MODELLING INDIVIDUAL TREE GROWTH OF NATURALLY OCCURING EVEN-AGED SHORTLEAF PINE

(*Pinus echinata* Mill.)

By

PRADIP SAUD

Bachelor of Science in Forestry
Tribhuvan University
Kathmandu, Nepal
2006

Master of Science in Forestry
West Virginia University
Morgantown, WV
2011

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2015

MODELLING INDIVIDUAL TREE GROWTH OF

NATURALLY OCCURING EVEN-AGED SHORTLEAF

PINE (*Pinus echinata* Mill.)

Dissertation Approved:

Dr. Thomas B. Lynch

Dissertation Adviser

Dr. Rodney Will

Dr. Steve W. Hallgren

Dr. Joshua D. Habiger

ACKNOWLEDGEMENTS

My special thanks and gratitude are to my advisor, Dr. Thomas B. Lynch, for his excellent guidance, continuous support, and patience throughout my graduate study and research work at Oklahoma State University. Thanks are also due to thank Dr. Rodney Will, Dr. Steve W. Hallgren and Dr. Joshua D. Habiger for helping and encouraging me as an advisory committee member. I am thankful to the Department of Natural Resource Ecology and Management (NREM) at OSU, its faculty and staff for providing all facilities and Graduate Research Assistantship that made my study possible. I would like to extend my thanks to USDA Forest Service Southern Research Station and the Ozark and Ouachita National Forests that provided data and financial support for this study. I appreciate the help and support of graduate students of NREM; especially Dipesh KC, Anup KC, Bharat Sharma Acharya, Anish Khanal, and Daniel Harrington.

Finally, I am thankful to my parents, for their support, encouragement and patience during my pursuit of the Doctorate degree. To my lovely wife Sarita Ayer, who inspired me and provided constant encouragement during the entire process. I also offer my regards to the Nepalese community of Stillwater and all of those who supported me in any respect in accomplishing my study.

Name: Pradip Saud

Date of Degree: DECEMBER, 2015

Title of Study: MODELLING INDIVIDUAL TREE GROWTH OF NATURALLY OCCURING EVEN-AGED SHORTLEAF PINE (*Pinus echinata* Mill.)

Major Field: Natural Resource Ecology and Management

Abstract: Modelling individual tree growth is important for the assessment and management of naturally occurring shortleaf pine, as its abundance (hectare) is reducing due to increased plantings of other commercially important southern pine species. Modelling was conducted using six-time measured mensurational data over a period of 25 years from permanent plots in western Arkansas and eastern Oklahoma. The major objectives of this study were to update and modify the existing nonlinear models that use the ordinary least square (OLS) technique to estimate parameters of an individual tree growth model: height-diameter relationship, crown ratio estimation, annual basal area growth, and individual tree mortality. The updated and modified model show that quadratic mean diameter better explained the height-diameter relationship and relative spacing index improved crown ratio estimation of an individual tree. The climate-based model with average minimum air temperature and average total monthly precipitation of the growing season optimized the prediction of annual basal area growth rate by lowering residual standard error by 1% compared to existing growth model. Simulations of climate change scenarios showed that under increased temperature scenarios the annual basal area growth rates of trees are less affected in older stands versus younger stands. The probability of annual mortality of an individual tree was influenced by its size with time (diameter/age), stand level productivity with time (basal area per hectare/age) and stand level competition (quadratic mean diameter). Annual mortality prediction model with a logistic function for binary response performed better than fitting a model with mixed-effects approach. Including a variable of cutting smaller diameter trees – "low thinning effect" – had little impact in the mortality model. OLS estimates of the nonlinear model with and without correction of autocorrelation and of heterogeneous error variance provided better predictions than a mixed-effects model with random effect set to have null effect in prediction. OLS model with the first order autoregressive (AR (1)) structure for modeling auto-correlation within individual observation than within stand in the repeated measurement provided smaller Akaike Information Criteria (AIC). These individual growth models would be very helpful in developing a growth and yield modelling system of shortleaf pine, which could be used to simulate forest growth under different management scenarios.

