, and then modify it by factors related to competition (Daniels and Burkhart 1975, Ek and Monserud 1974). Another approach is to relate diameter increment directly to variables related to the stand and individual tree sizes (Martin and Ek 1984, Huang and Titus 1995, Philips et al. 2003), such as dbh, height, mean crown radius, and crown length. We chose the latter approach, testing these variables and evaluating their performance. In addition to tree size variables, we also tested variables related to site quality (site index), stand density (stocking, total basal area per hectare), average tree size (quadratic mean diameter), and competition (using indices). Since tree age data were unavailable for all the trees in our dataset and to widen the transferability of our study, we did not include age as a covariate in our models.

## **Data**

This study utilized slash pine data collected from sites on the property of the Suwanee River Water Management District (SRWMD), which is one of the five regional

water management districts in Florida created by the Florida Legislature by the Water Resources Act of 1972[1]. The SRWMD (Figure 2-1) is located in north central Florida (USA), and includes a variety of growing conditions, origins and treatments. The SRWMD established a network of plots in 2000, which we re-measured in May and June 2008.

The sampling design and protocol used by the SRWMD to establish plots in 2000 closely followed the protocols of the USDA Forest Service Forestry Inventory and Analysis (Bechtold and Patterson 2005). A cluster of plots was established every 101.2 hectare (250 acres), with four circular 7.31 m (24 ft) radius tree plots; one plot was at the center and three others were located at 90, 120 and 240 degrees, 36.5 m (120 ft) from the center plot. For every plot, the following data were collected: current community; historical community; stand origin (natural, mixed or plantation); stand structure; stand age; slope; aspect; and GPS locations. For each plot, three dominant and co dominant trees were identified as site trees, and their height, age and species were recorded. In every 7.31 m radius tree plot, trees > 12.54 cm (5 inches) in diameter were identified to species and the following data were recorded: distance and azimuth from the plot center; diameter at breast height (dbh); total height; height to live crown; product class; crown class; and tree damage. All of the above described information were also collected for trees < 12.54 cm (5 inches) dbh, but on a 2.07 m (6.8 ft) radius subplot concentric within the larger plot. Along 15 m transects starting from the center of the tree plots, shrubs (woody plants <2.54 cm) and ground cover data were also collected. Two 2 x 2 m plots were established along this transect for shrubs, with two

0.25 x 0.50 m plots established for ground cover within each shrub plot. In addition to this, data were also available on fire and silvicultural history in each plot.

For the 2008 re-measurement, the original year 2000 set of plots were stratified on the basis of community type, age class (10 year increment) and origin, and plots were randomly selected from each stratum. Altogether, 70 plots were chosen for re-measurement. The community types sampled were: Sandhills (SH), Mesic Flatwoods (MF), Upland Mixed Forest (UMF), Bottomland Forest (BF), Baygall (BG), and Swamp (Dome Swamp, Floodplain Swamp). Stand ages ranged from 9 years to 72 years, and were natural, plantation and of mixed origin.

During the 2008 re-measurement, we collected data on all of the variables measured in 2000. In addition, we collected more explicit information on tree crown variables. Both the total crown radius and the exposed crown radius of each tree was measured along four cardinal directions, as well as the degree of crown exposure (1 = exposed, 0 = covered). Crown radius was measured by extending a tape from the center of the bole to the end of the branch. In 2008, we also collected data on trees in a 6.40 m (21 ft) buffer around each tree plot so that the competition indices described below could be calculated for trees within the 2000 measurement plot boundaries.

Since crown radius data was not collected in 2000, we developed linear allometric equations relating 2008 mean crown radius to 2008 dbh, height and crown length using the data collected in 2008. This relationship was developed separately for two types of pines and four hardwood groups. Groupings of hardwoods corresponded to *Quercus* sp., *Taxodium* sp., the Cornaceae family, and other hardwoods present in the plot. The pine tree data consisted overwhelmingly of slash pine, with relatively fewer loblolly pine

(*Pinus taeda* L.) and only a few longleaf pine (*Pinus palustris* Mill.) and pond pine (*Pinus serotina* Michx.). A separate predictive model for crown radius was developed for slash pine, with longleaf and pond pine grouped together with loblolly pine. We used this relationship to estimate 2000 crown radius from 2000 dbh, height and crown length, adding a random error to each prediction. For brevity, results from this preliminary data step are not presented.

Since dbh and height of buffer trees were not available in 2000, we also developed a linear allometric relationship relating dbh and height in 2000 to dbh and height in 2008. Both 2000 variables were predicted from their 2008 variables using the other variable as a covariate (e.g., dbh for height and height for dbh) via linear regression. This relationship was developed for both pine trees and hardwood trees present in the plot, and the species groupings as described earlier for crown radius were used. A random error was added to each prediction of dbh and height. Since these predictions were made only to ensure that buffer tree information was available for calculation of competition indices, results from this preliminary data step are not presented.

Total crown area (TCA) was obtained by assuming an ellipsoid shape and utilizing crown radii in four cardinal directions. Similarly, exposed crown area (ECA) was calculated by summing up the area of ellipses calculated from the exposed crown radii in each cardinal direction. The % ECA was calculated by dividing ECA by TCA and multiplying by 100. Degree of crown exposure was coded as binary; if the crown radii were overtopped in any direction they were regarded as shaded (code 0), and if they were fully exposed they were treated as exposed (code 1).

Annual dbh growth calculation took into account the number of days between the two measurements. The difference between the two measurements was divided by total number of days during the two measurement periods, which was then multiplied by 365 to come up with an annual growth rate. Stocking of each plot was calculated as a measure of crowding following May (1990). Total basal area per hectare of each plot was also calculated. Similarly, site index for each plot was calculated following Farrar (1973).

Various distance dependent and independent competition indices were calculated for use in the model. Choi et al. (2007) defined the competitive zone as an area 3.5 times the average crown radius of dominant and codominant trees in their study of northern hardwoods. We modified this slightly for our pine-dominated stands, using 3 times the average radius of the measured slash pine trees (2.13 m; 7 ft), resulting in a competitive zone with 6.40 m (21 ft) radius around the subject tree. All variables indicating competition (both distance dependent and independent, as defined in Cole and Lorimer 1994, Choi et al. 2001, Choi et al. 2007), as described below, were calculated considering trees within this competition zone. Two relative diameter indices were calculated as:

$$RD_i = \frac{dbh_i}{\frac{1}{n} \sum_{j=1}^n dbh_j} \quad \text{and} \quad RDgh_i = \frac{dbh_i}{\frac{1}{n'} \sum_{j=1}^{n'} dbh_j}$$

where:  $dbh_i$  is the diameter at breast height of the  $i$ th subject tree,  $n$ =number of all competitors within the 6.40 m radius competitive zone, and  $n'$ =number of trees within the 6.40 m ft radius competitive zone with diameters greater than or equal to the subject tree. Therefore,  $RD$  is the relative diameter calculated with all trees within the 6.40 m

radius competitive zone of the subject tree, whereas  $RDgh$  is calculated only with trees equal to or larger than the subject tree. Relative height and relative basal area were calculated similarly.  $RHgh$  is the relative height calculated with trees within the 6.40 m radius competitive zone with heights greater than or equal to that of the subject tree, whereas  $RH$  is the relative height calculated with all trees within the 6.40 m ft radius competitive zone.  $RBAGh$  is the relative basal area calculated by including trees within the 6.40 m radius competitive zone with diameter greater than or equal to the subject tree, whereas  $RBA$  is the relative diameter calculated with all trees within the 6.40 m radius competitive zone of the subject tree.

Other competition indices investigated included the modified Hegyi indices for diameter and height ( $sum\_si$  and  $Sum\_Clr$ ; Heygi 1974), calculated as:

$$sum\_si_i = \sum_{j=1}^n \frac{dbh_j}{dbh_i d_{ij}}$$

where:  $dbh_j$  is the diameter of competitor tree  $j$ , and  $d_{ij}$  is the distance between the subject tree and its competing neighbors.  $Sum\_Clr$  is similar to  $sum\_si$ , using height as the variable instead of diameter:

$$Sum\_Clr_i = \sum_{j=1}^n \frac{H_j}{H_i d_{ij}}$$

We also calculated other measures of stand level competition such as  $BAL$  (basal area of trees equal to or greater in diameter than the subject tree), and quadratic mean diameter for each plot ( $qmd$ ; Teck and Hilt 1990, Cao 2002, Adame et al. 2008).

## Mixed Modeling of Growth

Growth models are commonly constructed from data collected in fixed area plots, where plots might be from the same stand or from different stands located in different areas. Since multiple trees are typically measured within each fixed area plot, there is a lack of independence among the data. That is, observations from trees within the same plots and plots within same stands are likely to be more highly correlated than those that are located in different plots located farther away (West et al. 1984). While ordinary least square methods will give unbiased estimates of model parameters in this kind of correlated data structure, standard errors and test statistics associated with correlated data will be biased and inappropriate for interpretation (Calama and Montero 2004). Mixed models have been proposed as a way to properly account for correlated data. Mixed model estimation techniques estimate fixed and random parameters simultaneously, and give unbiased and efficient estimates of fixed parameters (Calama and Montero 2004). Incorporating random effects in the model also increases the accuracy of fixed effect prediction and testing (Pinheiro and Bates 2000, Tao and Littell 2002, Budhathoki 2008). Details about mixed model procedures can be found in Vonesh and Chinchilli (1997), Pinheiro and Bates (2000), Calama and Montero (2004) and Meng and Huang (2009).

A mixed model can be written in a general form

$$\mathbf{y}_i = f(\mathbf{x}_i, \boldsymbol{\beta}, \mathbf{b}_i) + \boldsymbol{\varepsilon}_i \quad (2-1)$$

where  $\mathbf{y}_i = [y_{i1}, y_{i2}, \dots, y_{in}]'$  is a vector of tree diameter growth measurements in plot  $i$ ,  $\mathbf{x}_i$  is a design matrix of the covariates (dbh, height, competition index, etc.),  $\boldsymbol{\beta}$  is a vector of fixed parameters, and  $\mathbf{b}_i$  is a vector of random effects parameters.  $\boldsymbol{\varepsilon}_i$  is a vector of errors

$[\boldsymbol{\varepsilon}_{i1}, \boldsymbol{\varepsilon}_{i2}, \dots, \boldsymbol{\varepsilon}_{in}]'$ , and the random effect parameter ( $\mathbf{b}_i$ ) and error ( $\boldsymbol{\varepsilon}_i$ ) are assumed to be normally distributed and uncorrelated. Their distribution can be written as:

$$E \begin{bmatrix} \mathbf{b}_i \\ \boldsymbol{\varepsilon}_i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{Var} \begin{bmatrix} \mathbf{b}_i \\ \boldsymbol{\varepsilon}_i \end{bmatrix} = \begin{bmatrix} \mathbf{D} & 0 \\ 0 & \mathbf{R}_i \end{bmatrix}$$

where  $\mathbf{D}$  is the variance-covariance matrix for the random effect and  $\mathbf{R}_i$  is the variance covariance matrix of the errors for plot  $i$ .  $\mathbf{R}_i$  is a matrix written as  $\sigma^2 \mathbf{I}_n$ , where  $\sigma^2$  is the error variance and  $\mathbf{I}_n$  is a  $n \times n$  identity matrix. SAS procedures such as PROC MIXED and PROC NL MIXED (SAS Institute Inc., 2008) fit models via maximum likelihood by integrating an approximation of likelihoods over the random effects. Various integral approximations of likelihood are available, and in the present study we used the first order method of Beal and Sheiner (1982). This method uses a Taylor series expansion of equation 2-1 around a  $\beta^*$  close to  $\beta$  and a  $b_i^*$  set to zero (the empirically best linear unbiased predictor) for linear approximations of marginal likelihood functions. The predictor,  $\hat{\mathbf{b}}_i$ , of  $\mathbf{b}_i$  can be estimated using the following (Davidian and Giltinan 1995, Vonesh and Chinchilli 1997):

$$\hat{\mathbf{b}}_i = \widehat{\mathbf{D}} \mathbf{Z}_i' (\mathbf{Z}_i \widehat{\mathbf{D}} \mathbf{Z}_i' + \widehat{\mathbf{R}}_i)^{-1} [\mathbf{y}_i - \mathbf{f}(\mathbf{x}_i, \widehat{\boldsymbol{\beta}}, \mathbf{0})] \quad (2-2)$$

where  $\widehat{\mathbf{D}}$  and  $\widehat{\mathbf{R}}_i$  are the estimates of  $\mathbf{D}$  and  $\mathbf{R}_i$ ; and  $[\mathbf{y}_i - \mathbf{f}(\mathbf{x}_i, \widehat{\boldsymbol{\beta}}, \mathbf{0})]$  is the residual where  $\mathbf{f}(\mathbf{x}_i, \widehat{\boldsymbol{\beta}}, \mathbf{0})$  is the prediction of the fixed effect part of the model.  $\mathbf{X}_i$  and  $\mathbf{Z}_i$  are partial derivative matrices of equation (2-1):

$$\mathbf{X}_i = \frac{\partial \mathbf{f}(\mathbf{x}_i, \boldsymbol{\beta}, \mathbf{0})}{\partial \boldsymbol{\beta}'}$$

$$\mathbf{Z}_i = \frac{\partial \mathbf{f}(\mathbf{x}_i, \boldsymbol{\beta}, \mathbf{0})}{\partial \mathbf{b}_i}$$

## Model Fitting

As a preliminary exploratory analysis, we fit a variety of linear and nonlinear models to the data. Among these models, three models were selected for detailed analysis. The first model, similar to Huang et al. (1995), takes into account the unimodal, right skewed relationship between the size of a tree (diameter) and diameter growth. The model form uses the Box Lucas function (Box and Lucas 1959, Huang et al. 1995) as its base (equation 2-3), and the parameters of the base model were related to other tree and stand level variables using a parameter prediction approach (equations 2-4 and 2-5; Clutter et al. 1983). This model is hereafter referred to as Model 1.

$$\Delta dbh_{ij} = \frac{\theta_2}{\theta_1 - \theta_2} * (\exp(-\theta_2 dbh_{ij}) - \exp(\theta_1 dbh_{ij})) \quad (2-3)$$

$$\theta_1 = f(X_1, X_2, X_3, \dots) \quad (2-4)$$

$$\theta_2 = f(X_1, X_2, X_3, \dots) \quad (2-5)$$

where  $\Delta dbh_{ij}$  is annual diameter growth and  $dbh_{ij}$  is the initial diameter of the  $j$ th tree in the  $i$ th plot, and  $X_1$ ,  $X_2$ , and  $X_3$  are the covariates associated with individual trees and plots. Thus, under the parameter prediction approach, both  $\theta_1$  and  $\theta_2$  are functions of covariates to be selected. In Model 1, we tested as covariates in (2-4)-(2-5) all distance independent and distance dependent competition indices described earlier. In addition, we included the size variables dbh, total tree height, mean crown radius and total crown area, and also tested stand level variables such as total basal area per hectare, quadratic mean diameter and percent stocking of plots.

We also fit a linear model to the square root of diameter growth using an approach similar to those of Martin and Ek 1984, Trasobares et al. 2004, and Weiskittel et al.

2007, hereafter referred as Model 2:

$$\sqrt{\Delta dbh_{ij}} = \mu + \beta_1 dbh_{ij} + \beta_2 H_{ij} + \beta_3 sum_{si_{ij}} + \beta_4 SPI_i + \beta_5 tbaha_i + b_i + e_{ij} \quad (2-6)$$

where  $\beta_1, \beta_2, \beta_3, \beta_4$  and  $\beta_5$  are parameters to be estimated,  $b_i$  is the random plot effect of the  $i$ th plot and  $e_{ij}$  is the error term for the  $j$ th tree in the  $i$ th plot. Also,  $H_{ij}$  is the total height of the  $j$ th tree in the  $i$ th plot,  $SPI_i$  is the site index of the  $i$ th plot,  $sum_{si_{ij}}$  is the competition index of the  $j$ th tree in the  $i$ th plot, and  $tbaha_i$  is the total basal area per hectare for the  $i$ th plot. In addition, various combinations and transformations of the independent variables were also tested in this model.

Finally, we fit a modified version of the loblolly pine model used by Amateis et al. (1989) and Cao et al. (2002). This model is hereafter referred as Model 3.

$$\Delta dbh_{ij} = \beta_1 CR_{ij}^{\beta_2} H_{ij}^{\beta_3} e^{(\beta_4(1-(qmd_i/dbh_{ij})))} SPI_i^{\beta_5} tbaha_i^{\beta_6} + e_{ij} \quad (2-7)$$

where  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  and  $\beta_6$  are model parameters to be estimated,  $CR_{ij}$  is the crown radius for the  $j$ th tree in the  $i$ th plot,  $qmd_i$  is the quadratic mean diameter in the  $i$ th plot, and all other variables are as described earlier. In the above model, we also included various transformations and combinations of the independent variables. In addition to the  $qmd/dbh$  ratio in the above model, which describes competition, we tested all the distance dependent and independent competition indices mentioned earlier. In addition to crown radius, we also tested  $dbh$  and total crown area as the tree size variable. For the measure of stand density, we also tested  $\%stocking$  and  $qmd$  of the plot instead of  $tbaha$ .

While it is straightforward to include a random plot effect ( $b_i$ ) into the linear model (Model 2), it is not obvious how a random effect should be incorporated into the non-linear models (Models 1 and 3). Indeed, the random effect could be incorporated to relate to any of the independent variables, e.g., crown radius, diameter, or basal area, by adding the effect to the independent variable's coefficient. For instance, Budhathoki et al. (2008) incorporated a random effect as related to tree basal area, whereas Adame et al. (2008) incorporated random effects related to the intercept and individual tree diameter. The choice of model form random effects is an additional step in the process of fitting these non-linear models. Moreover, since a primary purpose of this model is prediction, it is not clear if model selection procedures should include random effects. That is, given we want to predict diameter growth for a new plot, there is no random effect available to characterize that new plot data.

When prior observations on growth are not available (e.g., when predicting values for a new plot), it is very difficult to estimate plot-level random effects and predictions based on random effects are often inappropriate. In this situation, one may choose to calculate the population averaged responses (PA; overall mean response), or the typical mean responses (TM). PA responses are estimated when predictions are made from a model fit with fixed effects only, i.e. without including the random effects in the model. TM responses are estimated when predictions are made from a mixed model fit with both fixed and random effects, ignoring the random effect. TM responses for a non-linear mixed model are comparable to PA responses for a linear mixed model, because PA and TM responses in a linear mixed model are similar (Davidian and Giltinan 1995, 2003; Fitzmaurice et al. 2004). However, these TM responses are inferior

in prediction versus PA responses (Meng et al. 2009), and therefore, a fixed effects model should be used when comparing non-linear and linear models.

Models 1-3 were therefore first fit with fixed effects only using data from all seventy plots. The linear model (Model 2) was fit with fixed effects only using a generalized linear least squares estimation technique via the SAS procedure PROC MIXED. For the non-linear models (Models 1 and 3), a model including only the fixed effects was first fit using a non-linear least squares technique via the SAS procedure PROC NLIN. This preliminary step is also useful for obtaining a more parsimonious set of potential predictors.

To choose the best form of Models 1, 2, and 3, we started with the full suite of covariates. We then sequentially dropped non-significant ( $\alpha=0.05$ ) variables, each time examining Akaike's Information Criteria (AIC) and the Bayesian Information Criteria (BIC). Where dropping variables did not lower AIC and BIC, we retained non-significant variables. To obtain AIC and BIC, the fixed effects non-linear models were fit using maximum likelihood methods via the SAS procedure PROC NLMIXED.

Once the best form of each of Model 1, 2, and 3 was chosen, we compared the three models using the model evaluation criteria given below. To this best overall model we then introduced random effects, accounting for groupings of trees within plots.

### **Model Evaluation**

The model fit was compared between Models 1-3 by calculating the following statistics for each plot following Arabatzis and Burkhart (1992) and Calama and Montero (2004):

Mean residual (Bias),  $\bar{E}_i$ , for plot  $i$ :

$$\bar{E}_i = \frac{\sum_{j=1}^{n_i} e_{ij}}{n_i}$$

Model precision,  $V_i$ , for plot  $i$ :

$$V_i = \frac{\sum_{j=1}^{n_i} e_{ij}^2 - \left[ \sum_{j=1}^{n_i} e_{ij} \right]^2 / n_i}{(n_i - 1)n_i}$$

Mean square error,  $MS_i$ , for plot  $i$ :

$$MS_i = \bar{E}_i^2 + V_i$$

Mean square error of prediction,  $MSE_i$ , for plot  $i$ :

$$MSE_i = \frac{\sum_{j=1}^{n_i} (e_{ij})^2}{n_i - 1}$$

Percentage bias,  $pbias_i$ , for plot  $i$ :

$$pbias_i = \frac{\bar{E}_i}{\bar{y}_i} \times 100\%$$

where  $e_{ij}$  represents residual (observed dbh growth minus predicted dbh growth) of the  $j$ th tree in plot  $i$ ,  $\bar{y}_i$  is the mean of the observed values for plot  $i$ , and  $n_i$  represents the number of trees in the  $i$ th plot. Once plot-level values of each fit statistic were calculated, mean values for each statistic were calculated over all 70 plots. The mean values are hereafter referred to as E, V, MS, MSE, and PBIAS. In addition, we also calculated the root mean squared error (RMSE) of each model, and the standard deviation of bias. Finally, plots of residual versus predicted values were plotted to ensure satisfaction of the assumptions of normality and homoscedasticity of the residuals.

Independent data were not available for validation; neither were there enough data available to subdivide data into model fitting and validation datasets. Therefore, we used a leave-one-out cross validation approach (Temesgen et al. 2008) to evaluate the model predictions. Since our data are naturally grouped by plot, we modified this approach so that plots rather than individual trees were left out in each step; we deleted a plot from the data set, fit the model with the remaining data, and then used the estimated parameters to predict the growth of trees in the deleted plot. This was done for all 70 plots in the dataset. Observed and predicted values were then used to calculate model fit statistics. Based on those statistics, the predictive ability of those models was determined.

### **Growth among Natural Communities**

We grouped plots into natural communities using the palustrine ecosystem classifications of the Florida Natural Areas Inventory (1990). Bay gull (BA, Seepage wetlands), Basin Swamp (BS), Dome Swamp (DS) and Bottomland Forest (BF, Floodplain Wetlands) were grouped together into one category called Basin wetlands (BW). Our Sandhill category included both plots classified as Sandhill (SA), and Xeric Hammock (XH), as both classes belong to Xeric Upland community and the XH class contained only 8 trees. Mesic Flatwood (MF) and Upland Mixed Forest (UMF) were left as separate categories. Community types were included as a fixed class effect in the models selected in model fitting section; this effect was then examined for significance ( $\alpha=0.05$ ) to determine whether a specific model was required for each community or if communities could be combined into a single model.

## Results

Descriptive statistics describing the dependent and independent variables in the model are shown in Table 2-1. The size variables tested in the models included dbh, total height, mean crown radius, and total crown area. Among the size variables, total height had the highest absolute correlation with annual dbh growth (-0.37;  $P < 0.05$ ), followed by current dbh (-0.10;  $P < 0.05$ ) and mean crown radius (-0.03; not significant at 5% level). Stand level variables tested in the model included site index, quadratic mean diameter, total basal area per hectare, % stocking and stand age. Site index had a negative correlation with dbh growth (-0.25;  $P < 0.05$ ), as did variables indicating stand level competition such as total basal area per hectare (-0.28;  $P < 0.05$ ), quadratic mean diameter (-0.30;  $P < 0.05$ ), and % stocking (-0.28;  $P < 0.05$ ). The competition variables tested in the models were *RDgh*, *RD*, *sum\_si*, *Sum\_Clr*, *RHgh*, *RH*, *RBAgd*, and *RBA*. Among the competition variable, *RHgh* (relative height calculated based on the trees equal in height and taller than the subject tree) had the strongest (-0.25;  $P < 0.05$ ) correlation with growth followed by *sum\_si* (-0.15;  $P < 0.05$ ).

## Models

We tested Models 1-3 to predict diameter growth. Model 1 used the approach of Huang and Titus (1995) with the Box-Lucas base function and a parameter prediction approach. In equations 2-4 and 2-5, we tested different variables and their combinations. The best form of Model 1 included the following forms for equations (2-4) and (2-5):

$$\theta_1 = a_0 + a_1 H_{ij} + a_2 \text{sum\_si}_{ij} \quad (2-8)$$

$$\theta_2 = c_1 \text{SPI}_i \quad (2-9)$$

Estimated parameters for Model 1 with fixed effects only are shown in Table 2-2. Mean crown radius and total crown area (tree crown as a proxy of tree size) were not better predictors in the model compared to dbh and total height, as models including dbh and total height gave smaller AIC and BIC values than those that included mean crown radius or total crown area. Because the combination of height and age determines site index (Quicke et al. 1994, Lee et al. 2004), we also fit equation 2-7 without height. Since the model fit was worse, however, we kept height in the model. Several competition variables described above were tested in Model 1; the best model according to fit statistics such as AIC and BIC was obtained by using *sum\_si* as the competition index. When testing variables representing stand level competition, including *tbaha* gave smaller AIC and BIC values than including either %stocking or *qmd*, but since removing it resulted in an even lower AIC, it was removed from the model.

In model 2, we tested all of the independent variables listed in equation (2-6). We also included the log and inverse of the independent variables in an effort to capture the non-linear nature of these relationships with growth. However, the best model based on AIC and BIC values did not include any transformed variables. Interaction terms that made biological sense were also tested in the model. The terms *dbh \* sum\_si*, *dbh \* SPI*, and *sum\_si \* SPI* were significant and were therefore included in the model. Several competition indices were also tested in this model, and the best model (lowest AIC, BIC) was obtained by including *sum\_si*. Estimated parameters for Model 2 are shown in Table 2-3, which had the form:

$$\begin{aligned}
\text{Sqrt}(\Delta dbh_{ij}) = & \mu + \beta_1 dbh_{ij} + \beta_2 H_{ij} + \beta_3 sum\_si_{ij} + \beta_4 SPI_i + \beta_5 tbaha_i + \beta_6 dbh_{ij} \\
& * sum\_si_{ij} + \beta_7 dbh_{ij} * SPI_i + \beta_8 sum\_si_{ij} * SPI_i + e_{ij}
\end{aligned}
\tag{2-10}$$

For Model 3 fit with fixed effects only, the model form given in equation 2-11 gave the best fit statistics based on AIC and BIC. For example, mean crown radius was tested as the size variable instead of dbh, but the model with dbh gave lower AIC and BIC. Estimated parameters for Model 3 are shown in Table 2-4.

$$\Delta dbh_{ij} = dbh_{ij}^{\beta_1} H_{ij}^{\beta_2} e^{(\beta_3(1-sum\_si_{ij}))} SPI_i^{\beta_4} + e_{ij}
\tag{2-11}$$

The three models were compared using the various fit statistics described in Section 2.4 (Table 2-5). For comparison purposes, fit statistics were generated with models including only fixed effects. Among the three models, Model 3 (equation 11) outperformed the other models based on these statistics (Table 2-5). Its precision statistic (V) was similar to other models, but it had much lower bias (E; Table 2-5). The percentage bias was also lower for Model 3 than for the other models. Model 3 had lower MSE and RMSE than Model 1, but these values were slightly higher than that of Model 2 (Table 2-5). The standard deviation for bias calculated for each plot was also lower for Model 3. Plots of residual versus predicted values for all the three models showed no systematic pattern (Figure 2-2, Figure 2-3, and Figure 2-4), indicating homogeneity of variance and no systematic bias in prediction.

In order to confirm the above results, fit statistics were re-calculated using the cross-validation approach. The results indicated that Model 3 performed better than the other models (Table 2-6). Precision statistics (V) for Model 3 were similar to Model 2, but model 3 had significantly lower bias (E) and percentage bias (PBIAS) than the other

models. The root mean square error (RMSE) of Model 3 was slightly lower than that of Model 2.

Model 3 was chosen as our best model, and is recommended for use in slash pine stands in this region of Florida. As expected, initial diameter had a positive effect on the growth, and tree height had a negative effect on tree diameter growth (Table 2-4; equation 2-11). The parameter estimate for  $\beta_3$  was positive, which indicated that the higher the competition, the lower the tree growth; the parameter for site index ( $\beta_4$ ) was positive, indicating that better sites had faster growth.

### **Mixed Modeling of Growth**

Due to the hierarchical nature of the data (with trees grouped within plots), we needed to account for within-plot correlations. The best model (Model 3; equation 2-11) was further developed by incorporating a random effect for plots. A likelihood ratio test and fit statistics were used to determine whether incorporation of random effect improved the model fit. First, random effects were incorporated into the coefficients related to both tree level variables (dbh and height) in equation 2-11. When both random effects were included, the model failed to converge; thus, random effects were incorporated into the coefficient related to either height or dbh in equation 2-11. Between these two choices of random effects, incorporation into the coefficient related to tree height gave a better model fit than doing so with dbh. When the random effect was incorporated into the model (equation 2-12), the parameter estimate for the plot level site index variable (*SPI*) was insignificant, and this reduced model also had lower AIC and BIC values; therefore, it was removed from the final model (equation 2-12).

$$\Delta dbh_{ij} = dbh_{ij}^{\beta_1} H_{ij}^{\beta_2 + b_i} e^{(\beta_3(1 - sum\_si_{ij}))} + e_{ij} \quad (2-12)$$

where all other variables are as defined above and  $b_i$  is the random plot effect.

The likelihood ratio test was performed to justify the significance of the random effect in the model (Pinheiro and Bates 2004, Nothdurft 2006). The likelihood ratio (LR) was calculated as:

$$LR = -2(LL_{urm} - LL_{rm}) \quad (2-13)$$

where:  $LL_{urm}$  is log likelihood value of unrestricted model (fixed effects only) and  $LL_{rm}$  is log likelihood value of the restricted model (with random effects included). The resultant value (LR) was compared to a chi square distribution with degrees of freedom equal to the number of additional parameters estimated by the restricted model. The calculated likelihood ratio statistics was highly significant at 0.05 significance level ( $P < 0.0001$ ), justifying the inclusion of the random effect in the model. Plots of the residual versus predicted values for the random effects model also did not indicate heteroskedasticity or lack of normality. Parameter values for Model 3 with the random effect are shown in Table 2-7.

### **Cross-validation of Mixed Effects Model**

The model including both fixed and mixed effects was also cross validated. For cross validation of the mixed effects model, the predictor of the random parameter  $\hat{b}_i$  should be estimated for each deleted plot for prediction of growth for that plot. Since calculating a random effect for a new plot is an important step for using a mixed effect model (Huang et al. 2009a), we present a method in this section to estimate  $\hat{b}_i$  similar to other studies on mixed models (Calama and Montero 2004, Huang et al. 2009a). . First, the fixed effects parameters for the deleted plot were estimated using the SAS procedure PROC NLMIXED with the parameter for the random effect,  $b_i$ , set to zero. To

estimate the random parameter  $b_i$ , we then calculated the derivative of equation 2-12 with respect to parameter  $b_i$

$$\frac{\partial \Delta dbh}{\partial b_i} = \frac{\partial f(\mathbf{X}_i, \hat{\boldsymbol{\beta}}, 0)}{\partial b_i} = dbh_{ij}^{\beta_1} * H_{ij}^{\beta_2} * \ln(H_{ij}) * \exp(\beta_3(1 - sum_{s_{ij}})) \quad (2-14)$$

where  $\ln$  is the natural log, and all other variables are as described earlier. The derivatives thus calculated (equation 14) provided the  $\mathbf{Z}$  matrix (equation 2-2; see Section Mixed Modeling of Growth). The  $\mathbf{Z}$  matrix is shown for an example plot with six trees below. Since the model in the present study had only one random variable, the variance of random parameter is a scalar:  $D = \sigma^2_{b_i}$ . The  $\mathbf{R}_i$  matrix is an  $n \times n$  (where  $n$  is the number of trees in a plot) matrix as given below with error  $\sigma^2$  variance on the diagonals and zero elsewhere (since the inclusion of a random plot effect incorporates correlation between trees in a plot, covariance between trees is assumed to be zero).

$$\mathbf{z}_i = \begin{bmatrix} 1.340 \\ 1.149 \\ 1.181 \\ 1.175 \\ 1.497 \\ 1.069 \end{bmatrix} \quad \text{and} \quad \hat{\mathbf{D}} = 0.0344$$

$$\hat{\mathbf{R}}_i = \begin{bmatrix} 0.05365 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.05365 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.05365 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.05365 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.05365 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.05365 \end{bmatrix}$$

Once the  $\mathbf{Z}$  matrix was calculated, the random parameter for each deleted plot was calculated using equation 2-2 (see Section Mixed Modeling of Growth). The fixed effect prediction ( $\Delta dbh_{fix_{ij}}$ ) of the model was given by the following equation with  $b_i$  (random parameter) set to zero:

$$\Delta dbh_{fix_{ij}} = dbh_{ij}^{\beta_1} H_{ij}^{\beta_2 + b_i} e^{(\beta_3(1 - sum_{s_{ij}}))} \quad (2-15)$$

Values for  $\beta_1, \beta_2$ , and  $\beta_3$  were taken from Table 2-5, and values for  $\hat{D}$  and  $\hat{R}_i$  were estimated by fitting equation 2-12 to the data. The residual  $[y_i - f(x_i, \hat{\beta}, 0)]$  part of the equation 2-2 was calculated by subtracting  $\Delta dbh_{fix_{ij}}$  values from observed annual diameter growth values for the  $i$ th plot.

The comparison of Model 3 with fixed effects only and with the random effect also justified the inclusion of plot as a random effect in the model (Table 2-8). For both the model fit and the cross validation, the mixed effects model had better fit statistics than that of the model with fixed effects only. In terms of fit statistics, the fixed effects model had lower bias (E) and PBIAS than the mixed effects model, but all other statistics (precision, RMSE, SD bias) were better for the mixed effects model (Table 2-8). When the predictive ability (using the cross validation approach) of Model 3 with and without the random effect was compared, the mixed effects model performed better than the fixed effects model (Table 2-8). The mixed effects model had higher precision, lower bias, lower RMSE, lower PBIAS and lower SD bias than the fixed effects model (Table 2-8). Diameter predictions using the mixed effects model were closer to the observed values than the fixed effects only model (Figure 2-5).

### **Difference in Growth between Natural Communities**

In order to determine differences between communities with respect to diameter growth, community types were included as a fixed effect in equation (2-12). Three dummy variables ( $NC1$ ,  $NC2$ , and  $NC3$ ) were created to describe the four community types (Mesic flatwood, Sandhill, and Upland Mixed Forest):

$$\Delta dbh_{ij} = dbh_{ij}^{\beta_1} H_{ij}^{\beta_2 + b_i} e^{(\beta_3(1 - sum_{s_{ij}}))} + m_1 NC1 + m_2 NC2 + m_3 NC3 + e_{ij} \quad (2-16)$$

where:  $m_1$ ,  $m_2$ , and  $m_3$  are parameters to be estimated, and all other variables are as described earlier. As suggested by Calama and Montero (2004), we performed a likelihood ratio test to determine whether the full model including community type was superior to that of a model combining different community types together. A likelihood ratio test of the constrained model versus the full model tested the hypothesis of equal community effect. A non-significant likelihood ratio test would indicate that a constrained model was a better fit than a full model, whereas a significant likelihood ratio test would indicate that at least one community should have its own community-specific parameter.

When all dummy variables were included in the model, all parameters associated with the dummy variables were insignificant in the full model. Removing non-significant dummy variables had the effect of combining natural communities. A model combining MF (Mesic Flatwood), SA (Sandhill) and BW (Basin Wetlands), and a separate model for UMF (Upland Mixed Forest) was indicated (Table 2-9). Results of a likelihood ratio test indicated that this model was significantly different from the model with all communities combined. Moreover, all parameters in the model were significant, and the model had the lowest AIC and BIC values (Table 2-9).

The model for UMF is given by the following equation, the parameter values are presented in Table 2-10.

$$\Delta dbh_{ij} = dbh_{ij}^{\beta_1} H_{ij}^{\beta_2 + b_i} e^{(\beta_3(1 - sum_{s_{ij}}))} + m_3 NC3 + e_{ij} \quad (2-17)$$

## Discussion and Conclusion

Individual tree growth models allow for the inclusion of competitive effects of individual trees; therefore, they are suitable for forecasting tree growth regardless of species mixture, age distribution, and applied silvicultural systems (Hasenauer 2006). They are also more suitable for detailed descriptions of stand structure and dynamics than stand level models (Vanclay 1994, Chojnacky 1997). The model proposed here is suitable for simulating stand conversion processes, which is the interest of forest managers who have acquired plantations and want to convert them into more structurally diverse conditions.

Tree locations are useful for calculating competition indices – an important component of distance dependent models. Competition indices are important in explaining variation in stand structure and are useful tools to predict development of an individual tree (Daniels et al. 1986, Pukkala and Kolström 1991). In this study, various distance dependent and independent competition indices were tested. Relative height calculated with trees taller than the subject tree (*RHgh*) had the highest correlation with diameter growth, followed by the distance dependent competition index similar to Hegyi's index (Hegyi 1974), *sum\_si*. Although *sum\_si* did not contribute significantly in the model, AIC and BIC values associated with having it in the model were lower than those with *RHgh*. *Sum\_Clr* was a significant predictor in the model; however, AIC and BIC values for models including it were higher compared to those of the model with *sum\_si*. *BAL* in large trees, which is described as capturing the competition for light (Adame et al. 2008, Schwinning and Weiner 1998), was also tested in the model, but did not give a better fit than *sum\_si*. In terms of information criteria, the distance

dependent competition index *sum\_si* gave a better model fit for diameter growth in slash pine than the distance independent indices.

The relationship between diameter increment and tree size (dbh) is usually unimodal right skewed (Huang et al. 1995, Adame et al. 2008). Growth slowly peaks toward a maximum and starts to decrease slowly when trees gets older. We had only trees 12.7 cm (5 inches) or larger; therefore, we expected a concave downward relationship between growth and tree size. We tested this relationship in the linear model using dbh and its inverse (after Martin and Ek 1984), and also with Model 1. In Model 3, which was the best performing model form in this study, we tested a logistic relationship of dbh with growth, and also included diameter and inverse of diameter in the same model, but this did not result in a better fit. We concluded that diameter of pine in our study slowly increases until ~61 cm (24 inches) dbh (the largest dbh sampled in the study), and levels off around that point (Figure 2-6). If we had sampled trees with diameters greater than 61 cm (24 inches), we might have seen a downward growth trend in larger size classes, as seen in other studies (e.g., Martin and Ek 1984). Moreover, our data included thinned plots, and the increase in growth due to an increase in growing space after thinning also might have changed the relationship between growth and tree size. Cao et al. (2002) also reported a positive coefficient for dbh, indicating a positive effect of dbh on diameter growth of loblolly pine.

The best model (Model 3) showed a negative effect of tree height on diameter growth (Figure 2-6), indicating that taller (likely older) trees have slower growth compared to smaller trees, a result also reported by Adame et al. (2008). When comparing trees with similar site index, taller trees are found in more mature stands,

hence they have slower growth rate and there is an overall negative effect of height on diameter growth (Adame et al 2008, Calama and Montero 2005).

In all the models tested, as well as the selected model, total basal area was included as a measure of density. In most of the models and the final model, this variable was insignificant. We suspect that *sum\_si* (a distance dependent competition index), which already includes distance and size of neighbors, incorporated the effect of density on growth, hence eliminating the need for total basal area in the model.

The use of a mixed modeling approach in forest growth models is a recent phenomenon in forestry (Trasobares et al 2004, Weiskittel 2007, Adame et al. 2008, Budhathoki 2008). Mixed models account for the correlation between trees within the same plot and ensure the assumption of independence is met. The better fit statistics of the mixed effects model (Table 2-8) also justified the inclusion of the random effect in the model. The predictive ability of the mixed effects model (as determined by the cross validation approach) was also better than the fixed effects model (Table 2-8). Since random effects incorporate all other possible unaccounted effects of plot level variables that could affect the growth (e.g., topography, soil, microclimate, nutrient, moisture etc. [Huang et al. 2009b]), it gives better predictions than a fixed effects model. The insignificant coefficient of site index (a plot level variable), once the random effect was included in the model, also supports this conclusion. The random effect in essence incorporated the effect of site; hence site index was not required in the model.

Similar results were observed by Huang et al. (2009b) when they developed their height-diameter relationship and Yang et al. (2009) in their basal area increment model. The base model of Huang et al. (2009b) was based on dbh only, and additional models

were also developed including stand level variables in addition to dbh. When random parameters were included in the model, the addition of stand level variables did not improve the model predictions in comparison with the base model. They concluded that if enough prior observations were available to estimate random parameters for the plots used in prediction, a mixed effects model with dbh only was better than a model with dbh and stand level variables. Adame et al. (2008) also used diameter growth increment data for some of the trees from a plot in validation data to estimate the value of random parameters for that plot, and used those random parameters to predict growth of the rest of the trees in the plot. They suggested using at least two prior observations to estimate the value of random parameters. Thus, if we are interested in population averaged predictions (i.e., only using the fixed parameters of the non linear mixed models), then the addition of site index in the model might improve the accuracy of predictions. This is similar to linear least squares regression where additional variables will always improve model predictions. If we want to predict subject-specific prediction (plot level), then inclusion of the random parameter will improve accuracy, and there is no need to include site index, the effect of which is already accounted for by the random parameter.

Interestingly enough, Mesic flatwoods, Bottomland Swamps and Sandhill had similar and higher diameter growth than Upland Mixed Forest. Swamps and mesic flatwoods are seasonally inundated and have low nutrients and acidic soil. Slash pine is historically better adapted to these conditions (Abrahmson and Hartnett 1990), as the growth of slash pine is related to moisture, especially in the drier areas such as sandhill site (Foster and Brooks 2001). Foster and Brook (2001) found higher basal area

increment of slash pine in mesic flatwood sites where water table is less than 1 m compared with transition zones (areas between mesic flatwoods and sandhill) where water table is >2m. We expected that a model where MF and BW were combined together would be significant and that we would see higher growth rates there versus those in SA and UMF, but the parameters of this model were not significant. MF and BW are mesic sites and therefore not water limited; hence, they exhibit higher growth rates. Sandhills and Upland mixed forests have lower moisture content in the soil, greater depth to the water table, and higher nutrients compared to swamps (wet flatwoods) and mesic flatwoods; therefore, lower growth of slash pine is expected in these communities.

The distance dependent diameter growth model presented here will be further used for simulating conversion of even-aged stand of slash pine into uneven-aged stand. The model incorporates the effect of initial size (dbh and height) along with individual tree competition. Stand level variables, including measures of stand density, were not significant in the model and did not improve model fit. The effect of environment as measured by site index improved the fixed effects model, but once the random effect was included, this variable was not significant. This implies that the random effect incorporated a variety of both unmeasured and measured plot level effects. A variety of goodness of fit statistics confirmed the statistical precision and accuracy of our chosen model. Although we could not cross validate our model with an independent dataset, our results were confirmed via leave-one-out cross validation. In the future, the model should be validated using an independent validation data. This

model will help land managers desiring to promote uneven-aged stands with structural characteristics that are compatible with biodiversity conservation objectives.

Table 2-1. Descriptive statistics for slash pine dataset (70 plots).

Variables	N	Mean	Std Dev	Min	Max
Annual dbh growth ( $\Delta dbh$ ; cm)	264	0.51	0.368	0	1.96
Total basal area per hectare ( $tbaha$ ; m <sup>2</sup> /ha)*	264	32.64	16.7	5.99	103.79
Quadratic mean diameter ( $qmd$ ; cm)	264	18.09	3.67	11.55	32.57
Age	264	27	61	10	72
Size of trees (dbh; cm)	264	19.05	6.79	12.70	61.72
Height ( $H$ ; m)	264	15.49	4.90	7.31	32.0
Site Index ( $SPI$ ; m)	264	21.15	4.90	11.2	35.90

\*includes basal area of all the hardwoods and pine trees in the plot

Table 2-2. Parameter estimates and their respective standard errors (SE) and P-values for diameter growth Model 1 with fixed effects only.

Parameter	Estimate	SE	Pr> t
$a_0$	0.1398	0.0258	<0.0001
$a_1$	-0.0038	0.00063	<0.0001
$a_2$	-0.00632	0.00275	0.0223
$c_1$	0.00116	0.00046	<0.0001

Table 2-3. Parameter estimates and their respective standard errors (SE) and P-values for diameter growth Model 2 with fixed effects only.

Parameters	Estimate	SE	Pr> t
Intercept ( $\mu$ )	0.6411	0.2311	0.005
$\beta_1$	0.0268	0.0084	0.001
$\beta_2$	-0.0397	0.0049	<0.0001
$\beta_3$	0.0044	0.0489	0.928
$\beta_4$	0.0324	0.0125	0.01
$\beta_5$	-0.00271	0.00113	0.0181
$\beta_6$	0.0066	0.0021	0.0029
$\beta_7$	-0.00100	0.00044	0.024
$\beta_8$	-0.0066	0.0020	0.001

Table 2-4. Parameter estimates and their respective standard errors (SE) and P-values for Model 3 fit with fixed effects only.

Parameter	Estimate	SE	Pr>t
$\beta_1$	1.1738	0.1626	<0.0001
$\beta_2$	-1.9214	0.1692	<0.0001
$\beta_3$	0.0987	0.0289	0.0007
$\beta_4$	0.3895	0.01252	0.0021

Table 2-5. Fit statistics used to compare the three annual diameter growth models. Models are given in the text.

Fit Statistics	Model 1	Model 2	Model 3
V	0.010819	0.011042	0.010959
E	0.038825	0.070362	0.028613
MS	0.1438	0.11569	0.111
MSE	0.21797	0.17771	0.17696
RMSE	0.37049	0.31889	0.32506
PBIAS	7.15	12.97	5.27
SD(bias)	0.36	0.31	0.31

Note: V = Model Precision; E = Mean Residual (Bias); MS = Mean Square Error; MSE = Mean Square Error of Prediction, RMSE = Root Mean Square Error, PBIAS = Percentage bias, SD (bias) = Standard deviation of bias.

Table 2-6. Fit statistics used to compare the prediction ability of the three annual diameter growth models using the cross validation data. Models are given in the text.

Fit Statistics	Model 1	Model 2	Model 3
V	0.014436	0.011392	0.011184
E	0.060504	0.07329	0.027836
MS	0.21007	0.12722	0.11684
MSE	0.28645	0.19335	0.18549
RMSE	0.40543	0.33914	0.33482
PBIAS	11.15	13.51	5.13
SD(bias)	0.44	0.33	0.32

Note: V = Model Precision; E = Mean Residual (Bias) ; MS = Mean Square Error; MSE = Mean Square Error of Prediction, RMSE = Root Mean Square Error, PBIAS = Percentage bias, SD (bias) = Standard deviation of bias.

Table 2-7. Parameter estimates and their respective standard errors (SE) and P-values for selected mixed effects annual diameter growth model (Model 3; equation 2-12).

Parameter	Estimates	SE	t value	Pr>t
$\beta_1$	0.5845	0.1939	3.01	0.0038
$\beta_2$	-0.8231	0.2218	-3.71	0.0005
$\beta_3$	0.1159	0.0357	3.24	0.0020

Table 2-8. Comparison of Model 3 with various fit statistics using both the entire dataset and cross validation approach.

Fit Statistics	Model 3 with fixed effects only (all data)	Model 3 with random effect (all data)	Model 3 with fixed effects only (cross-validation data)	Model 3 with random effect (cross-validation data)
V	0.010959	0.009951025	0.011184	0.010792
E	0.028613	-0.057326	0.027836	0.01531
MS	0.111	0.021478	0.11684	0.036578
MSE	0.17696	0.053234	0.18549	0.070788
RMSE	0.32506	0.19311	0.33482	0.20001
PBIAS	5.27	-10.56	5.13	2.82
SD(bias)	0.31	0.091	0.32	0.16

Note: V = Model Precision; E = Mean Residual (Bias) ; MS = Mean Square Error; MSE = Mean Square Error of Prediction, RMSE = Root Mean Square Error, PBIAS = Percentage bias, SD (bias) = Standard deviation of bias.

Table 2-9. Likelihood ratio test results for comparison of a combined slash pine dbh growth model (a single model for all community types) with a separate model for UMF and a full model. See text for the description of community types.

Model Description	-2LL	Pr> $\chi^2$	sig	AIC	BIC
All combined (completely constrained model)	80.0			90	100.4
*UMF vs. MF-SA-BW	74.5	0.023	S	86.5	98.9
All separate (Full model)	73.6	0.093	NS	89.6	106.2

\*best model among all selected in the current study

Table 2-10. Parameter estimates and their respective standard errors (SE) and P-values for selected mixed effects annual diameter growth model for Upland Mixed Forest (equation 2-17).

Parameter	Estimates	SE	t value	Pr>t
$\beta_1$	0.5147	0.1718	2.99	0.0040
$\beta_2$	-0.7150	0.1989	-3.60	0.0007
$\beta_3$	0.1026	0.030	3.41	0.0012
$m_3$	-0.1887	0.0844	-2.23	0.0293

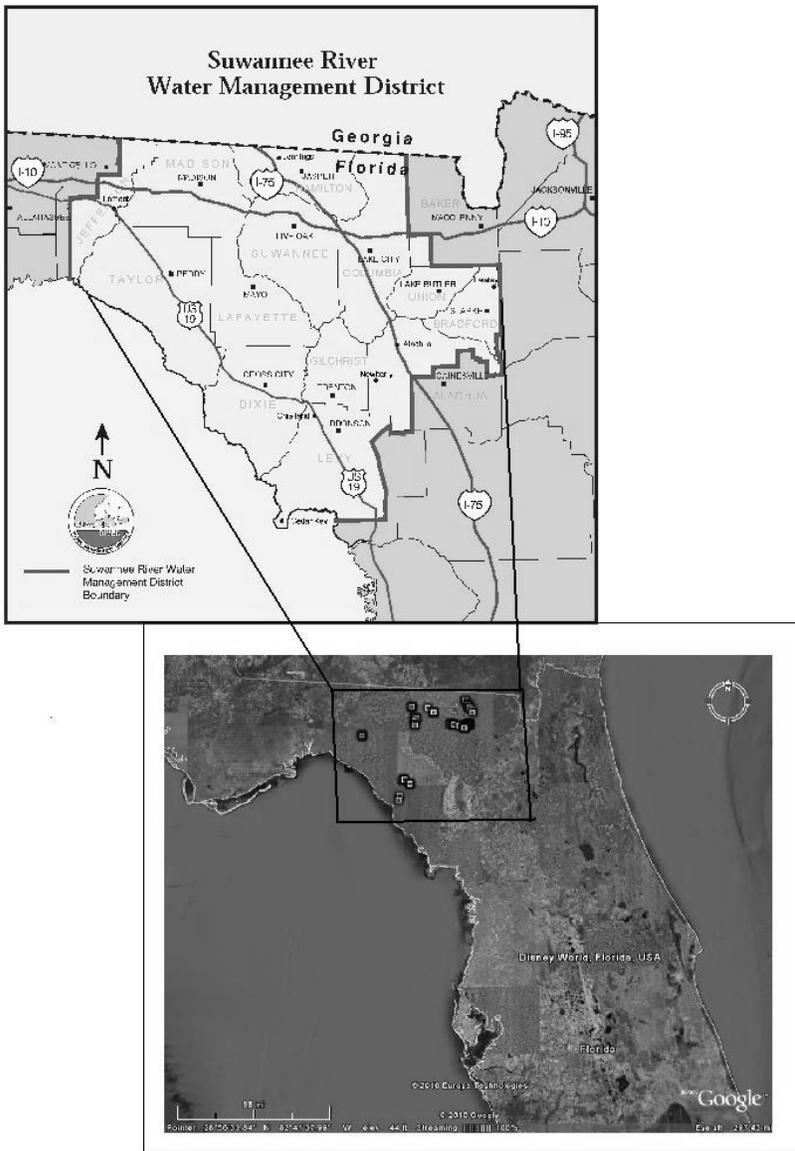


Figure 2-1. Map showing location of Suwannee River Water Management District (SRWMD) in Florida. The black rectangular box in the bottom map shows plot locations within SRWMD.

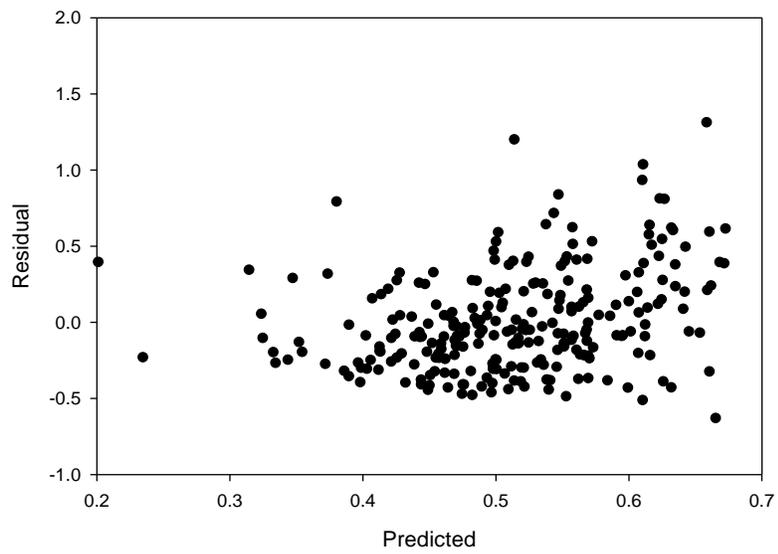


Figure 2-2. Plot of residual versus predicted values for Model 1.

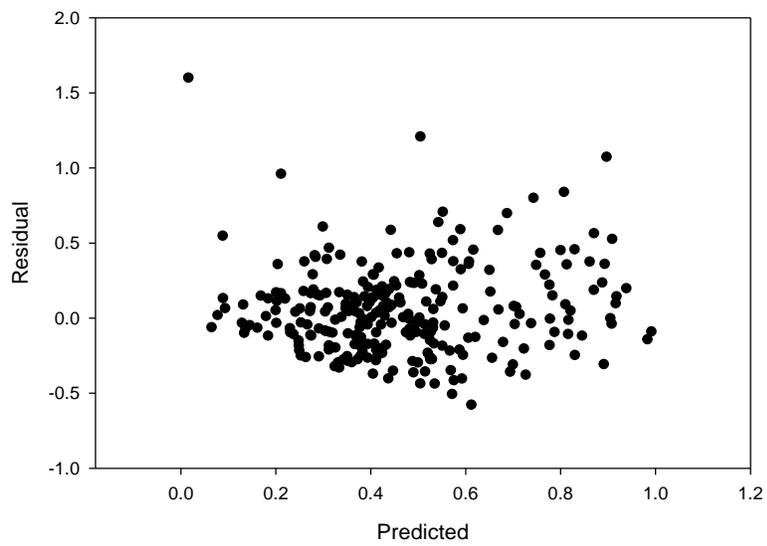


Figure 2-3. Plot of residual versus predicted values for Model 2.

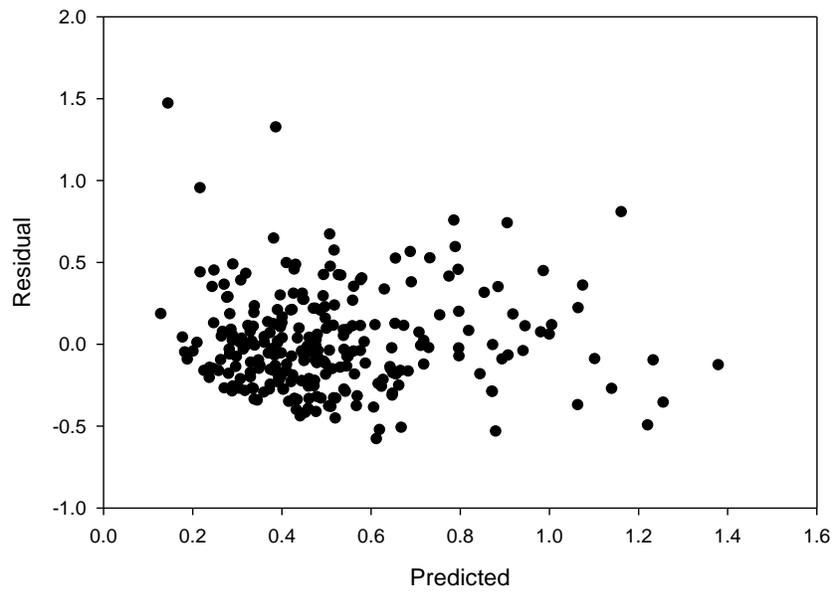
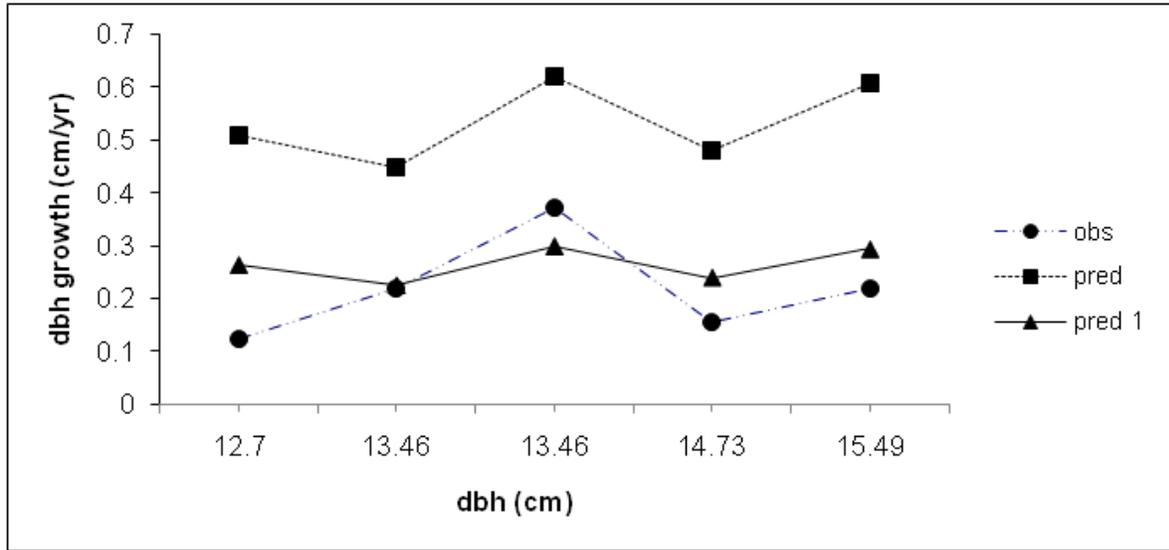


Figure 2-4. Plot of residual versus predicted values for Model 3.

a)



b)

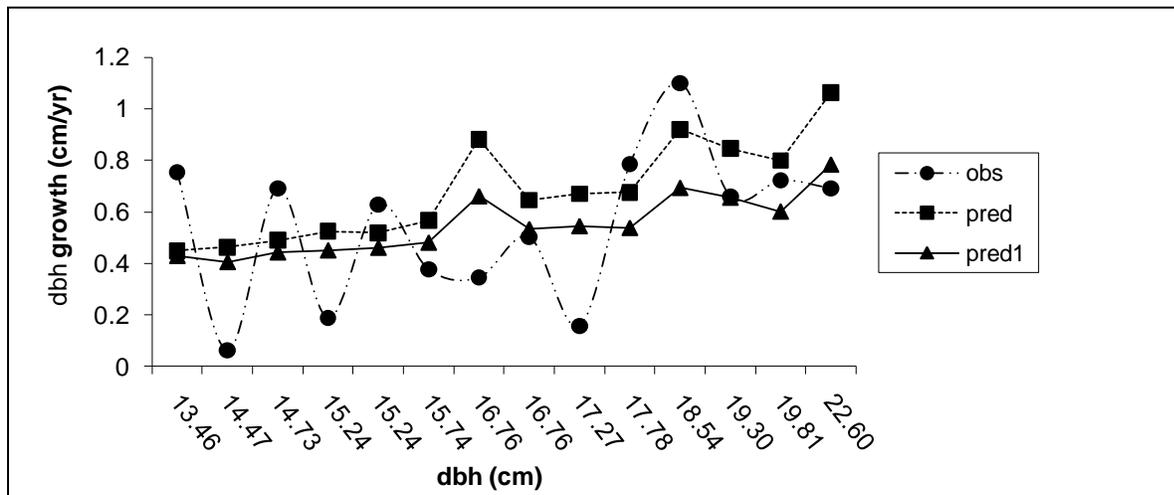


Figure 2-5. Leave-one-out cross validation predictions of annual diameter growth versus tree size (dbh) from Model 3 for two example plots. Graphs show predictions of (a) plot 10, and (b) plot 15, using parameter values obtained after fitting the model without these respective plots (obs = observed values, pred = predicted values using the fixed effect model, and pred1 = predicted values using the mixed effects model).

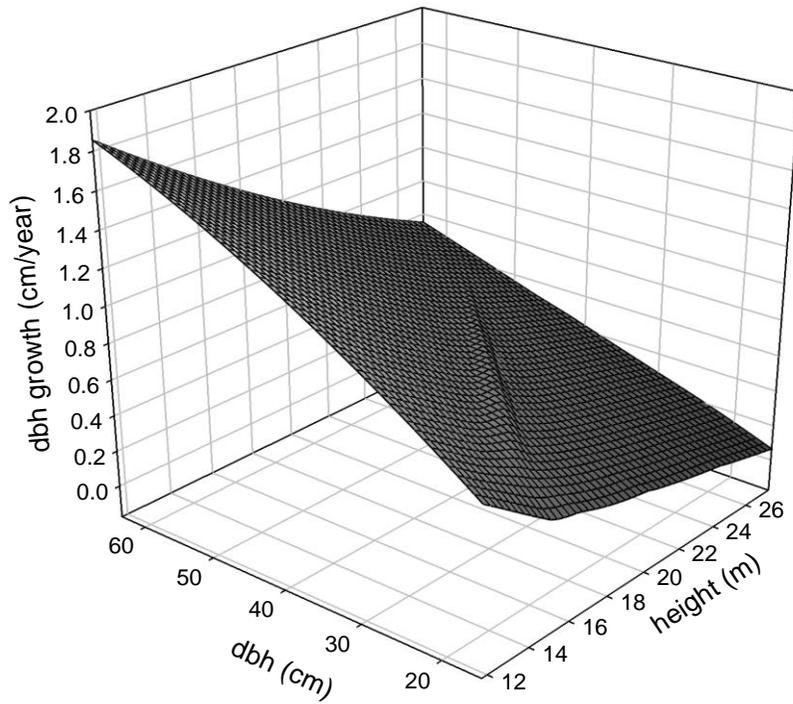


Figure 2-6. Predicted individual tree diameter growth for typical range of slash pine initial dbh-height data using Model 3(fixed effects only). Diameter growth increases as initial dbh increases, while diameter growth decreases as initial tree height increases.

## CHAPTER 3 INDIVIDUAL TREE HEIGHT GROWTH MODEL FOR SLASH PINE IN FLORIDA

### **Introduction**

Tree height growth models are an important component of growth and yield modeling. Height growth models are important to predict yield (Martin et al. 1999), estimate response to silvicultural treatments (Pienaar and Rheney 1995), and measure site potential for commercial tree species. It is also an important component of several individual tree-based growth and yield models such as FOREST (Ek and Monserud 1974), ORGANON (Hester et al., 1989) and PROGNOSIS (Stage 1973, Wykoff et al. 1982).

Tree height growth has been modeled using different approaches. In the potential-modifier approach, the potential height growth of each tree is calculated based on the height growth of free-growing dominant trees for that particular site. Individual tree potential height growth is then modified using a modifier that reflects inter-tree competition (Ritchie and Hann 1986, Weiskittel 2007, Nunifu 2009). Another approach is to relate height growth directly to tree, stand and site characteristics (Huang and Titus 1999, Uzoh and Oliver 2006, Condes and Sterba 2008). There is no specific reason to choose one approach over another, as both methods can produce reasonable results (Wykoff 1990).

Variables that affect tree height growth include tree size, site and competition between trees (Martin and Ek 1984, Huang and Titus 1999, Pretzsch et al. 2002, Cole and Lorimer 1994, Uzoh and Oliver 2006). Tree size has been described using tree diameter at breast height (dbh; Uzoh and Oliver 2006), total tree height (Cole and Lorimer 1994, Cao et al. 2002), and total crown area or mean crown radius (Cole and

Lorimer 1994). Site has been described using site index, especially for even-aged stands or plantations (Martin and Ek 1984, Uzoh and Oliver 2006, Condes and Sterba 2008). In addition to site index, some authors have also described site using elevation and slope (Uzoh and Oliver 2006), or other variables such as temperature, precipitation, and soil moisture and nutrients (Pretzsch et al. 2002).

Tree height growth models have been developed using a mixed modeling approach (Uzoh and Oliver 2006, Weiskittel 2007). The data used in these types of models are hierarchical in structure, as data are collected from trees within plots and plots are established within stands (Hall and Clutter 2004). Hierarchical data structures are associated with a lack of independence among observations (Calama and Montero 2004, Huang et al. 2009, Yang et al. 2009). For example, the correlation between trees within the same plot is higher than that of trees from separate plots. Linear and non-linear mixed modeling is the best approach to model such hierarchical correlated data (Vonesh and Chinchilli 1997, Pinheiro and Bates 2000), and there are several recent examples of their use for such types of data in forest growth and yield (Gregoire and Schabenberger 1996, Calama and Montero 2004, Adame et al. 2008).

There has been a growing interest in uneven-aged forest management, especially among forest managers in public lands. Uneven-aged forests are often described as having a reverse-J shaped diameter distribution, which is usually determined by a stand's  $q$  ratio. The  $q$  ratio is the ratio between tree frequencies in two adjacent diameter classes. The  $q$  ratio for an uneven-aged stand using 4 cm diameter classes, for example varies between 1.2 and 2.0 (Cancino and Gadow 2002). The desired  $q$  ratio for a particular stand depends on the management objectives, e.g., a higher  $q$  ratio

indicates a stand that has a large number of small trees, and therefore, lower economic returns (O'Hara 1996). The determination of the  $q$  ratio is important to achieve a desired structure that is compatible with management objectives, and provides information to managers who are interested in converting structurally homogenous even-aged plantations into the structurally diverse conditions usually found in uneven-aged stands.

Establishing field trials to gather information about how trees respond to treatments aimed at increasing structural diversity and promoting reverse-J distributions is intensive and time consuming, though it is the best approach to develop accurate models of these processes. Growth models are an alternative when field trials are not possible, and are also useful in determining different management scenarios to establish the field trials. Individual tree growth models are very useful in simulating different management scenarios, especially for uneven-aged stands (Groot et al. 2004, Pukkala et al. 2009). While there have been several models developed to simulate the dynamics of uneven-aged stands (e.g., Trasobares et al. 2004 for *Pinus sylvestris* L. and *Pinus nigra* Arn. in Spain, Palahi et al. 2008 for *Pinus brutia* in Greece, Pukkala et al. 2009 for several species in Finland), similar efforts to model slash pine (*Pinus elliottii* Engelm.) are lacking in the southeastern United States.

The objective of the current study is to create an individual tree height growth model. The model can be later used in simulating the stand conversion process. We related height increment directly to other tree- (size and individual tree competition) and stand-level variables (site and stand-level competition). The models were developed using a mixed modeling approach to account for the hierarchical structure of the data and also to account for the correlation existing within the hierarchy.

## Methods

### Data

The data used in this study were provided by the Suwannee River Water Management District (SRWMD). The SRWMD is one of the five water management districts in Florida (USA) (Figure 2-1) and is located in the north central part of the state. It includes a variety of growing conditions, origins and treatments. The network of plots established in 2000 by SRWMD was re-measured in May and June 2008.

SRWMD followed the same protocol used by the USDA Forest Service Forestry Inventory and Analysis (Bechtold and Patterson 2005) to establish plots in 2000. A cluster of plots was established every 101.2 ha (250 acres), with four circular 7.31 m (24 ft) radius tree plots; one plot was at the center and three others were located at 90, 120 and 240 degrees, 36.6 meters (120 ft) from the center plot. Information on current community, historical community, stand origin (natural, mixed or plantation), stand structure, stand age, slope, aspect, and GPS location was collected for every plot. In every 7.31 m plot, three dominant and co-dominant trees were identified as site trees, and their height, age and species were recorded. If there were no suitable site trees within the plot, trees adjacent to the plot were selected as site trees. In every 7.31 m radius tree plot, trees > 12.54 cm (5 inches) in diameter were identified to species and the following data were recorded: distance and azimuth to the tree from the plot center, diameter at breast height (dbh), total height, height to live crown, product class, crown class, and tree damage. All of the above described information were also collected for trees < 12.54 cm (5 inches) dbh, but on a 2.07 m (6.8 ft) radius subplot concentric within the larger plot.

For the 2008 re-measurement, plots from the original year 2000 set of plots that had more than 60% basal area in slash pine were considered. These plots were stratified on the basis of community type, age class (10 year increment) and origin, and plots were randomly selected from each stratum. After excluding harvested plots and those which were inaccessible, 70 plots were chosen for re-measurement. The community types sampled were: Sandhills (SH), Mesic Flatwoods (MF), Upland Mixed Forest (UMF), Bottomland Forest (BF), Baygall (BG), and Swamp (Dome Swamp, Floodplain Swamp). Stand ages ranged from 9 years to 72 years, and were natural, plantation and of mixed origin.

All tree and stand variables measured in 2000 were also re-measured in 2008. In addition, we also collected data on crown radius and degree of crown exposure. Crown radius was measured by extending a tape from the center of the bole to the end of the branch along four cardinal directions. Degree of crown exposure was coded as binary; if the crown radii were overtopped in any direction the tree was regarded as shaded (code 0), and if they were fully exposed the tree was regarded as exposed (code 1). Total crown area (TCA) was obtained by assuming an ellipsoid shape utilizing crown radii in the four cardinal directions. In 2008, we also collected data on trees in a 6.40 m (21 ft) buffer around each tree in the plot so that competition indices described below could be calculated for trees within the 2000 measurement plot boundaries.

To predict 2000 crown radius, a linear allometric relationship between crown radius in 2008 and dbh, height, and crown length in 2008 was developed by species groupings. Hardwoods were grouped as: *Quercus* spp.; *Taxodium* spp.; the Cornaceae family; and other hardwoods present in the plot. Pine species that were sparsely

present in the plots such as loblolly pine (*Pinus taeda* L.), longleaf pine (*Pinus palustris* Mill.), and pond pine (*Pinus serotina* Michx.) were grouped together. A separate predictive model for crown radius was developed for slash pine. The 2000 crown radius was then predicted from 2000 dbh, height, and crown length using the relationship developed from 2008 data. A random error term was added to each prediction. For brevity, results from this preliminary data step are not presented.

Dbh and height information of buffer trees were not available in 2000, but were necessary in order to correctly compute competition indices. To fill in this information, we also developed a linear allometric relationship relating dbh and height in 2000 to dbh and height in 2008. Both 2000 variables were predicted from their 2008 variables using the other variable as a covariate (e.g., dbh for height and height for dbh) via linear regression. Species groupings described earlier for crown radius were used in developing the relationship, and a random error was added to each prediction of dbh and height. As above, this preliminary data analysis step is not presented for brevity.

In order to calculate annual height growth rate, the difference between the two height measurements was divided by total number of days during the two measurement periods, and then multiplied by 365. Stocking of each plot was calculated as a measure of crowding following May (1990). Total basal area per hectare of each plot was also calculated. Similarly, site index for each plot was calculated following Farrar (1973).

Several distance dependent and independent indices were calculated to test in the models. Choi et al. (2007) defined a tree's competitive zone as an area 3.5 times the average crown radius of dominant and co-dominant trees in their study of northern hardwoods. In the present study, we modified this slightly, using 3 times the average

crown radius (2.13 m; 7 ft) of the measured slash pine trees. This resulted in a competitive zone of 6.40 m (21 ft) radius around each subject trees. All distance dependent and independent indices described below were calculated within the competitive zone following Cole and Lorimer (1994), Choi et al. (2001), and Choi et al. (2007). Two relative diameter indices were calculated for each  $i$ th tree as:

$$RD_i = \frac{dbh_i}{\frac{1}{n} \sum_{j=1}^n dbh_j} \quad \text{and} \quad RDgh_i = \frac{dbh_i}{\frac{1}{n'} \sum_{j=1}^{n'} dbh_j}$$

where:  $dbh_i$  is the diameter at breast height of the  $i$ th subject tree,  $n$ =number of all competitors within the 6.40 m radius competitive zone, and  $n'$ =number of trees within the 6.40 m ft radius competitive zone with diameters greater than or equal to the subject tree. Therefore,  $RD$  is the relative diameter calculated with all trees within the 6.40 m radius competitive zone of the subject tree, whereas  $RDgh$  is calculated only with trees equal to or larger than the subject tree. Relative height and relative basal area were calculated similarly.  $RHgh$  is the relative height calculated with trees within the 6.40 m radius competitive zone with heights greater than or equal to that of the subject tree, whereas  $RH$  is the relative height calculated with all trees within the 6.40 m ft radius competitive zone.  $RBAgh$  is the relative basal area calculated by including trees within the 6.40 m radius competitive zone with diameter greater than or equal to the subject tree, whereas  $RBA$  is the relative diameter calculated with all trees within the 6.40 m radius competitive zone of the subject tree.

Other competition indices investigated included the modified Hegyi indices for diameter and height ( $sum\_si$  and  $Sum\_Clr$ , respectively; Heygi 1974), calculated for each  $i$ th tree as:

$$sum\_si_i = \sum_{j=1}^n \frac{dbh_j}{dbh_i d_{ij}}$$

where:  $dbh_j$  is the diameter of competitor tree  $j$ , and  $d_{ij}$  is the distance between the subject tree and its  $j$ th competing neighbor.  $Sum\_Clr$  is similar to  $sum\_si$ , using height as the variable instead of diameter:

$$Sum\_Clr_i = \sum_{j=1}^n \frac{H_j}{H_i d_{ij}}$$

Following Cole and Lorimer (1994), we also modified and tested another competition index related to tree size. This competition index ( $CI$ ) was calculated as:

$$CI_i = \frac{dbh_i}{\sum_{j=1}^n dbh_j}$$

where:  $\sum dbh_j$  is the sum of the diameters of all  $n$  plot competitors, and  $dbh_i$  is the diameter of the subject tree. We also calculated other measures of stand level competition such as  $BAL$  and  $BAL1$  (basal area of trees greater than the subject tree using diameter or height for  $BAL$  and  $BAL1$ , respectively), and quadratic mean diameter for each plot ( $qmd$ ; Teck and Hilt 1990, Cao 2002, Adame et al. 2008).

### **Mixed Modeling of Growth**

Data used in forestry for creating growth models usually comes from permanent plots. These permanent plots are established within forest stands over different sites. Most of the time, trees in a plot are measured repeatedly. The data collected in this fashion are assumed to be both spatially and temporally correlated. Stands within a site are more correlated than stands in different sites. Similarly, plots within a stand and trees within a plot are more correlated than those in different stands and plots (West et al. 1984). This violates the assumption of independent observations required for various

statistical techniques. While ordinary least square methods will give unbiased estimates of model parameters in this kind of correlated data structure, standard errors and test statistics associated with correlated data will be biased and inappropriate for interpretation (Gregoire et al. 1995, Calama and Montero 2004). Also, traditional model fitting provides only population averaged responses (responses averaged over all plots), and does not account for specific plot level effects that could affect the response (Calama and Montero 2004, Huang et al. 2009).

In order to appropriately model such spatially and temporally correlated data, the covariance structure of the data must be appropriately specified (Gregoire et al. 1995). Mixed effects models can account for both the correlation and also the subject-specific (plot-specific) effects, and therefore they are becoming a popular choice in growth modeling (Fang and Bailey 2001, Calama and Montero 2004, Nothdurft et al. 2006, Huang et al. 2008). Mixed model estimation techniques estimate fixed and random parameters simultaneously, and give unbiased and efficient estimates of fixed parameters (Calama and Montero 2004). Incorporating random effects in the model also increases the accuracy of fixed effect prediction and testing (Pinheiro and Bates 2004, Tao and Littell 2002, Budhathoki et al. 2008). Details about mixed model procedures can be found in Vonesh and Chinchilli (1997), Pinheiro and Bates (2000), Calama and Montero (2004) and Meng and Huang (2009).

A linear mixed model can be written in the following form:

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i \quad (3-1)$$

where:  $\mathbf{y}_i = [y_{i1}, y_{i2}, \dots, y_{in}]'$  is a vector of tree height growth measurements in plot  $i$ ,  $\mathbf{X}_i$  is a design matrix for the covariates (dbh, height, competition index, etc.),  $\mathbf{Z}_i$  is a design

matrix for the random effects parameters,  $\boldsymbol{\beta}$  is a vector of fixed parameters,  $\mathbf{b}_i$  is a vector of random effects parameters, and  $\boldsymbol{\varepsilon}_i$  is a vector of errors  $[\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{in}]'$ .  $\mathbf{X}_i$  and  $\mathbf{Z}_i$  are the partial derivatives of the model with respect to fixed parameters  $\boldsymbol{\beta}$ , and random effects  $\mathbf{b}_i$ , respectively.

A non-linear mixed model can also be written in a general form:

$$\mathbf{y}_i = f(\mathbf{X}_i, \boldsymbol{\beta}, \mathbf{b}_i) + \boldsymbol{\varepsilon}_i \quad (3-2)$$

For the estimation of fixed and random effects of a non-linear model, equation 3-2 is linearized using Taylor series expansion (Vonesh and Chinchilli 1997). The random effects parameter ( $\mathbf{b}_i$ ) and error ( $\boldsymbol{\varepsilon}_i$ ) are assumed to be normally distributed and uncorrelated. Their distribution can be written as:

$$E \begin{bmatrix} \mathbf{b}_i \\ \boldsymbol{\varepsilon}_i \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \quad \text{Var} \begin{bmatrix} \mathbf{b}_i \\ \boldsymbol{\varepsilon}_i \end{bmatrix} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_i \end{bmatrix}$$

where:  $\mathbf{D}$  is the variance-covariance matrix for the random effects and  $\mathbf{R}_i$  is the variance covariance matrix of the errors for plot  $i$ .  $\mathbf{R}_i$  is a matrix written as  $\sigma^2 \mathbf{I}_n$ , where  $\sigma^2$  is the error variance and  $\mathbf{I}_n$  is an  $n \times n$  identity matrix. SAS procedures such as PROC MIXED and PROC NLMIXED (SAS Institute Inc., 2008) fit mixed effects models via maximum likelihood by integrating an approximation of likelihoods over the random effects. Various integral approximations of likelihood are available, and in the present study we used the first order method of Beal and Sheiner (1982). This method uses a Taylor series expansion of equation 3-1 around a  $\boldsymbol{\beta}^*$  close to  $\boldsymbol{\beta}$  with  $\mathbf{b}_i^*$  set to zero (the empirically best linear unbiased predictor) for linear approximations of the marginal likelihood functions. The predictor,  $\hat{\mathbf{b}}_i$ , of  $\mathbf{b}_i$  can be estimated using the following (Davidian and Giltinan 1995, Vonesh and Chinchilli 1997):

$$\hat{\mathbf{b}}_i = \hat{\mathbf{D}}\mathbf{Z}_i'(\mathbf{Z}_i\hat{\mathbf{D}}\mathbf{Z}_i' + \hat{\mathbf{R}}_i)^{-1} [\mathbf{y}_i - \mathbf{f}(\mathbf{X}_i, \hat{\boldsymbol{\beta}}, \mathbf{0})] \quad (3-3)$$

where:  $\hat{\mathbf{D}}$  and  $\hat{\mathbf{R}}_i$  are the estimates of  $\mathbf{D}$  and  $\mathbf{R}_i$ , respectively;  $[\mathbf{y}_i - \mathbf{f}(\mathbf{X}_i, \hat{\boldsymbol{\beta}}, \mathbf{0})]$  is the residual, and  $\mathbf{f}(\mathbf{X}_i, \hat{\boldsymbol{\beta}}, \mathbf{0})$  is the prediction of the fixed effect part of the model.  $\mathbf{X}_i$  and  $\mathbf{Z}_i$  are partial derivative matrices, computed as:

$$\mathbf{X}_i = \frac{\partial \mathbf{f}(\mathbf{X}_i, \boldsymbol{\beta}, \mathbf{0})}{\partial \boldsymbol{\beta}'}$$

$$\mathbf{Z}_i = \frac{\partial \mathbf{f}(\mathbf{X}_i, \boldsymbol{\beta}, \mathbf{0})}{\partial \mathbf{b}_i}$$

The error matrix  $\mathbf{R}_i = \sigma^2 \mathbf{I}_n$  is based on the assumption that the errors are independent. Sometimes, the trees within a plot could be spatially correlated and in this case, the assumption of independent errors might not hold. Although, the inclusion of a random plot effect might take care of this correlation, we fitted several spatial covariance structures of the residual to see whether accounting for this correlation would give a better fit. The spatial correlation structures were fit following Littell et al. (2006). This was done using SAS macro %NLINMIX for the nonlinear model. The spatial location of each tree was provided to fit the correlation structure. For example, the spatial power correlation structure is given by:

$$\mathbf{R}_i = \sigma^2 \begin{bmatrix} 1 & \rho^{d_{12}} & \rho^{d_{13}} & \dots & \rho^{d_{1n}} \\ \rho^{d_{21}} & 1 & \rho^{d_{23}} & \dots & \rho^{d_{2n}} \\ \rho^{d_{31}} & \rho^{d_{32}} & 1 & \dots & \rho^{d_{3n}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{d_{n1}} & \rho^{d_{n2}} & \rho^{d_{n3}} & \dots & 1 \end{bmatrix}$$

where:  $\rho$  is the correlation between two observations in a plot, and  $d_{ij}$  is the euclidean distance between the  $i$ th and  $j$ th trees in a plot.

## Model Fitting

We fit several linear and non-linear models to the data. In growth models, growth of a tree is usually related to initial size, competition, and stand and site characteristics (Huang and Titus 1995, Monserud and Sterba 1996). All the models tested in this study incorporated terms related to initial size, stand-level variables representing density and site, and individual tree competition terms representing both horizontal and vertical competition. While the density of a stand can be represented by stand basal area or tree count in a stand (Wykoff 1990, Yang et al. 2009), stand basal area was chosen as the measure of density in the present study because it is calculated from both tree size and count (Yang et al. 2009).

When prior observations on growth are not available (e.g., when predicting values for a new plot), it is very difficult to estimate plot-level random effects and predictions based on random effects are often inappropriate. In this situation, one may choose to calculate the population averaged responses (PA; overall mean response), or the typical mean responses (TM). PA responses are estimated when predictions are made from a model fit with fixed effects only, i.e. without including the random effects in the model. TM responses are estimated when predictions are made from a mixed model fit with both fixed and random effects, ignoring the random effect. TM responses for a non-linear mixed model are comparable to PA responses for a linear mixed model, because PA and TM responses in a linear mixed model are similar (Davidian and Giltinan 1995, 2003; Fitzmaurice et al. 2004). However, these TM responses are inferior in prediction versus PA responses (Meng et al. 2009). Therefore, both linear and non-linear models were first fit with fixed effects only. After choosing the best fitting most parsimonious model, random effects were added, and linear and non-linear mixed

effects models were fit so that comparisons between the two (fixed and mixed) models could be made.

First, we fit a linear model to the data (Model 1) with just fixed effects (equation 3-4). Then Model 1 was fit with an additional random effect (equation 3-5).

$$y_{ij} = \beta_0 + \beta_1 dbh_{ij} + \beta_2 H_{ij}^{-1} + \beta_3 CI_{ij} + \beta_4 BAL_{ij} + \beta_5 tbaha_i + \varepsilon_{ij} \quad (3-4)$$

$$y_{ij} = \beta_0 + \beta_1 dbh_{ij} + \beta_2 H_{ij}^{-1} + \beta_3 CI_{ij} + \beta_4 BAL_{ij} + \beta_5 tbaha_i + b_i + \varepsilon_{ij} \quad (3-5)$$

where: for the  $j$ th tree in the  $i$ th plot,  $y_{ij}$  is the annual tree height growth,  $dbh_{ij}$  is the diameter at breast height (dbh),  $H_{ij}$  is the initial total tree height,  $CI_{ij}$  is the competition index as described earlier,  $BAL_{ij}$  is the basal area in trees larger in dbh than the subject tree, and  $tbaha_i$  is the total basal area per hectare for a plot.  $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$  and  $\beta_5$  are fixed effects parameters,  $b_i$  is the random plot effect, and  $\varepsilon_{ij}$  is the error term.