TABLE OF CONTENTS

Chapter **Page**

LIST OF TABLES

Table **Page**

Table **Page**

LIST OF FIGURES

Figure **Page**

Figure **Page**

CHAPTER I

INTRODUCTION

Assessment of the growth and yield are important because it provides information on the present status of the stand characteristics to forecast future conditions and basis for formulating and implementing forest management decisions. The growth represents the increase in the size of an individual tree or population with time or over a given period (e.g. basal area growth in m^2ha^{-1} $year^{-1}$). Yield is the increase in the size of an individual tree or population at the end of the certain period (e.g. basal area in m^2ha^{-1} at age 70 years). This change in size of an individual or population with time is described by modeling an appropriate growth functions i.e. growth and yield model. The function used can be as either mechanistic or empirical to describe the behavior of the response variable. Mechanistic functions identify the causes and explain the phenomenon, but the empirical function doesn’t. In application, the mixture of both empirical and mechanistic features is of common choice to construct a model because such models provide insight into the theoretical or biological phenomenon (Weiskittel et al., 2011; Burkhart and Tomé, 2012). The form and characteristics of a model are often determined by the fitted data that represents the range of the observations and the attribute of individual sample trees from a stand used to estimate the equation’s coefficients. Regardless of the period of data collection and range of the data used for modelling, the growth and yield function should be consistent with the consistency with the fundamental principles of biological growth. The signs of the parameter also should exhibit a response curve compatible with biological growth (Weiskittel et al., 2011; Burkhart and Tomé, 2012).

Individual-tree based models are desirable when the detailed information about stand development and structure is required. They also allow flexibility to model co-occurring species, stand structures and different treatments. Individual-tree based models help to envision the dynamics of a stand by size and age distribution classes. This important feature allows forest manager to make predictions, formulate management policy and management prescriptions more reliable. Individual tree models could be distance dependent and distance independent (Weiskittel et al., 2011; Burkhart and Tomé, 2012). Models that takes into account of the spatial arrangement of individual trees are called distance dependent model otherwise distance independent models. Individual tree models are usually distance independent model. In individual tree models, present tree level and stand-level attributes are taken as an explanatory variable to predict the response variable at present as well as in the future. Individual trees grown using individual tree data are aggregated into stand level variables to develop stand-level information in the model (Waykoff, 1990; Monserud and Sterba, 1996; Lynch et al., 1999; Budhathoki et al., 2008b). So, the model provides estimates at individual tree level and also at stand level. However, the stand-level growth model cannot estimate individual tree level information. Therefore, individual tree models are useful for prediction of tree-level as well as stand-level information.

Growth and yield models are important for both natural and plantation forests to predict present as well as future productivity of an individual tree and as well as forest conditions. These models could be developed for either even-aged and uneven aged natural stands or thinned, and unthinned plantation stands (Davis et al., 2001, Weiskittel et al., 2011). Depending on the type and nature of the silvicultural treatment applied to the stand, the models could show either additive or multiplicative effect on productivity (Weiskittel et al., 2011). Growth models are developed from repeated measurements of the permanently established plots because they provide the information on a tree growth of long period and longitudinal aspect of the inventory data (Grégoire et al., 1995). These long-term data are useful sources for understanding the growth pattern, survivorship, and possible effect of the applied treatments on development of an individual tree and stands.

A complete growth model system is composed of different modelling components such as height-diameter model, crown ratio relationship model, basal area growth or diameter increment model, and mortality model. A forest growth simulator is one of the tools that assists in decision-making for developing forest management plans and undertaking different silvicultural systems. Annual tree growth is based on size, vigor, age and competitive stress experienced. Competition from neighboring trees may affect growth of the subject tree. So, individual tree level and stand level covariates are used as explanatory variables while modelling annual growth rate (Lynch et al., 1999; Budhathoki et al., 2008b; Sharma and Subedi, 2011). In conjunction to stand level covariates, the distance independent growth model assesses competition by using indices based on the size distribution of a tree within a given area (Waykoff, 1990; Monserud and Sterba, 1996; Temsegen et al., 2007), relative indices (Sharma and Patron 2007, Burkhart and Tomé, 2012) and relative spacing (Ducey, 2009; Zhao et al., 2012) are also often significant covariates in the growth modelling system.

Height, dbh, and crown ratio provide adequate information of a trees' growing condition. Individual tree height is one of the important variables used in the growth and yield modeling system. Height-diameter relationship models are often used to impute the missing heights observations because only the subsample of trees' height are recorded in the field measurement. It is commonly used in estimating current and future volume production, crown ratio and developing mortality models and predicting future heights (Lynch et al., 1991; Lynch et al., 1995; Garber et al., 2006). Similarly, crown ratio, the ratio of live crown length to total height also reflects the competition experienced by an individual tree. The crown ratio is also an important tree attribute that directly and greatly influences tapper, basal area increment, volume estimation of an induvial tree and survival (Monserud and Sterba, 1996; Jian et al., 2007).