In the linear model, we tested different variables representing tree size together and separately. Other size variables tested in the model were mean crown radius, total crown area and tree height. Similarly, several distance independent and dependent competition indices described in the Methods section were tested in the model. Site index was also included in the model as a proxy for growing condition. Several transformations and interaction terms were also tested. For the selection of the best fixed effects linear model, statistical significance of the covariates was assessed at  $\alpha=0.05$  level, and various fit statistics such as AIC, BIC, precision, bias and RMSE were compared. Random effects were then added to the best fitting most parsimonious linear model, and this model was again considered in terms of AIC, BIC, and significance of

the covariates. We also fitted various spatial covariance structures described earlier to account for spatial correlation of trees within the plots.

We also fit a height increment model provided by Huang et al. 1999 (equation 3-6), which is based on a two-parameter Box-Lucas function (Box and Lucas 1959). It is a typical height increment curve which starts at zero, steadily increases to a maximum and decreases smoothly and asymptotically towards zero (Huang et al. 1999). The parameters of equation 3-6 are then later related to other tree and stand level variables using a parameter prediction approach (equations 3-7a and 3-7b; Clutter et al. 1983).

$$y_{ij} = \frac{\theta_2}{\theta_1 - \theta_2} * (\exp(-\theta_2 H_{ij}) - \exp(\theta_1 H_{ij})) \quad (3-6)$$

$$\theta_1 = f(X_1, X_2, X_3, \dots) \quad (3-7a)$$

$$\theta_2 = f(X_1, X_2, X_3, \dots) \quad (3-7b)$$

This non-linear model (equation 3-6, 3-7a and 3-7b), however, did not give good results. Plots of the observed versus predicted values from the model were not biologically reasonable. Therefore, we excluded this model from further analysis.

A second non-linear equation of the following form was also fit to the data (after Cole and Lorimer 1994; Model 2).

$$y_{ij} = \beta_0 + \beta_1 X_{1ij} e^{\beta_2 X_{2ij}} \quad (3-8)$$

Several different forms of this base model were tested:

$$y_{ij} = \beta_1 H_{ij} \times e^{\beta_2 H_{ij}} \quad (3-8a)$$

$$y_{ij} = (\beta_1 H_{ij} + \beta_2 CI_{ij}) \times e^{\beta_3 H_{ij}} \quad (3-8b)$$

$$y_{ij} = \beta_0 + (\beta_1 H_{ij} + \beta_2 CI_{ij}) \times e^{\beta_3 H_{ij}} + \beta_4 dbh_{ij} + \beta_5 dbh_{ij}^2 \quad (3-8c)$$

$$y_{ij} = \beta_0 + (\beta_1 H_{ij} + \beta_2 CI_{ij}) \times e^{\beta_3 H_{ij}} + \beta_4 \ln(dbh_{ij}) \quad (3-8d)$$

$$y_{ij} = \beta_0 + (\beta_1 H_{ij}) \times e^{\beta_2 H_{ij}} + CI_{ij}^{\beta_3} + \beta_4 \ln(dbh_{ij}) \quad (3-8e)$$

where: all the variables are as described above and ln denotes the natural logarithm.

The various forms of Model 2 were evaluated based on fit statistics such as AIC, BIC, precision, bias and RMSE. The best model was then evaluated with the addition of random effects. First, all parameters in the model were considered with both fixed and random components (mixed), but this approach led to convergence problems.

Therefore, random effects were added to only one parameter at a time, and the resulting AIC and BIC values of the models along with the fit statistics were compared.

In the various forms of Model 2 (equations 3-8a to 3-8e), tree height growth was related to tree level variables that represent size and individual tree competition. After the addition of random effects, the best model among them was then expanded to include stand level variables. We used two different methods: 1) the random parameter of the model was fit to stand-level covariates (M1; Calama and Montero 2004), and 2) stand-level variables were directly added to the model with mixed effects (M2; Yang et al. 2009).

Stand-level variables considered in equations 3-8a to 3-8e included measures of site quality, crowding, and stand-level competition. For a given tree size, growth is usually significantly higher in a rich site than a poor site (Wykoff 1990). Therefore, we tested site index in these models. To represent overall crowding effects, stand basal area, was also included in the model (Wykoff 1990, Yang et al. 2009). Both one-sided and two-sided stand-level competition can affect tree growth (Cannell et al. 1984, Weiner 1990). For example, basal area in trees larger than a subject tree represents vertical competition for light as a one-sided competition measure; it does not include

competition from trees smaller than the subject tree for resources other than light. Per unit basal area and quadratic mean diameter ( $qmd$ ), on the other hand, consider all the trees in the stand and represent two-sided competition. Both one-sided and two-sided competition measures were considered to arrive at a final form of Model 2.

We grouped plots into natural communities using the palustrine ecosystem classifications of the Florida Natural Areas Inventory and Florida Department of Natural Resources (1990). Bay gall (BA, Seepage wetlands), Basin Swamp (BS), Dome Swamp (DS) and Bottomland Forest (BF, Floodplain Wetlands) were grouped together into one category called Basin wetlands (BW). Our Sandhill category included both plots classified as Sandhill (SA) and as xeric hammock (XH), as both classes belong to the Xeric Upland community and the XH class contained only 8 trees. Mesic Flatwood (MF) and Upland Mixed Forest (UMF) were left as separate categories. Community types were included as a fixed class effect in the best model selected in this study. We examined the significance ( $\alpha=0.05$ ) of the effect, and also performed likelihood ratio tests to examine whether a specific model was required for each community or if communities could be combined into a single model.

### **Model Evaluation**

The final forms of Models 1 and 2 were compared by calculating the following statistics for each plot following Arabatzis and Burkhart (1992) and Calama and Montero (2004):

Mean residual (Bias),  $\bar{E}_i$ , for plot  $i$ :

$$\bar{E}_i = \frac{\sum_{j=1}^{n_i} e_{ij}}{n_i}$$

Model precision,  $V_i$ , for plot  $i$ :

$$V_i = \frac{\sum_{j=1}^{n_i} e_{ij}^2 - \left[ \sum_{j=1}^{n_i} e_{ij} \right]^2 / n_i}{(n_i - 1)n_i}$$

Mean square error,  $MS_i$ , for plot  $i$ :

$$MS_i = \bar{E}_i^2 + V_i$$

Mean square error of prediction,  $MSE_i$ , for plot  $i$ :

$$MSE_i = \frac{\sum_{j=1}^{n_i} (e_{ij})^2}{n_i - 1}$$

Percentage bias,  $pbias_i$ , for plot  $i$ :

$$pbias_i = \frac{\bar{E}_i}{\bar{y}_i} \times 100\%$$

where:  $e_{ij}$  represents residual (observed height growth minus predicted height growth) of the  $j$ th tree in plot  $i$ ,  $\bar{y}_i$  is the mean of the observed values for plot  $i$ , and  $n_i$  represents the number of trees in the  $i$ th plot. Once plot-level values of each fit statistic were calculated, mean values for each statistic were calculated over all 70 plots. The mean values are hereafter referred to as E, V, MS, MSE, and PBIAS. In addition, we also calculated the root mean squared error (RMSE) of each model, and the standard deviation of bias. Finally, plots of residual versus predicted values were plotted to ensure satisfaction of the assumptions of normality and homoscedasticity of the residuals.

Independent data were not available for validation; neither were there enough data available to subdivide data into model fitting and validation datasets. Therefore, we used a leave-one-out cross validation approach (Temesgen et al. 2008) to evaluate the model predictions. Since our data are naturally grouped by plot, we modified this approach so that plots rather than individual trees were left out in each step; we deleted

a plot from the data set, fit the model with the remaining data, and then used the estimated parameters to predict the growth of trees in the deleted plot. This was done for all 70 plots in the dataset. Observed and predicted values were then used to calculate model fit statistics. Based on those statistics, the predictive ability of those models was determined.

## Results

### Model Selection

In the linear model (equation 3-4), we included various combinations, substitutions, and transformations of the listed variables. Because crown size is assumed to be correlated with leaf area and is an indicator of potential gross production in trees (Cole and Lorimer 1994), we tested total crown area (*TCA*) for the size variable instead of *dbh*. This variable was not significant, however, and fit statistics were worse than when *dbh* was included as the size variable. We also investigated variables that would characterize height increment as a peaking function over size. While larger trees have a competitive advantage over smaller trees, tending to grow faster than the smaller trees, larger trees have higher maintenance costs. Therefore, growth peaks and declines when a tree gets to a certain size (Oliver and Larson 1996). In order to capture this relationship, we included initial total height and inverse of initial total height together in the model, but initial total height was not significant. Other variables related to this peaking shape were also investigated, including the square of *dbh*, inverse of *dbh* and log transformations of the covariates; however, better fit was obtained using the original formulation of the model.

In order to better describe both vertical and horizontal stand-level competition, and stand condition, we included basal area in trees larger in *dbh* than the subject tree

(*BAL*), basal area in trees taller than the subject tree (*BAL1*), total basal area (*tbaha*, and also site index, but the model fit was better when only *BAL* and *tbaha* were included in the model. The inclusion of two interaction terms resulted in improved fit statistics, and thus they were included in the model. Finally, the intercept term was not significant; however, keeping this term gave a better fit than removing it from the model.

Parameter estimates and tests of fixed effects for the model with fixed effects only and the mixed effects model are listed in Table 3-1. The parameter value for *CI* was insignificant, but removing it from the model increased the AIC and BIC values (in the fixed effect model AIC increased from 52 to 57 and BIC increased from 56 to 61). Therefore, it was kept in the model. Among all the distance independent and distance dependent individual tree competition indices, *CI* gave a better fit. The best linear fixed effects and mixed effects models were:

$$y_{ij} = \beta_0 + \beta_1 dbh_{ij} + \beta_2 H_{ij}^{-1} + \beta_3 CI_{ij} + \beta_4 BAL_{ij} + \beta_5 tbaha_i + \beta_6 dbh_{ij} \times H_{ij}^{-1} + \beta_7 BAL_{ij} \times tbaha_i + \varepsilon_{ij} \quad (3-9a)$$

$$y_{ij} = \beta_0 + \beta_1 dbh_{ij} + \beta_2 H_{ij}^{-1} + \beta_3 CI_{ij} + \beta_4 BAL_{ij} + \beta_5 tbaha_i + \beta_6 dbh_{ij} \times H_{ij}^{-1} + \beta_7 BAL_{ij} \times tbaha_i + b_i + \varepsilon_{ij} \quad (3-9b)$$

Both fixed and mixed effects linear models had similar results in terms of magnitude and direction for all parameters except  $Height^{-1}$  (Table 3-1). There was a negative effect of increasing initial height on height growth in the mixed effects model (indicated by the positive parameter estimate for  $Height^{-1}$  in the mixed column in Table 3-1), but the height variable appeared to have the opposite magnitude in the fixed

effects model. This effect, however, was confounded by its dependence on the significant interaction term between *dbh* and *Height*<sup>1</sup>. Height growth prediction curves for different *dbh* were similar for both fixed and mixed effect model (Figure 3-1a; no result presented for mixed model). Due to the interaction between *dbh* and *Height*<sup>1</sup>, for similar *dbh*, height growth is lower for taller trees (Figure 3-1a) which may be due to old age. Surprisingly, *BAL* and *tbaha* had positive effects on growth, but again this effect is confounded by the significant negative effect of the interaction term *BAL x tbaha*. Due to the interaction, for similar *BAL*, there was a slight decline in height growth as total basal area increased (Figure 3-1b). We also observed that an increase in stand density (as represented by *tbaha*) decreased height. We do not present prediction curves for the mixed effects model, because similar prediction curves were observed for the mixed and fixed effects models.

Individual tree competition (*CI*) had a negative effect on tree height growth; higher values of *CI* (lower competition) resulted into lower growth. This result is contradictory to the general assumption about competition indices; however, this index was not significant. Other covariates such as *BAL* and *tbaha* might have incorporated the competition effect; hence the effect of *CI* was not as important.

Using the general form given in Equation (3-8), models including several covariates and their transformations were fit to the data (equations 3-8a – 3-8e). The first step in fitting these models used only the tree level variables in the model. When evaluating the models in equations 3-8a – 3-8e based on AIC, BIC and other fit statistics such as bias, precision, and RMSE, the best model was 3-8e. This model was then expanded by adding random effects to the parameters.

AIC and BIC values (Table 3-2) suggested that parameters  $\beta_0$ ,  $\beta_1$ , and  $\beta_4$  associated with intercept, height, and dbh, respectively, should be treated as having random effects, as all had similar AIC and BIC values. Therefore, we also calculated goodness of fit statistics for the model with each of the parameters as random. The results (Table 3-3) suggested that the model with *dbh* ( $\beta_4$ ) having a random effect was the best due to its lower bias (E), lower RMSE and lower percentage bias (pbias). The model treating initial height ( $\beta_2$ ) as random did not converge and therefore was not included in the comparison.

The final base non-linear mixed model was of the following form:

$$y_{ij} = \beta_0 + (\beta_1 H_{ij}) \times e^{\beta_2 H_{ij}} + CI_{ij}^{\beta_3} + (\beta_4 + b_i) \ln(dbh_{ij}) \quad (3-10)$$

### **Comparison of Fixed Versus Mixed Effects Models**

AIC and BIC values of the different models (Table 3-4) clearly indicate that there is substantial evidence that a mixed model is more appropriate than a fixed effects model both in the linear and non-linear cases. The other fit statistics (Table 3-5) also clearly show that a mixed model fits the data better than a fixed effects model. For the linear model, goodness of fit statistics indicated that the mixed model had better model precision (V) and lower bias (E) versus the fixed effects model (Table 3-5). Percentage bias (PBIAS) was also lower for the linear mixed model; however, the standard deviation of bias (SD bias) and root mean square error (RMSE) were very similar for both linear models (Table 3-5). Similar results were obtained for the non-linear model. The non-linear mixed model had better precision (V), lower bias (E) and lower RMSE than the non-linear fixed effects model. Also, PBIAS and SD bias were lower for the non-linear mixed model.

## Base and Expanded Non-linear Model

The best non-linear model (base model; equation 3-10) was expanded later with stand level variables and the random effects. The expanded models with method 1 (M1) and method 2 (M2) are given in equations 3-11 and 3-12, respectively:

$$y_{ij} = \beta_0 + \beta_1 H_{ij} e^{(\beta_2 H_{ij})} + (\beta_3 + \beta_4 tbaha_i + \beta_5 qmd_i + b_i) * \ln(dbh_{ij}) + \varepsilon_{ij} \quad (3-11)$$

$$y_{ij} = \beta_0 + \beta_1 H_{ij} e^{(\beta_2 H_{ij})} + \beta_3 tbaha_i + \beta_4 qmd_i + (\beta_5 + b_i) * \ln(dbh_{ij}) + \varepsilon_{ij} \quad (3-12)$$

The estimates of the intercept for both of the models in equations 3-11 and 3-12 were negative, and both equations gave negative predictions for height growth using *dbh*, height, *tbaha* and *qmd* values in the range observed in the real data). Therefore, we removed the intercept term from both models, re-fit the models (equations 3-13 and 3-14) and calculated fit statistics (Table 3-6 and Table 3-7). When the intercept was removed, the model expanded with method 2 (equation 3-14; M2) performed better than the model expanded with method 1 (equation 3-13; M1). For both fixed effects and mixed effects models, equation 3-14 had better precision (V), lower bias, lower RMSE and lower percentage bias (PBIAS) than equation 3-13 (Table 3-6 and Table 3-7).

Among the non-linear models, we considered equation 3-14 as our best model.

$$y_{ij} = \beta_1 H_{ij} e^{(\beta_2 H_{ij})} + (\beta_3 + \beta_4 tbaha_i + \beta_5 qmd_i + b_i) * \ln(dbh_{ij}) + \varepsilon_{ij} \quad (3-13)$$

$$y_{ij} = \beta_1 H_{ij} e^{(\beta_2 H_{ij})} + \beta_3 tbaha_i + \beta_4 qmd_i + (\beta_5 + b_i) * \ln(dbh_{ij}) + \varepsilon_{ij} \quad (3-14)$$

Comparison of the base non-linear mixed model (equation 3-10) with the expanded non-linear mixed models (equation 3-13 and 3-14) showed that the expanded models were better than the base model. The expanded models had smaller bias (E),

better precision (V), smaller RMSE, smaller PBIAS, and smaller SD bias than the base model (Table 3-7). Similar results were observed after using only the fixed effects part of the models (Table 3-6).

### **Fitting Covariance Structure to Non-linear Mixed Models**

We tested several structures for the residual covariance matrix to account for the correlation of trees within the same plot. This procedure was done for all of the nonlinear models (eqns. 3-10, 3-11, 3-12, 3-13 and 3-14), and for the linear model. Several spatial covariance structures provided in Littell et al. (2006) were tested and the model that did not take into account the correlation among trees within plots provided the best fit based on the AIC and BIC values (Table 3-8). Since similar results were obtained for all model forms, we only present results for the non-linear model given in eqn. 3-14. We also performed a likelihood ratio test to determine whether it was necessary to model the spatial correlation of the trees within a plot. The likelihood ratio statistic (LRS) was compared to a Chi square distribution with one degree of freedom, since the spatial covariance models have two covariance parameters (e.g., spherical model has  $\rho$  and  $\sigma^2$ ) and the independent error model has only one ( $\sigma^2$ ). The likelihood test failed to reject ( $P > 0.1$ ) the null hypothesis of zero correlation.

Finally, we plotted residual versus predicted values of all the models to ensure model assumptions were met. None of these plots indicated heteroscedasticity or non-normality.

### **Comparison between Linear and Non-linear Model**

There should be prior growth information available to calculate random effects when using a non-linear mixed model to predict subject-specific (plot-specific) growth for a new plot. Sometimes, this information is not available, however, and practitioners

often resort to using a fixed effects model to predict growth. As described in the methods section, predictions carried out by using fixed effects parameters from a non-linear mixed model are not accurate. Therefore, comparisons between linear and non-linear models were made using both fixed and mixed effect models.

The results (Table 3-6) showed that the linear fixed effect model (eqn. 3-9a) had better fit statistics than the non-linear fixed effect base model (eqn. 3-10). The linear fixed effect model, which included stand level variables as well, had better precision (V), similar bias (E) and PBIAS, and lower RMSE than the non-linear base model with fixed effect only (eqn. 3-10; Table 3-6). Once the non-linear models were expanded by adding stand level variables, we found that the fixed effects non-linear model expanded with M1 (eqn. 3-13) had inferior fit statistics than the linear model, but the nonlinear model expanded with M2 (eqn. 3-14) had better fit statistics than the linear model (Table 3-6). The non-linear model expanded with M2 (fixed effect only) had lower bias (E), better precision (V), lower percentage bias and lower SD bias than the linear fixed effects model. The RMSE of the linear and non-linear expanded models were similar (Table 3-6).

When comparing linear and non-linear mixed models, the linear mixed model (eqn. 9b) had better precision (V) than the non-linear models (Table 3-7). It also had lower bias (E) and lower percentage bias (pbias) than the non-linear mixed base model (eqn. 3-10) and non-linear expanded model without intercept (eqn. 3-13; Table 3-7). The linear mixed model was inferior, however, when compared to the best non-linear expanded mixed model (eqn. 3-14). The best non-linear mixed model had lower bias, lower RMSE, lower pbias and lower SD bias than the linear mixed model. Based on all

the observations described above, we selected the model without intercept expanded with method 2 (eqn. 3-14) as the best annual height growth model.

### Cross-validation of Models

Prediction ability of both linear and non-linear mixed models (base and expanded) was also tested employing the cross-validation approach. For cross-validation of the mixed effects model, the predictor of the random parameter,  $\hat{b}_i$ , should be estimated for each deleted plot for prediction of growth for that plot. We present a method in this section to estimate  $\hat{b}_i$  similar to other studies on mixed models (Calama and Montero 2004, Huang et al. 2009a). A similar procedure was employed for all estimated linear and non-linear mixed models, but for brevity, we present the result of our best non-linear mixed model (eqn. 3-14) only. First, the fixed effects parameters for the deleted plot were estimated using the SAS procedure PROC NLMIXED with the parameter for the random effect,  $\mathbf{b}_i$ , set to zero. To estimate the random parameter  $b_i$ , we then calculated the derivative of equation 3-14 with respect to parameter  $b_i$ :

$$\frac{\partial \Delta H}{\partial b_i} = \frac{\partial f(\mathbf{X}_i, \hat{\boldsymbol{\beta}}, 0)}{\partial b_i} = \ln(dbh) \quad (3-15)$$

The derivatives thus calculated (eqn. 3-15) provided the  $\mathbf{Z}$  matrix. The  $\mathbf{Z}$  matrix is shown for an example plot with seven trees below. Since the model in the present study had only one random variable, the variance of the random parameter is a scalar:  $D = \sigma^2_{b_i}$ . The  $\mathbf{R}_i$  matrix is an  $n \times n$  (where  $n$  is the number of trees in a plot) matrix as given below with error variance  $\sigma^2$  on the diagonals and zero elsewhere. Since no spatial covariance structure was better than the model without structure, we assumed zero covariance between trees within a plot.

$$\mathbf{Z}_i = \begin{bmatrix} 2.819 \\ 2.947 \\ 3.011 \\ 2.906 \\ 2.863 \\ 2.788 \\ 3.348 \end{bmatrix} \quad \text{and} \quad \hat{D} = 0.00290$$

$$\hat{R}_i = \begin{bmatrix} 0.0379 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.0379 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0379 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0379 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0379 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.0379 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.0379 \end{bmatrix}$$

Once the  $\mathbf{Z}$  matrix was calculated, the random parameter for each deleted plot was calculated using equation 3-3. The fixed effect prediction ( $y_{fix_{ij}}$ ) of the model was given by the following equation with  $b_i$  (random parameter) set to zero:

$$y_{fix_{ij}} = \beta_1 H_{ij} e^{(\beta_2 H_{ij})} + \beta_3 tbaha_i + \beta_4 qmd_i + (\beta_5 + \mathbf{b}_i) * \ln(dbh_i) \quad (3-16)$$

The residual [ $\mathbf{y}_i - f(\mathbf{x}_i, \hat{\boldsymbol{\beta}}, 0)$ ] part of the equation 3-2 was calculated by subtracting  $y_{fix_{ij}}$  values from observed annual height growth values for the  $i$ th plot.

The fit statistics (Table 3-9) showed that the non-linear mixed models were better than the linear mixed model, except for the base model (which does not include stand level variables). The base non-linear mixed model had lower precision, higher bias (E), lower RMSE and percentage bias (PBIAS) than the linear mixed model. The expanded non-linear mixed models without intercept (eqns. 3-13 and 3-14) had better precision (V), lower bias, lower RMSE and lower PBIAS than the linear mixed model (Table 3-9). Among the non-linear mixed models, the expanded models had better precision (V), lower bias (E), lower RMSE and lower percentage bias (pbias) than the base model. Among the expanded non-linear mixed models, eqn. 3-14 had higher precision, lower

bias, lower RMSE, and lower PBIAS than all other models. Therefore, we selected eqn. 3-14 as our best annual height growth model for slash pine in this study.

### **Comparing Growth among Different Natural Communities**

The data for this study was collected from different natural communities, and they differ in growth due to influences and interactions of microclimate, depth to water table, soil nutrients, etc. To incorporate these differences into our model, we fit a full model including community type effects as a set of dummy variables. The combined model (a single model for all community types) given in equation 14 was compared with the full model using a likelihood ratio test. A non significant likelihood ratio test (Table 3-10) indicated that a full model was not required. Moreover, parameters for the community effects were not significant when all communities were separated in the full model. To judge if a subset of community types might be different, we used a backwards stepwise procedure, sequentially removing unique parameter for the community type that was most insignificant until all effects were significant. The final result was a separate model for UMF and a combined model for MF, SA, and BW (Table 3-10). This model was also compared to the combined model using the likelihood ratio test. The test was significant, indicating that a separate model for UMF would be required (equation 3-17).

$$y_{ij} = \beta_1 H_{ij} e^{(\beta_2 H_{ij})} + \beta_3 tbaha_i + \beta_4 qmd_i + (\beta_5 + b_i) * \ln(dbh_{ij}) + m_1 * NC3 + \varepsilon_{ij} \quad (3-17)$$

where:  $m_1$  is the coefficient for UMF, and  $NC3$  is a dummy variable indicating UMF community type.

## Interpretation of the Parameters of the Best Model

Parameter estimates from the best model (eqn. 3-14; Table 3-11, Table 3-12) show initial height has both a positive and negative effect on growth. The positive  $\beta_1$  coefficient in equation 3-13 indicated that taller trees have higher growth rates, but negative  $\beta_2$  coefficient indicated that the growth decreases as trees grow taller. The height growth model was applied to data covering the range of height and diameter present in our dataset (Figure 3-2) to show that it followed a unimodal, positively skewed shape typical of tree growth processes (Assmann 1970, Wykoff 1990). Similarly, a positive  $\beta_3$  coefficient indicated that the basal area per hectare had a positive effect on height growth. The coefficient was not significant at the 5% significance level when the random effect was added to the model (Table 3-12), but when predictions were compared with and without the covariate, fit statistics were better for the model including basal area per hectare. The density of the plot as represented by basal area per hectare had a minimal effect on tree growth; tree growth for three different basal areas per ha was very similar (Figure 3-2). The quadratic mean diameter of a plot had negative effect on the growth of trees ( $\beta_4$  negative; Table 3-11, Table 3-12). The log of initial diameter had a positive coefficient, indicating that as tree diameter increased the height growth also increased, but at a slowing rate as tree diameter became larger (Figure 3-3).

Usually, trees with larger dbh values have larger crowns. So, for the same height, trees with larger dbh values also have higher height growth (Figure 3-4), and trees that are taller in height but smaller in size have reduced height growth (Figure 3-4). Taller trees grew faster in height due to their canopy position yet failed to allocate more

resources to crown growth. Height growth declines and levels off for taller trees with large dbh (older trees) (Figure 3-4).

### **Discussion and Conclusions**

The non-linear model performed better than the linear model in this study. Both linear and non-linear models included variables related to tree size, individual tree competition, and stand level variables such as total basal area and quadratic mean diameter. None of the competition indices that incorporate size and distance between trees (distance dependent indices) performed better than *CI*. *CI* is considered distance independent; although it takes into account the distances between trees indirectly, because only trees within a 6.4 m (21 ft) radius of the subject tree are utilized. Our initial intention was to develop a spatially explicit model with inclusion of a distance dependent competition index; however, in our data set the distance independent index performed better. Past studies have reported mixed results on the utility of distance dependent versus distance independent indices in growth models (Biging and Dobbertin 1995).

Relative size of the tree is important for height growth. The coefficient for *CI* was negative for both the linear models (Table 3-1) and non-linear base model (result not presented) indicating lower competition has lower growth. This result is contradictory to general assumption of competition; however, coefficient of *CI* was not significant in the linear models. When the non-linear base model was expanded with stand level variables, *CI* was not a significant predictor. This may be due to the fact that the total basal area incorporates the effect of overall competition; hence *CI* is not required in the model once *tbaha* is included. Site index was not a significant predictor of height increment in both linear and non-linear expanded models. Our model already has initial

height of a tree as an explanatory variable; the height of individual trees also might have represented the site effect.

Coefficients in the best model (equation 3-14) conform to accepted biological principles, and are consistent with the growth modeling literature. Trees that are larger in diameter have a competitive advantage over trees that are smaller in size. Therefore, the large-sized trees have higher height increments; similar observations of height growth in white spruce has been reported (Huang and Titus 1999), and Uzoh and Oliver (2006) also had a positive coefficient for the natural logarithm of diameter in their height growth model.

The decrease in height growth rate as trees grow taller is typical of height growth processes and has been observed in Sugar maple (*Acer saccharum*) in Wisconsin (Cole and Lorimer 1994). Huang and Titus (1999) also reported this unimodal right skewed relationship between height growth and initial tree height, and used this model form in their height growth model for white spruce (*Picea glauca* (Moench) Voss) and aspen (*Populus tremuloides* Michx.) in Alberta. MacFarlane et al. (2002) observed a decline in height growth of older loblolly pine trees compared to younger ones for the same crown area and the total light capture. They also observed a decrease in the value of light for growth in older trees. They reasoned that the older trees are usually taller but that they are inefficient in utilizing the higher amount of light they receive by being tall. This could be attributed to the increase in cost of maintenance (in terms of required site resources) for larger trees (Oliver and Larson 1996). Other authors have attributed the decline in growth in larger trees to the increase in cost of respiration, lower efficiency to utilize water and nutrients, and deterioration of tissue quality (Gower

et al. 1996, Hubbard et al. 1999). Decreased height increment with increasing tree height has been reported in black spruce as well (Robichaud and Methven 1991).

In our model, basal area had a positive but negligible effect on tree height growth (Table 3-11, Table 3-12, Figure 3-2). Also, the distance dependent individual competition index, CI, was not significant in the model after it was expanded by adding stand level variables such as basal area and quadratic mean diameter. Our result conforms to expectations of stand density effects on tree height growth. Previous studies modeling top height growth have disregarded competition effect, because it is believed to be relatively independent of stand density (Clutter et al. 1983, Smith et al. 1997). In slash pine, Pienaar and Shiver (1984) did not find a discernible effect of stand density on height growth in a 36 year-old slash pine plantation; however, other studies have found positive effects of density on juvenile Douglas-fir (*Pseudotsuga menziesii*) height growth, though this positive effect transformed to a negative effect after year 7 (Woodruff et al. 2002). MacFarlane et al. (2000) found a strong negative correlation between site index predictions and stand density for a 16 year old loblolly pine stand. Although, the dominant trees in a stand are selected for site index calculation, stand density affects the height growth of all trees within a stand. In another study, BAL, a variable that includes trees larger in size than the subject tree, had a greater influence on height growth (Uzoh and Oliver 2006) than the total basal area. SDI (Stand density Index), which is also a measure of stand density, had the smallest influence among all the variables in the height growth model (Uzoh and Oliver 2006).

Quadratic mean diameter of the stand had a negative effect on height growth (Table 3-11 and Table 3-12). A higher quadratic mean diameter means that there are

more large sized trees. The bigger trees are usually taller, and there will be more competition for light and other resources than when smaller trees are present. Also the stand with higher *qmd* will have large old trees, which also have a slower growth rate. Competition could also be one-sided or two-sided (Cannell et al. 1984). Basal area in trees larger than the subject tree (BAL) is one-sided competition, because it reflects only the vertical competition for light. *Qmd* is a two-sided competition term as basal area per hectare of a stand (Yang et al. 2009), because it incorporates vertical competition for light as well as the horizontal competition for water and nutrients between trees.

There is always stochastic variation in growth of trees due to genetics, microclimate, and overall climate. This stochastic variation has been explicitly incorporated into growth models (Philips et al. 2003). Mixed models allow for modeling the stochastic variation due to climate, genetics and environment by partitioning the variation into different components. Yang et al. (2009) modeled the variation due to microclimate and environment by including a plot level random effect, and genetic variation by including a tree level random effect. Calama and Montero (2005) modeled climatic variation by including random effects for each growth period used to collect data. Therefore, incorporation of a mixed modeling approach in the present study allowed us to include stochastic variation. Because the mixed models incorporate subject specific variation as well, they are better for prediction than the fixed effects model. This has been clearly shown by our study as well. Even in mixed modeling, the addition of a random effect does not account for all plot level variation. This is shown by the improved prediction of the expanded non-linear model from the base model. The expansion of the model by adding stand and site variables to the intercept term will not

change the form of the relationship, but can improve the predictive power of the model (Uzoh and Oliver 2006).

Yang et al. (2009) were able to remove serial correlation in their BA increment model when both plot level and tree level random effects were incorporated. They had repeated measurements of trees within the same plot; therefore, the incorporation of only a plot level random effect in the model did not remove as much of the correlation as compared to adding random effects at both levels. In our data set, we did not have repeated measurements; we were only concerned about the spatial correlation of trees within the same plot. None of the residual covariance structures that account for spatial correlation, however, performed better in our model. This result suggests that the incorporation of a plot level random effect might mitigate this problem.

We tested whether a separate model was required for different community types. The likelihood ratio test indicated that a separate model was needed for UMF. Similar results for a diameter growth model were observed for slash pine (Chapter 2).

The height growth model presented in this study showed that the tree level variables such as initial tree height and diameter, competition terms, and stand level variables such as basal area and quadratic mean diameter are important for predicting the height growth of a slash pine. This study also indicated that a non-linear model is better than a linear model for predicting height growth in slash pine. We also showed that a non-linear mixed model, including random effect, gives better prediction than a fixed effect model. Once the random plot effect is incorporated, we did not find any improvement in the model prediction by modeling the correlation of trees within a plot through several residual covariance structures. Our study did not suggest a separate

model for different communities except for UMF. The model predictions are biologically consistent. The model presented in the study will be useful for simulating growth of slash pine.

Table 3-1. Parameter estimates with standard errors for linear fixed and mixed effects height growth model for slash pine in north central Florida

Effect	Estimate		SE		Pr> t	
	fixed	mixed	fixed	mixed	fixed	mixed
Intercept	-0.1811	-0.0955	0.1488	0.117	0.2247	0.59
<i>dbh</i>	-0.0349	-0.0265	0.0084	0.0089	<0.0001	0.0034
<i>Height</i> <sup>1</sup>	-1.3479	0.0989	2.8211	3.0441	0.6332	0.9741
<i>CI</i>	-0.1603	-0.1362	0.1457	0.1442	0.2722	0.346
<i>BAL</i>	1.4054	0.5294	0.3956	0.3742	0.0005	0.1585
<i>tbaha</i>	0.0057	0.0044	0.0015	0.0021	0.0003	0.0389
<i>dbh x Height</i> <sup>1</sup>	1.14	0.9079	0.1902	0.1953	<0.0001	<0.0001
<i>BAL x tbaha</i>	-0.025	-0.0152	0.0055	0.0053	<0.0001	0.0047

Table 3-2. AIC and BIC values of the non-linear model (equation 3-8e), treating various parameters as having both fixed and random effects.

Random parameters	AIC	BIC
<i>Intercept</i> ( $\beta_0$ )	-30.2	-15.9
<i>Height</i> ( $\beta_1$ )	-28.7	-14.3
<i>CI</i> ( $\beta_3$ )	-24.3	-10.0
<i>dbh</i> ( $\beta_4$ )	-28.6	-14.3

Table 3-3. Goodness of fit statistics for non-linear model (equation 3-8e), treating various parameters as having both fixed and random effects.

Fit statistics	<i>Intercept</i> ( $\beta_0$ )	<i>Height</i> ( $\beta_1$ )	<i>dbh</i> ( $\beta_4$ )
MMP(V)	0.005830	0.005727	0.005828
MMR (E)	-0.01541	-0.01593	-0.01453
MMSQ	0.012231	0.014969	0.011755
MMSQP (MSE)	0.03289	0.036553	0.03225
RMMQSP (RMSE)	0.16672	0.17539	0.16408
pbias	-2.68	-2.77	-2.52
SDbias	0.079	0.090	0.070

Note: MMP = Model Precision; MMR = Mean Residual (Bias); MMSQ = Mean Square Error; MMSQP = Mean Square Error of Prediction, pbias = Percentage bias, SD bias = Standard deviation of bias.

Table 3-4. Comparison of AIC and BIC values when including fixed versus mixed effects, for linear and non-linear models (Models 1 and 2, respectively).

Models	AIC	BIC
<u>Linear model (Model 1)</u>		
Fixed effects (eqn. 3-9a)	53.7	57.3
Mixed effects (eqn. 3-9b)	-10.3	-6.3
<u>Non-linear model (Model 2)</u>		
Fixed effects (eqn. 3-8e)	11.7	32.9
Mixed effects (eqn. 3-10)	-28.6	-14.3

Table 3-5. Comparison of various fit statistics when including fixed versus mixed effects, for linear and non-linear (Models 1 and 2, respectively).

Fit statistic	Linear model (Model 1)		Non-linear model (Model 2)	
	Fixed effects (eqn. 3-9a)	Mixed effects (eqn. 3-9b)	Fixed effects (eqn. 3-8e)	Mixed effects (eqn. 3-10)
MMP(V)	0.00568	0.00513	0.006496	0.005828
MMR (E)	-0.02721	-0.01026	-0.02647	-0.01453
MMSQ	0.04079	0.04093	0.040828	0.011755
MMSQP (MSE)	0.07081	0.06945	0.071096	0.03225
RMMQSP (RMSE)	0.23375	0.23385	0.24066	0.16408
pbias	-4.73	-1.78	-4.6	-2.52
SD bias	0.18	0.19	0.18	0.07

Note: MMP = Model Precision; MMR = Mean Residual (Bias); MMSQ = Mean Square Error; MMSQP = Mean Square Error of Prediction, pbias = Percentage bias, SD bias = Standard deviation of bias.

Table 3-6. Fit statistics for fixed effects linear model, non-linear base model, and non-linear model expanded with different methods without intercept term.

Fit statistic	Expanded without intercept			
	Linear (eqn. 3-9a)	Non-linear base (eqn. 3-8e)	M1 (eqn. 3-13)	M2 (eqn. 3-14)
MMP(V)	0.00568	0.00649	0.00612	0.00548
MMR (E)	-0.02721	-0.02647	-0.02371	-0.02164
MMSQ	0.04079	0.040828	0.03919	0.03762
MMSQP (MSE)	0.07081	0.07109	0.07288	0.06837
RMMQSP (RMSE)	0.23375	0.24066	0.24045	0.23564
Pbias	-4.73	-4.6	-4.12	-3.76
SD bias	0.18	0.18	0.18	0.17

Table 3-7. Fit statistics for mixed effects linear model, non-linear base model, and non-linear model expanded with different methods without intercept.

Fit statistic	Expanded without intercept			
	Linear (eqn. 3-9b)	Non-linear base (eqn. 3-10)	M1 (eqn. 3-13)	M2 (eqn. 3-14)
MMP(V)	0.00513	0.005828	0.00558	0.00532
MMR (E)	-0.01026	-0.01453	-0.01038	-0.00924
MMSQ	0.04093	0.011755	0.01064	0.01036
MMSQP (MSE)	0.06945	0.03225	0.03271	0.03063
RMMQSP (RMSE)	0.23385	0.16408	0.16053	0.15501
pbias	-1.78	-2.52	-1.80	-1.60
SD bias	0.19	0.07	0.07	0.07

Note: MMP = Model Precision; MMR = Mean Residual (Bias); MMSQ = Mean Square Error; MMSQP = Mean Square Error of Prediction, pbias = Percentage bias, SD bias = Standard deviation of bias.

Table 3-8. AIC and BIC for expanded non-linear mixed model (equation 3-14) fit with residual covariance structures to account for spatial correlation between trees within a plot.

Spatial Covariance Structure	AIC	BIC
None	-28.2	-13.7
Power [sp(pow)]	-22.8	-6.2
Exponential [sp(exp)]	-22.8	-6.2
Spherical [sp(sph)]	-22.8	-6.2
Linear [sp(lin)]	-22.8	-6.2
Log linear [sp(linl)]	-22.8	-6.2
Gaussian [sp(gau)]	-22.8	-6.2
Matern [sp(matern)]	-20.8	-2.1

Table 3-9. Fit statistics for linear and non linear mixed base model, and non linear mixed model expanded with different methods using the cross validation approach.

Fit statistics	Non-linear mixed model without intercept			
	Linear Mixed Model	Non-linear mixed Base model (eqn 3-10)	Expanded with M1 (eqn 3-13)	Expanded with M2 (eqn 3-14)
MMP(V)	0.00587	0.00631	0.00566	0.00540
MMR (E)	-0.00892	-0.01763	-0.00811	-0.00718
MMSQ	0.01240	0.01395	0.00777	0.00778
MMSQP (MSE)	0.03448	0.03497	0.02907	0.02748
RMMQSP (RMSE)	0.16421	0.17101	0.14429	0.14166
pbias	-1.55	-3.06	-1.41	-1.24
SD bias	0.08	0.08	0.04	0.04

Note: MMP = Model Precision; MMR = Mean Residual (Bias); MMSQ = Mean Square Error; MMSQP = Mean Square Error of Prediction, pbias = Percentage bias, SD bias = Standard deviation of bias.

Table 3-10. AIC and BIC values, and likelihood ratio test results of constrained height growth model with full model, and a separate model for the UMF.

Model Description	-2LL	Pr> $\chi^2$	AIC	BIC
All combined (completely constrained model)	-42.2		-28.2	-13.7
*UMF vs. MF-SA-BW	-46.8	0.031*	-30.8	-14.2
All separate (Full model)	-48.4	0.102	-28.4	-7.7

\*significant at  $\alpha < 0.10$  level.

Table 3-11. Parameter estimates and their respective standard errors and p-values for the selected annual height growth model (equation 3-14) with fixed effects only.

Parameter	Estimate	SE	Pr> t
$\beta_1$	0.6608	0.0867	<0.0001
$\beta_2$	-0.2023	0.0157	<0.0001
$\beta_3$	0.0035	0.0009	0.0004
$\beta_4$	-0.0348	0.0053	<0.0001
$\beta_5$	0.2427	0.0391	<0.0001

Table 3-12. Parameter estimates and their respective standard errors and p-values for the selected annual height growth model (equation 3-14) with random effect.

Parameter	Estimate	SE	Pr> t
$\beta_1$	0.6594	0.1401	<0.0001
$\beta_2$	-0.2199	0.0245	<0.0001
$\beta_3$	0.0030	0.0016	0.0745
$\beta_4$	-0.0451	0.0069	<0.0001
$\beta_5$	0.3423	0.0439	<0.0001
$\mathbf{b}_i$	0.0029	0.0007	0.0004
$\epsilon_{ij}$	0.0379	0.0037	<0.0001

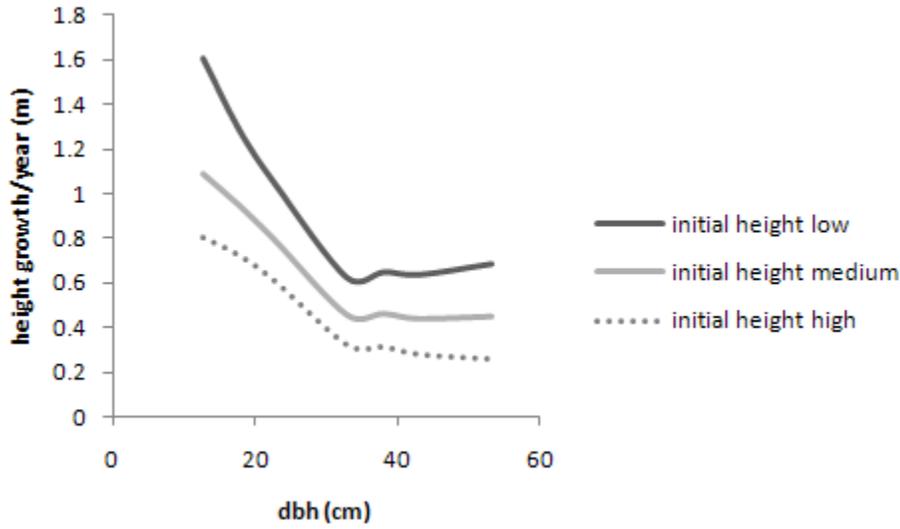
Table 3-13. Parameter estimates and their respective standard errors and p-values for the selected annual height growth model for UMF (equation 3-17) with random effect.

Parameter	Estimate	SE	Pr> t
$\beta_1$	0.8170	0.1264	<0.0001
$\beta_2$	-0.2274	0.0179	<0.0001
$\beta_3$	0.0022	0.0010	0.0258
$\beta_4$	-0.0312	0.0052	<0.0001
$\beta_5$	0.2327	0.0369	<0.0001
$m_1$	0.1524	0.0354	<0.0001

Table 3-14. Parameter estimates and their respective standard errors and p-values for the selected annual height growth model for UMF (equation 3-17) with fixed effects only.

Parameter	Estimate	SE	Pr> t
$\beta_1$	0.7462	0.1657	<0.0001
$\beta_2$	-0.2325	0.0252	<0.0001
$\beta_3$	0.0024	0.0016	0.1379
$\beta_4$	-0.0429	0.0068	<0.0001
$\beta_5$	0.3250	0.0433	<0.0001
$m_1$	0.1264	0.0575	0.0320
$\mathbf{b}_i$	0.0024	0.0007	0.0004
$\epsilon_{ij}$	0.0382	0.0037	<0.0001

a)



b)

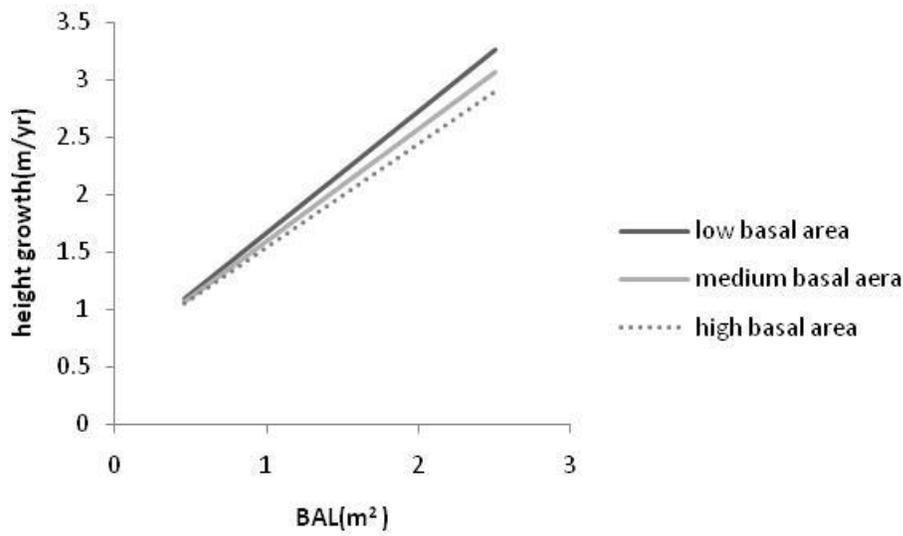


Figure 3-1. Height growth per year for slash pine in Florida using linear fixed effects model (equation 3-9a): a) height growth versus dbh for different initial tree heights b) height growth versus BAL (basal area in larger trees) for different stand basal areas.

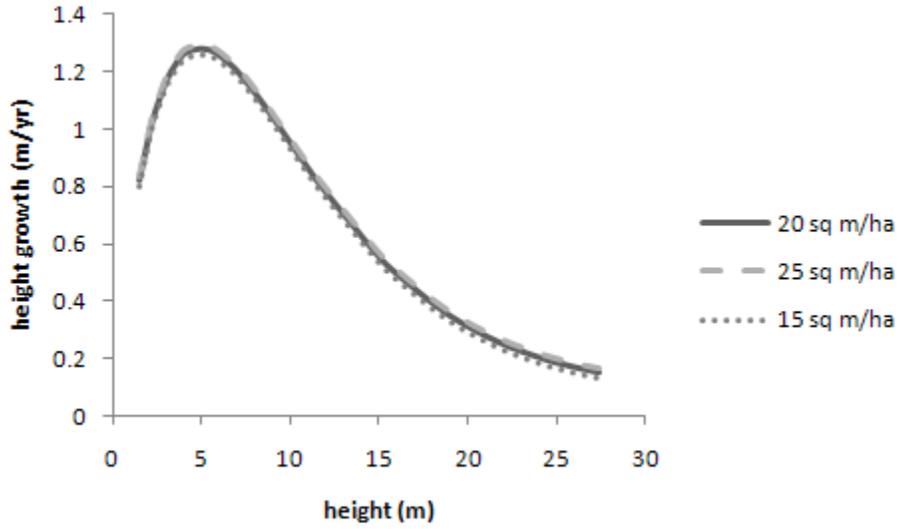


Figure 3-2. Annual height growth of slash pine in Florida against total tree height for different basal areas per hectare

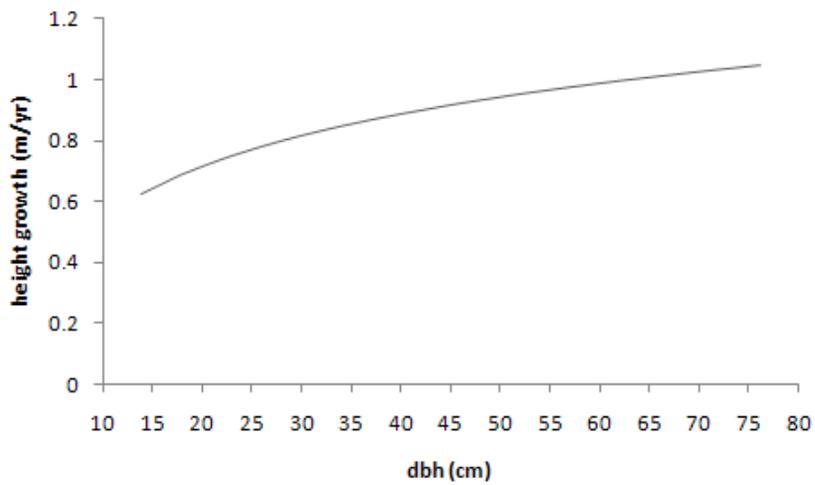


Figure 3-3. Annual height growth of slash pine in Florida against dbh

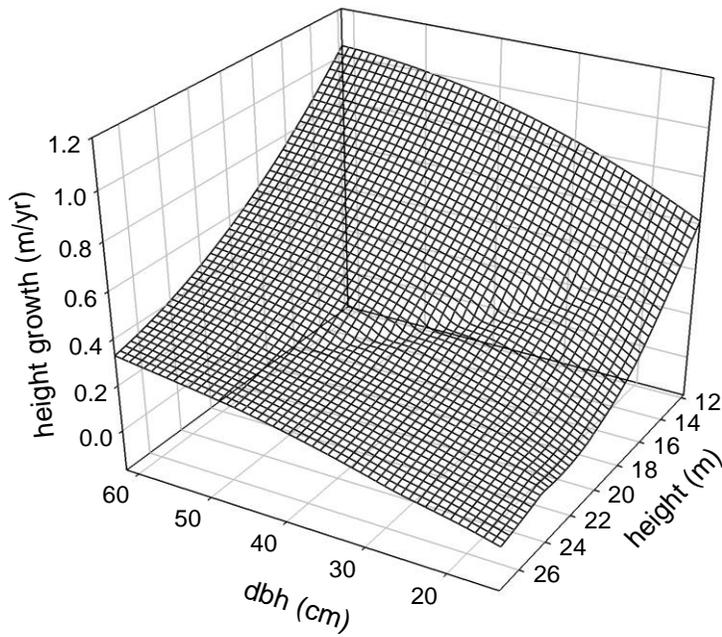


Figure 3-4. Annual height growth of slash pine in Florida against total tree height and dbh.

## CHAPTER 4 INDIVIDUAL TREE MORTALITY MODEL FOR SLASH PINE IN FLORIDA: A MIXED MODELING APPROACH

### **Introduction**

Mortality is an important component of all growth and yield models (Avila and Burkhardt 1992). When a tree dies, the reduced competition benefits the trees near the dead tree and positively impacting their growth (Yang et al. 2003). Similarly, gaps created by dead canopy trees are later filled by lateral growth of surrounding trees or new recruits (Oliver and Larson 1996). These gap dynamics are important determinants of structure and composition of a forest (McCarthy 2001). To improve the accuracy and biological authenticity of simulation models, the mortality process should be modeled and considered in stand simulation models.

Mortality is the least understood component of growth and yield estimation (Hamilton 1986, Monserud and Sterba 1999). This process is difficult to model due to the stochastic nature of mortality events. A tree's death is related to both its genetic makeup and the environment (Monserud and Sterba 1999). Tree death has been categorized into irregular and regular mortality. Irregular mortality is death due to catastrophic disturbances such as fire, windthrow, and diseases. Regular mortality is death due to suppression and competition for resources with other individuals (Lee 1971, Kiernan et al. 2009). In modeling mortality, only regular mortality is taken into account and modelers disregard genetic status, and other environmental extremes such as wind, drought, and disease (Monserud and Rehfeldt 1990, Monserud and Sterba 1999). Monserud and Sterba (1999) argue that the mortality will appear less stochastic if we can include these variables along with variables related to individual tree size, competition and stand characteristics. One way to overcome this problem is to include

stand or plot level random effects in mortality models in order to address some of the unexplained variation in mortality process due to climatic or environmental fluctuations.

Mortality has been modeled at the stand level using functions relating the number of trees per unit area at different ages to make predictions about future populations (Somers et al. 1980, Clutter et al. 1983). In contrast, individual tree mortality models estimate the probability of survival of individual trees throughout the growth period (Clutter et al. 1983). Using trees as the basic unit in individual tree models provides detailed information on stand composition and dynamics. Individual tree models are essential in modeling mixed species and uneven-aged stands (Avery and Burkhart 2002).

In the previous studies, mortality has been successfully modeled using logistic regression approaches (Monserud 1976, Hamilton 1986, Monserud and Sterba 1999, Cao 2002). The variables that have been found to be useful in modeling mortality are: tree size or vigor, individual tree competition, stand characteristics (density and site quality), and individual tree growth rate (Eid and Tuhus 2001, Yang et al. 2003, Chen et al. 2008). For example, Monserud (1976) used predicted diameter growth, competition index and predicted diameter in their survival model for northern hardwood forests. Cao (2002) used tree diameter, total height, crown ratio, and stand level variables such as quadratic mean diameter and basal area in their survival model for loblolly pine. There are several variables that can be used to represent these stand and tree level characteristics, e.g., tree vigor can be represented by dbh, height or crown size. Competition can be represented by relative measures such as ratio of subject tree diameter to average tree diameter or basal area in larger trees or other distance

dependent competition indices. The effects of stand density can be represented by basal area or trees per unit area. Site quality can be represented using site index, topography, soil characteristics, vegetation type, elevation etc. The utility of all these variables and their transformation varies among species due to their physiognomy, and growth and habitat characteristics. For example, survival probability of loblolly pine (*Pinus taeda*) decreased with increasing site index for the Piedmont/Upper Coastal Plain whereas the opposite was true for the Lower Coastal Plain (Zhao et al. 2007).

Most mortality models developed to date have ignored the multi-level data structure inherent in the common designs used in collecting forestry data. For example, trees are usually nested within plots and repeated measurements are often available for individual trees within the plots (Rose et al. 2006). Considering the multi-level data structures in models allows for efficient estimation of regression coefficients and reduces biases in the parameter estimates (Goldstein 1991, Rodriguez and Goldman 2001). Although several studies exist in forestry that consider multi-level structures when the response is continuous (Hall and Bailey 2001, Fang and Bailey 2001, Yang et al. 2009), only a few (e.g., Rose et al. 2006) have considered the hierarchical data structure where the response is binary as in mortality data.

Few individual tree mortality models for slash pine (*Pinus elliotti* var. *elliotti*) in Florida exist. Bailey et al. (1985) developed stand level survival model for slash pine plantations and related survival to density, stand age, site index, and thinning intensity. Lee and Coble (2002) developed a stand level survival model for unthinned slash pine plantations in east Texas. They used simultaneous nonlinear regression to model mortality and found the surviving planted slash pine declining with increasing density of

non-planted trees and increasing site quality. A post-fire survival analysis of south Florida slash pine (*Pinus elliotti* var. *densa*) was carried out by Menges and Deyrup (2001). Using path analysis, they concluded that the important determinants of mortality after fire include: season of burn, fire intensity, presence of beetle attack after fire, and presence of scrubby vegetation in the understory.

The objective of this study is to develop an individual tree mortality model for slash pine that accounts for both irregular and regular mortality. We tested variables suggested in the literature to model individual tree survival for slash pine in Florida, and selected the best suite of variables. We first fit a model with only fixed effects using a logistic function, and then added a random effect (to account for the multi-level nature of the data) using a generalized linear mixed model to compare the outcomes of the two fitting processes. The models were also validated using a cross validation approach. The data used for model development was collected from both natural and managed stands representing different natural communities, so we expect the model to be representative of slash pine grown in north-central Florida. The ultimate goal is to use the mortality model in a simulation model to predict the stand dynamics of even-aged slash pine stands converted into uneven-aged. We also investigated the cut off probability that should be used to assign trees into dead or live categories from continuous mortality probability values.

## **Methods**

### **Data**

The data used in this study came from property owned and managed by the Suwanee River Water Management District (SRWMD) in north central Florida, USA. SRWMD is among the five water management district created in Florida by the Water

Resource Act of 1972. SRWMD established permanent plots in 2000, and we re-measured those plots in 2008.

SRWMD followed a sampling design and protocol similar to that used by the USDA Forest Service to collect Forest Inventory and Analysis data (Betchold and Patterson 2005). A cluster of four circular 7.31 m (24 ft) plots were established every 101.15 ha (250 acres). Each cluster included a center plot and three other plots at 90, 120 and 240 degrees and at a distance of 36.57 m (120 ft) from the center plot. For every plot, the following data were collected: current community; historical community; stand origin (natural, mixed or plantation); stand structure; stand age; slope; aspect; and GPS locations. For each plot, three dominant and co dominant trees were identified as site trees, and their height, age and species were recorded. In every 7.31 m radius tree plot, trees > 12.54 cm (5 inches) in diameter were identified to species and the following data were recorded: distance and azimuth from the plot center; diameter at breast height (dbh); total height; height to live crown; product class; crown class; and tree damage. All of the above described information were also collected for trees < 12.54 cm (5 inches) dbh, but on a 2.07 m (6.8 ft) radius subplot concentric within the larger plot. In addition to this, data were also available on fire and silvicultural history in each plot.

To select plots for 2008 re-measurement, all slash pine plots measured in 2000 were stratified on the basis of community type, age class (10 year increment) and origin. Plots were randomly selected from each stratum. Altogether, 70 plots were selected for re-measurement. In addition to all the variables collected in 2000, we also measured crown radius in four cardinal directions and used this information to estimate mean crown radius and total crown area (assuming an ellipsoid shape) for each tree. Trees in

each plot were identified as dead (1) or survived (0), and we also determined the cause of death wherever possible. In 2008, data was collected for trees in a 6.40 m (21 ft) buffer around each tree plot so that the competition indices described below could be calculated for trees within the 2000 measurement plot boundaries.

Crown radius was not measured in 2000, and a linear allometric relationship was developed relating 2008 mean crown radius to dbh, height and crown length in 2008. The relationship was used to predict mean crown radius in 2000. Dbh and height information were not available for buffer trees in year 2000. Therefore, we developed linear allometric equations relating dbh and height in 2008 to dbh and height in 2000, and used this relationship to estimate dbh and height of buffer trees in 2000. This was done only to calculate competition indices using the buffer trees. A random error term was added to all the predictions. The relationships were developed separately for two types of pines and four hardwood groups. Groupings of hardwoods corresponded to *Quercus* sp., *Taxodium* sp., the Cornaceae family, and other hardwoods present in the plot. The pine tree data consisted overwhelmingly of slash pine, with relatively fewer loblolly pine (*Pinus taeda* L.) and only a few longleaf pine (*Pinus palustris* Mill.) and pond pine (*Pinus serotina* Michx.). A separate predictive model was developed for slash pine, with longleaf and pond pine grouped together with loblolly pine.

Stocking of each plot was calculated as a measure of crowding following May (1990). Total basal area per hectare and quadratic mean diameter of each plot was also calculated. Site index for each plot was calculated following Farrar (1973).

Various distance dependent and independent competition indices were calculated for use in the model. Choi et al. (2007) defined the competitive zone as an area 3.5

times the average crown radius of dominant and codominant trees in their study of northern hardwoods. We modified this slightly for our pine-dominated stands, using 3 times the average radius of the measured slash pine trees (2.13 m; 7 ft), resulting in a competitive zone with 6.40 m (21 ft) radius around the subject tree. All variables indicating competition (both distance dependent and independent, as defined in Cole and Lorimer 1994, Choi et al. 2001, Choi et al. 2007) were calculated considering trees within this competition zone. Two relative diameter indices were calculated as:

$$RD_i = \frac{dbh_i}{\frac{1}{n} \sum_{j=1}^n dbh_j} \quad \text{and} \quad RDgh_i = \frac{dbh_i}{\frac{1}{n'} \sum_{j=1}^{n'} dbh_j}$$

where:  $dbh_i$  is the diameter at breast height of the  $i$ th subject tree,  $n$ =number of all competitors within the 6.40 m radius competitive zone, and  $n'$ =number of trees within the 6.40 m radius competitive zone with diameters greater than or equal to the subject tree. Therefore,  $RD$  is the relative diameter calculated with all trees within the 6.40 m radius competitive zone of the subject tree, whereas  $RDgh$  is calculated only with trees equal to or larger than the subject tree. Relative height and relative basal area were calculated similarly.  $RHgh$  is the relative height calculated with trees within the 6.40 m radius competitive zone with heights greater than or equal to that of the subject tree, whereas  $RH$  is the relative height calculated with all trees within the 6.40 m ft radius competitive zone.  $RBAgd$  is the relative basal area calculated by including trees within the 6.40 m radius competitive zone with diameter greater than or equal to the subject tree, whereas  $RBA$  is the relative diameter calculated with all trees within the 6.40 m radius competitive zone of the subject tree.

Other competition indices investigated included the modified Hegyi indices for diameter and height ( $sum\_si$  and  $Sum\_Clr$ ; Heygi 1974), calculated as:

$$sum\_si_i = \sum_{j=1}^n \frac{dbh_j}{dbh_i d_{ij}}$$

where:  $dbh_j$  is the diameter of competitor tree  $j$ , and  $d_{ij}$  is the distance between the subject tree and its competing neighbors.  $Sum\_Clr$  is similar to  $sum\_si$ , using height as the variable instead of diameter:

$$Sum\_Clr_i = \sum_{j=1}^n \frac{H_j}{H_i d_{ij}}$$

We also calculated other measures of stand level competition such as *BAL* (basal area of trees equal to or greater in diameter than the subject tree) and quadratic mean diameter for each plot (*qmd*; Teck and Hilt 1990, Cao 2002, Adame et al. 2008).

### **Selection of Variables**

In modeling mortality, the variables that have been commonly used are related to tree size and vigor, individual tree competition, age, stand level competition, growth rate, and effect of site (Hamilton 1986, Monserud and Sterba 1999).

DBH as a representation of tree size has been widely used in mortality models (Hamilton 1986, Monserud and Sterba 1999, Cao 2002, Crecente-Campo et al. 2009). DBH is directly related to a tree's ability to compete for scarce resources; therefore, the larger the tree, the more it can compete for resources and less chance of its mortality. Besides dbh, tree height has also been used as a proxy for tree size for loblolly pine mortality (Cao 2002). In forest trees, mortality is higher when trees are smaller, and it declines rapidly when a tree reach a certain size. Mortality again increases once the tree becomes very large and old; thereby rendering a U shaped relationship between size and mortality (Lorimer and Frelich 1984). To incorporate this relationship in the

model, various transformations of dbh such as  $dbh^{-1}$ ,  $dbh^2$ , and square root of dbh have been used (Hamilton 1986, Monserud and Sterba 1999, Eid and Tuhus 2001). Another variable that has been used as the measure of size in mortality models is the crown ratio (Avila and Burkhart 1992, Zhang et al. 1997, Cao 2002). Crown ratio provides information about crown size, which is also an indicator of tree vigor. We tested these three variables related to tree size and vigor in our models. For dbh and height, we also tested various transformations singly or in combination to describe the nonlinear relationship of tree size with mortality.

Competition is another variable that affects tree mortality. There are many competition indices available, both distance-dependent and independent, that can be tested in mortality models. For including the effect of competition on individual trees, Wykoff et al. (1982) and Wykoff (1990) introduced a tree-specific measure of density, basal area in larger trees (BAL), which has been very useful in predicting stand dynamics. The variable along with the crown competition factor (Krajicek et al., 1961) has also been used as a measure of competition by Monserud and Sterba (1999). CCF is the maximum possible area per acre occupied by tree crowns if they are grown in open, if it is less than 100 that means there are some growing spaces available. The ratio of the quadratic mean diameter of the stand to a subject tree diameter (Cao 2002) or vice versa (Crecente-Campo et al. 2009) has also been used as a measure of competition. Relative terms such as relative diameter (Hamilton 1986, Chen et al. 2008), relative height, and relative basal area can also be used to represent distance independent competition. We tested each of the distance dependent and independent competition indices described earlier (i.e., *RD*, *RDgh*, *RH*, *RHgh*, *RBA*, *RBAgd*, *BAL*,

*sum\_si*, and *Sum\_Clr*) in our models and selected the competition index that best explained our data.

Annual diameter increment of the previous growing period also has been used as the independent variable, but usually where tree growth has been measured multiple times (Hamilton 1986, Yao et al. 2001). We did not use growth increment in our model; we did not have repeated measurement data, and therefore diameter increment data were not available for dead trees.

Stand level variables such as basal area per unit area (acre or hectare) have also been widely used in growth models, since this is a measure of overall competition (Hamilton 1986, Yao et al. 2001, Cao 2002). Hamilton (1986) also tested various transformations of basal area, such as the inverse, square root and logarithm, but the choice of transformation did not improve model fit. Hamilton (1986) selected the square root of basal area because it was hypothesized that mortality rate increased at a decreasing rate with increases in stand basal area. We tested basal area in our model along with its transformations singly or in combination.

Monserud and Sterba (1999) did not include site variables such as soil type, vegetation type, slope, aspect, and elevation in their mortality models because of the risk of over-parameterization. Age and site index were also not included because they wanted to generalize their model for both even and uneven-aged stands. Yao et al. (2001) found conflicting evidence (different signs of the coefficients for different species) for using a site productivity index in their mortality model. Other studies have used site index to describe site conditions in their models (Eid and Tuhus 2001, Diéguez-Aranda et al. 2005). Since our ultimate goal is to use this mortality model in a simulation

converting even-aged stands (where site index information are easily available) into uneven-aged stands, we used site index as one of the variables to describe our sites. Age information was not available for all the trees; therefore, our model did not include this variable.

### Model Development

Mortality in the data was coded as 0 or 1 (0 = survived; 1 = dead). The standard method for analyzing such binary responses is by using a logistic function (Hamilton 1986, Avila and Burkhart 1992, Eid and Tuhus 2001, Diéguez-Aranda et al. 2005). Weighted non-linear regression and multivariate maximum likelihood methods are the commonly used methods for parameter estimation of such non-linear functions bounded by 0 and 1. In likelihood methods, parameters are estimated iteratively until parameters that maximize the log of the likelihood are obtained. We define probability of survival as  $P_s$  and probability of mortality as  $P_m$ . In the logistic model,  $P_m = P(Y=1 | \mathbf{X} = \mathbf{x}) = 1 - P(Y=0 | \mathbf{X} = \mathbf{x}) = 1 - P_s$ , where  $Y$  and  $\mathbf{X}$  are response and explanatory variables, respectively, is given by:

$$P_m = \frac{1}{1+e^{-\mathbf{x}\boldsymbol{\beta}}} \quad (4-1)$$

$$P_s = \frac{e^{-\mathbf{x}\boldsymbol{\beta}}}{1+e^{-\mathbf{x}\boldsymbol{\beta}}} \quad (4-2)$$

Equivalently, the log odds or logit of the probability has a linear relationship with the explanatory variables, and is given by:

$$\log\left(\frac{P_m}{1-P_m}\right) = \mathbf{X}\boldsymbol{\beta} \quad (4-3)$$

where  $\mathbf{X}$  is the matrix of independent variables related to tree size, competition, and stand level variables,  $e$  is the base of the natural logarithm, and  $\boldsymbol{\beta}$  is the vector of parameters to be estimated.

The assumption of independence is violated when data are temporally or spatially correlated. In the present study, the grouping of trees within a sample plot creates a hierarchical data structure. Trees within the same plot are more correlated than trees in different plots. We considered this hierarchical data structure with within-plot correlation as well as the binary nature of the data by fitting a generalized linear mixed model (GLMM) using the SAS procedure PROC GLIMMIX. GLMMs are an extension of GLMs (generalized linear models) to incorporate correlation among the responses (Schabenberger 2005). This is accomplished by adding a random component to the linear predictor or by directly modeling the correlation among the responses. For more information, see for example, Schabenberger (2005), SAS Institute Inc. (2008), or Schabenberger and Pierce (2001). A general linear mixed model for plot  $i$  can be written in the following form:

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i \quad (4-4)$$

where  $\mathbf{X}_i$  and  $\mathbf{Z}_i$  are design matrices for fixed and random effects parameters, respectively.  $\boldsymbol{\beta}$  and  $\mathbf{b}_i$  are vectors of fixed effects parameters and random effects, respectively.  $\boldsymbol{\varepsilon}_i$  is a vector of errors.  $\mathbf{X}_i$  and  $\mathbf{Z}_i$  are the partial derivatives of the model with respect to fixed effects parameters, and random effects parameters, respectively. The random effects,  $\mathbf{b}_i$ , are normally distributed with mean 0 and variance-covariance matrix  $\mathbf{G}$ , and the errors,  $\boldsymbol{\varepsilon}_i$ , are also normal with mean 0 and variance-covariance matrix  $\mathbf{R}$ . The correlation among the observations can be modeled either by specifying columns of the  $\mathbf{Z}_i$  matrix and the structure of the  $\mathbf{G}$  matrix (modeling  $\mathbf{G}$  side effects) or by specifying the covariance structure of the  $\mathbf{R}$  matrix directly (modeling  $\mathbf{R}$  side effects).

With generalized linear mixed models, a similar approach is used to model **G** side random effects, where a random effect is added to the linear predictor giving the following model form:

$$g(E[\mathbf{Y}|b_i]) = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}b_i \quad (4-5)$$

where  $\mathbf{Y}$  is a vector ( $n \times 1$ ) of observed data, and  $b_i$  is a vector of random effects.  $\mathbf{X}$  ( $n \times p$ ) is a matrix of rank  $k$  and  $\mathbf{Z}$  ( $n \times r$ ) design matrix for the random effects.  $g(\cdot)$  is a link function that relates the mean of the data to the linear predictor. Here the interest is in the distribution of  $\mathbf{Y}$  conditioned on the random effects  $b_i$ .  $\mathbf{Y}$  can be from any member of the exponential family of distributions, e.g., binomial, Poisson, negative binomial. The model with **G** side random effects is the conditional model form. When only the **R** side random effects are modeled (the marginal model), the model form is given by:

$$g(E[\mathbf{Y}]) = g(\mu) = \mathbf{X}\boldsymbol{\beta} \quad (4-6)$$

In the marginal model only the mean and a model including covariance among observations is specified. The variance of  $\mathbf{Y}$  is given by:

$$\text{var}[\mathbf{Y}] = \mathbf{A}^{1/2} \mathbf{R} \mathbf{A}^{1/2} \quad (4-7)$$

where  $\mathbf{A}$  is a diagonal matrix containing variance functions. We can also model **G** side random effects and **R** side effects together. For example, in the present study we can model plots as the **G** side effect and spatial correlations among trees within a plot using the **R** side effect. However, sometimes modeling the **G** side random effect will take care of correlation among trees within a plot. When both effects are modeled, the model form and the variance are given by the following expressions:

$$\mathbf{E}[\mathbf{Y} | b_i] = g^{-1}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}b_i) = g^{-1}(\boldsymbol{\eta}) = \boldsymbol{\mu} \quad (4-8)$$

$$\text{var}[b_i] = \mathbf{G} \quad (4-9)$$

$$\text{var}[\mathbf{Y}|\mathbf{b}_i] = \mathbf{A}^{1/2} \mathbf{R} \mathbf{A}^{1/2} \quad (4-10)$$

where  $g^{-1}(\cdot)$  is the inverse link function.

Using the generalized linear mixed model, a population averaged (PA) response for the mortality of the  $j$ th tree in the  $i$ th plot,  $P_{m_{ij}}$ , without considering the subject specific random effect is given by the following equation:

$$P_{m_{ij}} = 1 / \left[ 1 + \exp \left( - \left( \beta_0 + \beta_1 X_{1ij} + \dots + \beta_n X_{nij} \right) \right) \right] \quad (4-11)$$

The subject specific (SS) response incorporates the random effect of individual subjects and is given by:

$$P_{m_{ij}} = 1 / \left[ 1 + \exp \left( - \left( \beta_0 + \beta_1 X_{1ij} + \dots + \beta_n X_{nij} + b_i \right) \right) \right] \quad (4-12)$$

In the present study, we modeled 8 year probability of mortality for each tree. The available data for all the plots were from the same period. If annual probability of mortality is of interest in a simulation model, the 8 year mortality can be converted to annual mortality using the algebraic approach following Flewelling and Monserud (2002). This approach considers that a tree can only die once in its lifetime, and therefore, mortality is not a Markov process. However, survival is. Thus, all algebraic manipulation should be done using probability of survival. A compound interest formula is used to convert 8 year survival into annual mortality. In year one, the probability of survival  $P_{s_1} = 1 - P_{m_1}$ . In year two,  $P_{s_2} = P_{s_1} P_{s_1} = (1 - P_{m_1})^2$ . In year  $n$ ,  $P_{s_n} = P_{s_1} P_{s_1} \dots P_{s_1} = (P_{s_1})^n = (1 - P_{m_1})^n$ . Therefore,  $P_{m_1} = 1 - P_{s_n}^{1/n}$ .

### Model Evaluation and Validation

We fit both a logistic function, and a generalized linear mixed model including a random plot effect. For comparing models fit with a logistic function, the  $R^2$  goodness of

fit statistics used to assess the best model in linear regression is not appropriate due to the binary nature of the response variable (Ryan 1997, Diéguez-Aranda et al. 2005). Therefore, we used a generalization of the coefficient of determination ( $\bar{R}^2$ ) following Nagelkerke (1991) to compare competing models. Similar to the Akaike Information Criteria (AIC) and Schwarz Criterion (SC; also known as the Bayesian Information Criteria) statistics, the  $\bar{R}^2$  value is useful in comparing competing models that are not necessarily nested (SAS Institute Inc. 2008). The value of  $\bar{R}^2$  lies between 0 and 1, and higher value indicates better model.

$$\bar{R}^2 = \frac{R^2}{\max(R)^2} = \frac{1 - \left(\frac{L(0)}{L(\hat{\beta})}\right)^{2/n}}{1 - L(0)^{2/n}}$$

where:  $L(0)$  is the likelihood of the intercept only model,  $L(\hat{\beta})$  is the likelihood of the specified model, and  $n$  is the number of samples. In addition to  $\bar{R}^2$ , we also used AIC and SC statistics to compare competing models (SAS Institute Inc. 2008). Lower values of AIC and SC indicate better models. In addition, statistical significance of each model parameter was also considered in model evaluation and comparison.

We also used the Hosmer-Lemeshow goodness of fit test to evaluate the overall model (Hosmer and Lemeshow 2000). To perform the test, approximately  $g=10$  groups, each including roughly equal number of subjects, are created based on the distribution of estimated probabilities. Within those groups, the discrepancies between observed and predicted numbers of observations are summarized by calculating a Pearson chi-square statistic. This statistic is compared with a chi-square distribution with  $g-2$  degrees of freedom. A smaller  $p$ -value suggests that the model is not a good fit.

For evaluating generalized linear mixed models, we also compared AIC and BIC values. In addition, we calculated the ratio of the Pearson chi-square statistic to its

degrees of freedom to assess whether the model displayed good fit to the data. The closer this value is to 1 the better the model explains the variation (SAS Institute Inc. 2008).

Independent data was not available for model validation; therefore, we used a cross validation approach (Temesgen et al. 2008, Crecente-Campo et al. 2009). For cross validation, each plot was deleted from the data set and remaining plots were used to fit the models. Mortality probabilities of all trees in a deleted plot were estimated using the parameters from the fitted model. This was done for each plot in succession in the data. For calculating SS responses, the random effect calculated after fitting the generalized linear mixed model to the whole dataset was added to PA responses.

We also estimated area the under the receiver operating characteristic (ROC) curve to compare the models (Hosmer and Lemeshow 2000, Chen et al. 2008). The area under the curve was estimated for both model fitting data and cross validation data. ROC depends on the true positivity (sensitivity) and false positivity (1- specificity) of a test. Sensitivity is the proportion of observed positive events identified as positive, and specificity is the proportion of observed negative events identified as negative. In ROC curve evaluation, sensitivity is plotted against 1- specificity, and the area under the ROC curve is used to determine the discriminatory power of a curve. If the area is less than 0.5, then a model has no discriminatory power, if it is between 0.7 and 0.8, a model has acceptable discriminatory power, and if it is between 0.8 and 0.9, a model has excellent discriminatory power (Hosmer and Lemeshow 2000, Chen et al. 2008, Crecente-Campo et al. 2009).

Several methods have been used to convert continuous prediction probabilities from mortality models into a dichotomous variable that indicates whether a tree is dead (1) or alive (0). The cut-off approach is implemented by identifying a threshold probability above which a tree is considered dead (Monserud and Sterba 1999). A second method termed the stochastic approach is implemented by generating uniform random numbers between 0 and 1; a tree is considered dead if the estimated probability is above the generated value (Weber et al. 1986, Crecente- Campo et al. 2009). A third method termed the expansion factor approach uses the probability of mortality of each tree to reduce the number of trees per hectare (the expansion factor) represented by that particular tree (Wykoff et al. 1982, Hann et al. 1997, Crecente- Campo et al. 2009). Crecente-Campo et al. (2009) tested these three methods to determine their utility in classifying mortality status and found that the cut-off approach gave the best result based on the correct classification rate (percentage of correctly classified trees) and numbers of observed versus predicted mortalities for different classes of independent variables. Therefore, we also decided to use a cut-off approach and determined the threshold value that gave the best prediction of mortality. We used the predictions from cross validation, tested different cut off values, and selected the cut-off value that gave the best classification rate, agreement between observed and predicted number of dead trees, and maximized sensitivity and specificity percentages.

## **Results**

Since the time between the initial measurement and second measurement was eight years, we chose to model 8 year mortality directly. This can be converted to annual survival or mortality using the formula given in the Methods section. We first fit

the model using linear logistic regression and chose the best, most parsimonious model. The best model was given by the following (equation 4-13).

$$P_{m_{ij}} = 1 / \left[ 1 + \exp \left( -(\beta_0 + \beta_1 dbh_{ij} + \beta_2 H_{ij} + \beta_3 RBAgd_{ij} + \beta_4 SPI_i + \beta_5 tbaha_i + \beta_6 H_{ij} \times tbaha_i) \right) \right] \quad (4-13)$$

where  $\beta_0$ -  $\beta_6$  are estimated parameters of the model and all other variables are as described earlier.  $dbh_{ij}$  is the diameter of the  $j$ th tree in the  $i$ th plot,  $RBAgd_{ij}$  is the relative basal area,  $SPI$  is the site index, and  $tbaha_i$  is the stand basal area. The overall model was significant based on the likelihood ratio test ( $P < 0.0001$ ). All parameters in the model were significant based on their Wald chi-square statistics (Table 4-1), and the generalized coefficient of determination ( $\bar{R}^2$ ) was 0.1514. The Hosmer and Lemeshow test was insignificant ( $P = 0.26$ ) indicating that there was no significant difference between the observed and predicted mortality at  $\alpha = 0.05$ . The area under the receiver operating characteristic curve (ROC) was 0.70 for the model.

We then fit the model using a generalized linear mixed modeling framework to address the hierarchical structure of the data and account for correlations among trees within a plot. The generalized linear mixed model had the same model form as the linear logistic regression model, with the addition of a random effect (equation 4-14). All parameters in the model were significant and had similar signs and values compared to the linear logistic regression (Table 4-2). Although site index ( $SPI$ ) was not significant in the model, removing it from the model increased the AIC value.

$$P_{m_{ij}} = 1 / \left[ 1 + \exp \left( -(\beta_0 + \beta_1 dbh_{ij} + \beta_2 H_{ij} + \beta_3 RBAgd_{ij} + \beta_4 SPI_i + \beta_5 tbaha_i + \beta_6 H_{ij} \times tbaha_i + b_i) \right) \right]$$

(4-14)

where  $P_{m_{ij}}$  is the probability of mortality for  $j$ th tree in the  $i$ th plot, and  $b_i$  is the plot level random effect. The ratio of the Pearson chi-square statistic to its degrees of freedom was close to 1 (0.83). This is an indication of proper modeling of variability by the model, and lack of residual overdispersion. The generalized coefficient of determination ( $\bar{R}^2$ ) was 0.21. The area under the ROC curve was 0.81.

In the above models, both with and without random effects, we tested variables that indicate tree vigor and size such as dbh, height and crown ratio. These variables were tested separately or in combination. When crown ratio was introduced as a proxy of tree size in the model, it produced unreasonable results compared to dbh and height. All the competition indices, both distance dependent and distance independent were tested in the model. The best fit was provided by the model including *RBAgd* (a distance independent competition index) as the competition index. We tried to address the U shape relationship between size and mortality using various transformation of *dbh* such as  $dbh^2$ ,  $dbh^{-1}$  and  $\sqrt{dbh}$ , but the transformations did not improve the model fit and were also not significant at  $\alpha = 0.05$ . These variables were tested separately or in combination as done by other studies. Several transformations of basal area were also tested in the model.

The accuracy of the various models was compared using the ROC curve. The ROC curve for the logistic function and population averaged response of the generalized linear mixed model were the same for model fitting data. The area under

the curve was 0.70 for both models (Figure 4-1). When the subject-specific response of the generalized linear model was used, the area under the curve was 0.813 (Figure 4-1). This showed that the accuracy of the mortality model increased after adding a random plot effect. Similar results were obtained using the cross validation data. The area under the ROC curve was 0.631, 0.651, and 0.78 for the logistic function, population averaged response and subject-specific responses, respectively. This gives evidence suggesting that when periodic observations are available to calculate the random plot effect, subject-specific responses from a generalized linear mixed model give better predictions than prediction from linear logistic regression or population averaged functions from the generalized mixed model.

We estimated the probability of mortality for different values of the independent variables to test the biological consistency of the model. The probability of mortality decreased as the size of a tree increased (Figure 4-2). The probability of mortality is close to zero as trees reach 30 cm dbh. This probability should increase as the tree becomes older, but the data in our model did not show this relationship. The largest tree in our sample was 61 cm, and we failed to capture this relationship in our model. We also estimated mortality, using  $\text{dbh}^2$  in the model (although insignificant), but the shape of probability of mortality against dbh was virtually the same.

The probability of mortality decreased with tree height; however, the decreasing rate slowed once the tree reached around 15 m in height (Figure 4-3). This trend varied with basal area; as expected, higher basal area increased the probability of mortality in smaller trees (Figure 4-3), but the effect of basal area disappeared as trees reached approximately 17 m in height.

For similar sized trees (dbh and height), an increase in site index also increased the probability of mortality (Figure 4-4). The effect of site index on the mortality probability also decreased as trees grew larger (Figure 4-4). Similarly, for similar height and dbh, as the relative position of a tree in a stand improved (higher *RBAgd*), the probability of mortality decreased (Figure 4-5). This indicated that as individual tree competition increased, the mortality probability also increased.

When periodic data are available to calculate random effects, then we can use subject-specific (SS) responses to evaluate predictions. However, when periodic data are unavailable, we have to use population averaged responses. Therefore, we determined a threshold value using both. For the SS response, when the cutoff threshold was set to 0.46, the correct classification rate was the highest (82%), but the percentage of predicted dead trees was 4.7% compared to 19.3% dead trees observed in the data. While the model underestimated mortality in every dbh class (Figure 4-6a), it predicted no mortality for trees above 20 cm dbh. When the threshold value was changed to 0.303, the correct classification rate was reduced to 79%, but the percentage of predicted dead trees was 17.5%, and the distribution of dead trees over different dbh classes was similar to the observed percentages (Figure 4-6b). Also, using this threshold value maximized percentage sensitivity and percentage specificity.

Similarly, for the population averaged response, a threshold value of 0.59 gave an 81% correct classification rate. However, the predicted mortality rate was only 1.39%, and it also underestimated the mortality severely for larger trees (Figure 4-7a). When the threshold value was changed to 0.292, the correct classification rate decreased to 75%, but the predicted mortality over different dbh classes was closer to the observed

(Figure 4-7b) and the overall predicted percent mortality was 18.6%, much closer to the observed rate of 19.3% percent. While using this methodology did overestimate mortality in the lowest dbh class, using this value maximized percentage sensitivity and percentage specificity.

## Discussion

The mortality model presented in this study used dbh and height as a measure of tree vigor, relative basal area calculated using trees equal to or greater in size than the subject tree as a measure of individual competition, basal area per hectare as a measure of overall competition, and site index representing the growing condition. Other measures of tree vigor such as crown width and live crown ratio have been previously used in the mortality models (Wykoff et al. 1982, Monserud and Sterba 1999); these variables were not significant predictors of mortality of slash pine in our model. Since we did not have repeated measurements on dead trees, we could not use the diameter increment as our predictor. Diameter increment had been used by several authors as a predictor of mortality (Hamilton 1986, Yao et al. 2001, Palahi et al. 2003); however, Yao et al. (2001) did not find any effect of decrease in diameter increment on survival probability of white spruce (*Picea glauca* (Moench) Voss) and lodgepole pine (*Pinus contorta* Dougl. ex Loud. var. *latifolia* Engelm.).