Growth models that consider growth and mortality rates in the same period are slightly better than those which used the constant growth rate approach (Crecente-Campo et al., 2010). However, such models are usually complex and challenging to implement. Models that easily accounts for morality rate and irregular interval of re-measurement are more desirable. So, mortality models are commonly developed using the logistic or logit cumulative distribution function (Woodall et al., 2005; Zhao et al., 2006; Crecente-Campo et al., 2010). Tree mortality rate is usually interpreted to a common length of either 1-year or a 5-year growth period. Tree mortality could be irregular (insects, diseases, fire, and snow damage) or regular (competition for light, moisture, and nutrients) (Groom et al., 2012). Irregular tree mortality events are episodic and difficult to predict, but regular morality is predictable. However, predicting individual tree mortality is highly variable because it depends on the nature of the data collected to model over the length of the study period.

An individual tree's growth depends on several fundamental factors regardless of any silvicultural treatments. Tree stem increase at breast height may be expressed as basal area growth or diameter increment and can be used to simulate the effects of intensive management scenarios. Annual basal area growth of individual trees is influenced by site productivity, stand characteristics, and climate experienced by a tree over the period (Pokharel and Forese, 2009). The variation in the monthly temperature and precipitation of the growing seasons also affect the annual diameter or radial growth rate of a tree species (Pilcher and Gray, 1982). The extreme climatic events i.e. temperature and precipitation are significantly correlated with the annual growth rate (Graumlich, 1993). Therefore, tree response to climatic stress over an extended period significantly affects forest stand productivity (Subedi and Sharma 2013). It is easier to detect such changes at large scale but hard to decode at finer scales (Boisvenue and Running, 2006).

Besides the influence of climate, thinning and the number of trees stocked at the time of measurement also affect the individual productivity of residual trees. Thining is a desirable treatment to maintain stand productivity because it maintains a healthy balance between individual trees position and enhances radial growth. Forest management requires a growth model that can handle the management or treatment effects such as thinning to address variability in individual tree growth models (Crecente-Campo et al., 2009, Turcotte et al. 2012). Thinning reduces stand level competition and often has a pronounced effect on the increase of the stand's quadratic mean diameter and on overall tree biomass (Zhang et al. 1997; Cain and Shelton 2003; Zhang et al. 2006; Crecente-Campo et al., 2010; Uzoh and Mori 2012). However, the timing of response to the thinning effects varies among the tree species (Cutter et al. 1991). However, the thinning methods i.e. thinning from above (crown thinning) and or thinning from below (low thinning) in the control stock also influence the volume, basal area and biomass growth of a tree. In general, great growth in a stand is achieved greater in thinning from above than from below thinning in the control experimental (Bradford and Palik, 2009) but volume growth from below thinning is strongly site dependent (Skovsgarrd, 2009).

Shortleaf pine (*Pinus echinata* Mill.) is second to loblolly pine (*Pinus taeda* L.) in terms of volume of the southern pines in the United States. It has the widest adaptation of the range of geography, soils, topography and habitats, covering 22 states over more than 1,139,600 km^2 (440,000 mi^2), ranging from southeastern New York to eastern Texas. The most prominent natural shortleaf communities are found in the Ouachita Highlands (Guldin, 1986), which ranges in elevation from almost sea level to 1,006 m (3,300 ft). Shortleaf pine naturally grows well in the areas having an average air temperature from 9^0 to 21^0C, with minimums of -30^0C and maximums of 39^0C and rainfall ranging from 101 to 140 inches per year (Williston and Balmer, 1980). Naturally regenerated shortleaf pine forests of the Ouachita Mountains of Arkansas and Oklahoma cover almost 12,950 km^2 (5000 mi^2) (Smith, 1986), where it is an important timber species.

In a long-term study, with an objective to understand the growth and yield system of shortleaf pine, over 200 permanent plots were established on the Ozark and Ouachita National Forest during the period from 1985-1987. The re-measured data has served many studies on growth and yield of naturally occurring shortleaf pine. These include: Lynch et al. (1991), Murphy et al. (1992), Lynch and Murphy (1995), Lynch et al. (1999), Budhathoki et al. (2008a; 2008b). Though previous studies dealt with all essential components such as height prediction, missing height, crown ratio estimation, basal area growth and mortality, all these studies were conducted with fewer re-measurements data i.e. Lynch et al. (1999) with two measurements and Budhathoki et al. (2008a; 2008b) with three measurements.

Budhathoki et al. (2008a; 2008b) used the advantages of a mixed-effects model for estimating parameters of annual basal area growth of an individual tree and height-diameter relationship over the least square methods. The advantage of the mixed-effects models is that we can use the random effect to account for factor level covariate effects in this data structure.

Previous studies based on subsets of the current dataset did not fully account for temporal correlation among measurement periods while modelling growth and yield components using the mixed-effects approach. The thinning that occurred after the third measurement was also not analyzed to understand its' effect on the growth and mortality. Modelling mortality is a crucial part of growth and yield component because it changes the stand structure and influences the stand productivity by reducing resource competition between individuals. Over the period of 25 years, it is expected that climate also influenced individual basal area growth. Besides height and dbh, the crown ratio is also an important tree attribute that directly and greatly influences taper and volume estimation of an induvial tree. Six measurements on these plots spanning a period of, 25 years provided an opportunity to update the existing models with improved parameter estimates for better prediction. Additionally, it served as the basis for evaluating the different model forms of the growth and yield component and utilize the statistical tools to account heterogeneity and correlation understanding of growth and yield system of shortleaf pine. It was important to develop an individual basal area growth model, the height-dbh model for missing height, crown ratio model and mortality model simultaneously to make a complete and updated growth and yield system.

Objectives

This dissertation aims to update, improve, modify and incorporate the influence of thinning effect on the growth and yield modelling system of naturally occurring shortleaf pine stand with data from plots measured six times spanning a period of 25 years. Specifically, the objectives are:

1) Improve and modify the height-dbh relationship and crown ratio estimation models
2) Develop annual mortality models with and without thinning effect
3) Improve annual basal area growth models and develop a climate-based growth model.

References

Bradford, J.B., and Palik, B.J. 2009. A comparison of thinning methods in red pine: consequences for stand-level growth and tree diameter. Can. J. For. Res. 39:489-496.

Boisvenue, C., and Running, S.W. 2006. Impacts of climate change on natural forest productivity – evidence since the middle of the 20th century. Glob. Change. Biol. 12:862-882.

Budhathoki, C.B., Lynch, T.B., and Guldin, J.M. 2008a. A Mixed-Effects Model for the dbh–Height Relationship of Shortleaf Pine (*Pinus echinata* Mill.). South. J. Appl. For. 32:5-11.

Budhathoki, C.B., Lynch, T.B., and Guldin., J.M. 2008b. Nonlinear mixed modeling of basal area growth for shortleaf pine. For. Ecol. Manage. 255:3440-3446.

Burkhart, H.E., and Tomé, M. 2012. Modeling forest trees and stands. Springer Science & Business Media.

Cain, M. D., and Shelton, M.G. 2003. Effects of Alternative Thinning Regimes and Prescribed Burning in Natural, Even-Aged Loblolly-Shortleaf Pine Stands: 25 Year Results. South. J. Appl. For. 27:18-29.

Crecente-Campo, F., Pommerening, A., and Rodríguez-Soalleiro, R. 2009. Impacts of thinning on structure, growth and risk of crown fire in a *Pinus sylvestris* L. plantation in northern Spain. For. Ecol. Manage. 257:1945-1954.

Crecente-Campo, F., Soares, P.,Tomé, M., and Diéguez-Aranda, U. 2010. Modelling annual individual-tree growth and mortality of Scots pine with data obtained at irregular measurement intervals and containing missing observations. For. Ecol. Manage. 260:1965-1974.

Cutter, B. E., Lowell, K.E., and Dwyer, J.P. 1991. Thinning effects on diameter growth in black and scarlet oak as shown by tree ring analyses. For. Ecol. Manage. 43:1-13

Davis, L.S., Johnson, K.N., Bettinger, P.S., and Howard. T.E. 2001. Forest Management, 4th Edition. McGraw-Hill, New York. 804 p.

Ducey, M.J. 2009. Predicting Crown Size and Shape from Simple Stand Variables. J. Sustain. For. 28:5-21.

Jiang, L., Brooks, J.R. and Hobbs, G.R. 2007. Using Crown Ratio in Yellow-Poplar Compatible Taper and Volume Equations. North. J. Appl. For. 24:271-275.

Graumlich, L.J. 1993. Response of tree growth to climatic variation in the mixed conifer and deciduous forests of the upper Great Lakes region. Can. J. For. Res. 23: 133-143.

Gregoire, T.G., Schabenberger, O., and Barrett, J. 1995. Linear Modelling of Irregularly Spaced, Unbalanced, Longitudinal Data from Permanent-Plot Measurements. Can. J. For. Res. 25:137–156.

Groom, J.D., Hann, D.W., and Temesgen, H. 2012. Evaluation of mixed-effects models for predicting Douglas-fir mortality. For. Ecol. Manage. 276:139-145.