Our model did not capture the decrease in survival or the increase in mortality for older trees, because the transformations of diameter (for e.g.,  $dbh^2$ ) were not significant in the model. This could be due to the unavailability of data on older stands. The oldest stand in our sample was 72 years, and the largest *dbh* tree was 60.96 cm (24 inches). Had we sampled older stands, we might have captured this relationship. The reason for declining survival probability in older trees has been attributed to increase in carbon

allocation to maintenance respiration and to fine roots, decrease in photosynthetic efficiency, and genetic traits of individuals (Ryan and Yoder 1997). Eid and Tuhus (2001) also did not find significant  $dhb^2$  term in their survival models for Norway spruce (*Picea abies* (L.) Karst.) and Scots pine (*Pinus sylvestris* L.). As in this study, the mortality rate declined rapidly for smaller trees, and started to decline slowly when trees increased in size.

Spatially explicit distance dependent competition indices were not a better predictor of mortality, and a distance independent index fared better in our model. *RBAgd* and *tbaha* represented competitive interactions between the trees. The higher value of *RBAgd* means relatively lower competition, therefore, the negative sign of the coefficient is biologically consistent. The probability of mortality decreased as the competitive pressure decreased (Figure 4-5). *RBAgd* as *BAL* (basal area in large trees) represents vertical competition for light. Slash pine is a shade intolerant tree; therefore, its growth is affected by shade (Lohrey and Kossuth 1990). Since *RBAgd* as a measure of competition considered trees only within 6.4 m (21 ft) radius of the subject tree, it did not take into account basal area in taller trees that are farther away from the subject tree and have no shading influence on the subject.

Basal area (*tbaha*) is a measure that takes into account both the vertical competition for light and the horizontal competition for rooting space, water, and nutrients (Yang et al. 2003). This measure takes into account both size and number of trees. Our results showed that the mortality increased with higher density (Figure 4-3); this has been supported by other studies as well. For a given age and site index, slash pine survival was lower at higher initial density (Bailey et al. 1985). Diéguez-Aranda et

al. (2005) also observed lower survival at higher stand density in Scots pine (*Pinus sylvestris* L.). Their model had interaction between age and stand density. For similar stand density, the survival probability decreased with increasing age. In our model, the effect of density also depended on tree heights. For trees with similar height, the probability of mortality increased with higher density, but this relationship changed as trees reached certain height. The reason might be due to the fact that for a shade intolerant species such as slash pine light must be the limiting factor for growth and mortality than competition for other resources. When trees grow taller, they can compete with other individuals for light, and also for other resources because they will have a well developed rooting system. Therefore, the effect of basal area disappears as a tree grows to a certain size.

The probability of mortality in our model increased with the site index (positive sign of the coefficient; Figure 4-4). This is consistent with findings of several other studies. For a given age and density, Bailey et al. (1985) also found lower survival of slash pine in stands of higher site index. A negative sign for site productivity index (SPI) indicating higher mortality for less shade-tolerant lodgepole pine on better sites has been reported by Yao et al. (2001), and by Crow and Hick (1990) for mixed oak stands. The reason for this is due to the faster growth and more competition leading to death, especially of smaller trees (Yao et al. 2001). There are conflicting results on the effect of site index on survival. Yao et al. (2001) showed higher mortality for aspen on poorer sites, and attributed this to lower tolerance to water and nutrient limitations. Similarly, Zhao et al. (2007) found an increased probability of survival for loblolly pine on better sites of the lower coastal plain. Due to conflicting effects of site index on mortality, several others

have recommended considering site index with caution in mortality models (Jutras et al. 2003, Diguez-Aranda et al. 2005, Crecente-Campo 2009). Site index was a non-significant parameter in our mortality model after the random plot effect was incorporated.

In a simulation model, mortality can be made stochastic by comparing the predicted probability of mortality to a random number generated from a uniform distribution. This can be implemented in the simulation model in the future; however, for the predictions to be deterministic, a cut-off value should be determined. Averaged observed survival rate as suggested by Monserud and Sterba (1999), and the value that gives maximum correct classification rate as suggested by Ryan (1997) did not work in our model. Using the observed mortality rate as cut off overestimated mortality in the lower dbh classes, and maximizing the correct classification rate underestimated mortality; therefore, we selected the cut off values considering both the correct classification rate, and the close agreement between observed mortality and survival with predicted mortality and survival.

SS Predictions from generalized linear mixed models were superior to the PA predictions or the predictions from the linear logistic regression. The comparison of area under ROC curve obtained from the model fitting and the cross validation procedures also supported this conclusion. Mortality also occurs in trees that grow in less suitable microsites – a natural variability found within forest stands (Greenwood and Weisberg 2008). Since the inclusion of random effects also incorporates the influence of other unmeasured plot level variables on mortality (e.g., topography, soil, microclimate, nutrient, moisture etc. [Huang et al. 2009]), it improves the accuracy of the model. The

insignificant coefficient of site index (a plot level variable) once the random effect was included in the model, also supported this conclusion. Also in loblolly pine plantation, SS predictions were better than PA predictions when survival was modeled using multilevel approach (Rose et al. 2006). The study used complementary log-log (CLL) approach to model survival instead of a generalized linear mixed model used in this study. It is possible to estimate PA predictions (using equation 4-13) using parameters from a generalized linear mixed models; however, SS predictions cannot be estimated using linear logistic function. Therefore, it is better to use a mixed effect model rather than using only a fixed effect model. However, to make good use of random effects model, we need data on at least two periods so that random effects can be calculated, at the same time most growth data are available for several periods.

This kind of individual tree mortality model does not exist for slash pine, but have been developed for several other species in USA and other parts of the world. This model provides information about the variables that are important in determining individual tree mortality in slash pine, which is useful for management or development of models in other areas of slash pine distribution. The model presented might be useful in simulating different silvicultural regimes (such as cutting and thinning), because it includes variables that are affected by tree removal. The effect of harvesting treatments on the mortality is incorporated into models by using terms like relative dbh and basal area (Hamilton 1986). For e.g., when a stand is thinned from below, its average stand diameter increases thereby reducing the relative dbh. Similarly, thinning also reduces stand basal area and subsequently affects tree mortality (Hamilton 1986). We did not have enough data to separate into model fitting or validation data set or an independent

data to validate the model; we recognize the importance of validating the model with an independent data set and this will be one of the area of future research. The oldest stand we have in our data set was 72 years and the largest tree was 61 cm (24 inches), caution should be used while implementing the model beyond the data range, because the model was unable to capture the U shaped relationship. It would be interesting to validate the model performance with data from the older stands.

Table 4-1. Parameter estimates, their standard errors, and statistical significance of the linear logistic regression model for slash pine mortality in north Florida.

Variable	Parameter	Estimate	SE	Pr> ChiSq
Intercept	$\beta_0$	-4.2889	1.2502	0.001
dbh	$\beta_1$	-0.1022	0.0460	0.026
H	$\beta_2$	0.2146	0.1014	0.034
RBAgd	$\beta_3$	-11.54	5.6496	0.040
SPI	$\beta_4$	0.0699	0.0350	0.045
tbaha	$\beta_5$	0.1312	0.0438	0.002
Hxtbaha	$\beta_6$	-0.0079	0.0029	0.008

Table 4-2. Parameter estimates, their standard errors, and statistical significance of the generalized linear logistic regression model for slash pine mortality in north Florida.

Variable	Parameter	Estimate	SE	Pr> t
Intercept	$\beta_0$	-5.2725	1.7700	0.0042
dbh	$\beta_1$	-0.1369	0.0572	0.017
H	$\beta_2$	0.2934	0.1352	0.030
RBAgd	$\beta_3$	-14.64	6.2979	0.020
SPI	$\beta_4$	0.0854	0.0487	0.080
tbaha	$\beta_5$	0.1804	0.0648	0.005
Hxtbaha	$\beta_6$	-0.0113	0.0043	0.010

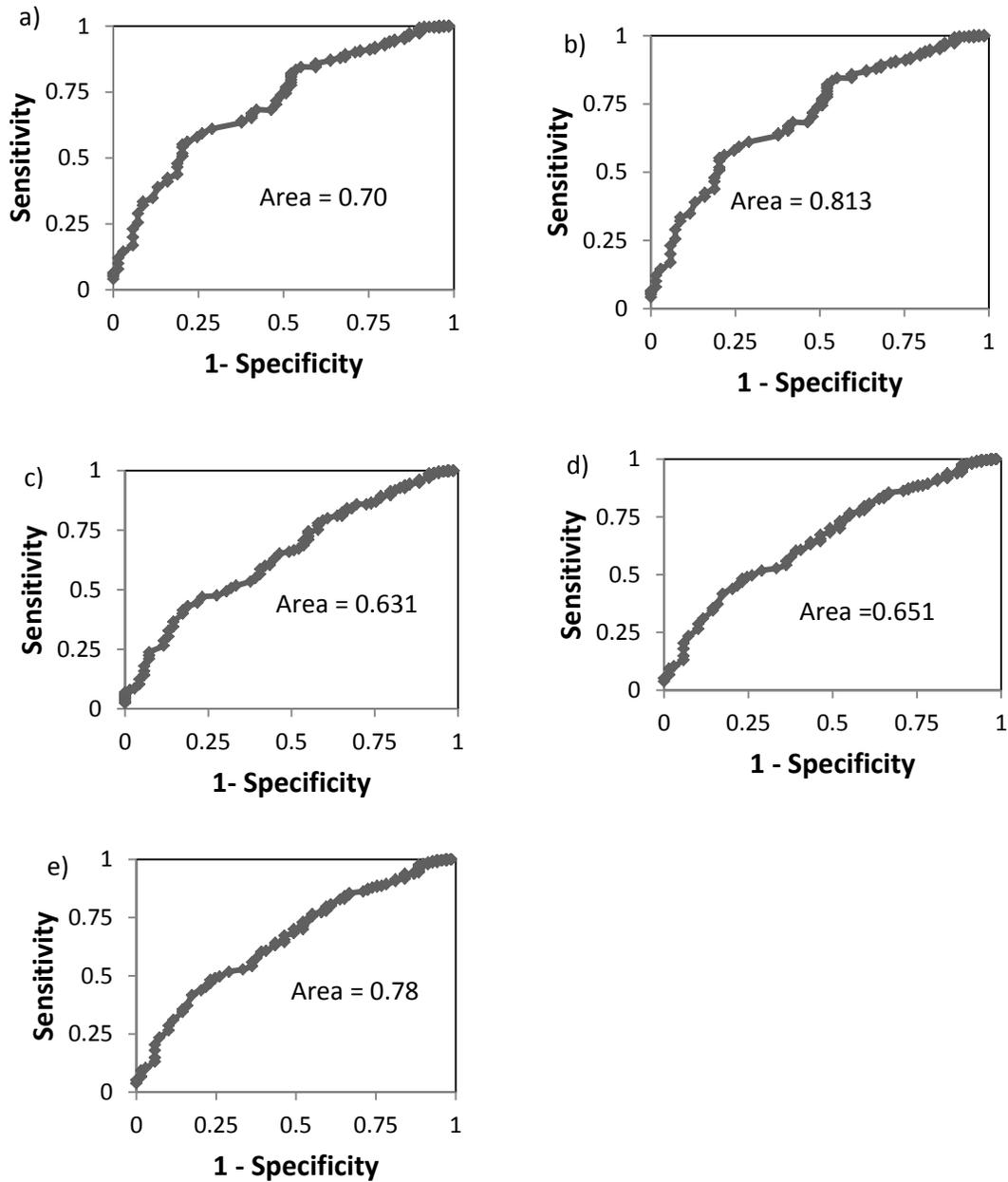


Figure 4-1. Receiver operating characteristic (ROC) curve produced by using: a) predictions from logistic function and population averaged predictions of generalized linear mixed model, b) subject-specific predictions from the generalized linear mixed model, c) prediction from cross validation data using the logistic function, d) population averaged predictions from cross validation data, and e) subject specific predictions using cross validation data.

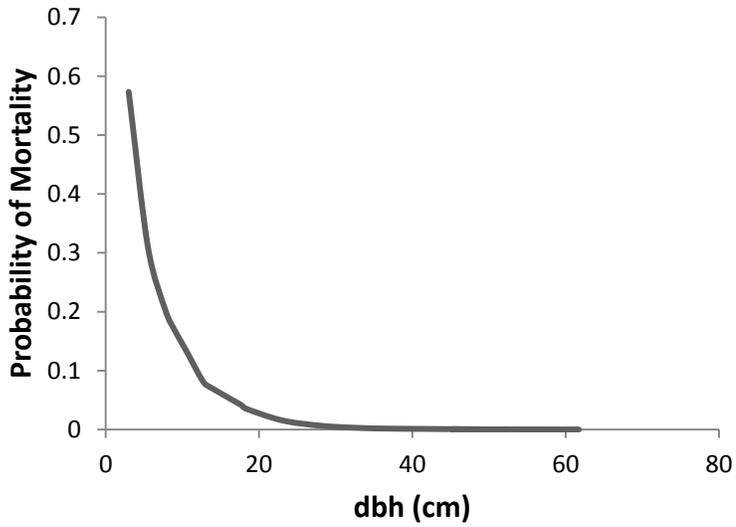


Figure 4-2. Probability of mortality of slash pine trees versus dbh (cm) as predicted by the mortality model.

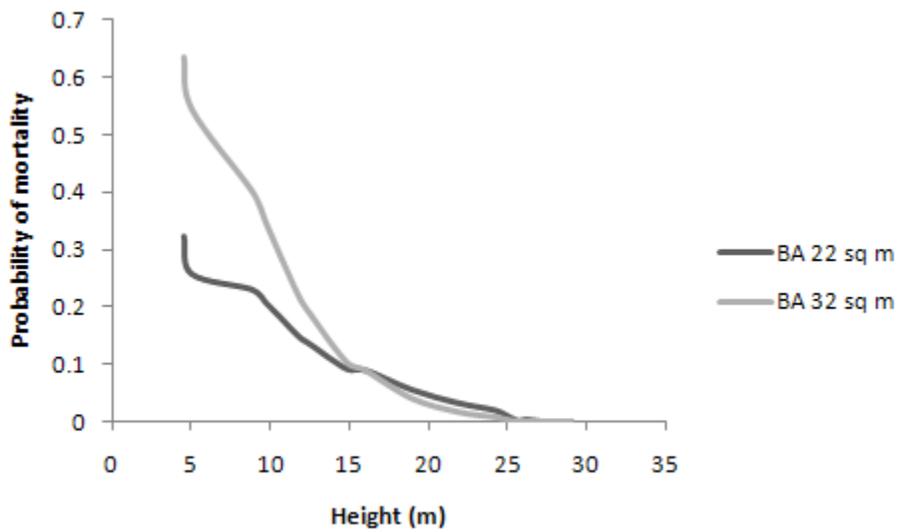


Figure 4-3. Probability of mortality of slash pine trees versus height (m) for different values of basal area per hectare (BA) as predicted by the mortality model.

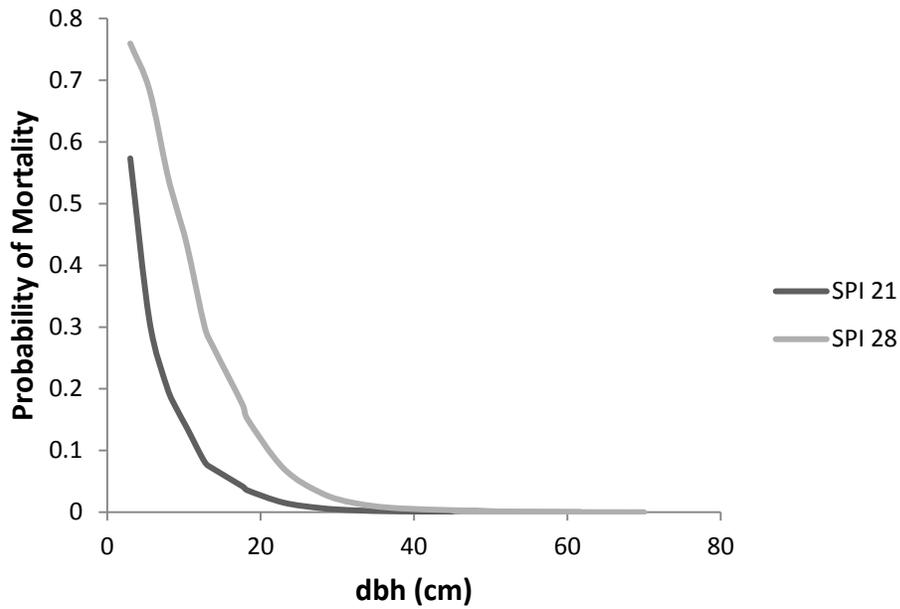


Figure 4-4. Probability of mortality of slash pine trees versus dbh (cm) for different values of site index (SPI in m) as predicted by the mortality model.

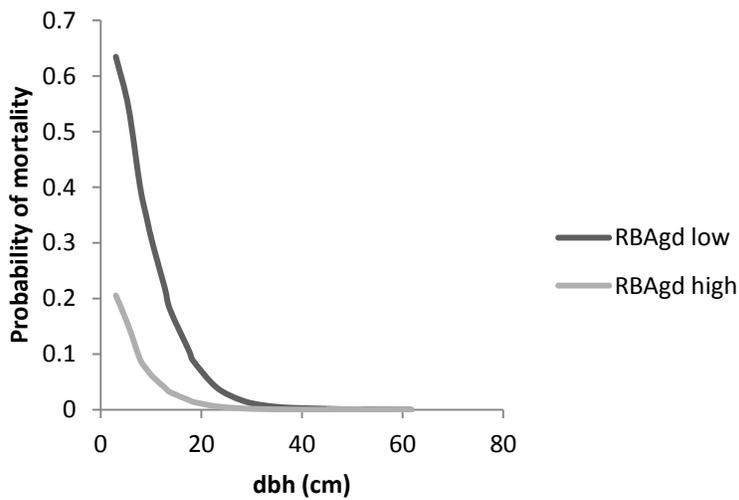


Figure 4-5. Probability of mortality of slash pine trees versus dbh (cm) for different values of competition index (RBAgd) as predicted by the mortality model.

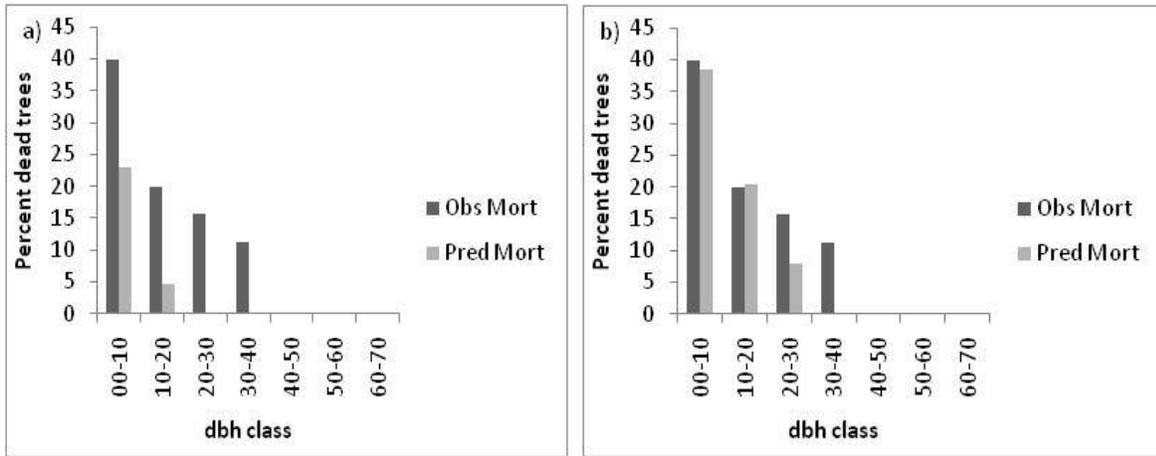


Figure 4-6. Subject-specific predicted (Pred mort) and observed (Obs mort) probability of mortality for different threshold values of assigning dichotomous survival status: a) 0.460 and b) 0.303.

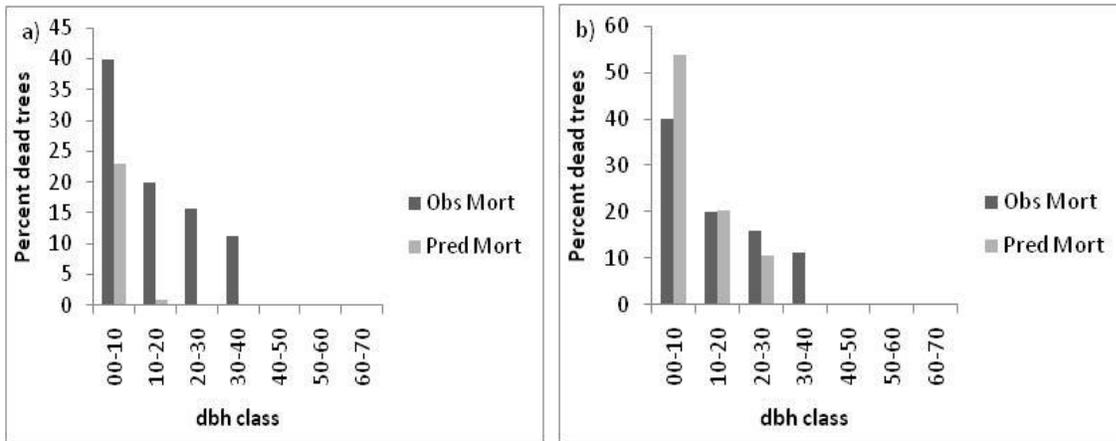


Figure 4-7. Population averaged predicted (Pred mort) and observed (Obs mort) probability of mortality for different threshold values of assigning dichotomous survival status: a) 0.590 and b) 0.292.

## CHAPTER 5 RECRUITMENT MODEL FOR SLASH PINE IN FLORIDA

### **Introduction**

Regeneration and recruitment are the processes that sustain a forest. If adequate conditions are available for germination and growth, seeds from an adult grow into seedlings, saplings, and adults. The seed and seedling stages are those which are most vulnerable to mortality, and changes occurring at these stages determine the population structure of adults (Harcombe 1987, Shibata and Nakashizuka 1995). Successful regeneration is determined by several factors: sufficient seed source, dispersal, successful germination, and seedlings survival and growth (Silvertown and Lovett Doust 1993). All these processes are affected by existing biotic (e.g., interactions with other individuals and species) and abiotic (e.g., soil, moisture, light, climate and disturbances) factors (Reader 1992, Ostfeld et al. 1997, Price et al. 2001).

Due to growing interest in sustainable forest management, many simulation models have been developed to test different harvest regimes for managing stands toward ecological objectives or for maintenance of uneven-aged stand structure (Peng 2000, Porte and Bartelink 2002, Trasobares et al. 2004, Sterba and Lederman 2006). For successful uneven-aged management, there should be continuous or periodic addition of new individuals to maintain a steady state or equilibrium condition (Eerikäinen et al. 2007). This requires a regeneration and recruitment module in simulation models.

For incorporating new individuals in growth simulation models, Vanclay (1994) suggested two different approaches: 1) regeneration models that simulate the development from seed to seedlings, and 2) recruitment models that simulate

establishment of saplings into a predefined threshold height or diameter during the monitoring period. A regeneration model incorporates processes such as seed production, seed germination, seedling establishment, seedling survival, and growth (Vanclay 1994, Eerikäinen et al. 2007). A large amount of data on these processes is required to adequately model regeneration. One of the problems in creating regeneration models with inventory data is that most inventory data starts with relatively large trees (usually 12.7 cm in dbh), but for regeneration models, smaller tree data are required. Because of the uncertainty involved in modeling regeneration, and also the lack of data, modeling recruitment (which is more predictable) is a viable alternative (Vanclay 1992, Lexerod 2005, Bravo et al. 2008).

A recruitment model should incorporate processes such as availability of seed source, dispersal, germination, and subsequent seedling survival (Price et al. 2001); however, the complexity of modeling regeneration has prompted to use simpler approaches for modeling recruitment. Traditional gap models such as JABOWA (Botkin et al. 1972) and FORET (Shugart 1984) directly recruit saplings within gaps based on the number of growing degree days and shading by other trees. The later variable determines whether the sapling of a particular species would have survived the simulation period or not. If a species is surviving, recruitment parameter for the species are used as an input to a random number generator to determine the number of saplings. Similarly, in FORSKA (Prentice et al. 1993), maximum sapling recruitment rate in optimal growing condition is modified using environmental constraints to determine actual recruitment rate for a particular site. The actual recruitment rate then enters as the input to a random number generator to determine the number of recruits for a

particular simulation time step. This approach is potentially effective given that sapling establishment is stochastic, and no improvement in model will necessarily be achieved by including seed production (Shugart 1984, Price et al. 2001).

Recruitment modeling can be performed with static or dynamic approaches. A static approach predicts the same amount of recruitment for different projection periods (e.g., a proportion of trees in large diameter classes) irrespective of the stand conditions affecting the projection. Stand table projection or matrix models are examples of static approaches (Usher 1966, Vanclay 1992). Recent matrix models (e.g. Cropper and Loudermilk 2006) incorporate density dependence and variable recruitment. In the dynamic approach, recruitment is modeled as a function of stand conditions. The probability of recruitment and number of recruits varies according to the stand conditions during the projection period (Vanclay 1992, Vanclay 1994). Because recruitment is affected by stand conditions prior to the recruitment period, this approach is more realistic.

Modeling recruitment is very difficult because of variability in regeneration, which is due to species' genetic traits, seed availability, and stand and climatic conditions (Vanclay 1992). Regeneration is also stochastic in nature; at any given period, regeneration may or may not occur even though other conditions are favorable for regeneration. This has led to a two-way approach to model regeneration or recruitment. In the first step, probability of regeneration or recruitment within an area is predicted using logistic regression of the presence/absence data. Conditional on the probability of regeneration, in the second step, the number of recruits is predicted using linear or other regression techniques (Vanclay 1992, Lexerød 2005).

The two-way approach has been used in the Prognosis model (Stage 1973, Wykoff et al. 1982, Wykoff 1986), and also by other studies (e.g., Vanclay 1992, Lexerod 2005). The recruitment model used in Prognosis predicts the probability of recruitment in 0.001 ha subplots based on environmental variables (habitat, slope, aspect, and elevation), distance to seed source, residual basal area, and time since disturbance. If regeneration is known to occur in a subplot, the amount of regeneration is predicted using a pseudo-random number generator. Vanclay (1989) directly predicted number of recruits into a 20 cm diameter class as a linear function of stand basal area and site quality. Vanclay (1988) also estimated the amount of one-year-old established seedlings based on basal area and site quality. Regeneration was modeled as cohorts of different height classes until they reached breast height (4.5 meters), after which they entered the main model for tree growth.

Uneven-aged forest management is considered a practice that mimics natural disturbance regimes. For this reason, it has gained popularity among public land managers and small land owners around the world (Lexerod 2005), and also among managers of former slash pine plantations in Florida. For modeling the dynamics of forest stands under various types of uneven-aged management, recruitment of new individuals in a stand is necessary; hence, recruitment modeling is an important part of the process of modeling stands under uneven-aged management. In slash pine, most modeling work has been done in plantation growth and yield, and these models predict dominant height growth, basal area, and volume under even-aged management (Benett 1980, Pienaar and Rheney 1995, Logan and Shiver 2006). We know of no models that use stand and tree level variables to model recruitment. The Forest Vegetation

Simulator has a regeneration module, but no equations to predict slash pine regeneration (Dixon 2002). Natural regeneration must be scheduled by the user; based on their experience with the regeneration of a particular species, users must supply the value for number of recruits (Dixon 2002). The objective of this study is to use a two-stage dynamic approach to model recruitment in slash pine. We develop models for predicting the probability of recruitment, and conditional on that probability, predict the number of recruits. We also test several tree and stand level variables in the model and provide a suite of variables that can influence recruitment in slash pine. Ultimately, the model will be used to test uneven-aged management of slash pine in Florida.

## **Methods**

### **Data**

This study used data from the property of the Suwanee River Water Management District (SRWMD) located in north central Florida, USA. The plots were established in 2000 by the Suwanee River Water Management District and re-measured in 2008.

Data were collected following a sampling procedure similar to that used by the USDA Forest Service for collecting Forest Inventory and Analysis data (Betchold and Patterson 2005). For data collection, a cluster of four circular 7.31 m (24 ft) radius plots were established approximately every 101.15 ha (250 acres). Each cluster had four plots, including a center plot and other plots at 90, 120, and 240 degrees from the center plot and at a distance of 36.57m (120 ft). For every plot, the following data were collected: current community, historical community, stand origin (natural, mixed or plantation), stand structure, stand age, slope, aspect, and GPS location. At each plot, three dominant and codominant trees were identified as site trees, and their height, age, and species were recorded. In every 7.31 m radius tree plot, trees > 12.54 cm (5

inches) in diameter were identified to species and the following data were recorded: distance and azimuth from the plot center, diameter at breast height (dbh), total height, height to live crown, product class, crown class, and tree damage. All of the above described information were also collected for trees < 12.54 cm (5 inches) dbh, but on a 2.07 m (6.8 ft) radius subplot concentric within the larger plot. In addition to this, data were also available on fire and silvicultural history for some of the plots.

Among the plots established in 2000, 70 plots that had high basal area in slash pine were randomly selected for re-measurement. All data collected in 2000 were collected in 2008 as well. In addition we also measured crown radius in 2008. The allometric relationship between crown radius and tree size (dbh) created from data collected in 2008 were used to come up with the crown radius in 2000. Total basal area per hectare and quadratic mean diameter per ha were also calculated for each plot. The site index for each plot was calculated using Farrar (1973), with 50 as the reference year. Plots were also classified into four different community types: Mesic flatwood (MF), Bay Swamp (BS), Sandhill (SA), and Upland Mixed Forest (UMF). The former two represent mesic and the latter represent drier sites.

For all plots, percentage cover (percentage area of the plot covered by tree crowns) was calculated using ArcGIS 9. The mean crown radius of each tree in each plot was used to calculate the crown area. A buffer was created around each tree using the mean crown radius information. The area of the buffer gave the crown area information; we did not consider the opacity of the crowns. Since the average height of recruits (12.54 cm) was 8m, we only used crown area of trees taller than 8m to calculate

percentage cover. We also considered overlap of crowns to increase the accuracy of our estimates of percentage cover for each plot.

Ingrowth for this study was defined as individuals that were not recorded in 2000 in the larger plot (due to their size being less than the minimum required 12.54 cm), but had dbh greater than or equal to 12.54 cm in 2008. The trees that were recorded in the concentric subplot in 2000, but were larger than the minimum dbh (12.54 cm) were also considered as recruits. Plots were classified as having recruited (1) or not recruited (0) based on the presence or absence of recruits. The number of recruits recorded in each plot were converted into number per ha for modeling purpose.

### **Model Development and Evaluation**

The variables included in the models were selected based on the understanding of the biology of the species, and studies conducted previously in other species around the world. For example, the shade intolerant behavior of slash pine was incorporated by testing percentage cover in the model. Slash pine regeneration and recruitment is affected by soil and moisture conditions (Lohrrey and Kossuth 1992); therefore, site index and different natural communities were included as independent variables in the model. On the other hand, there are conflicting results about the effect of overstory density on slash pine natural regeneration and seedling survival. For example, Feduccia (1978) reported a negative effect of density on regeneration and growth of slash pine seedlings whereas McMinn (1981) reported no effect of density on seedling establishment. We included stand density variables to test in our model.

We also included other variables based on the published literature of other species. Vanclay (1992) in his two-way approach to modeling recruitment included explanatory variables such as stand basal area, presence or absence of parent species,

treatment response, site quality, and soil parent material in his probability model. Indicator species were also tested in the model as a proxy for moisture availability (e.g., presence of palms indicated a moist site). Lexerød and Eid (2005) and Lexerød (2005) also used a two-way approach to modeling recruitment. Their models predicted probability of recruitment and number of recruits using information about location (altitude and latitude), site conditions (site index and vegetation types), and stand characteristics. Stand characteristics tested in the model included age, dominant height, basal area, number of trees, basal area mean diameter, stand volume, and percentage basal area of different species. Different vegetations types were included as a proxy for soil fertility, moisture and competition from herbs to predict probability of recruitment. We incorporated all applicable variables that were available in our dataset. We tested basal area, number of trees, quadratic mean diameter, site index, percentage cover, and community types and their transformations. Because our data included some naturally regenerated stands however, we did not have age information available for all our plots.

We followed a two-step procedure to model the number of recruits as done in several studies in the past (Vanclay 1992, Lexerød 2005, Hasenauer and Kindermann 2006, Bravo et al. 2008). Recruitment was considered a stochastic process, and recruitment data consisted of many zeros. Fitting traditional binomial models and poisson distribution to this kind of data can underestimate the occurrence of zeros and overestimate the occurrence of larger counts (Fortin and DeBlois 2007). One way to overcome this problem is to model recruitment as a consequence of two distinct processes. First, the probability of recruitment is modeled, and conditional on that

probability the number of recruits is predicted. Recently, zero inflated models have also been used to model tree recruitment, which can mitigate this problem (Fortin and DeBlois 2007).

The probability of recruitment was modeled using a logistic regression. Since the plots were classified as recruited (1) and not recruited (0), the dependent variable is binary. Logistic regression is the best procedure to model binary variables (Hamilton 1986, Avila and Burkhart 1992, Agresti 1996). The probability of recruitment ( $P_r$ ) for the 8 year period for each plot was modeled using the following logistic regression model:

$$P_r = \frac{e^{\mathbf{X}\boldsymbol{\beta}}}{1+e^{\mathbf{X}\boldsymbol{\beta}}} = \frac{1}{1+e^{(-\mathbf{X}\boldsymbol{\beta})}} \quad (5-1)$$

where  $\mathbf{X}$  is the matrix of independent variables, and  $\boldsymbol{\beta}$  is a vector containing parameters. The log odds or the logit of the probability has a linear relationship with the explanatory variables:

$$\log\left(\frac{P_r}{1-P_r}\right) = \mathbf{X}\boldsymbol{\beta} \quad (5-2)$$

The parameters for the model were estimated using maximum likelihood methods in the SAS procedure Proc logistic (version 9.2, SAS Institute Inc. 2008). Significance of the parameters was evaluated using Wald statistics (Agresti 1989). Similarly, Hosmer and Lemeshow goodness of fit tests (Hosmer and Lemeshow 1989) were also used to evaluate models. Competing models were evaluated using a generalization of the coefficient of determination ( $\bar{R}^2$ ) following Nagelkerke (1991). Similar to the Akaike Information Criteria (AIC) and Schwarz Criterion (SC; also known as the Bayesian Information Criteria) statistics, the  $\bar{R}^2$  value is useful in comparing competing models that are not necessarily nested (SAS Institute Inc. 2008). The value of  $\bar{R}^2$  lies between 0 and 1, and a higher value indicates a better model.

We also estimated the area under the receiver operating characteristic (ROC) curve to compare the models (Hosmer and Lemeshow 2000, Chen et al. 2008). The area under the curve was estimated for both model fitting data and cross validation data. ROC depends on the true positivity (sensitivity) and false positivity (1- specificity) of a test. Sensitivity is the proportion of observed positive events identified as positive, and specificity is the proportion of observed negative events identified as negative. In ROC curve evaluation, sensitivity is plotted against 1- specificity, and the area under the ROC curve is used to determine the discriminatory power of a curve. If the area is less than 0.5, then a model has no discriminatory power, if it is between 0.7 and 0.8, a model has acceptable discriminatory power, and if it is between 0.8 and 0.9, a model has excellent discriminatory power (Hosmer and Lemeshow 2000, Chen et al. 2008, Crecente-Campo et al. 2009).

The number of recruits was first modeled using a Poisson regression. The Poisson distribution is a member of a family of generalized linear models that allow the mean of a population to depend on a linear predictor through a non-linear link function. The response probability distribution is commonly defined to be from an exponential family of distributions (McCullagh and Nelder 1989). Suppose  $Y_i$  is a count variable from a population. In Poisson regression it is modeled as:

$$\log(E(Y_i)) = \beta' x_i + e_i \quad (5-3)$$

where  $\beta'$  is a vector of regression coefficients, and  $x_i$  is a vector of covariates for subject  $i$ . If a variable has a Poisson distribution, its expected value is equal to its variance ( $\mu = E(Y_i) = Var(Y_i)$ ).

Since we found significant over-dispersion in the data, we modified our model to fit a generalized Poisson regression (Famoye et al. 2004), which is an extension of Poisson regression. Overdispersion occurs when the variance of count data is higher than its mean, which is a violation of the basic assumption of Poisson regression. Generalized Poisson regression mitigates the overdispersion problem by introducing a dispersion parameter in the relationship between the mean and the variance.

$$Var(Y_i) = \emptyset\mu$$

where  $\emptyset$  is a dispersion parameter. If  $\emptyset = 1$ , we have a regular Poisson model, and if  $\emptyset > 1$ , we have an overdispersed Poisson model. The model fit was evaluated with deviance and Pearson Chi-square statistics.

We evaluated the model using a cross validation approach. For cross validation, each plot was deleted from the data set, and the model was fit with the remaining plots to estimate the parameters. The probability of recruitment for the deleted plot was predicted using the parameters estimated without the plot. This process was repeated for all plots. The ROC curve was then calculated to evaluate the predictions from this process. Since the number of recruits was conditional on the probability of recruitment, the validation estimate of the number of recruits was deterministically estimated by multiplying the expected number of recruits for a particular plot by the probability of recruitment. The predictions from the cross validation were compared with the observed using statistics following Huang et al. (2003) as given below:

$$E(\text{mean residual bias}) = \sum_{i=1}^k \frac{y_i - \hat{y}_i}{k}$$

$$\text{PBIAS (percentage bias)} = 100 \times \frac{E}{\bar{y}}$$

$$\text{MAD (mean absolute deviation)} = \sum_{i=1}^k \frac{|y_i - \hat{y}_i|}{k}$$

$$\text{MSEP}(\text{mean square error of prediction}) = \sum_{i=1}^k \frac{(y_i - \hat{y}_i)^2}{k}$$

$$R_p^2(\text{prediction coefficient of determination}) = \frac{1 - [\sum_{i=1}^k (y_i - \hat{y}_i)^2]}{\sum_{i=1}^k (y_i - \bar{y})^2}$$

In addition, we compared observed and predicted values using a paired t-test.

## Results

### Model for Predicting the Probability of Recruitment

The final model to predict probability of recruitment ( $P_r$ ) was given by (Model 1):

$$P_r = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 tbaha + \beta_2 SPI + \beta_3 qmd + \beta_4 \ln tbaha + \beta_5 \ln N + \beta_6 \ln SPI + \beta_7 tbaha \times qmd))} \quad (5-4)$$

where *tbaha* is the basal area per ha, *SPI* is the site index, *qmd* is the quadratic mean diameter, and *N* is the number of trees per hectare. The best model selected was based on the AIC and BIC values, and significance of the parameters at  $\alpha=0.05$ . The coefficient of determination ( $\bar{R}^2$ ) for the model was 0.70. The Hosmer and Lemeshow test was highly insignificant ( $P > 0.99$ ), suggesting that the model fits the data well. The area under ROC curve was 0.94, which indicated that the model classified the plots well into recruitment categories (recruited or not recruited). Preliminary graphs of the data indicated that basal area had curvilinear relationship with the probability of growth; therefore, both *tbaha* and its logarithmic transformation were included in the model (Table 5-1). Plots of recruitment versus basal area show that as basal area (*tbaha*) increases the probability of recruitment increases to a maximum, and then declines when basal area becomes too large (Figure 5-1). Similarly, both site index and its logarithmic transformation were included in the model (Table 5-1). As with basal area, plots of recruitment versus *SPI* indicate that the probability of recruitment increases with site index, reaches a maximum and then declines for higher site index (*SPI*; Figure 5-2).

Although the logarithmic transformation of site index was marginally insignificant, it was kept in the model, because it gave a more biologically realistic relationship. Removing this variable from the model would lead to the decline in the probability of recruitment with increasing site index. Quadratic mean diameter had a negative effect on the probability of recruitment, whereas the number of trees in a plot had a positive effect on the probability of recruitment (Table 5-1). There was also a significant interaction between basal area and quadratic mean diameter in the plot. For similar basal area, probability of recruitment was higher for plots with a small quadratic mean diameter (Figure 5-3).

Slash pine seed viability and seedling development has been shown to depend on soil moisture, depth to fine textured horizon, soil aeration, and silt and clay content of the soil (Lohrey and Kossuth 1990). Since we did not have data on soils, we decided to incorporate different natural communities in the model as a proxy for different growing conditions. Incorporation of natural communities as a categorical explanatory variable, however, did not improve the model fit, and the variable was insignificant. When the proportion of plots with regeneration was observed in the raw data, mesic communities such as BW and MF had higher recruitment than drier communities such as SA and UMF. Therefore, a separate model was fit, combining mesic and drier communities. This also did not improve the model fit, and the variable specifying mesic and drier sites was insignificant. To account for gap openings and light availability, we also included the percentage overstory cover for each plot as a proxy of competition light competition; however, this did not improve the model fit and the variable was not included in the final model.

The model was validated using the cross validation approach, and the ROC curve (Figure 5-4) was generated using the predicted values from this approach. The area under the curve was 0.95, and the model gave 91% sensitivity and 83% specificity.

### **Model for Conditional Number of Recruits**

The number of recruits ( $R_n$ ) is given by the conditional model (Model 2) :

$$\log(R_n) = \beta_0 + \beta_1 tbaha + \beta_2 SPI + \beta_3 qmd + \beta_4 percov + \beta_5 N + \beta_5 tbaha \times qmd \quad (5-5)$$

where *percov* is the percentage cover of a plot and all other variables are as previously defined. This model was selected based on AIC and BIC values, and all parameters were significant at a 5% level. We also compared predictions from the cross validation approach, and several fit statistics were calculated in order to select the best model. When the model was fit assuming a Poisson distribution, the ratio of the Pearson-Chi square to its degrees of freedom was much larger than 1, which indicated lack of fit due to overdispersion. To avoid this problem, we fit the model using generalized Poisson (SAS Institute Inc. 2008); the ratio of Pearson chi square to degrees of freedom for this model was 0.91, which indicated much better model fit.