Guldin, J.M. 1986. Ecology of Shortleaf Pine. In: Proc. Symp. on the Shortleaf Pine Ecosystem.Arkansas Coop. Ext. Ser., University of Arkansas, AR, 25 – 40.

Lynch, T.B., Murphy, P.A., and Lawson, E.R. 1991. Stand volume equations for managed natural even-aged shortleaf pine in eastern Oklahoma and western Arkansas. Research report P-Oklahoma Agricultural Experiment Station.

Lynch, T.B. and Murphy, P.A. 1995 A Compatible Height Prediction and Projection System for Individual Trees in Natural, Even-Aged Shortleaf Pine Stands. For. Sci. 41:194-209.

Lynch, T.B., Hitch, K.L., Huebschmann, M.M., and Murphy, P.A., 1999. An Individual-Tree Growth and Yield Prediction System for Even-Aged Natural Shortleaf Pine Forests. South. J. Appl. For. 23:203–211.

Murphy, P.A., Lawson, E.R., and Lynch, T.B. 1992. Basal area and volume development of natural even-aged shortleaf pine stands in the Ouachita Mountains. South J Appl. For. 16:30-34.

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

Pilcher, J.R. and Gray, B. 1982. The Relationships between Oak Tree Growth and Climate in Britain. J. Ecol. 70:297-304.

Pokharel, B., and Froese, R.E. 2009. Representing site productivity in the basal area increment model for FVS-Ontario. For. Ecol. Manage. 258: 657-666.

Skovsgaard, J.P. 2009. Analysing effects of thinning on stand volume growth in relation to site conditions: A case study for even-aged Sitka spruce (*Picea sitchensis* (Bong.) Carr.). Forestry. 82:87-104.

Smith, K.L. 1986. Historical Perspective. In: Proc. Symp. on the Shortleaf Pine Ecosystem. Arkansas Coop. Ext. Ser., University of Arkansas, AR, 1–8.

Subedi, N., and Sharma, M. 2011. Individual-tree diameter growth models for black spruce and jack pine plantations in northern Ontario. For. Ecol. Manage. 261: 2140-2148.

Temesgen, H., Hann, D.W., and Monleon, V.J. 2007 Regional Height– Diameter Equations for Major Tree Species of Southwest Oregon. West. J. Appl. For. 22:213-219.

Turcotte, R.M., Elliott, T.R., Fajvan, M.A., Park, Y.L, Snider, D.A., and Tobin, P.C. 2012. Effects of Ice Storm Damage on Hardwood Survival and Growth in Ohio. North. J. Appl. For. 29:53-59.

Uzoh, F.C.C., and Mori, S.R. 2012. Applying survival analysis to managed even-aged stands of ponderosa pine for assessment of tree mortality in the western United States. For. Ecol. Manage. 285:101-122.

Weiskittel, A.R., Hann, D.W., Kershaw Jr, J.A., and Vanclay, J.K. 2011. Forest growth and yield modeling. Wiley.

Woodall, C.W., Grambsch, P.L., and Thomas, W. 2005. Applying survival analysis to a large-scale forest inventory for assessment of tree mortality in Minnesota. Ecol. Model. 189:199-208.

Wykoff, W.R. 1990. A Basal Area Increment Model for Individual Conifers in the Northern Rocky Mountains. For. Sci. 36:1077-1104.

Zhang, S., Burkhart, H.E., and Amateis, R.L. 1997. The Influence of Thinning on Tree Height and Diameter Relationships in Loblolly Pine Plantations. South. J. Appl. For. 21:199-205.

Zhang, S.Y.,Chauret, G., Swift, D.E., and Duchesne, I. 2006. Effects of precommercial thinning on tree growth and lumber quality in a jack pine stand in New Brunswick, Canada. Canadian Journal of Forest Research 36:945-952.

Zhao, D., Borders, B., Wilson, M., and Rathbun, S.L. 2006. Modeling neighborhood effects on the growth and survival of individual trees in a natural temperate species-rich forest. Ecol. Model. 196:90-10.

Zhao, D., Kane, M., and Borders, B.E. 2012. Crown Ratio and Relative Spacing Relationships for Loblolly Pine Plantations. Open Journal of Forestry, 2:110-115.

CHAPTER II

USING QUADRATIC MEAN DIAMETER AND RELATIVE SPACING INDEX TO ENHANCE HEIGHT-DIAMETER AND CROWN RATIO MODELS FITTED TO LONGITUDINAL DATA

Abstract

The inclusion of quadratic mean diameter (*QMD*) and relative spacing index (*RSI*) substantially improved the predictive capacity of height-diameter at breast height (dbh) and crown ratio models (*CR*), respectively. Data were obtained from 208 permanent plots established in western Arkansas and eastern Oklahoma during 1985-1987 and remeasured for the sixth time (2012-14). Existing height-dbh, and *CR* estimation models for naturally-occurring shortleaf pine forests (*Pinus echinata* Mill.) were updated and modified for improved performance. Additionally, eight height-dbh relationship models that use only dbh (fundamental local models) were modified using covariates. The model performance was evaluated using fit statistics (root mean square error (*RMSE*), *Fit index* and Akaike information criteria (*AIC*)). The results showed that the best model form which was an extended nonlinear model (ENM) with autoregressive first order AR (1) structure and power variance function performed better than extended mixed-effects models (EMEMs) and predicted well as an ordinary least squares (OLS) nonlinear model. The autocorrelation within individual trees was larger for the height-dbh relationship than for *CR* estimation. The addition of *QMD*, to mean dominant height (H_D) improved height-dbh relationship with a reduction of 8% in *RMSE* greatly, compared to the use of basal area per hectare (*BAH).* Similarly, multiplying a fundamental local model by *QMD* raised to a parametric power reduced *RMSE* by 16%, and improved *Fit index* by 12%, and decreased the *AIC* value by 7%. Dbh, H_D and *RSI* best explained the crown ratio relationship with an improved *Fit index* by 6.7% compared to alternative nonlinear models without *RSI*. The logistic model for *CR* also provided prediction accuracy similar to that of a commonly-used nonlinear model. Nonlinear model with an application of remedial measures to enhance adherence to modeling assumptions can provide better parameter estimates over mixed-effects modeling approach.

Keywords: mixed-effects model; autocorrelation; height-dbh relationship; crown ratio; quadratic mean diameter; relative spacing index.

This article has been submitted to Forestry An International Journal of Forest Research for publication and it's under major revision status.

Introduction

Two fundamental mensurational quantities in forest inventory, i.e. height and diameter, are frequently used to characterize forest productivity in forest growth and yield models. Diameter measurement at breast height (dbh) is less costly and requires less effort than the total height measurement. Therefore, height-dbh relationship models have been used in the growth and yield models to predict “missing” tree heights (Lynch et al., 1995), to predict future tree heights (Lynch et al., 1995; Lynch et al., 1999), and also to impute height to estimate volume production (Garber, 2006). The height-dbh relationship is nonlinear from the biological perspective with a curve that is asymptotic to a maximum possible total height at upper ranges of diameter. So, various nonlinear models with respect to parameters have been proposed to model the height-dbh relationship of different tree species. The height-dbh relationship has sometimes been modeled using only dbh as single independent variable (Meyer, 1940; Richards, 1959; Burkhart and Strub, 1974; Stage, 1975; Bates and Watts, 1980; Wykoff et al., 1982; Ratkowsky, 1990; Schmidt et al., 2011; Sharma and Briedenbach, 2015). Here we term this type of model a “fundamental local model” because these models are: primarily developed at local or at a regional level, specific to a tree species and site, developed when stand level covariate or competition index is difficult or inconvenient to obtain, and can be easily extended to incorporate additional covariates.

Height-dbh relationships are also modelled using plot or stand level covariates (Lynch et al., 1999; Sharma and Patron, 2007; Budathoki et al., 2008; Temesgen et al., 2007; Temesgen et al., 2008; Arcangeli et al., 2014; Temesgen et al., 2014; Sharma and Briedenbach, 2015) to demonstrate the influence of stand density or effect of competition. Therefore, fundamental local models are expected to have large error variance (Haung et al., 1992; Fang and Bailey, 1998; Sharma and Briedenbach, 2015) compared to models that use covariates in addition to dbh such as: stand structures (dominant height) (Lappi, 1997; Lynch et al., 1999; Sharma and Briedenbach, 2015), relative dimensions at stand level (ratio of tree per hectare to basal area per hectare) (Sharma and Patron, 2007), competition index (basal area in larger trees (*BAL*)) (Temesgen et al., 2007), and stand density (basal area per hectare (*BAH*)) (Sharma and Zang, 2004; Budathoki et al., 2008).