Parameter values for Model 2 (Table 5-2) indicated that basal area (*tbaha*) had a positive effect on the number of recruits (perhaps on the seed production), whereas quadratic mean diameter (*qmd*) had a negative effect on the number of recruits. However, the effects were confounded by the significant interaction between the variables. Since the number of recruits was conditional on the probability of recruitment, the number of recruits predicted by Model 2 was multiplied by the probability of recruitment to obtain the final number of recruits per hectare. The conditional number of

recruits increased as the stand basal area increased, reached a maximum, and declined when basal area became larger (Figure 5-5). Due to the interaction between *tbaha* and *qmd*, for similar basal area, recruitment was higher in stands where quadratic mean diameter was lower, however, when *qmd* was 18, the highest number of recruits occurred around 35 to 45 m<sup>2</sup>/ha of basal area (Figure 5-5). Similarly, the conditional number of recruits increased with increasing number of trees (N), reached a peak, and then declined (Figure 5-6). Site index had a negative effect on the number of recruits (Table 5-2), as did percent cover of the plot, which is a proxy for the amount of light on the forest floor (Table 5-2). Percent cover might also reflect size and number of trees, and competition. Natural communities as described earlier were also tested as a proxy for growing conditions such as soil and moisture, but as in Model 1, the variable did not improve the model fit, and was also highly insignificant.

Fit statistics (Huang et al. 2003) were calculated using the cross validation approach. The fit statistics (Table 5-3) shows that the model has reasonable predictive power. The mean residual error was 20 recruits per hectare. Similarly, the prediction coefficient of determination was also high (0.54). The model over predicted the number of recruits as shown by the positive percentage bias (9.29%). The paired t-test between the predicted and observed number of recruits was not significant (P = 0.57), indicating that there was no significant difference between the two. The observed and predicted number of recruits by site index (Figure 5-7) and dbh classes (Figure 5-8) were also similar.

## **Discussion**

To our knowledge, the slash pine model presented in this study is unique in Florida. This model will be very useful in growth simulations. The final objective of this

model is to use it in a growth simulator for testing different harvest regimes for the conversion of even-aged stands into uneven-aged conditions. Our model predicts both the probability of recruitment, and number of recruits. Since recruitment is a stochastic process, the probability model can be used to incorporate stochasticity into the simulator (Lexerød and Eid 2005).

Our model assumes that there is sufficient regeneration of slash pine. For this to occur there should be seed source available for sufficient seed production. Seed viability is usually good in slash pine (Lohrrey and Kossuth 1990). In addition, adequate soil moisture, exposed mineral soil, and competition control is required for seedling regeneration and survival (Russel and Hebb 1972, Hebb and Clewell 1976, Lohrrey and Kossuth 1990). In slash pine, frequent fire helps in controlling competition from hardwoods and exposing the mineral soil. Soil texture and moisture are also important for early growth and survival of seedlings (Lohrrey and Kossuth 1992). The model assumes that these conditions for the regeneration and survival are available.

The influence of some of the environmental variables described above has been indirectly incorporated in the model by using site index as one of the explanatory variables. In addition, effects of soil and moisture conditions were incorporated after adding different natural communities as dummy variables in the model. The effects of drier versus mesic sites were also incorporated in the model. These variables did not improve the model fit, and were not significant. Our model indicated that there was no difference in probability of recruitment or the number of recruits between these various community types that differ in soil and moisture conditions. Inclusion of site index in the model might have indirectly incorporated these effects, and hence the variables were

insignificant. Lexerød and Eid (2005) also found similar result for their recruitment models for Norway Spruce (*Picea abies* (L.) H. Karst.), Scots Pine (*Pinus sylvestris* L.), and Birch (*Betula* sp.). They also found a significant effect of site index, but vegetation types as a proxy of environmental conditions were not significant. In another study (Lexerød 2005), vegetation type was a significant predictor for Norway spruce and Birch, but not for Scots pine. Vanclay (1992) also did not find a significant effect of indicator species in models for predicting probability of recruitment and number of recruits for tropical rainforests in northern Queensland; however, Vanclay (1992) included soil type and site quality as significant predictors in the model.

The model presented in this study incorporated variables commonly used in recruitment models such as site productivity and stand density (Vanclay 1992, Vanclay 1994). Other variables such as time since and nature of harvest are also important, but including these variables will make it difficult to apply the model, as these variables may not be available for all stands (Vanclay 1992).

In the present study, the model predicting the probability of recruitment included both basal area and logarithm of basal area. This ensured a low probability of recruitment for low stand basal area. This result is consistent with the findings of other studies (Vanclay 1992). As the basal area increases, seed production also increases due to the increase in large sized old trees, but at the same time density dependent mortality occurs at higher density (Zaval et al. 2007, Hubbell 1980). The combined effect of these two opposing forces determines recruitment, and is species and site specific (Hubbell 1980, Zaval et al. 2007). Other studies have reported negative effect of basal area on probability of recruitment (Lexerød 2005, Liang et al. 2005). The

probability model in this study includes an interaction between basal area and quadratic mean diameter, which causes low recruitment in plots with higher quadratic mean diameter given the same basal area (Figure 5-3). When quadratic mean diameter is smaller, there will be higher number of smaller trees and less larger trees. In this condition, there will more open spaces in the crown, and also less competition for other resources; therefore, the probability of recruitment will be higher. This finding is consistent with others, which found that quadratic mean diameter had negative effects on both probability of recruitment and number of recruits in Mediterranean pine forests (Bravo et al. 2008).

In our model, probability of recruitment and the conditional number of recruits increased with site index, reached a peak, and declined (Figure 5-2). This result is contrary to other studies where probability of recruitment increased with site index (Lexerød 2005, Lexerød and Eid 2005), but site index had no influence on probability of recruitment for some species (for example, Scots pine). Other studies have reported a mixed response of site index on number of recruits; a positive effect of site index on number of recruits of Norwegian species and Birch have been reported, but site index had a negative effect on number of recruits for Scots pine (Lexerød 2005). This result can be explained by the fact that as site index increases, the conditions required for growth and survival also improves and numbers of recruits increase; however, at higher site index, the growth in height and crown closure will also be higher. This will close the canopy and limit the amount of light reaching the lower canopy and ground. This light limitation could be the reason for decline in probability and number of recruits. This is also consistent with a significant percent cover variable in our model predicting the

number of recruits. Our model indicated that the increase in percent cover had a negative effect on the number of recruits; however, the same variable did not affect the probability of recruitment. This indicated the importance of light for the recruitment of slash pine.

Since slash pine is a shade intolerant species (Hebb and Clewell 1976, Lohrrey and Kosuth 1992), we expected percent cover to influence both recruitment models. In the present study, only the conditional model predicting the number of recruits had a percent cover effect. As expected for shade intolerant species, the increase in percent cover had a negative effect on number of recruits.

The parameter value signs (Table 5-2) indicated that the number of recruits increased with basal area, and decreased with number of trees, site index, and quadratic mean diameter. The predicted number of recruits was conditional on the probability of recruitment; therefore, the predicted number of recruits should reflect this probability as well. We incorporated this effect by multiplying the predicted number of recruits by the probability of recruitment to deterministically estimate expected number of recruits for each plot. The effect of the variables on the number of recruits was interpreted using this expected value. Vanclay (1992) and Lexerød (2005) used a similar approach to evaluate their recruitment models. In this study, the number of recruits increased with basal area, reached a peak, then declined (Figure 5-5). In a model developed for a tropical rainforest, the number of recruits decreased with basal area for some species groups, whereas it increased to a maximum and declined for other species groups (Vanclay 1992). As in the present study, the increase in site index had a negative effect on number of recruits of Scots pine (Lexerød 2005). Number of

recruits increased as the number of trees increased, reached a peak, and declined due to increased competition and mortality (Figure 5-6). This result agrees with the pattern of competition and tree mortality in forests and similar result had been reported in an earlier study (Lexerød and Eid 2005).

The presented ingrowth model can be used for both stochastic and deterministic prediction. If the ingrowth model is used deterministically, the probability of recruitment should be multiplied with the number of predicted recruits. The number thus generated is used to determine the ingrowth per hectare. To use the model stochastically, the probability of recruitment is compared with a random number generated from a uniform distribution between 0 and 1. If the estimated probability is greater than the generated number, then a plot is classified as recruited. The number of recruits should then be predicted for the plot classified as having new recruitment using the conditional model.

Although we did not validate our model with an independent dataset as in other studies (Vanclay 1992, Lexerød 2005), we used a cross validation approach to validate our model. The area under the ROC curve and significance tests of the probability model indicated very good discriminatory power of the model. The fit statistics calculated using the cross validation approach, such as mean residual error and bias, were within reasonable values. The model over-predicted the number of recruits as indicated by the positive percentage bias; however, the difference between observed and predicted values was not statistically different. In addition to that, the predictions from the model gave biologically meaningful results; therefore, the models should be useful for predicting recruitment in slash pine.

Table 5-1. Parameter estimates and their associated standard errors and p-values of the model predicting the probability of recruitment (equation 5-4) for slash pine in north Florida.

Variable	Parameter	Estimate	SE	P value
<i>Intercept</i>	$\beta_0$	222.8	129.5	0.085
<i>tbaha</i>	$\beta_1$	-1.5625	0.5468	0.004
<i>SPI</i>	$\beta_2$	-1.7579	0.8711	0.043
<i>qmd</i>	$\beta_3$	-7.1433	3.0925	0.020
$\ln(tbaha)$	$\beta_4$	47.08	20.13	0.019
$\ln(N)$	$\beta_5$	-42.15	20.78	0.042
$\ln(SPI)$	$\beta_6$	28.81	16.08	0.073
<i>tbahaxqmd</i>	$\beta_7$	0.0702	0.0234	0.002

Table 5-2. Parameter estimates of the model predicting conditional number of recruits (equation 5-5) for slash pine in north Florida

Variable	Estimate	SE	P value
<i>Intercept</i>	14.3950	2.5462	0.0001
<i>tbaha</i>	0.6153	0.1411	0.0011
<i>SPI</i>	-0.1147	0.0031	0.0053
<i>qmd</i>	-0.5069	0.1573	0.0081
<i>percov</i>	-0.0405	0.0099	0.0018
<i>N</i>	-0.0053	0.0015	0.0045
<i>tbahaxqmd</i>	-0.0131	0.0032	0.0020

Table 5- 3. Fit statistics for the cross validation of model for predicting number of recruits (Model 2) of slash pine in north Florida.

E	PBIAS	MAD	MSEP	$R_p^2$
20.02	9.29	104.69	15698.52	0.54

Note: E = Mean Residual (Bias); MAD = Mean Absolute Deviation; MSE = Mean Square Error of Prediction; PBIAS = Percentage bias;  $R_p^2$  = Prediction coefficient of determination.

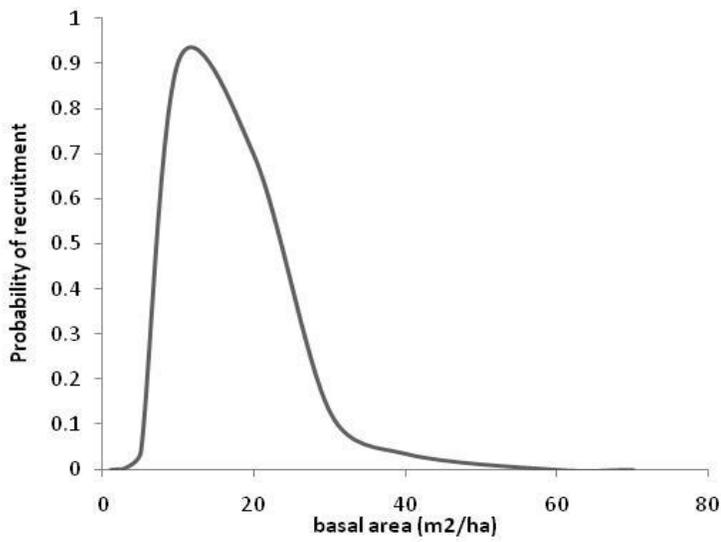


Figure 5-1. Probability of recruitment at 12.7 cm (5 inches) dbh for slash pine in north Florida at different values of stand basal area.

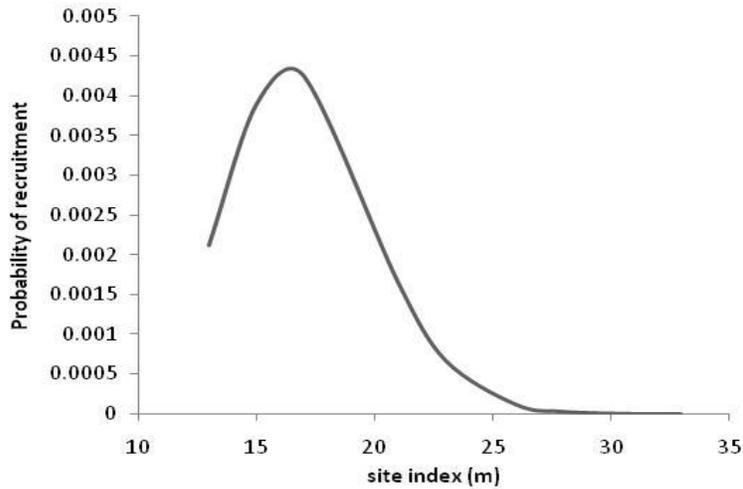


Figure 5-2. Probability of recruitment at 12.7 cm (5 inches) dbh for slash pine in north Florida at different values of site index.

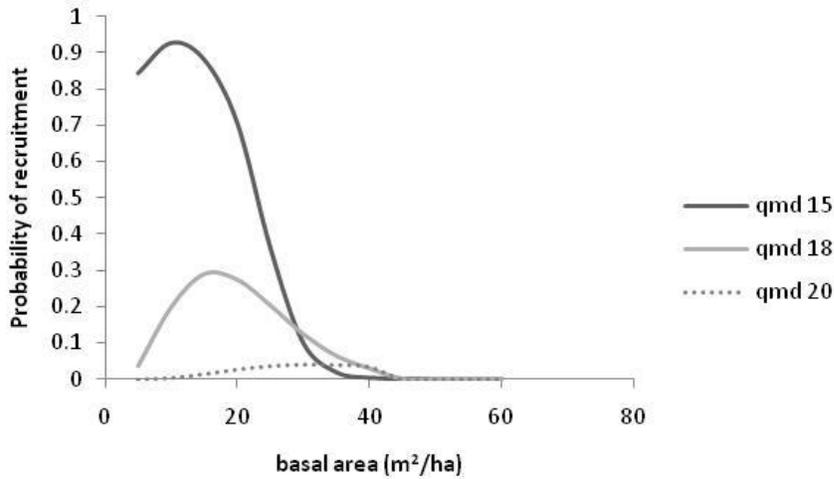


Figure 5-3. The relationship between probability of recruitment at 12.7 cm (5 inches) dbh and basal area for different values of stand quadratic mean diameter (qmd) of slash pine in north Florida.

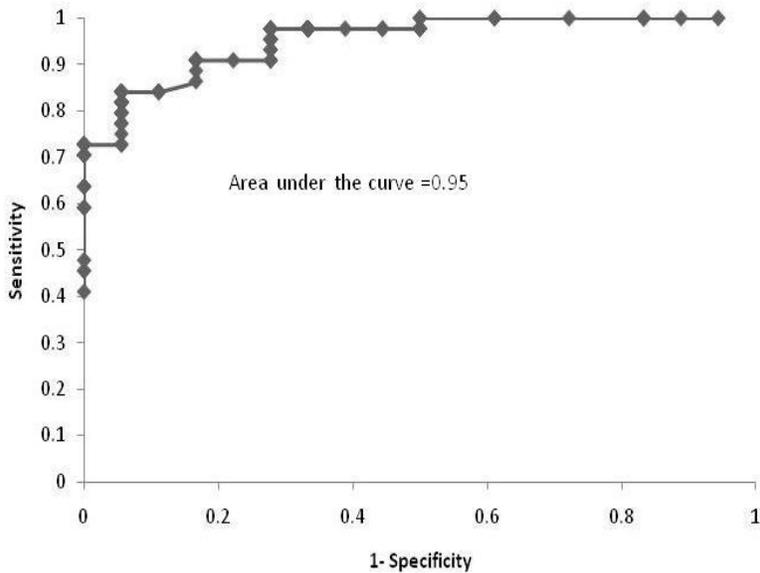


Figure 5-4. Receiver Operating Characteristics curve of the recruitment model using the predicted values from the cross validation approach.

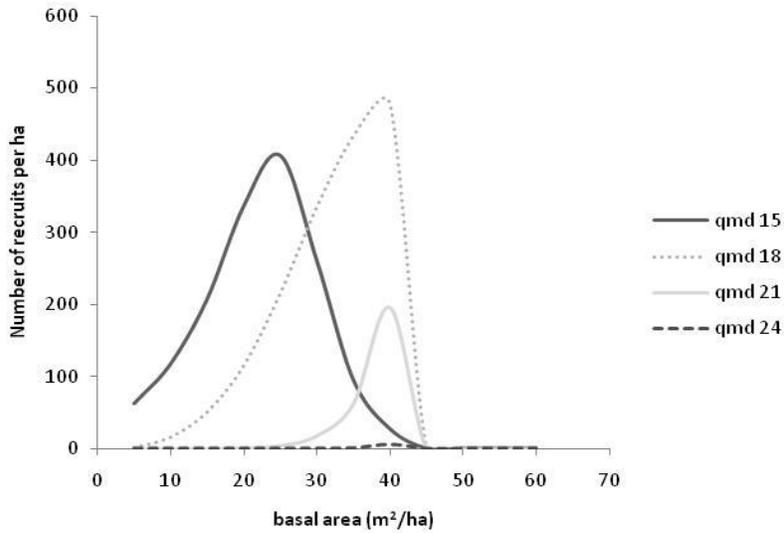


Figure 5-5. Relationship between the conditional number of recruits and stand basal area at different values of quadratic mean diameter for slash pine in north Florida.

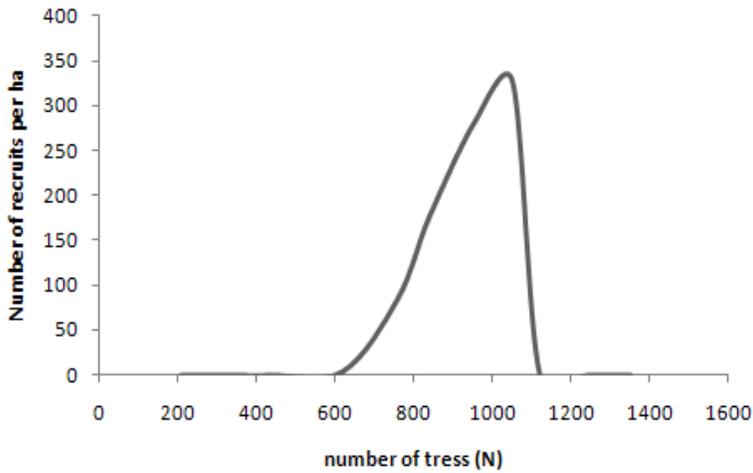


Figure 5-6. Relationship between the conditional number of recruits and number of trees in a stand for slash pine in north Florida.

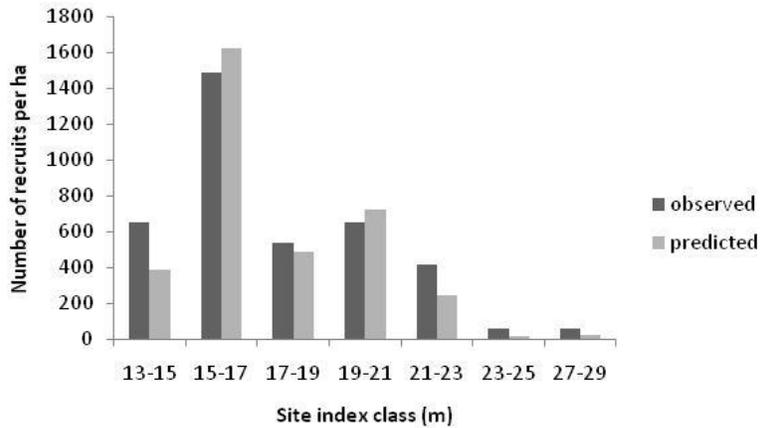


Figure 5-7. Observed and predicted number of recruits for site index classes from cross validation of Model 2 for slash pine in north Florida.

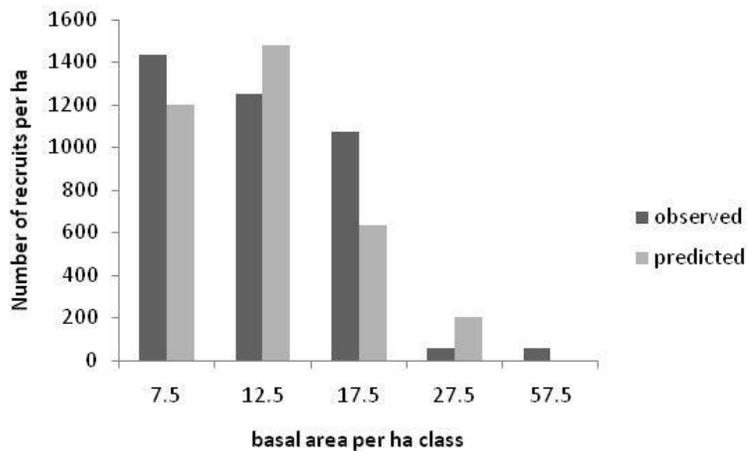


Figure 5-8. Observed and predicted number of recruits for basal area classes from cross validation of Model 2 for slash pine in north Florida.

## CHAPTER 6 SIMULATION OF DIFFERENT HARVEST REGIMES FOR STAND CONVERSION OF SLASH PINE IN FLORIDA

### **Introduction**

Simulation models are important tools to project the future of natural systems given current conditions and management. Over the last decade, global warming and its effect on various ecosystems have attracted much scientific attention. Many scientists are predicting species distribution and other ecosystem changes in relation to increased warming of the earth. This has also increased the importance of simulation models, because models can be used to project future changes under various climate change scenarios. Similarly, natural resource managers are interested in knowing the potential future consequences of their current management actions. One approach to help managers may be to use a simulation model of the system to evaluate various management scenarios.

Public perception of natural resource management is changing. In addition to timber resources, other ecosystem services from the forest such as biodiversity conservation, aesthetics, and maintenance of water quality are being increasingly valued (Brag 2004). Clearcutting and establishing a monoculture plantation compromise ecosystems benefits (except timber), and therefore, this practice is a matter of widespread public disagreement (Brag 2004, Franklin et al. 2007). This has increased the interest of public land managers in managing their lands for overall ecosystem benefits.

One example of managing forests for ecosystem benefits is uneven-aged forest management. Uneven-aged forests have continuous cover and vertical and horizontal diversity, which provide diverse wildlife habitat and are also aesthetically pleasing

(Ambuel and Temple 1983, Guldin 1996, Liang et al. 2005). The reason for structural diversity in uneven-aged stands is the constant regeneration in gaps created by natural disturbances – in a single stand, seedlings, saplings, and different sized trees are present at once. In managed forests such as former plantations, natural processes have been altered, and they do not contain gaps and adequate environment for regeneration to occur. Therefore, regeneration environments should be created artificially by forest management.

One way to achieve uneven-aged forest structure is to modify harvest techniques to mimic natural disturbances (Murphy et al. 1993, Palik et al. 2002, Seymour et al 2002, Franklin et al. 2002). For example, harvest techniques such as low thinning with a combination of group selection and competition control (fire) have been used for old growth restoration of mature loblolly and shortleaf pine (Brag 2004). Similarly, single tree selection and irregular shelterwood has been used to generate timber and maintain structural diversity in longleaf pine stands (Palik et al. 2002). Thinning has also been used in old growth feature restoration (Choi et al 2007).

Several silvicultural prescriptions have been discussed in the literature and also tested in field experiments for the conversion of homogenous even-aged plantations into structurally diverse uneven-aged conditions (Smith et al. 1997, Brockway et al. 2005). Single-tree selection and group selection are commonly used methods (Baker et al. 1996, Smith et al. 1997) that open gaps in the canopy to allow regeneration (Smith et al. 1997). In single tree selection, individual trees are selected and cut based on their form, vigor, maturity, quality, and growth rate. Cutting is spread over the stand uniformly to achieve regeneration at different locations. As opposed to single tree selection, group

selection identifies trees for cutting in groups at various locations within a stand, thereby creating larger openings. Other methods such as irregular shelterwood and diameter limit cutting can also be used to achieve an uneven-aged structure (Smith et al. 1997, Brockway et al. 2005). There is no “one size fits all” type of prescription, and each method can be modified to fit the site, management objectives and species of interest. In some situations, rather than a single method, a combination of different methods along with thinning has been used for the conversion process (Guldin and Farrar 2002, Nyland 2003).

Thinning has also been prescribed as a treatment method to promote structural diversity and old growth structure (Franklin et al. 2007). In a forest, small-scale canopy disturbances from wind, lightning, insects or fire create spatial heterogeneity in stand structure (Franklin et al. 2002). Thinning for economic purposes creates uniformly distributed crop trees, and the main goal is to increase the size of remaining trees by removing competition (Franklin et al. 2007). However, ecological thinning should follow the pattern of small scale disturbances and gap forming processes (Franklin et al. 2007). Variable density thinning (VDT) leaves strips of unthinned patches along with some trees removed in groups, and intermediately thinned areas will follow the pattern of competitive tree mortality and small-scale disturbances (Carey 2001, Franklin and Lindenmayer 2002). Thinning to a residual basal area of 18.4 sq m/ha (80 sq ft per acre) has been implemented in loblolly and shortleaf pine stand conversion (Guldin and Farrar 2002).

Group selection and single-tree selection has been successfully used in uneven-aged management for loblolly and shortleaf pine forests (Baker et al. 1996). In another

study, single-tree selection using the volume guided diameter limit cut (VGDL) method and residual basal area, diameter and q ratio (BDq) method were tested for loblolly pine stand conversion in Arkansas (Guldin and Farrar 2002). The same study also used thinning to a residual basal area of 18.4 sq m/ha (80 sq ft per acre) for stand conversion (Guldin and Farrar 2002). The successful management of loblolly and shortleaf pine using uneven-aged techniques in the “Good” and “Poor” Farm Forestry Forties in Crossett Experimental Forest (Crossett, Arkansas) surprised many people because loblolly and shortleaf pine are considered relatively shade intolerant (Bragg 2008). Although these methods have been recommended and tested for longleaf, loblolly and shortleaf pine, their effectiveness has not been tested for the conversion prescriptions in slash pine in Florida.

A good way to test conversion techniques is to establish field trials and monitor changes over time. The monitoring data will help in evaluating the progress and changes can be made accordingly. Another way to evaluate uncertainty in using conversion techniques is to use a simulation model to guide the process of establishing field trials. Methods that are useful in simulation can be tested in the field. Simulation models have been used in the past to evaluate different harvest regimes. Hanewinkel and Pretzsch (2000) simulated different harvest regimes for the conversion of even-aged stands of Norway spruce into uneven-aged structure in southwest Germany. In another example, Sterba and Ledermann (2006) simulated single tree selection and natural regeneration using the PROGNAUS model in a mixed broadleaf and conifer forest in Austria, and evaluated their performance in achieving stable harvest and growth, structural diversity, and species diversity.

Land managers in Florida are also interested in conversion of slash pine plantations into more structurally diverse forests. While there are efforts to establish field trials to test conversion, to date, there are no published records of slash pine conversion using methods that has been tested elsewhere. The objective of this study is to assist this process by creating a simulation model for testing some of the prescribed techniques for conversion of slash pine plantations. Our goal is to demonstrate the utility of the model; therefore, we ran a single simulation for each scenario. However, for more comprehensive results, many runs (on the order of 1000) of each scenario should be performed and average values including the uncertainty of these values should be reported. We expect that some of these harvest techniques will be useful in converting slash plantations into structurally diverse forests. Toward this purpose, first we describe a simulation model, created using height growth, diameter at breast height (dbh) growth, mortality and regeneration submodels described in earlier chapters. The model is then evaluated, comparing its predictions with prior studies on slash pine growth. Finally, we use the model to test different intensities and timings of thinning, single tree selection using the VGDL method, and group selection along with thinning. There are limited records on the structure of natural old growth slash pine stands to compare our results to; therefore, we followed a standard protocol of measuring structural diversity in forests. These regimes were compared based on some of the ecological indicators such as tree size diversity and height diversity and number of large trees (Franklin et al. 2002). We also compared the volume of trees harvested in different product categories, and their economic returns in order to give information to land owners who are concerned with economic as well as ecological aspects of stand conversion.

## Methods

### Description of the Model and Variables

The diameter growth, height growth, mortality and recruitment submodels presented in Chapters 2, 3, 4, and 5, respectively, were coded in Python 2.6.5 for the simulation of forest structural change over time. Each of these submodels was developed with a random effect in order to better describe the multi-level nature of the underlying data. However, since the use of a mixed effects model requires periodic data to calculate the random stand or plot effect, we used fixed effects models to run the simulations.

The initial stand used in the simulations was created using diameter information from an 18 year old plantation described in Gholz and Fisher (1982). Trees heights were estimated using the equation relating dbh to height from Gholz and Fisher (1982). The stands simulated were 200 m x 60 m (1.2 ha), following similar study in western Washington (Sprugel et al. 2009). Since the aim of this study was to convert even-aged plantations to uneven-aged, we created a regular spatial distribution of trees as in plantations, and 2 m x 2 m spacing between trees.

The simulated initial stand was created by randomly choosing a dbh value following the distribution in Gholz and Fisher (1982), estimating height, and randomly assigning (X, Y) coordinates on a 2m x 2m grid. The program then calculated basal area per hectare and quadratic mean diameter per hectare of the initial stand. The various distance dependent and independent competition indices required to run all the submodels (see Chapters 2, 3, 4 and 5) were also calculated. Crown radii for each tree were calculated using the equation relating tree size and average crown radius (eqn. 6-

1). Crown radius information was used to calculate crown area for each tree, which was finally used to calculate percentage cover.

$$meancr = -1.2836 + 0.8360 * dbh \quad (6-1)$$

The first step in the simulation was to calculate the eight-year probability of mortality of each tree, probability of recruitment and number of recruits, using the current information on individual trees and the stand. After mortality and recruitment was estimated, the dbh and the height growth submodels calculated annual growth, which was added to the current height and dbh of the trees. The calculated probability of mortality was converted to an annual probability of mortality, which was compared with a cutoff probability (0.292 as described in Chapter 4). If the calculated annual probability of mortality was greater than the annual cutoff, the tree was tagged dead (1) otherwise it was tagged as surviving (0). All dead trees were removed from the stand at the end of each year of simulation.

For assigning ingrowth to the stand, the eight year probability of ingrowth was converted to an annual probability using the compound interest formula as described in Chapter 4. This annual probability was compared with a uniform random number between 0 and 1. If the calculated annual probability was greater than the uniform random number, the stand was considered to have recruitment. If there was recruitment, the conditional model (number recruited, given there is recruitment) estimated the eight year number of recruits per hectare, which was also converted to an annual number of recruits. The numbers of recruits were then assigned to the stand with the help of a moving 2 m radius circular window. It was assumed that a 2m radius (approximately the mean crown radius of dominant and codominant trees in our data)

would be the size of a gap created by the death of a medium sized tree. If there was only one recruit within that window, it was considered a gap suitable to assign the recruits. The number of recruits was assigned proportional to the area. If all the moving windows had trees or more than one recruits, and the criterion was not satisfied, the recruits were assigned randomly within the stand.

The initial age of the stand was 18 years, and we simulated growth, recruitment, and mortality with a one-year time step until the stand was 68 years (close to the oldest stand age in our data). Site index for all simulations was 24 m. We first performed simulations without implementing any harvest, hereafter referred to as the control.

### **Model Evaluation and Sensitivity Analysis**

Model performance was evaluated by comparing predicted basal area, total volume, and volume increments with the values of the variables in prior studies on slash pine growth. Also the final values of height and dbh obtained after implementing harvests were compared to the prior studies to determine if the model predictions after harvests were also within the natural range reported.

We performed sensitivity analysis for the site index input variable, and several of the model parameters. Initially, we used site index 24 m for our model simulation. We changed site index values to 18 m and 28 m to perform the sensitivity analysis. For all other parameters, plus and minus one standard error from the estimated value was used for sensitivity analyses (Table 6-1). We changed the following parameters (Table 6-1):  $\beta_1$  parameter of the dbh growth model, which is related to starting individual tree diameter (see Chapter 2);  $\beta_2$  parameter of the height growth model, which is related to starting individual tree height (see Chapter 3);  $\beta_3$  parameter of the height growth model,

which is related to stand basal area (see Chapter 3);  $\beta_1$  parameter of the conditional model predicting number of recruits, which is related to stand basal area (see Chapter 5). The model sensitivity was evaluated based on the percentage change relative to the base value (estimated value from the model fitting procedure) of the final stand basal area, volume, number of recruits, and number of trees. The sensitivity analysis was performed changing one parameter or input variable at a time.

## **Harvest Simulation**

### **Thinning**

The thinning regime implemented simulates low thinning, removing trees less than 16 m in height, and avoids trees less than 13.5 cm in dbh to avoid removing early recruits from the stand. Two types of low thinning regimes were utilized that are typically implemented in slash pine to improve pine growth. Low and frequent thinning brings the basal area of stand to 17.21 sq m/ha (75 sq ft/acre) and is implemented every 5 years in practice. Heavy and infrequent thinning brings the basal area of the stand to 11.47 sq m/ha (50 sq ft/acre) and is implemented every 10 years in practice. The thinning algorithm selected trees randomly from the stand, and the tree was removed if its height and dbh were within the thinning range. The algorithm stopped if the residual basal area was within the limit specified for the regime. The way this algorithm worked, it could remove trees in groups, leave some areas with trees, and perform intermediate thinning in other areas thereby simulating the VDT described earlier. The first low and frequent thinning was performed when the stand was 23 years (five years after the initiation of the simulation). Because there was not enough basal area to be removed every five years, subsequent harvests were performed every 10 years. Altogether five low and

frequent thinnings were performed. For heavy and frequent thinnings, four harvests were performed with the first harvest at age 28.

### **Single-tree selection (VGDL)**

We also simulated a single tree selection method. The single tree selection was regulated by using a volume guided diameter control (VGDL) method. For the implementation of the VGDL method, the stand was first grown for ten years. The annual volume increment was calculated, and the periodic volume increment in the next ten years was predicted by multiplying the cutting cycle length (10 years) with the annual volume increment. In each harvest, the amount of the predicted periodic volume increment was removed from the stand. To prevent removing all of the large trees and ingrowth from the stand, we did not allow removal of trees that were within 3 cm of the smallest diameter tree in the stand or trees that were within 5 cm of the largest diameter tree in the stand. Trees that were within the range were randomly selected and removed from the stand until the harvested volume equaled the periodic volume increment. Four harvests at a 10-year cutting cycle, starting 10 years after the initiation of the simulation (stand age 28 yrs), were implemented.

### **Group selection**

Finally we simulated group selection. Since there was no information available on the size of gaps in slash pine forests, we used information from longleaf pine forests (Gagnon et al. 2004). Since 80% of gaps in the study were less than 600 sq m (~13 m radius), we used a gap size of 13 m for the simulation of group selection openings. Keeping in mind the stand dimensions, points that were used to create gaps were provided manually. All trees within a 13 m radius of the points were removed. The usual practice in group selection is to thin outside the gap openings; therefore we also

simulated this by performing thinning outside the gaps during each harvest. The harvest was performed every ten years starting at age 28. The first and second harvest created three random openings each. To avoid overlap of the openings, the last harvest created only two openings.

### **Comparison of Harvests**

We only simulated one run of each scenario; more comprehensive results on the performance of the different harvest scenarios would be achieved by running the simulation many times and examining the distribution of end values. However, since our goal is to demonstrate the potential utility of the model, we present a suite of methods for evaluating the performance of different management regimes in achieving structural diversity. First, the tree diameter and the height distributions of the stand at the end of the simulation were visually compared for different harvest regimes. We also calculated Shannon-Wiener diversity index ( $H'$ ; Shannon and Weaver 1949, Buongiorno et al. 1994) for height and dbh. Since the index depends on the defined height and dbh classes, we estimated the index using two different height and dbh classes. We also evaluated the distribution by estimating standard deviation, skewness and kurtosis. Furthermore, the diameter distribution was also compared using a shape index (Siipilehto 1999, Rouvinen and Kuuluvainen 2005). For the calculation of shape index, the basal area of the median tree is multiplied by the observed number of trees to calculate an estimate of basal area. The estimated basal area divided by the observed basal area is the shape index (Siipilehto 1999). Siipilehto (1999) described the following characteristics of the shape index: 1) the value of the index is 1 for peaked unimodal distribution, 2) unimodal distributions have values greater between 0.54 and 1.0, 3) the

index decreases below 1 with increasing deviation in diameter and increasing right-skewness, 4) the reverse-J shaped distribution has a value between 0.48 and 0.54.

One of the characteristics of old growth forest and uneven-aged stands is a multilayered canopy (Franklin et al. 1981, Malcolm et al. 1999). Height diversity was also characterized using a measure of multilayered forest known as the Berger-Parker Index ( $d$ ), which is a measure of evenness (Magurran 1988, Malcolm et al. 1999). The index, which is given by the following, was calculated using 2 m height classes:

$$d = \frac{N_{Total}}{N_{max}}$$

where  $N_{Total}$  is the total number of trees,  $N_{max}$  is the maximum number of trees in the height class with most trees. This is a rough measure of vertical diversity in forest stands, and higher values indicate more evenness and less dominance (Carey et al. 1992, Malcolm et al. 1999).

### **Economic Analysis**

Outside bark volume of each tree was calculated using the following slash pine volume equation (eqn. 2; Pienaar et al. 1988):

$$Predicted\ Volume = 0.00482\ dbh^{1.9853} H^{0.8598} \quad (2)$$

where  $dbh$  (inches) is the diameter at breast height and  $H$  (ft.) is the tree height. The equation predicts volume in cubic feet per acre, which was converted to cubic meter per hectare. Individual tree volumes were summed up to calculate stand volume. The volume harvested in each harvest and total volume harvested in each regime were estimated and compared. The volume harvested was divided into different product categories using appropriate current market ranges (Timber Mart-South 2010). The product categories were pulpwood (15.24 cm to 20.52 cm dbh), chip-n-saw (20.32 cm to

27.94 cm dbh), and sawtimber (>30.48 cm dbh). Furthermore we also calculated the net present value (*NPV*) of the harvest using the following:

$$NPV = \frac{R_t}{(1+i)^t}$$

where  $R_t$  is the net cash inflow (amount of cash inflow minus outflow),  $t$  is the time of the cash flow, and  $i$  is the discount rate. In this case, cash inflow is the amount received for the harvest, and cash outflow is the harvest cost. The 2010 prices for pulpwood (\$10 per ton), chip-n-saw (\$17/ton), and sawtimber (\$26 per ton) along with the cost per ton of harvest were taken from current market sources (Timber Mart-South 2010). The mean inflation rates from the last 10 years were averaged to come up with a rate of 2.53 % (<http://forecast-chart.com/forecast-inflation-rate.html>), which was used to convert the present rate of harvest income and cost to the rate at the year when the harvest was performed. After calculating net cash inflow using income from harvest and cost for each harvest, we calculated NPV using a 4% discount rate as suggested in Matta et al. (2009).

## Results

### Evaluation of Model Prediction

We first evaluated model performance without simulating any harvest regime. The initial stand density was high; therefore the basal area growth declined very slowly for the base site index (24; Figure 6-1a). This could be because of mortality due to competition. Over the 50 years of simulation, the basal area declined approximately 13 sq m/ha (Figure 6-1a). At 68 years of age, the stand had approximately 39 sq m/ha. For site index 18 (Figure 6-1b), the basal area increased initially, reached a peak around 28 years and declined.

The net volume growth also followed an expected biological pattern of dry matter accumulation in forests (Figure 6-2; Smith et al. 1997). It reached a peak around age 45 and then started to decline and level off around 65 years of age. The outside bark volume at 68 years was 530 m<sup>3</sup>/ha. The periodic annual volume increment for the first 10 years (up to a stand age of 28 years) for the control stand was 7.83 m<sup>3</sup>/ha, but later the volume increment declined to approximately 3 m<sup>3</sup>/ha. The mean annual volume increment at 68 years was 7.8 m<sup>3</sup>/ha. The trend in mean annual increment was similar, declining after 18 years (Figure 6-3). The mean annual increment reached a peak and declined as the stand grew older and denser (Figure 6-3), consistent with the principles of stand development (Smith et al. 1997).

### **Sensitivity Analysis**

The model was very sensitive to the site index value. A decrease in the site index value increased the volume, whereas an increase in site index decreased the volume (Table 6-2). The same result was seen for basal area, total number of trees and total number of recruits (Table 6-2). There was no significant change in the model outputs after changing the  $\beta_1$  parameter of the dbh growth model, which is related to initial dbh (Table 6-2). Changing the  $\beta_2$  parameter of the height growth model, which is related to initial height only slightly influenced the volume, but did not show any effects on other outputs. A similar result was observed for the  $\beta_3$  parameter of the height growth model, which is related to basal area.

The simulation results were extremely sensitive to the parameter  $\beta_1$  related to stand basal area in the conditional model predicting number of recruits. Decreasing this parameter greatly decreased the volume, the basal area, the total number of trees, and

number of recruits, whereas increasing the parameter value increased the number of recruits by a very large number, and also significantly increased the total number of trees, volume and basal area (Table 6-2).

### **Diameter Distribution of Trees**

We first analyzed the diameter distribution for all the simulations. The diameter distribution of the control was unimodal (Figure 6-4a). It was positively skewed for other harvesting scenarios (Figure 6-4), though somewhat less so for single tree selection (Figure 6-4d) as it did not have a wide range of diameters in comparison to group selection, light thinning, and heavy thinning. Visually, the diameter distribution for heavy thinning was similar to the reverse J shaped defined for uneven-aged stands.

We further analyzed the size distribution of trees under various harvest regimes using various statistics (Table 6-3). The control had negative skewness, indicating many small trees and few large trees. Group selection had the highest positive skewness indicating a longer right tail, and it also had the highest range of tree sizes available (Table 6-3). Similarly, heavy thinning and light thinning had the highest values for skewness after group selection. Also, group selection, light thinning and heavy thinning had highly positive kurtosis, indicating distribution with a sharp peaks and flattened right tails.