Individual tree height (H_i) and dbh (D_i) can be used as independent variables with other plot or stand level attributes to model another important tree characteristic, the 'crown ratio' (*CR*) of an individual tree which is the ratio of live crown length to the total height. The crown ratio is an important measure of tree vigor that reflects competition experienced by an individual tree because stand density over the period reduces the crown length (Smith et al., 1992; Hynyen, 1995; Temesgen et al., 2005). The distance between trees determines the crown shape and size, which is related to the crown length, total height and diameter increment attained by an individual tree (Smith et al., 1992; Monserud and Sterba, 1996). In many growth and yield models, *CR* is used for improved prediction of forest attributes. For example, height increment (Daniels and Burkhart, 1975), basal area increment (Wykoff, 1992; Moserud and Sterba, 1996; Leites et al., 2009), taper and volume of individual trees (Valentini and Cao, 1986; Jiang et al., 2007; Jiang and Liu, 2011), and survival of an individual tree (Saud et al., 2015). It is also useful in estimating crown biomass for energy production (Tahvanainen and Fross, 2006). The crown ratio has been modeled using either a variety of nonlinear functions

(Holdaway, 1986; Dyer and Burkhart, 1987; Lynch et al., 1999) or the logistic function (Hasenauer and Monserud, 1996; Temesgen et al., 2005) which is also a nonlinear function. Both approaches utilize model forms that restrict crown ratio predictions to the feasible range of 0 to 1.

However, the height-dbh relationship can vary over time due to differences in stand age, productivity, and competition (Lappi, 1997; Peng et al., 2001; Sharma and Patron, 2007; Budhathoki et al., 2008) and also differences in a geographical region (Calama and Montero, 2003; Arcangeli et al., 2014). This also applies in the case of *CR* modeling. Such variation could be reduced by using distance independent variables as covariates. Relative dimensions (ratios) are distance independent indices, which measure the hierarchical position of the subject tree within plot, e.g. ratio of D_i to max dbh; ratio of quadratic mean diameter (*QMD*) to dbh (*RAQD*), *BAL*, and crown competition factor in larger trees (*CCFL*). These dimensionless ratios when used as covariates help to assure better prediction and tend to make the fundamental relationships among tree components more stable (Burkhart and Tomé, 2012, pg. 202). Ducey (2009) and Zhao et al. (2009) demonstrated that the relative spacing index (*RSI*) accounted for the effects of space between individual trees on crown ratio more effectively than other covariates. Interestingly, in addition to *RSI*, Ducey (2009) also suggested the inverse of *RAQD* as an important variable in height prediction and *CR* estimation. Temesegen et al. (2005, 2007, 2008, and 2014) suggested that distance independent variables including *BAL* and *CCFL* also improve fits for tree height-dbh and crown relationships.

Height-dbh relationship and *CR* models are often developed using mensurational records of permanent plots from repeated measurements. The tree attributes such as height, dbh, and crown length measured at different time intervals are auto-correlated and also exhibit heterogeneous errors either at a tree or stand level. As a result, the use of nonlinear ordinary least square (OLS) estimation is often not reliable, because an assumption of random samples and independent observations is violated and the presence of autocorrelation does not conform to the assumptions of OLS. Therefore, mixed-effects modeling as an alternative to OLS for repeated measurements and grouped data has been widely used in forestry growth and yield models (Lappi, 1991; Lynch et al., 2005; Budhathoki et al., 2008; Lynch et al., 2012; Temesgen et al., 2014). This helps to address a possible source of subject-specific variation that the OLS approach does not consider because the fixed-effect parameters represent population average responses while random effects parameters represent response specific to each sampling unit (Lappi, 1991; Lynch et al., 2005). One advantage of the mixed-effects model is that random effects can be calibrated for an unsampled location (new data outside the estimation data) to improve the predictive accuracy of the parameter estimates from a mixed-effects model (Lappi, 1991; Peng et al., 2001; Lynch et al., 2005; Sharma and Patron, 2007; Lynch et al., 2012).

Mixed-effects models or hierarchical mixed-effects models easily account for spatial autocorrelation by using a plot specific or group specific random effects, but not so for temporal autocorrelation within observation. However, shorter study time period (lag) has been a limiting factor in the analysis of longitudinal data in growth modeling that sometimes does not allow us to specify appropriate autocorrelation structures. Patterns of heterogeneous errors can often be associated with the covariates (Phinero and Bates, 2000). Many investigators have used graphical methods to investigate the assumption of constant error variance for the dbh-height relationship model (Fang and

Bailey, 1998; Peng et al., 2001; Sharma and Parton, 2007; Budhathoki et al., 2008). But perhaps few have done a formal statistical test on this issue for nonlinear model forms (Temesgen et al., 2007; Lynch et al., 2012). Others have used weighted regression or logarithmic transformations that tend to stabilize variance (Huang et al., 1992; Fang and Bailey, 1998; Temesgen et al., 2007; Temesgen et al., 2014).