The Shannon-Wiener diversity index for tree sizes showed similar patterns for both 2 cm and 4 cm size classes. They were the highest for the heavy thinning, followed by low thinning and group selection (Table 6-3). The shape index also suggested similar patterns, with heavy thinning having a shape value closest to the value (0.54) for a reverse J shaped diameter distribution, followed by group selection and light thinning (Table 6-3).

## **Height Distribution of Trees**

We also examined histograms of the distribution of tree heights at the end of the simulations (Figure 6-5) and through statistics (Table 6-4). The control had the highest range of tree heights (Table 6-4), but the range of tree heights in the other four harvest treatments was similar. All the harvest treatments including the control had higher numbers of trees in the larger height classes (Figure 6-5). Heavy thinning and light thinning had trees more evenly distributed in the different height classes, followed by group selection (Figure 6-5).

We also calculated Shannon-Wiener Index ( $H'$ ) for height diversity for both 1 m and 2 m classes. The indices showed similar patterns in both classes (Table 6-4). The index was the highest for heavy thinning followed by low thinning. Control and group selection had similar  $H'$  values, while single tree selection had the lowest value for height diversity. To confirm the height diversity for different regimes, we also evaluated Berger-Parker Index ( $d'$ ; Table 6-4). The index was highest for the heavy thinning and low thinning followed by the group selection. This indicated that the vertical structural diversity was higher in the treatments versus the control, except for when considering single tree selection.

## **Harvest Volume and Economic Analysis of Harvest**

We also compared the different regimes in terms of total volume harvested, volume in different product classes, and revenue generated from harvests. The total volume harvested was the highest for heavy thinning, followed by light thinning and group selection (Table 6-5). The volume in different product classes was also the highest for heavy thinning (Table 6-5). The pulpwood volume was higher for the heavy

thinning followed by single-tree selection, group selection, and light thinning (Table 6-5). There was no chip-n-saw and sawtimber in group selection, and only a small amount of saw timber was obtained from single-tree selection. Heavy thinning also earned the highest income from the harvests (Table 6-5). Light thinning was the second highest revenue generator, followed by group selection and single-tree selection (Table 6-5).

### **Discussion and Conclusion**

Long-term predictions from a simulation model should be evaluated in terms of real management applications. Although individual models were evaluated through cross validation and examinations of their consistency with prior knowledge of stand growth and development, the predictions of stand growth using all submodels together should also be evaluated because errors from each submodel over multiple years could accumulate and produce results that are not compatible with prior knowledge of stand growth. Interactions between the parameters of the submodels can only be understood within the context of the full model. The basal area prediction of the control stand at 68 years (without any management and disturbances) was within the range published by other studies. For example, basal area of an old growth slash pine (85 yrs) stand was 34 m<sup>2</sup>/ha (Hebb and Clewell 1976). Similarly, a 50 year old unthinned slash pine plantation had 36 m<sup>2</sup>/ha basal area with site index 22 m (Stanley et al. 1991), and an unthinned 45 year old slash pine plantation (planted at 3048 trees per ha) had 55 m<sup>2</sup>/ha basal area when the site index was 25 m (Pienaar and Shiver 1984). The pattern of basal area growth also did not contradict previous observations of stand growth. The basal area reached a peak and declined to reach a steady state as in other studies (Liang et al. 2005). In sites with higher site index and higher density, the basal area has

been observed to reach the maximum quickly and start to decline as density dependent mortality occurs (Dickens and Will 2004).

The final volume of the control stand was also within the range reported in other studies. Hebb and Clewell (1976) reported 687 m<sup>3</sup>/ha (outside bark) for an old growth stand in the Florida panhandle. Another study reported 307.8 m<sup>3</sup>/ha (outside bark) for 50 year old unthinned natural slash pine stand with site index 23 and 870 trees per ha (Schreuder et al 1979). Benett (1980) also reported 405 m<sup>3</sup>/ha of volume outside bark for unthinned natural stand of slash pine at 50 years of age with a site index of 24 and basal area at 36 m<sup>2</sup>/ha. The periodic annual volume increment for the control scenario was higher (7.83 m<sup>3</sup>) for the first year, and declined in later years. This result is also consistent with other observations of stand growth, where volume growth is faster in the initial years in denser stands, and declines later when leaf area declines (Will et al. 2001). The mean annual increment of 7.83 m<sup>3</sup>/ha at 68 years is also within the range of 5 to 10 m<sup>3</sup>/ha reported for slash pine stands in flatwood soils (Sarigumba 1984, Borders and Bailey 1985). The mean annual increment has been reported to culminate early (around age 20 to 25 years; Sarigumba 1984, Borders and Bailey 1985). In the present study, the culmination was also early (because the stand density was higher) and started declining later (Figure 6-3).

Among the different harvest scenarios simulated, heavy thinning, low thinning, and group selection performed well in achieving a diameter distribution with trees in a wide range of size classes. Although this evaluation was done with only one simulation run per scenario in order to show the utility of the model, the results are still of interest in demonstrating the possible effects of different harvesting regimes. The positive

skewness of the diameter distribution for these harvest regimes indicated a heavier left tail with lots of smaller trees, and a flattened right tail with tree number decreasing as diameter increased. The diameter distribution of heavy thinning was the closest towards achieving a reverse J diameter distribution of an uneven-aged stand; this was also confirmed by the lower shape index for this harvest regime. Group selection had a wider range of tree sizes, and the largest trees among all scenarios. Our simulation indicated that heavy thinning, light thinning and group selection could be used to achieve a tree size distribution found in an uneven-aged stand, and all of these regimes could be useful for the conversion process. On the other hand, single-tree selection – always touted as one of the methods to be used in conversion (Farrar 1984, Baker et al. 1996) – did not perform well. One reason for this could be the removal of large trees during the harvest. For single-tree selection, the imposed size restriction to prevent harvesting of bigger trees during each harvest was not useful, and this led to the absence of large sized trees. If we could preserve some large trees during each harvest, this regime might also be able to achieve a wider dbh distribution. The risk of high grading and absence of large trees in single-tree selection has been seen by others (Franklin et al. 2007); therefore, in practice some of the larger trees within the diameter limit are kept and the required volume is fulfilled by cutting trees in lower size classes. In our simulations, however, this did not occur.

The range of height found in all scenarios including the control was visually similar, but heavy thinning, light thinning and group selection had trees more evenly distributed among different height classes. The Shannon-Wiener index also showed higher height diversity for heavy thinning and light thinning, with the control having similar diversity as

group selection. Further, the Berger-Parker index revealed that heavy thinning, light thinning, and group selection were able to achieve higher height diversity compared to single-tree selection and control. The range of height found in all scenarios was within the range reported in old growth slash pine forest (30-36 m at 85 years; Hebb and Clewell 1976). This result also confirmed that vertical structural diversity could be increased with harvesting treatments.

While public land managers (e.g national forests) are less interested in revenue than meeting established allowable cuts (Schulte and Buongiorno 1998), many land owners and managers are also interested in revenue generated from harvests while making a conversion from even-aged to uneven-aged structure; It is therefore useful to analyze products generated from harvests and their economic value. We did not analyze the amount of different product classes produced annually; however, we presented the amount of different products harvested and their economic returns overall. We reiterate that this evaluation was done with only one simulation run per scenario in order to show the utility of the model. However, the results are still of interest in demonstrating the possible effects of different harvesting regimes and giving a framework for comparisons for more complete analyses. Most of the volume harvested was in pulpwood. Chip-n-saw was available in light thinning, heavy thinning and group selection only, and sawtimber was available only in heavy thinning and light thinning. In the southeastern US, a large part of the wood harvested goes to the production of pulp and paper (62% of the total manufactured product sales), and pulp wood is the main product sold by landowners (Hodges et al. 2005). Therefore, the pulpwood production could be an important source of revenue.

Regarding income from the harvest, heavy thinning and light thinning were more profitable. Group selection also performed better than single tree selection. The revenue from harvest in the present study was lower compared to \$3035 per ha reported in other studies (Matta et al. 2009). The price for pulpwood used in the 2009 study was the same as used in this study, but the saw timber price was almost two times what we used. Even after using the same price for saw timber, our income was lower (difference was >\$1300 per ha). The reason for this is the higher amount of pulpwood from the harvests and the lower amount of chip-n-saw and saw timber. The regimes simulated here are geared towards achieving structural diversity rather than maximizing profit; this could lead to lower revenue from harvests. If a landowner wants to maximize revenue, he could selectively cut large size trees to maximize his or her income or increase the cutting cycle to get larger trees. However, this practice could lead to decreasing the amount of larger trees.

Our economic analysis included only revenue from wood, it did not include the recent market of waste for bioenergy production or using wood for electricity production (Hodges et al. 2005), or carbon markets (Stainback and Alavalapati 2002). Also income from pine straw (Dickens and Will 2004) could add to the revenue. On the other hand, the value derived from other ecosystem benefits such as water quality and aesthetics (Hodges et al. 2005) could make harvesting toward uneven-aged structure more profitable.

Thinning treatments are not usually discussed in the conversion literature (see for example Farrar 1984, Baker et al. 1996, Farrar 1996, Guldin and Baker 1998), except for old growth restoration (Choi et al. 2007) and to manipulate stands for ecological

values (Sprugel et al. 2009). Our simulation indicated that thinning treatments could be useful to achieve structural diversity, and our simulation of both heavy thinning and light thinning were the best among all scenarios. One reason could be that the thinning algorithm implemented in this study performed as a variable density thinning, which is different than commercial thinning. Another reason is that the thinning regimes reduced competition and allowed trees to grow into larger size classes. In most of the conifers including slash pine, higher stand density decreases tree diameter growth (Jones 1987, Dickens and Will 2004). In our ingrowth submodel as well, both probability of recruitment and number of recruits increased with medium basal area, and both decreased when basal area was too high or too low.

We simulated only a few harvest treatments from a range of harvest treatments available (Farrar 1984, Schulte and Buongiorno 1998, Guldin and Baker 1998). Different harvests could be tested by modifying the cutting cycle length and the amount and sizes of trees to be harvested. This is one of the important usages of a simulation model like the one in this study. Many different scenarios could be simulated with lower cost, though they require a lot of computational time. One possible area of future research would be to link the model with an optimization algorithm, and find the best intensity and timing of harvest to achieve a predefined target size structure while maximizing wood revenue.

Due to the high amount of computational cost, we did not perform a global sensitivity analysis, and only a few parameters that influenced the model output were tested for sensitivity. In the future, a global sensitivity analysis of all the parameters in the model should be done. Among the input variable and parameters tested, model

outputs were very sensitive to the site index and a parameter from the recruitment submodel. The increase in site index might increase competition and therefore, decrease the amount of recruits and total number of trees (Table 6-2). This influenced the final volume and basal area (Table 6-2). Increases in the basal area parameter of the recruitment model increased the number of recruits by an exorbitant amount, so care must be given to calibrate this parameter in the future use of this model.

In the present study, we used the model deterministically, except for the ingrowth model which had a stochastic component. In future, stochasticity could be incorporated in dbh growth, height growth and mortality models using mixed models with the random component selected from the distribution of the random effects used in data fitting procedure.

Our model behaved well within the biological limits of slash pine growth, and predictions were within the range described for slash pine in prior studies. Harvest regimes such as heavy thinning, light thinning and group selection were all able to shift the diameter distribution towards the distribution similar to that described for uneven-aged stands. These scenarios also had higher vertical structural diversity as indicated by the distribution of tree heights. Although light thinning was used every five years, there was not enough basal area to be removed; therefore, the algorithm also performed light thinning every ten years except for the first thinning which was done five years after the starting period. In all our simulations heavy thinning was the best regime, followed by light thinning and group selection. This was shown by the structural diversity and volume harvested for each of these different regimes. Single-tree selection did not perform well in our simulation; efforts to preserve larger sized trees during harvest could

result in higher tree size diversity. In single-tree selection the initial harvest did not remove large amounts of volume; bringing the stand to a lower residual basal area [17.21 m<sup>2</sup>/ha (75 ft<sup>2</sup>/acre)], and then following VGDL regulation could improve the result.

Slash pine growth and regeneration is also influenced by interaction with hardwoods. Scarification of the soil and control of hardwood competition is necessary for slash pine growth and regeneration (Feduccia 1978, McMinn 1981). Our model did not include this interaction; therefore, it assumes that there are no barriers to regeneration and growth from interaction with hardwoods. The study does demonstrate the potential application of the model for comparing various harvest scenarios.

Table 6-1. Variables and parameters used for sensitivity analysis of simulation model.

Variables/parameters	Base value	SE	Low	High	Parameters related to variable
Site index	24	-	18	28	-
$\beta_1$ (dbh growth model)	1.1738	0.1626	1.0112	1.3364	Initial individual tree diameter
$\beta_2$ (height growth model)	-0.2023	0.0157	0.2180	-0.1866	Initial individual tree height
$\beta_3$ (height growth model)	0.0035	0.0009	0.0026	0.0044	Stand basal area
$\beta_1$ (recruitment model/conditional number of recruits)	0.6153	0.1411	0.4742	0.7564	Stand basal area

Table 6-2. Percentage change in model output values relative to the base output value after changing the input variable and parameters.

Variables/parameters	Volume		Basal Area		Total number of trees		No of recruits	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Site index	+ 19%	-8%	+20%	-8%	+28.92%	-16.7%	>50%	-50.7%
$\beta_1$ (dbh growth model)	-0.9%	+1.2%	-0.4%	+2%	-0.8%	1.3%	-3.1%	+3%
$\beta_2$ (height growth model)	-4.2%	+5.4%	0.2%	0.6%	-0.2%	-0.1%	-0.3%	0%
$\beta_3$ (height growth model)	-4%	8.4%	0.7%	0.6%	-0.1%	-0.1%	0%	0%
$\beta_1$ (recruitment model/conditional number of recruits)	-16.5%	>100%	-17.3%	>100%	-26.3%	>100%	-100%	>+200%

Table 6-3. Various statistics describing the final diameter distributions of the simulated stand at age 68 under various harvest regimes with base site index 24.

Variables	Control	Light thin	Heavy thin	Single-tree selection	Group selection
Range of dbh (cm)	12.14-19.88	12.4-49	12.40-67.39	12.5-25.87	12.41-78.26
No of large trees (>30 cm dbh)	0	21	55	0	14
SD	1.67	0.156	0.2933	1.67	4.84
Skewness	-0.3187	2.20	2.20	0.98	6.48
Kurtosis	-1.0416	8.95	7.25	3.06	62.90
Shannon-Wiener	0.6918	1.3004	1.595	0.677	1.054
Index (H'; 4 cm)	(0.6931)	(2.1972)	(2.397)	(1.3862)	(2.3025)
H' (2 cm)	1.1883	1.94837	2.2663	1.2475	1.691
	(1.3862)	(2.7725)	(2.890)	(1.9459)	Max(2.8903)
Shape Index	1.04	0.90	0.72	0.98	0.86

Note: Values in parentheses are the maximum possible for the given size class.

Table 6-4. Various statistics describing the final height distributions of the simulated stands at age 68 under various harvest regimes.

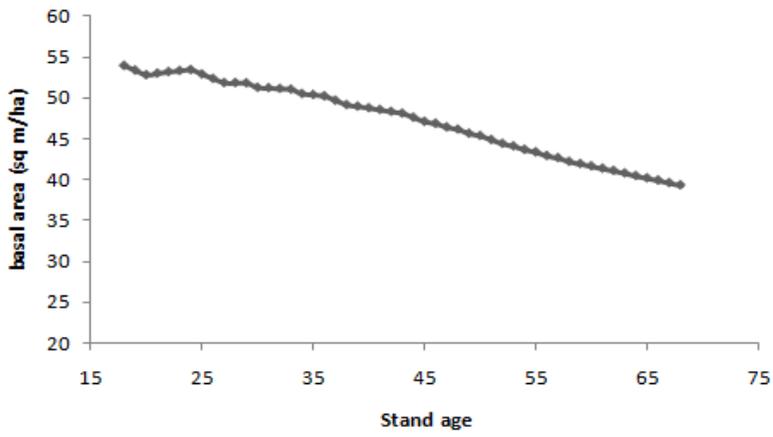
Variables	Control	Light thinning	Heavy thinning	Single-tree selection	Group selection
Range of ht (m)	11.96-32.6	12.01-29.28	11.6-28.7	12.75-30.35	12.38-29.35
SD	3.84	4.08	4.84	2.55	2.82
Skewness	-2.51	-1.64	-0.62	-3.18	-2.72
Kurtosis	6.017	1.64	-0.95	11.23	7.31
Shannon-Wiener	1.8265	2.1134	-2.4117	1.3722	1.4631
Index 1m (H')	(3.0910)	(2.8903)	(2.8332)	(2.9444)	(2.8903)
H' (2 m)	1.2087 (2.4849)	1.7098 (2.1972)	1.9737 (2.3025)	0.8719 (2.3025)	1.1988 (2.1972)
Berger-Parker Index (2 m classes)	1.4105	2.8300	2.84	1.2665	1.6029

Note: Values in parentheses are the maximum possible for the given height class.

Table 6-5. Volume of trees harvested (cubic meter outside bark) during each harvesting scenario simulated.

Number of harvests	Light thinning	Heavy thinning	Single-tree selection	Group selection
1 <sup>st</sup>	331.61	424.81	180.94	356.38
2 <sup>nd</sup>	26.85	37.65	108.88	28.88
3 <sup>rd</sup>	25.25	39.17	78.81	20.13
4 <sup>th</sup>	23.69	42.52	62.90	-
5 <sup>th</sup>	29.93	-	-	-
Total	437.34	544.15	431.53	405.38
<u>Product Classes</u>				
Volume pulpwood	354.92	424.52	430.14	387.33
Volume chip-n-saw	17.14	21.46	1.39	-
Volume saw timber	6.42	5.67	-	-
Net present value of harvest (US \$)	\$1392.01	\$1697.85	\$1140.33	\$1251.67

a)



b)

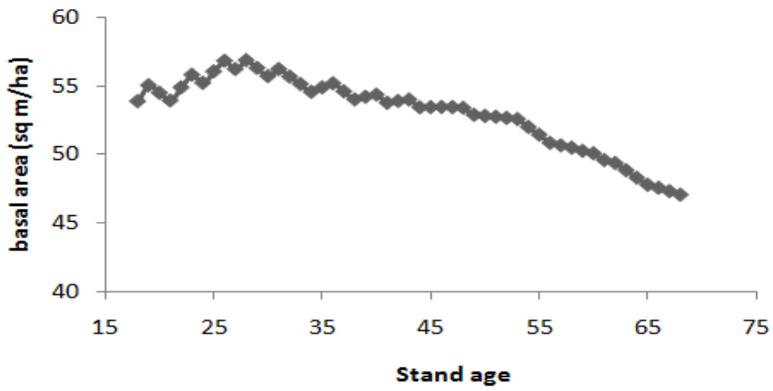


Figure 6-1. Simulated basal area development over time using the model for control stand (no harvest) for (a) base site index 24 (b) and site index 18.

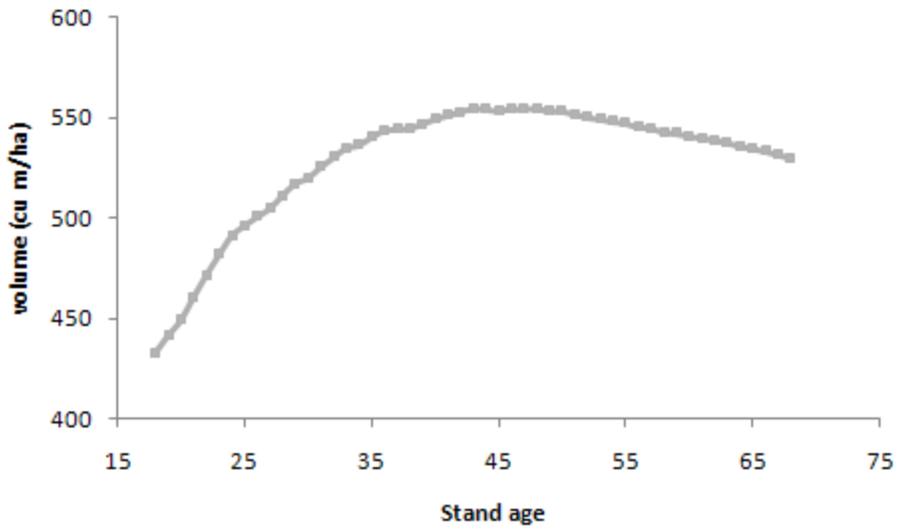


Figure 6-2. Simulated net stand volume over time for control stand (no harvest) with base site index 24.

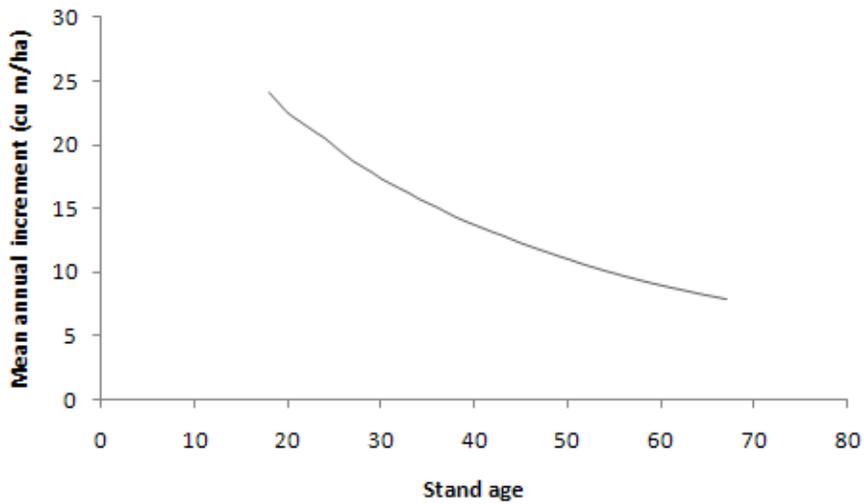


Figure 6-3. Simulated mean annual volume increment over time for the control stand (no harvest) with base site index 24.

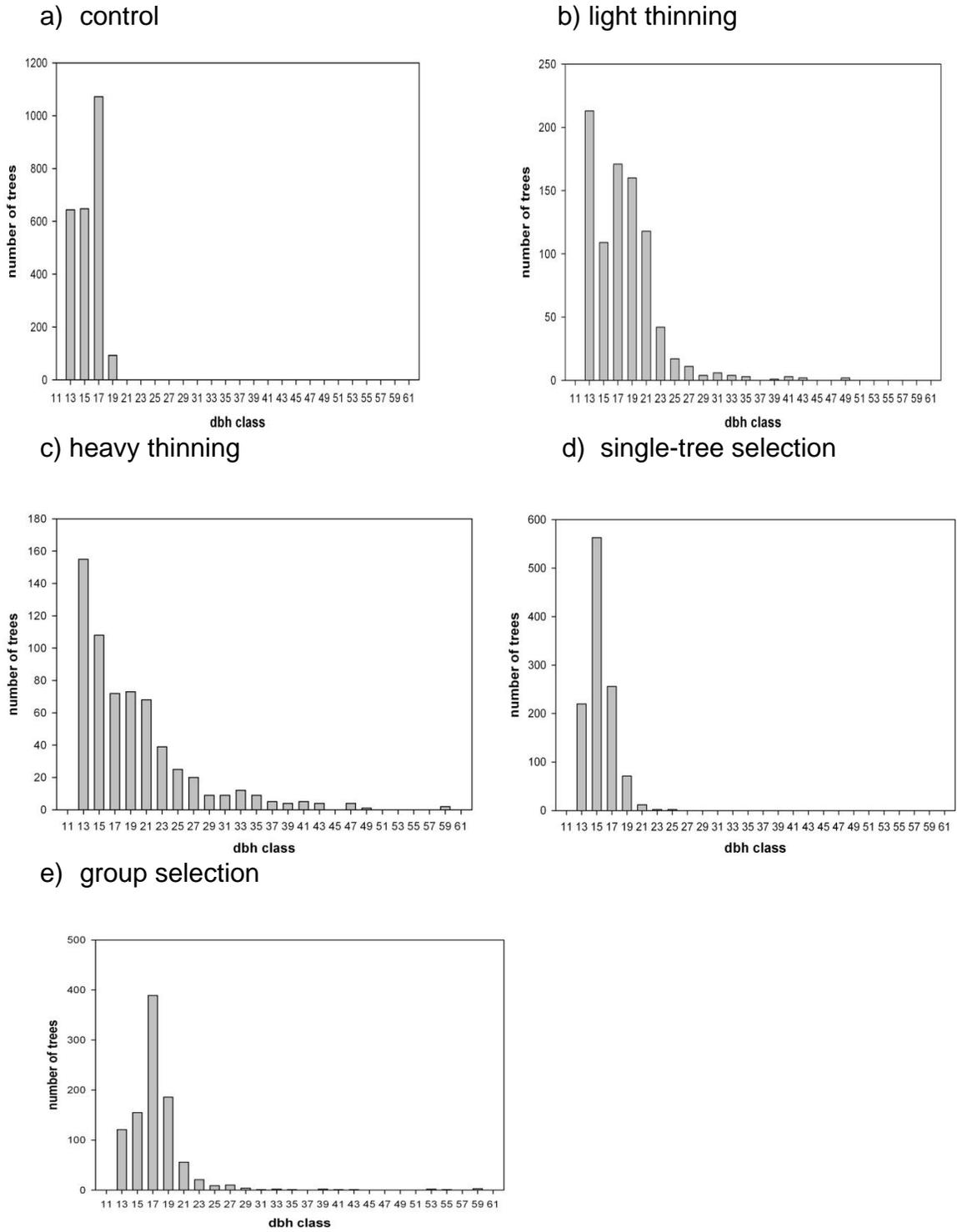


Figure 6-4. Diameter distribution (number of trees/ha) at the end of the simulation (stand age 68 years) under different harvest techniques. Dbh distribution is for 2 cm dbh classes (e.g., dbh class 13 includes trees >12 cm and <=14 cm).

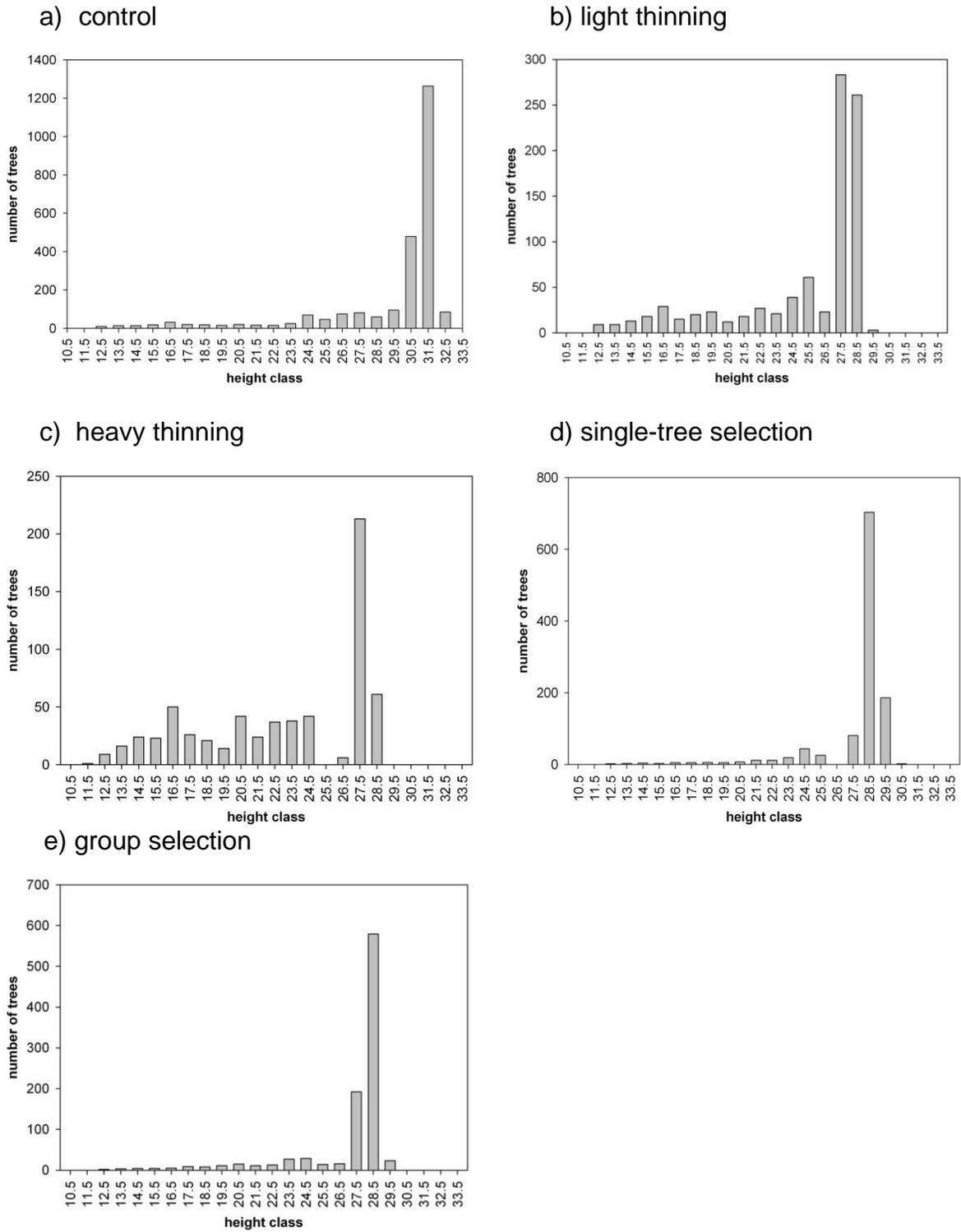


Figure 6-5. Height distribution (number of trees/ha) at the end of the simulation (stand age 68 years) under different harvest techniques. Height distribution is for 1 m classes (e.g., 11.5 m class includes trees > 11 m and ≤12m).

## CHAPTER 7 CONCLUSION

We developed diameter growth, height growth, mortality, and recruitment models for slash pine in Florida. The models were later used in a simulation model to test different harvest regimes. The diameter growth model in this study incorporates the effect of initial size (dbh and height) along with individual tree competition. Stand level variables, including measures of stand density, were not significant in the model and did not improve model fit. The effect of environment as measured by site index improved the fixed effects model, but once the random effect was included, this variable was not significant. This implies that the random effect incorporated a variety of both unmeasured and measured plot level effects. We also showed that a mixed effects model is better than a fixed effects model in predicting diameter growth. A variety of goodness of fit statistics confirmed the statistical precision and accuracy of our chosen model. The height growth model presented in this study showed that the tree level variables such as initial tree height and diameter, competition terms, and stand level variables such as basal area and quadratic mean diameter are important for predicting the height growth of a slash pine. This study also indicated that a non-linear model is better than a linear model for predicting height growth in slash pine. We also showed that a non-linear mixed model, including a random effect, gives better prediction than a fixed effects model. Once the random plot effect is incorporated, we did not find any improvement in the model prediction by modeling the correlation of trees within a plot through several residual covariance structures. Our study did not suggest a separate model for different communities except for UMF. The model predictions are biologically consistent.

An individual tree mortality model did not previously exist for slash pine, but has been developed for several other species in USA and other parts of the world. The model developed in our study provides information about the variables that are important in determining individual tree mortality in slash pine, which is useful for management or development of models in other areas of slash pine distribution. The model presented will be useful in simulating different silvicultural regimes (such as cutting and thinning), because it includes variables that are affected by tree removal. The effect of harvesting treatments on the mortality is incorporated into models by using terms like relative dbh and basal area (Hamilton 1986). For example, when a stand is thinned from below, its average stand diameter increases, thereby reducing the relative dbh. Similarly, thinning also reduces stand basal area and subsequently affects tree mortality (Hamilton 1986). The oldest stand we have in our data set was 72 years and the largest tree was 61 cm (24 inches), caution should be used while implementing the model beyond the data range, because the model was unable to capture the U shaped relationship between tree size and mortality. It would be interesting to validate the model performance with data from the older stands. Our study also showed that including a random plot effect in the mortality using a generalized linear mixed model improved the predictions.

We used a two step approach to model recruitment. The first model predicted the probability of recruitment, and conditional on that probability the second model predicted the number of recruits. The presented recruitment model can be used in a growth simulator for both stochastic and deterministic prediction. If the recruitment model is used deterministically, the probability of recruitment should be multiplied with the

number of predicted recruits. The number thus generated is used to determine the recruitment per hectare. To use the model stochastically, the probability of recruitment is compared with a random number generated from a uniform distribution between 0 and 1. If the estimated probability is greater than the generated number, then a plot is classified as recruited. The number of recruits should then be predicted for the plot classified as having new recruitment using the conditional model.

Although we did not validate our model with an independent dataset as in other studies (Vanclay 1992, LexerØd 2005), we used a cross validation approach to validate our model. The area under the ROC curve and significance tests of the probability model indicated very good discriminatory power of the model. The fit statistics calculated using the cross validation approach, such as mean residual error and bias, were within reasonable values. The model over-predicted the number of recruits as indicated by the positive percentage bias; however, the difference between observed and predicted values was not statistically different. In addition to that, the predictions from the model gave biologically meaningful results; therefore, the models should be useful for predicting recruitment in slash pine.

We used all the developed submodels in a simulation model. Our model behaved well within the biological limits of slash pine growth, and predictions were within the range described for slash pine in prior studies. The model was run once with each of several different harvest scenarios to demonstrate its utility and present a framework for evaluating different management scenarios. In order to more appropriately evaluate scenarios, however, the model should be run many times and the distribution of end values should be examined. Our simulation suggested that the harvest regimes such as

heavy thinning, light thinning and group selection were all able to shift the diameter distribution towards the distribution similar to that described for uneven-aged stands. They also had higher vertical structural diversity as indicated by distribution of tree heights. Although light thinning was used every five years, there was not enough basal area to be removed; therefore, the algorithm also performed light thinning every ten years except for the first thinning which was done five years after the start of the simulation. In all our simulations heavy thinning was the best performing regime, followed by light thinning and group selection. This was shown by the structural diversity statistics and volume harvested for each of these different regimes. Single-tree selection did not perform well in our simulation; though with effort to prevent larger sized trees from being removed during harvest could result in higher tree size diversity. In single-tree selection the initial harvest did not remove a large amount of volume. Bringing the stand to a lower residual basal area [17.21 sq m/ha (75 sq ft/acre)], and then following VGDL regulation could improve the result.

Slash pine growth and regeneration is also influenced by interactions with hardwoods. Scarification of the soil and control of hardwood competition is necessary for slash pine growth and regeneration (Feduccia 1978, McMinn 1981). Our model did not include this interaction; therefore, it assumes that there are no barriers to regeneration and growth from interaction with hardwoods.

We did not have enough data to make separate model fitting and validation datasets or an independent data to validate the models; we recognize the importance of validating our models with an independent data set and this will be one of the areas of future research. Finally, we want to reiterate the limitations of the simulation analysis.

Due to the large amount of computing time necessary, we did not replicate each scenario and the simulation was run only once per scenario. For more comprehensive results, each scenario should be run a large number (100 or 1000) times. This is an important future research area. On the other hand, our study does demonstrate the potential utility of the model, and how it can be used to compare various harvest scenarios. Moreover, we have presented a comprehensive method to evaluate structural diversity and economic returns.

## LIST OF REFERENCES

- ABRAHAMSON, W.G., AND D. C. HARTNETT. 1990. Pine flatwoods and dry prairies. P. 103-149 in *Ecosystems of Florida*, Myers, R. L., and J. J. Ewel (eds.). University of Central Florida Press, Orlando.
- ADAME, P., M. DEL RÍO AND I. CAÑELLAS. 2008. A mixed nonlinear height–diameter model for pyrenean oak (*Quercus pyrenaica* Willd.). *For. Ecol. Manage.* 256(1-2):88-98.
- AGRESTI, A. 1996. *An introduction to categorical data analysis*. John Wiley & Sons, New York. 290 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.
- AMBUEL, B., AND T.A. STANLEY. 1983. Area dependent changes in the bird communities and vegetation of southern Wisconsin forests. *Ecology.* 64:1057-1068.
- ARABATZIS, A.A., AND H.E. BURKHART. 1992. An Evaluation of Sampling Methods and Model Forms for Estimating Height-Diameter Relationships in Loblolly-Pine Plantations. *For. Sci.* 38(1):192-198.
- ASSMANN, E. 1970. *The principles of forest yield study*. Pergamon Press, New York.
- AVERY, T.E., AND H.E. BURKHART. 2002. *Forest measurements*. McGraw Hill, Boston, Massachusetts.
- 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., B.E. BORDERS, K.D. WARE, AND E.P. JONES. 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.
- BAKER, J. B., M.D. CAIN, J. M. GULDIN, P. A. MURPHY, AND M. G. SHELTON. 1996. *Uneven-aged silviculture for the loblolly and shortleaf pine forest cover types*. USDA Gen. Tech. Rep. SO-1 18., Asheville, NC. 65 p.
- BEAL, S. L., AND L. B. SHEINER. 1982. Estimating population kinetics. *Crit. Rev. Biomed. Eng.* 8(3): 195-222.
- BECHTOLD, W.A., AND P.L. PATTERSON (eds.). 2005. *The enhanced Forest Inventory and Analysis program—national sampling design and estimation procedures*. USDA For. Serv. Gen. Tech. Rep. SRS-80. 85 p.

- BENETT, F. A. 1970. *Variable-density yield tables for managed stands of natural slash pine*. USDA For. Serv. Res. Note. SE-141. 7p.
- BENETT, F. A. 1980. *Growth and Yield in natural stands of slash pine and suggested management alternatives*. USDA For. Serv. Res. Pap. SE-211. 8 p.
- BENETT, F. A., AND G. L. CLUTTER. 1968. *Multiple-product yield estimates for unthinned slash pine plantations-pulpwood, sawtimber gum*. USDA For. Serv. Res. Pap. SE-35. 21p.
- BENNETT, F. A. 1980. *Growth and yield in natural stands of slash pine and suggested management alternatives*. USDA For. Serv. Res. Pap. SE-211., Asheville, NC. 8 p.
- BIGING, G.S., AND M. DOBBERTIN. 1995. Evaluation of competition indices in individual tree growth models. *For. Sci.* 41(2):360-377.
- BOISVENUE, C., H. TEMESGEN, AND P. MARSHALL. 2004. Selecting a small tree height growth model for mixed-species stands in the southern interior of British Columbia, Canada. *For. Ecol. Manage.* 202(1-3):301-312.
- BORDERS, B. E., AND R. L. BAILEY. 1985. *Stand density effects in slash and loblolly pine plantations*. PMRC Res. Pap. 1985-1., Univ. of Georgia, Athens, GA. 149 p.
- BOTKIN, D.B. 1993. *Forest dynamics, an ecological model*. Oxford University Press, Oxford.
- BOX, G.E.P., AND H.L. LUCAS. 1959. Design of Experiments in Non-Linear Situations. *Biometrika.* 46(1-2):77-90.
- BRAGG, D. C. 2004. A prescription for old-growth-like characteristics in southern Pines. P. 80-92 in *Proc. of conf. on Silviculture in special places*, Shepperd, W. D., L. G. Eskew, and G. Lane (compils.). USDA For. Serv. RMRS-P-34., Fort Collins, CO. 255 p.
- BRAVO, F., C. ORDONEZ, AND I. LIZARRALDE. 2008. Modelling ingrowth in mediterranean pine forests: a case study from scot pine (*Pinus sylvestris* L.) and mediterranean maritime pine (*Pinus pinaster* Ait.) stands in Spain. *Investigacion Agraria: Sistemas y Recursos Forestales.* 17(3):250-260.
- BROCKWAY, D.G., K.W. OUTCALT, J. M. GULDIN, W. D. BOYER, J. L. WALKER, D. C. RUDOLPH, R. B. RUMMER, J. P. BARNETT, S. JOSE, AND J. NOWAK . 2005. *Uneven-aged management of longleaf pine forests: a scientist and manager dialogue*. USDA For. Serv. Gen. Tech. Rep. SRS-78., Asheville, NC. 38 p.
- BUDHATHOKI, C.B., T.B. LYNCH, AND J.M. GULDIN. 2008. Nonlinear mixed modeling of basal area growth for shortleaf pine. *For. Ecol. Manage.* 255(8-9):3440-3446.

- BUONGIORNO, J., S. DAHIR, H. C. LU, AND C. R. LIN. 1994. Tree size diversity and economic returns in uneven-aged forest stands. *For. Sci.* 40:83-103.
- BURKHART, H. E. 1990. Status and future of growth and yield models. P. 409-412 in *Proc. of a symposium on State of the art methodology of forest inventory*. USDA For. Serv. Gen. Tech. Rep. PNW GTR-263.
- BURNS, R. M., AND E. A. HEBB. 1972. *Site preparation and reforestation of droughty acid sands*. USDA Agriculture Handbook 426, Washington, DC. 61 p.
- CALAMA, R., AND G. MONTERO. 2004. Interregional nonlinear height-diameter model with random coefficients for stone pine in Spain. *Can. J. For. Res.* 34(1):150-163.
- CALAMA, R., AND G. MONTERO. 2005. Multilevel linear mixed model for tree diameter increment in stone pine (*Pinus pinea*): a calibrating approach. *Silva Fenn.* 39(1):37-54.
- CANCINO, J., AND K. V. GADOW. 2002. Stem number guide curves for uneven-aged forests –development and limitations. P 163-174 in *Continuous Cover Forestry- Assesment, Analysis, Scenarios*, Gadow, K. V., J. Nagel, and J. Saborowski (eds.). Kluwer Academic Publishers, The Netherlands. 351 p.
- CANNELL, M.G.R., P. ROTHERY, AND E.D. FORD. 1984. Competition within stands of *Picea sitchensis* and *Pinus contorta*. *Ann. Bot.* 53:349-362.
- CAO, Q.V., S.S. LI, AND M.E. MCDILL. 2002. Developing a system of annual tree growth equations for the loblolly pine - shortleaf pine type in Louisiana. *Can. J. For. Res.* 32(11):2051-2059.
- CAREY, A.B., S.P. HORTON, AND B.L. BISWELL. 1992. Northern spotted owls. Influence of prey base and landscape character. *Ecol. Monogr.* 62:223-250.
- CAREY, J.B. 2001. Experimental manipulation of spatial heterogeneity in Douglas-fir forests: effects on squirrels. *For. Ecol. Manage.* 152:13-30.
- CHEN, H.Y.H., S. FU, R.A. MONSERUD, AND I.C. GILLIES. 2008. Relative size and stand age determine *Pinus banksiana* mortality. *For. Ecol. Manage.* 255(12):3980-3984.
- CHOI, J., C.G. LORIMER, AND J.M. VANDERWERKER. 2007. A simulation of the development and restoration of old-growth structural features in northern hardwoods. *For. Ecol. Manage.* 249(3):204-220.
- CHOI, J., C.G. LORIMER, J. VANDERWERKER, W.G. COLE, AND G.L. MARTIN. 2001. A crown model for simulating long-term stand and gap dynamics in northern hardwood forests. *For. Ecol. Manage.* 152(1-3):235-258.

- CHOJNACKY, D.C. 1997. Modeling diameter growth for pinyon and juniper trees in dryland forests. *For. Ecol. Manage.* 93(1-2):21-31.
- CLUTTER, J.L., J.C. FORSTON, L. V. PIENAAR, G. H. BRISTER, AND R. L. BAILEY. 1983. *Timber management—a quantitative approach*. Wiley, New York. 333 p.
- COLE, W.G., AND C.G. LORIMER. 1994. Predicting tree growth from crown variables in managed northern hardwood stands. *For. Ecol. Manage.* 67(1-3):159-175.
- CONDES, S., AND H. STERBA. 2008. Comparing an individual tree growth model for *Pinus halepensis* Mill. in the Spanish region of Murcia with yield tables gained from the same area. *Eur.J.For.Res.* 127:253-261.
- CRECENTE-CAMPO, F., P. MARSHALL, AND R. RODRÍGUEZ-SOALLEIRO. 2009. Modeling non-catastrophic individual-tree mortality for *Pinus radiata* plantations in northwestern Spain. *For. Ecol. Manage.* 257(6):1542-1550.
- CROPPER, W. P. JR. 1998. Modeling the potential sensitivity of slash pine (*Pinus elliottii*) stem growth to increasing temperature and CO<sub>2</sub>. P. 353-366 in *The productivity and sustainability of southern forest ecosystems in a changing environment*, Mickler, R. A., and S. Fox (eds.). Springer, New York.
- CROPPER, W. P. JR., AND E. L. LOUDERMILK. 2006. The interaction of seedling density-dependence and fire in a matrix population model of longleaf pine (*Pinus palustris*). *Ecol. Model.* 198: 487-494.
- CROPPER, W. P. JR., AND H. L. GHOLZ. 1993. Simulation of the carbon dynamics of a Florida slash pine plantation. *Ecol. Model.* 66: 231-249.
- CROW, G.R., AND R.R. HICKS. 1990. Predicting mortality in mixed oak stands following spring insect defoliation. *For. Sci.* 36:831-841.
- DANIELS, R. F., AND H. E. BURKHART. 1975. *Simulation of individual tree growth and stand development in managed loblolly pine plantations*. Div. of For. and Wild. Resour., Virginia Polytech Inst. and State Univ. FWS-5-75. 69 p.
- DANIELS, R.F., H.E. BURKHART, AND T.R. CLASON. 1986. A comparison of competition measures for predicting growth of loblolly pine trees. *Can. J. For. Res.* 16(6):1230-1237.
- DAVIDIAN, M., AND D. GILTINAN. 2003. Nonlinear Models for Repeated Measurement Data: An Overview and Update. *J. Agricult. Biol. Env. Stat.* 8(4):387-419.
- DAVIDIAN, M., AND D. M. GILTINAN. 1995. *Nonlinear models for repeated measurement data*. Chapman & Hall, London, UK. 360 p.

- DICKENS, E. D., AND R. E. WILL. 2004. Planting density impacts on slash pine stand growth, yield, product class distribution, and economics. P 36-44 in *Proc. of slash pine symp. on Slash pine: still growing and growing!*, Dickens, E. D., J. P. Barnett, W. G. Hubbard, and E. J. Jokela (eds.). USDA Gen. Tech. Rep. SRS-76., Asheville, NC. 148 p.
- DIEGUEZ-ARANDA, U., F. CASTERDO-DORADO, G. ALVAREZ-GONZALEZ, AND R. RODRIGUEZ-SOALLERIO. 2005. Modeling mortality of Scots pine (*Pinus sylvestris* L.) plantations in the northwest of Spain. *Eur. J. For. Res.* 124:143-153.
- DIXON, G. E. 2002. *Essential FVS: A user's guide to the Forest Vegetation Simulator*. USDA For. Serv. For Mgmt. Serv. Cent. Fort Collins, CO. 240 p.
- EERIKÄINEN, K., J. MIINA, AND S. VALKONEN. 2007. Models for the regeneration establishment and the development of established seedlings in uneven-aged, Norway spruce dominated forest stands of southern Finland. *For. Ecol. Manage.* 242(2-3):444-461.
- EID, T., AND E. TUHUS. 2001. Models for individual tree mortality in Norway. *For. Ecol. Manage.* 154(1-2):69-84.
- EK, A. R., AND R. A. MONSERUD. 1974. *FOREST: A computer model for simulating the growth and reproduction of mixed species forest stands*. Univ. Wis-Madison, Coll. Agri. and Life Sci, Res Rep R2635. 85 p.
- FAMOYE, F., J.T.J. WULU, AND K.P. SINGH. 2004. On the generalized poisson regression model with an application to accident data. *J. Data. Sci.* 2:287-295.
- FANG, Z.X., 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.
- FARRAR, R. M. JR. 1984. Density control-natural stands. P. 20-22 in *Proc. of symp. on The loblolly pine ecosystem*, Karr, B. L., J. B. Baker, and T. Monaghan (eds.). Jackson, MS.
- FARRAR, R. M. JR. 1996. *Fundamentals of uneven-aged management in southern pine*. Misc. Publ. No. 9. Tall Timbers Res. Stn., Tallahassee, Fl.
- FARRAR, R.M. 1973. Southern Pine Site Index Equations. *J. For.* 71(11):696-697.
- FEDUCCIA, D.P. 1978. Regenerating slash pine naturally without suppression. *South. J. Appl. For.* 2(3):86-89.
- FITZMAURICE, G. M., N. M. LAIRD, AND J. H. WARE. 2004. *Applied longitudinal analysis*. John Wiley & sons, New York.

- FLEWELLING, J. W., AND R. A. MONSERUD. 2002. Comparing methods for modeling tree mortality. P.168-177 in *Proc. RMRS-P-25. Second Forest Vegetation Simulator Conference*, Crookston, N. L., and R. N. Harvis. (comps.). U. S. Dept. Agri. For. Serv. Rocky Mountain Res. Stn. Fort Collins, Co.
- FLORIDA NATURAL AREAS INVENTORY AND FLORIDA DEPARTMENT OF NATURAL RESOURCES. 1990. *Guide to the natural communities of Florida*. 116 p.
- FORTIN, M., AND J. DEBLOIS. 2007. Modeling tree recruitment with Zero-Inflated models: the example of hardwood stands in Southern Quebec, Canada. *For. Sci.* 53(4):529-539.
- FOSTER, T.E., AND J.R. BROOKS. 2001. Long-term trends in growth of *Pinus palustris* and *Pinus elliottii* along a hydrological gradient in central Florida. *Can. J. For. Res.* 31(10):1661-1670.
- FRANKLIN, J. F., AND D. B. LINDENMAYER. 2002. *Conserving forest biodiversity: a comprehensive multiscaled approach*. Island Press, Washington, DC.
- FRANKLIN, J. F., K. CROMACK, W. D. A. MCKEE, C. MASER, J. SEDEII, F. SWANSON, AND G. JUDAY. 1981. *Ecological characteristics of old-growth Douglas-fir forests*. USDA For. Serv. Gen. Tech. Rep. PNW-118, Portland, OR. 48 p.
- FRANKLIN, J.F., D. B. LINDENMAYER, J. A. MACMAHON, A. MCKEE, J. MAGNUSSON, D. A. PERRY, R. WAIDE, AND D. R. FOSTER. 2000. Threads of continuity: ecosystem disturbances, biological legacies and ecosystem recovery. *Conserv. Bio. in Prac.* 1: 8-16.
- FRANKLIN, J.F., R.J. MITCHELL, AND B. J. PALIK. 2007. *Natural disturbance and stand development principles for ecological forestry*. USDA For. Serv. Gen. Tech. Rep. NRS-19. 44 p.
- FRANKLIN, J.F., T.A. SPIES, R.V. PELT, A.B. CAREY, D.A. THORNBURGH, D.R. BERG, D.B. LINDENMAYER, M.E. HARMON, W.S. KEETON, D.C. SHAW, K. BIBLE, AND J. CHEN. 2002. Disturbances and structural development of natural forest ecosystems with silvicultural implications, using Douglas-fir forests as an example. *For. Ecol. Manage.* 155(1-3):399-423.
- FROST, C. 2006. History and future of the longleaf pine ecosystem. P. 9-42 in *The longleaf pine ecosystem: ecology, silviculture and restoration*, Jose, S., E. J. Jokela, D. L. Miller (eds.). Springer, New York. 438 p.
- GHOLZ, H.L., AND R.F. FISHER. 1982. Organic matter production and distribution in slash pine (*Pinus elliottii*) plantations. *Ecology*. 63(6):1827-1839.

- GOLDSTEIN, H. 1991. Nonlinear multilevel models, with an application to discrete response data. *Biometrika*. 78:45-51.
- GOLSER, M., AND H. HASENAUER. 1997. Predicting juvenile tree height growth in uneven-aged mixed species stands in Austria. *For. Ecol. Manage.* 97(2):133-146.
- GOURLET-FLEURY, S., AND F. HOULLIER. 2000. Modelling diameter increment in a lowland evergreen rain forest in French Guiana. *For. Ecol. Manage.* 131(1-3):269-289.
- GOWER, S.T., R.E. MCMURTIE, AND D. MURTY. 1996. Aboveground net primary productivity decline with stand age: potential causes. *Trends Ecol. Evol.* 11:378-382.
- GREENWOOD, D.L., AND P.J. WEISBERG. 2008. Density-dependent tree mortality in pinyon-juniper woodlands. *For. Ecol. Manage.* 255:2129-2137.
- GREGOIRE, T.G., AND O. SCHABENBERGER. 1996. A non-linear mixed-effects model to predict cumulative bole volume of standing trees. *J. of Appl. Stat.* 23(2):257-271.
- GREGOIRE, T.G., O. SCHABENBERGER, AND J. BARRETT. 1995. Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. *Can. J. For. Res.* 25:137-156.
- GROOT, A., S. GAUTHIER, AND Y. BERGERON. 2004. Stand dynamics modelling approaches for multicohort management of eastern Canadian boreal forests. *Silva Fenn.* 38(4):437-448.
- GULDIN, J.M. 1996. The role of uneven-aged silviculture in the context of ecosystem management. *Western. J. Appl. For.* 11(1):4-12.
- GULDIN, J. M., AND R. M. FARRAR JR. 2002. The plantation conversion demonstration at the crossett experimental forest- implications for converting stands from even-aged to uneven-aged structure. P 281-286 in *Proc. of Eleventh biennial southern silvicultural conference*, Outcalt, K. W. (ed.). USDA Gen. Tech. Rep. SRS-48., Asheville, NC.
- GULDIN, J.M., AND J.B. BAKER. 1998. Uneven-aged silviculture, southern style. *J. For.* 96(7):22-26.
- HALL, D.B., AND M.L. CLUTTER. 2004. Nonlinear mixed effects models for forest growth and yield prediction. *Biometrics*. 60(1):16-24.
- HALL, D.B., AND R.L. BAILEY. 2001. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. *For. Sci.* 47:311-321.

- HAMILTON, D.A. 1986. A logistic model of mortality in thinned and unthinned mixed conifer stands of northern Idaho. *For. Sci.* 32:989-1000.
- HANEWINKEL, M., AND H. PRETZSCH. 2000. Modelling the conversion from even-aged to uneven-aged stands of Norway spruce (*Picea abies* L. Karst.) with a distance-dependent growth simulator. *For. Ecol. Manage.* 134:55-70.
- HANN, D.W., C.L. OLSEN, AND A.S. HESTER. 1997. *ORGANON User's Manual*. Dept. of For. Res., Oregon State Univ., Corvallis, OR. 133 p.
- HARCOMBE, P.A. 1987. Tree life tables: simple birth, growth and death data encapsulate life histories and ecological roles. *BioScience*. 37:557-568.
- HASENAUER, H. 2006. Concepts within tree growth modeling. P. 3-14 in *Sustainable forest management: Growth model for Europe*, Hasenauer, H. (ed.). Springer-Verlag, New York.
- HASENAUER, H., AND G. KINDERMANN. 2006. Modeling regeneration in even and uneven-aged mixed species forests. P. 167-193 in *Sustainable forest management: Growth models for Europe*. Hasenauer, H. (ed.). Springer, Berlin, Germany.
- HASENAUER, H., AND R.A. MONSERUD. 1997. Biased predictions for tree height increment models developed from smoothed 'data'. *Ecol. Model.* 98(1):13-22.
- HEBB, E.A., AND A.F. CLEWELL. 1976. A remnant stand of old-growth slash pine in the Florida Panhandle. *Bull. Torrey Bot. Club.* 103(1):1-9.
- HEGYI, F. 1974. A simulation model for managing jack pine stands. P. 74-90 in *Growth models for tree and stand simulation*, Fires, J. (ed.). Royal College of Forestry, Stockholm, Sweden.
- HESTER, A. S., D. W. HANN, AND D. R. LARSEN. 1989. *ORGANON: Southwest Oregon growth and yield model*. User Manual. Version 2.0. For. Res. Lab., Oregon State Univ., Corvallis, OR.
- HODGES, A. W., W. D. MULKEY, J. R. ALAVALAPATI, D. R. CARTER, AND C. F. KIKER. 2005. *Economic impacts of the forests industry in Florida, 2003. Final report to the Florida Forestry Association*. Univ. of Florida, Inst. Food and Agri. Sci., Gainesville, FL. 47 p. Available online at <http://edis.ifas.ufl.edu/fe538>; last accessed Sept. 21, 2010.
- HOSMER, D.W., AND S. LEMESHOW. 2000. *Applied Logistic Regression*. John Wiley & Sons, New York.
- HUANG, S., AND S.J. TITUS. 1999. An individual tree height increment model for mixed white spruce–aspen stands in Alberta, Canada. *For. Ecol. Manage.* 123(1):41-53.

- HUANG, S., D.P. WIENS, Y. YANG, S.X. MENG, AND C.L. VANDERSCHAAF. 2009. Assessing the impacts of species composition, top height and density on individual tree height prediction of quaking aspen in boreal mixedwoods. *For. Ecol. Manage.* 258(7):1235-1247.
- HUANG, S., S. X. MENG, AND Y. YANG. 2009a. Using nonlinear mixed model technique to determine the optimal tree height prediction for black spruce. *Mod. Appl. Sci.* 3(4): 3-18.
- HUANG, S., Y. YANG, AND Y. WANG. 2003. A critical look at procedures for validating growth and yield models. P 271-294 in *Modelling forest ecosystems*, Amaro, A., D. Reed, and P. Soares (eds.). CABI Publishing, UK.
- HUANG, S.M., AND S.J. TITUS. 1995. An Individual Tree Diameter Increment Model for White Spruce in Alberta. *Can. J. For. Res.* 25(9):1455-1465.
- HUBBARD, R.M., B.J. BOND, AND M.G. RYAN. 1999. Evidence that hydraulic conductance limits photosynthesis in old Pinus ponderosa trees. *Tree Physiol.* 19(165):172.
- HUBBELL, S.P. 1980. Seed predation and the coexistence of tree species in tropical forests. *Oikos.* 35:214-229.
- Jones, E. P. 1987. Slash pine plantation study – age 30. P. 45-49 in *Proc. of the Fourth Biennial Southern Silvicultural Conf.*, Phillips, D. R. (ed.). USDA For. Serv. Gen. Tech. Rep. SE-42., Asheville, NC. 608 p.
- JUTRAS, S., H. HOKKA, V. ALENIOUS, AND H. SALMINEN. 2003. Modeling mortality of individual trees in drained peatland sites in Finland. *Silva Fenn.* 37(2):235-251.
- KRAJICEK, J.E., K.E. BRINKMAN, AND S.F. GINGRICH. 1961. Crown competition - a measure of density. *For. Sci.* 7:35-42.
- LEE, W., K.V. GADOW, D. CHUNG, J. LEE, AND M. SHIN. 2004. DBH growth model for Pinus densiflora and Quercus variabilis mixed forests in central Korea. *Ecol. Model.* 176(1-2):187-200.
- LEE, Y., AND D.W. COBLE. 2002. Modeling survival for unthinned slash pine plantations in East Texas under the influence of non-planted tree basal area and incidence of fusiform rust. *Texas Journal of Science.* 54(4):325-338.
- LEE, Y.J. 1971. Predicting mortality for even-aged stands of lodgepole pine. *For. Chron.* 47:29-32.
- LEXEROD, N., AND T. EID. 2005. Recruitment models for Norway Spruce, Scots Pine, Birch and Other broadleaves in young growth forests in Norway. *Silva Fenn.* 39(3):391-405.

- LEXERØD, N.L. 2005. Recruitment models for different tree species in Norway. *For. Ecol. Manage.* 206(1-3):91-108.
- LIANG, J., J. BOUONGIORNO, AND R.A. MONSERUD. 2005. Estimation and application of a growth and yield model for uneven-aged mixed conifer stands in California. *Int. For. Rev.* 7(2):101-112.
- LITTELL, R. C., G. A. MILLIKEN., W. W. STROUP., R. D. WOLFINGER., AND O. SCHABENBERGER. 2006. *SAS for Mixed Models, Second Edition*. SAS Institute Inc., Cary, NC, USA. 814 p.
- LIU, C., AND S.Y. ZHANG. 2005. Equations for predicting tree height, total volume, and product recovery for black spruce (*Picea mariana*) plantations in northeastern Quebec. *For. Chron.* 81(6):808-814.
- LIU, C., AND S.Y. ZHANG. 2005. Equations for predicting tree height, total volume, and product recovery for black spruce (*Picea mariana*) plantations in northeastern Quebec. *For. Chron.* 81(6):808-814.
- LOGAN, S. R., AND B. D. SHIVER. 2006. Adjusting slash pine growth and yield for silvicultural treatments. P 328-332 in *Proc. of the 13<sup>th</sup> biennial southern silvicultural research conference*, Connor, K. F (ed.). USDA For. Serv. Southern. Res. Stn. Gen. Tech. Rep. SRS-92. 640 p.
- LOHREY, R. E., AND S. V. KOSSUTH.1992. *Pinus elliottii* Engelm. In *Silvics of North America: 1. Conifers; 2. Hardwoods*, Burns, R. M., and Honkala, B. M. (tech. cords.). USDA For. Serv., Washington, DC. 877 p.
- LORIMER, C.G., AND L.E. FRELICH. 1984. A simulation of equilibrium diameter distributions of sugar maple (*Acer saccharum*). *Bull. Torrey Bot. Club.* 111:193-199.
- MACFARLANE, D.W., E.J. GREEN, A. BRUNNER, AND H.E. BURKHART. 2002. Predicting survival and growth rates for individual loblolly pine trees from light capture estimates. *Can. J. For. Res.* 32:1970-1983.
- MACFARLANE, D.W., E.J. GREEN, AND H.E. BURKHART. 2000. Population density influences assessment and application of site index. *Can. J. For. Res.* 30:1472-1475.
- MAGURRAN, A. E. 1988. *Ecological diversity and its measurement*. Princeton University Press, Princeton, NJ. 179 p.
- MALCOLM, P.N., J.F. FRANKLIN, A.B. CAREY, E.D. FORSMAN, AND T. HAMER. 1999. Forest stand structure of the northern spotted owl's foraging habitat. *For. Sci.* 45(4):520-527.

- MARTIN, G.L., AND A.R. EK. 1984. A Comparison of Competition Measures and Growth-Models for Predicting Plantation Red Pine Diameter and Height Growth. *For. Sci.* 30(3):731-743.
- 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. *South. J. Appl. For.* 23:39-45.
- MATTA, J.R., J.R.R. ALAVALAPATI, AND G.A. STAINBACK. 2009. Effect of conserving habitat for biodiversity on optimal management of non-industrial private forests in Florida. *J. For. Econ.* 15:223-235.
- MAY, D. M. 1990. *Stocking, forest type, and stand size class- the Southern Forest Inventory and Analysis unit's calculation of three important stand descriptors.* USDA For. Serv. Gen. Tech. Rep. SO-77. 7 p.
- MCCARTHY, J. 2001. Gap dynamics of forest trees: a review with particular attention to boreal forests. *Environ. Rev.* 9:1-59.
- MCCULLAGH, P., AND J.A. NELDER. 1989. *Generalized Linear Models.* Chapman Hall/CRC Press, Boca Raton, Fl. 511 p.
- MCMINN, J.W. 1981. Site preparation for natural regeneration of Slash Pine. *South. J. Appl. For.* 5(1):10-12.
- MENG, F.R., C.H. MENG, S. TANG, AND P.A. ARP. 1997. A new height growth model for dominant and codominant trees. *For. Sci.* 43(3):348-354.
- MENG, S.X., AND S. HUANG. 2009. Improved Calibration of Nonlinear Mixed-Effects Models Demonstrated on a Height Growth Function. *For. Sci.* 55(3):238-248.
- MENG, S.X., S. HUANG, Y. YANG, G. TRINCADO, AND C.L. VANDERSCHAAF. 2009. Evaluation of population-averaged and subject-specific approaches for modeling the dominant or codominant height of lodgepole pine trees. *Can. J. For. Res.* 39(6):1148-1158.
- MENGES, E.S. AND M.A. DEYRUP. 2001. Postfire survival in south Florida slash pine: interacting effects of fire intensity, fire season, vegetation, burn size, and bark beetles. *Int. J. Wildland Fire.* 10:53-63.
- MONSERUD, R.A. 1976. Simulation of forest tree mortality. *For. Sci.* 22:438-444.
- MONSERUD, R.A., AND G.E. REHFELDT. Genetic and environmental components of variation of site index in inland Douglas-fir. *For. Sci.* 36(1):1-9.
- MONSERUD, R.A., AND H. STERBA. 1996. A basal area increment model for individual trees growing in even- and uneven-aged forest stands in Austria. *For. Ecol. Manage.* 80(1-3):57-80.

- MONSERUD, R.A., AND H. STERBA. 1999. Modeling individual tree mortality for Austrian forest species. *For. Ecol. Manage.* 113(2-3):109-123.
- MURPHY, P. A., M. G. SHELTON, AND D. L. GRANEY. 1993. Group selection – problems and possibilities for the more shade-intolerant species. P 229-247 in *Proc. of 9th conf. on central hardwood forests*, Gillespie, A. R., G. R. Parker, P. E. Pope, and G. Rink (eds.). USDA For. Serv. Gen. Tech. Rep. NC-161, St Paul, MN.
- NAGELKERKE, N.J.D. 1991. A note on a general definition of the coefficient of determination. *Biometrika.* 78:691-692.
- NOTHDURFT, A., E. KUBLIN, AND J. LAPPI. 2006. A non-linear hierarchical mixed model to describe tree height growth. *Eur. J. For. Res.* 125(3):281-289.
- NUNIFU, T.K. 2008. Compatible diameter and height increment models for lodgepole pine, trembling aspen, and white spruce. *Can. J. For. Res.* 39(1):180-192.
- NYLAND, R.D. 2003. Even- to uneven-aged: the challenges of conversion. *For. Ecol. Manage.* 172(2-3):291-300.
- NYLAND, R.D. 2003. Even- to uneven-aged: the challenges of conversion. *For. Ecol. Manage.* 172(2-3):291-300.
- O'HARA K.L. 1996. *Dynamics and stocking-level relationships of multi-aged ponderosa pine stands*. Forest Science Monograph 33.
- OGAWA, K. 2009. Theoretical analysis of the interrelationships between self-thinning, biomass density, and plant form in dense populations of hinoki cypress (*Chamaecyparis obtusa*) seedlings. *Eur. J. For. Res.* 128(5):447-453.
- OLIVER, C. D., AND B. C. LARSON. 1996. *Forest stand dynamics: Update edition*. John Wiley & Sons, Inc., New York. 540p.
- OSTFELD, R.S., R.H. MASON, AND C.D. CANHAM. 1997. Effects of rodents on tree invasion of old fields. *Ecology.* 78:1531-1542.
- PALAHÍ, M., T. PUKKALA, D. KASIMIADIS, P. KONSTANTINOS, AND A.C. PAPAGEORGIOU. 2008. Modelling site quality and individual-tree growth in pure and mixed *Pinus brutia* stands in north-east Greece. *Ann. For. Sci.* 65(5):501p1-501p14.
- PALAHÍ, M., T. PUKKALA, J. MIINA, AND G. MONTERO. 2003. Individual-tree growth and mortality models for Scots pine (*Pinus sylvestris* L.) in north-east Spain. *Ann. For. Sci.* 60:1-10.
- PALIK, B.J., R.J. MITCHELL, AND J.K. HIERS. 2002. Modeling silviculture after natural disturbance to sustain biodiversity in the longleaf pine (*Pinus palustris*) ecosystem: balancing complexity and implementation. *For. Ecol. Manage.* 155(1-3):347-356.

- PENG, C. 2000. Growth and yield models for uneven-aged stands: past, present and future. *For. Ecol. Manage.* 132(2-3):259-279.
- PHILLIPS, P.D., T.E. BRASH, I. YASMAN, P. SUBAGYO, AND P.R. Van GARDINGEN. 2003. An individual-based spatially explicit tree growth model for forests in East Kalimantan (Indonesian Borneo). *Ecol. Model.* 159(1):1-26.
- PIENAAR, L. V., AND J. W. RHENEY. 1995. *Modeling stand growth and yield response to silvicultural treatments*. PMRC Tech. Rep 1995-2. Warnell Sch. of For. Nat. Res. Univ. of Georgia, Athens. 11 p.
- PIENAAR, L. V., W. M. HARRISON, T. R. BURGAN, AND J. W. RHENEY. 1988. *Yield prediction for site-prepared slash pine plantations in the coastal plain*. PMRC Tech. Rep. 1988-1. Univ. of Georgia, Athens, GA.
- PIENAAR, L.V., AND B.D. SHIVER. 1984. An analysis and models of basal area growth in 45-year-old unthinned and thinned slash pine plantation plots. *For. Sci.* 30(4):933-942.
- PIENAAR, L.V., AND B.D. SHIVER. 1984. The effect of planting density on dominant height in unthinned slash pine plantations. *For. Sci.* 30(4):1059-1066.
- PIENAAR, L.V., AND J.W. RHENEY. 1995. Modeling stand level growth and yield response to silvicultural treatments. *For. Sci.* 41(3):629-638.
- PIENAAR, L.V., B.D. SHIVER, AND G.E. GRIDER. 1985. Predicting basal area growth in thinned Slash pine plantations. *For. Sci.* 31(3):731-741.
- PINHEIRO, J. C., AND D. M. BATES. 2004. *Mixed effects models in S and S-PLUS*. Springer, New York. 528 p
- PLOCHMAN, R. 1992. The forests of central Europe: a changing view. *J. For.* 90:12-16.
- PORTÉ, A., AND H.H. BARTELINK. 2002. Modelling mixed forest growth: a review of models for forest management. *Ecol. Model.* 150(1-2):141-188.
- PRETZSCH, H., P. BIBER, AND J. DURSKÝ. 2002. The single tree-based stand simulator SILVA: construction, application and evaluation. *For. Ecol. Manage.* 162(1):3-21.
- PRICE, D.T., N.E. ZIMMERMANN, P. J. VAN DER MEER, M.J. LEXER, P. LEADLEY, I.T.M. JORRITSMA, J. SCHABER, D.F. CLARK, P. LASCH, S. MCNULTY, J. WU, AND B. SMITH. 2001. Regeneration in gap models: priority issues for studying forest responses to climate change. *Climate Change.* 51:475-508.
- PUKKALA, T., AND T. KOLSTROEM. 1991. Effect of spatial pattern of trees on the growth of a Norway spruce stand. A simulation model. *Silva Fenn.* 25(3):117-131.

- PUKKALA, T., E. LÄHDE AND O. LAIHO. 2009. Growth and yield models for uneven-sized forest stands in Finland. *For. Ecol. Manage.* 258(3):207-216.
- QUICKE, H.E., R.S. MELDAHL, AND J.S. KUSH. 1994. Basal Area Growth of Individual Trees - a Model Derived from a Regional Longleaf Pine Growth Study. *For. Sci.* 40(3):528-542.
- READER, R.J. 1993. Control of Seedling Emergence by Ground Cover and Seed Predation in Relation to Seed Size for Some Old-Field Species. *J. Ecol.* 81(1):pp. 169-175.
- RITCHIE, M.W., AND D.W. HANN. 1986. Development of a tree height growth model for Douglas-fir. *For. Ecol. Manage.* 15(2):135-145.
- ROBICHAUD, E., AND I.R. METHVEN. 1991. Tree vigor and height growth in black spruce. *Trees.* 5(158):163.
- RODRIGUEZ, G., AND N. GOLDMAN. Improved estimation procedures for multilevel models with binary response: A case study. *J. R. Statist. Soc. A.* :339-355.
- ROSE, C.E., D.B. HALL, B.D. SHIVER, M.L. CLUTTER, AND B. BORDERS. 2006. A multilevel approach to individual tree survival prediction. *For. Sci.* 52(1):31-43.
- ROUVINEN, S. AND T. KUULUVAINEN. 2005. Tree diameter distributions in natural and managed old *Pinus sylvestris*-dominated forests. *For. Ecol. Manage.* 208(1-3):45-61.
- RYAN, M.G., AND B.Y. YODER. 1997. Hydraulic limits to tree height and tree growth. *BioScience.* 47:235-242.
- RYAN, T.P. 1997. *Modern Regression methods.* Wiley, New York.
- SARIGUMBA, T. I. 1984. Sustained response of planted slash pine to spacing and site preparation. P. 79-84 in *Proc. of the 3rd biennial southern silviculture research conf.* USDA For. Serv. Gen. Tech. Rep. SO-34., Asheville, NC.
- SAS INSTITUTE INC. 2008. *SAS 9.2 Help and Documentation: Your complete guide to, Syntax, How to, Examples, Procedures, Concepts, What's New, and Tutorials.* SAS Institute Inc., Cary, NC.
- SCHABENBERGER, O. 2005. *Introducing the Glimmix procedure for Generalized Linear Mixed Models.* SAS Global Forum Proceedings. Paper 196-30. Available online at <http://www2.sas.com/proceedings/sugi30/196-30.pdf>; last accessed on Sept. 9, 2010. 20 p.
- SCHABENBERGER, O., AND F.J. PIERCE. 2001. *Contemporary Statistical Models for the Plant and Soil Sciences.* CRC Press, Boca Raton, FL.

- SCHREUDER, H.T., W.L. HAFLEY, AND F.A. BENNETT. 1979. Yield Prediction for Unthinned Natural Slash Pine Stands. *For. Sci.* 25(1):25-30.
- SCHREUDER, H.T., W.L. HAFLEY, AND F.A. BENNETT. 1979. Yield predictions for unthinned natural slash pine stands. *For. Sci.* 25(1):25-30.
- SCHULTE, B.J., AND J. BUONGIORNO. 1998. Effects of uneven-aged silviculture on the stand structure, species composition, and economic returns of loblolly pine stands. *For. Ecol. Manage.* 111(1):83-101.
- SCHWINNING, S., AND J. WEINER. 1998. Mechanisms determining the degree of size asymmetry in competition among plants. *Oecologia.* 113(4):447-455.
- SEDJO, R. A. 1996. Toward an operational approach to public forest management. *J. For.* 94:24-27.
- SEYMOUR, R., AND M. L. HUNTER. 1999. Principles of ecological forestry. P. 22-61 in *Maintaining biodiversity in forest ecosystems*, Hunter, M.L. (ed.). Cambridge University Press, Cambridge.
- SEYMOUR, R.S., A.S. WHITE, AND P.G. DEMAYNADIER. 2002. Natural disturbance regimes in northeastern North America—evaluating silvicultural systems using natural scales and frequencies. *For. Ecol. Manage.* 155(1-3):357-367.
- SHANNON, C. E., AND W. WEAVER. 1949. *The mathematical theory of communication*. University of Illinois Press, Urbana.
- SHIBATA, M., AND T. NAKASHIZUKA. 1995. Seed and Seedling Demography of Four Co-Occurring *Carpinus* Species in a Temperate Deciduous Forest. *Ecology.* 76(4):pp. 1099-1108.
- SHUGART, H.H. 1984. *A theory of forest dynamics. The ecological implications of forest succession models*. Springer-Verlag, New York. 278 p.
- SIIPILEHTO, J. 1999. Improving the accuracy of predicted basal-area diameter distribution in advanced stands by determining stem number. *Silva Fenn.* 33(4):281-301.
- SILVERTOWN, J.W., AND J. LOVETT DOUST. 1993. *Introduction to Plant Population Biology*. Blackwell Science, Oxford.
- SMITH, D. M., B. C. LARSON, M. J. KELTY, AND M. S. ASHTON. 1997. *The practice of silviculture: applied forest ecology, Ninth edition*. John Wiley & Sons, Inc., New York. 537p.
- SOMERS, G.L., R.G. ODERWALD, W.R. HARMS, AND O.G. LANGDON. 1980. Predicting mortality with a Weibull distribution. *For. Sci.* 26:291-300.

- SPRUGEL, D.G., K.G. RASCHER, R. GERSONDE, M. DOVČIAK, J.A. LUTZ, AND C.B. HALPERN. 2009. Spatially explicit modeling of overstory manipulations in young forests: Effects on stand structure and light. *Ecol. Model.* 220(24):3565-3575.
- STAGE, A. R. 1973. *Prognosis model for stand development*. USDA For. Serv. Res. Pap. INT-137. 32 p.
- STAINBACK, G.A., AND J.R.R. ALAVALAPATI. 2002. Economic analysis of slash pine forest carbon sequestration in the southern U. S. *J. For. Econ.* 8:105-117.
- STANLEY, J. Z., D. P. FEDUCCIA, V. C. BALDWIN JR., AND T. R. DELL. 1991. *Growth and yield predictions for thinned and unthinned slash pine plantations on cutover sites in the West Gulf region*. USDA For. Serv. Res. Pap. SO-264. New Orleans, Louisiana. 32 p.
- STERBA, H., A. BLAB, AND K. KATZENSTEINER. 2002. Adapting an individual tree growth model for Norway spruce (*Picea abies* L. Karst.) in pure and mixed species stands. *For. Ecol. Manage.* 159(1-2):101-110.
- STERBA, H., AND T. LEDERMANN. 2006. Inventory and modelling for forests in transition from even-aged to uneven-aged management. *For. Ecol. Manage.* 224(3):278-285.
- STERBA, H., AND T. LEDERMANN. 2006. Inventory and modelling for forests in transition from even-aged to uneven-aged management. *For. Ecol. Manage.* 224(3):278-285.
- TAO, J., AND R. LITTELL. 2002. *Mixed model analysis using the SAS system: course notes*. SAS Institute Inc, Cary, NC. 492 p.
- TECK, R.M., AND D. E. HILT. 1990. *Individual-tree Probability of Survival Model for the Northeastern United States*. USDA For. Serv. Res. Pap. NE-642. 10 p.
- TEMESGEN, H., V.J. MONLEON, AND D.W. HANN. 2008. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. *Can. J. For. Res.* 38(3):553-565.
- TIMBER MART-SOUTH. 2010. *Timber Mart-South market news quarterly. Abridged Version*. The Norris Foundation, Univ. of Georgia, Athens, GA. 17 p.
- TRASOBARES, A., AND T. PUKKALA. 2004. Optimising the management of uneven-aged *Pinus sylvestris* L. and *Pinus nigra* Arn. mixed stands in Catalonia, north-east Spain. *Ann. For. Sci.* 61(8):747-758.
- TRASOBARES, A., T. PUKKALA, AND J. MIINA. 2004. Growth and yield model for uneven-aged mixtures of *Pinus sylvestris* L. and *Pinus nigra* Arn. in Catalonia, north-east Spain. *Ann. For. Sci.* 61(1):9-24.

- USHER, M.B. 1966. A matrix approach to the management of renewable resources, with special reference to selection forests. *J. Appl. Eco.* 3:355-367.
- UZOH, F.C.C., AND W.W. OLIVER. 2006. Individual tree height increment model for managed even-aged stands of ponderosa pine throughout the western United States using linear mixed effects models. *For. Ecol. Manage.* 221(1-3):147-154.
- VANCLAY, J.K. 1988. A stand growth model for cypress pine. P. 310-312 in *Modelling trees, stand and forests*, Leech, J.W., P.W. McMurtrie and P.W. West(eds.). School of Forestry, University of Melbourne.
- VANCLAY, J.K. 1989. A growth model for north Queensland rainforests. *For. Ecol. Manage.* 27:245-271.
- VANCLAY, J.K. 1992. Modelling regeneration and recruitment in a tropical rain forest. *Can. J. For. Res.* 22:1235-1248.
- VANCLAY, J.K. 1994. *Modelling forest growth and yield: applications to mixed tropical forests*. CABI International, New York, USA. 312 p.
- VONESH, E. F., AND V. M. CHINCHILLI. 1997. *Linear and nonlinear models for the analysis of repeated measurements*. Marcel Dekker, New York. 560 p.
- WEBER, L.A., A.R. EK, AND T.D. DROESSLER. 1986. Comparison of stochastic and deterministic mortality estimation in an individual tree based stand growth model. *Can. J. For. Res.* 16(5):1139-1141.
- WEINER, J. 1990. Asymmetric competition in plant populations. *Trends in Ecology & Evolution.* 5(11):360-364.
- WEISKITTEL, A.R., S.M. GARBER, G.P. JOHNSON, D.A. MAGUIRE, AND R.A. MONSERUD. 2007. Annualized diameter and height growth equations for Pacific Northwest plantation-grown Douglas-fir, western hemlock, and red alder. *For. Ecol. Manage.* 250(3):266-278.
- WEISKITTEL, A.R., S.M. GARBER, G.P. JOHNSON, D.A. MAGUIRE, AND R.A. MONSERUD. 2007. Annualized diameter and height growth equations for Pacific Northwest plantation-grown Douglas-fir, western hemlock, and red alder. *For. Ecol. Manage.* 250(3):266-278.
- WEST, P.W., D.A. RATKOWSKY, AND A.W. DAVIS. 1984. Problems of hypothesis testing of regressions with multiple measurements from individual sampling units. *For. Ecol. Manage.* 7(3):207-224.
- WIENS, J. A. 2008. Allerton Park 1983: the beginnings of a paradigm for landscape ecology? *Land. Ecol.* 23:125-128

- WILL, R.E., G.A. BARRON, E. C. BURKES, B. SHIVER, AND R.O. TESKEY. 2001. Relationship between intercepted radiation, net photosynthesis, respiration, and rate of stem volume growth of *Pinus taeda* and *Pinus elliottii* stands of different densities. *For. Ecol. Manage.* 154(1-2):155-163.
- WILL, R.E., N.V. NARAHARI, B.D. SHIVER, AND R.O. TESKEY. 2005. Effects of planting density on canopy dynamics and stem growth for intensively managed loblolly pine stands. *For. Ecol. Manage.* 205(1-3):29-41.
- WOODRUFF, D.R., B.J. BOND, G.A. RITCHIE, AND W. SCOTT. 2002. Effects of stand density on the growth of young Douglas-fir trees. *Can. J. For. Res.* 32:420-427.
- WYKOFF, W. R. 1986. *Supplement to the user's guide for the stand prognosis model-version 5.0*. USDA For. Serv. Gen. Tech. Rep. INT-208. 36 p.
- WYKOFF, W.F., N. L. CROOKSTON, AND A. R. STAGE. 1982. *User's guide to the Stand Prognosis model*. USDA For. Serv. Gen. Tech. Rep. INT-133. 112 p.
- WYKOFF, W.R. 1990. A basal area increment model for individual conifers in the northern Rocky Mountains. *For. Sci.* 36(4):1077-1104.
- YANG, Y., S. HUANG, S.X. MENG, G. TRINCADO, AND C.L. VANDERSCHAAF. 2009. A multilevel individual tree basal area increment model for aspen in boreal mixedwood stands. *Can. J. For. Res.* 39(11):2203-2214.
- YANG, Y., S.J. TITUS, AND S. HUANG. 2003. Modeling individual tree mortality for white spruce in Alberta. *Ecol. Model.* 163(3):209-222.
- 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.
- ZAVALA, M.A., O. ANGULO, R. B. DE LA PARRA, AND J.C. LOPEZ-MARCOS. 2007. An analytical model of stand dynamics as function of tree growth, mortality and recruitment: the shade tolerance-stand structure hypothesis revisited. *Journal of Theoretical Biology.* 244:440-450.
- ZHANG, S., H.E. BURKHART, AND R.L. AMATEIS. 1996. Modeling individual tree growth for juvenile loblolly pine plantations. *For. Ecol. Manage.* 89(1-3):157-172.
- ZHANG, S., R.L. AMATEIS, AND H.E. BURKHART. 1997. Constraining individual tree diameter increment and survival models for loblolly pine plantations. *For. Sci.* 43:414-423.
- ZHAO, D., B. BORDERS, M. WANG, AND M. KANE. 2007. Modeling mortality of second-rotation loblolly pine plantations in the Piedmont/Upper Coastal Plain and Lower Coastal Plain of the southern United States. *For. Ecol. Manage.* 252(1-3):132-143.

## BIOGRAPHICAL SKETCH

Nilesh Timilsina is a citizen of Nepal, and was born in Kathmandu, Nepal. He finished his master's degree in zoology (ecology) in Nepal, and his thesis was on status of a species of freshwater Dolphin. After his degree, he worked as a biology teacher for a year and later as a Conservation officer for Bird Conservation, Nepal – a national non-governmental organization dedicated to conserving birds and biodiversity. He has also successfully completed several research works on the status of endangered bird and its habitat in Nepal.

In fall 2002, he started master's in environmental studies at Florida International University (FIU), Miami, and graduated in 2005. His thesis work was on the forests of western Nepal. After his master's, he worked for the Tropical Ecosystem lab at FIU, and was involved in several projects in Everglades National Park. He has also worked as an Adjunct lecturer teaching an online course. He started his PhD in the School of Forest Resources and Conservation, University of Florida in fall 2006.