Shortleaf pine (*Pinus echinata* Mill.) forests contain standing volume in the southern USA second only to loblolly pine (*Pinus taeda* L.) among the four southern pines (Lawson, 1990). However, relatively few quantitative studies of the height-dbh relationship, CR or other aspects of growth and yield o natural stands of shortleaf pine have been published compared to other southern pines in the USA (Budhathoki et al., 2008). Graney and Burkhart (1973) provided a polymorphic system of site index curves to estimate dominant stand height for shortleaf pine using nonlinear ordinary least square (OLS) method. The relationship between height-dbh for naturally occurring even-aged stands of shortleaf pine was fitted using seemingly unrelated regression by Lynch and Murphy (1995), nonlinear OLS by Lynch et al. (1999), and mixed-effects estimation by Budhathoki et al. (2008). The studies of Lynch et al. (1999) for height-dbh relationship and CR estimation, and of Budhathoki et al. (2008) for height-dbh relationship used only the first two and three measurements respectively of the Oklahoma State University (OSU) and USDA Forest Service Southern Research Station (USFS) naturally occurring shortleaf pine growth study.

Although, Budhathoki et al. (2008) updated height-dbh relationship by adding basal area per hectare (BAH) as an independent variable and by fitting a mixed-effects model, they did not fit a crown ratio model. But now six measurements of shortleaf plot data spanning a 25-year period is available that allows us to update existing models and incorporate the most recent developments in modeling the height-dbh and crown relationships. This also provides an opportunity to test other model forms and independent covariates that may provide improved fits to the data from other studies as well. Therefore, the current study aims to update and improve the existing height prediction and CR estimation models suitable for practical application by resolving the issues of autocorrelation and heteroscedasticity of errors in repeated measurements. The specific aims are: to modify and improve the model performance by introducing plot or stand level covariates; to evaluate the performance of different sets of OLS nonlinear models, mixed-effects models, and their extended forms while correcting for autocorrelation and stabilizing heterogeneity of errors. In addition to this, we will evaluate available fundamental local models for the height-dbh prediction that do not use dominant height and modify them with plot or stand level covariates to minimize the prediction error. Moreover, also we will test homoscedasticity assumptions independently for each model. It is expected that the best model for both the height prediction and the crown ratio estimation will be useful for estimating volume, biomass and other tree attributes of the natural stand of shortleaf pine.

Materials and methods

Data

In 1985-87, the Department of Forestry (now part of the Department of Natural Resource Ecology and Management) at Oklahoma State University (OSU) and the United States Forest Service (USFS) at Monticello, Arkansas collaboratively established growth and yield plots in even-aged natural shortleaf pine stands. The plots were established as permanent plots in the Ozark and Ouachita National Forests in western Arkansas and southeastern Oklahoma. Prior to the establishment of this study in 1985, the major sources of data for shortleaf growth and yield were from fully stocked plots or unmanaged shortleaf pine stands (Lynch et al., 1999). These plots were designed to represent a range of ages, basal area levels, and site qualities, which were designated as design variables so that plots were thinned to specific residual densities at their establishment (for a detailed description see Lynch et al., 1999). At plot establishment, woody understory vegetation stems with dbh greater than 2.54 centimeters (cm) in were controlled using herbicide.

These plots have been re-measured in every 4 to 6 years, with the latest (sixth) measurement made during the period from 2012 to 2014. At each measurement, D_i of all trees from plot were measured but H_i in meters (m), and crown length (height to base of live crown) in m were recorded for selected subsample trees from each plot to represent the range of tree diameters and crown classes of dominant, co-dominant and intermediate trees in the plot. Individual trees were classified as the dominant and codominant tree based on the definitions given by Avery and Burkhart (2002 p. 163). The number of height and crown length subsample trees within plots was increased after the first measurement to continue a good representation of samples within plot dbh classes. The measurement plots were circular with a radius of 17.4 m (57.2 feet) and area of 0.0809 hectare (ha) (0.2 acres). The measurement plots were surrounded by a buffer strip 10 m (33 feet) wide that received the same silvicultural treatments as plots at the establishment. The total sample consisted of 208 plots (for details see Lynch et al., 1999). Plot age

(*PAG*) was obtained by averaging the ages of dominant and codominant trees, which was determined from the ring counts of increment cores.

Ice damaged trees occurred on a total of 101 plots in the year 2002, just before the fourth measurement. It was expected that the plots with significant numbers of ice damaged trees could influence the growth characteristics of individual trees on these plots. Therefore, in the model development process, plots with more than 30% of ice damage trees were excluded. Any individual trees having ice damage was also removed from the model development process. Trees having forks or other significant defects were excluded from the model development dataset. Many plots were re-thinned to their original basal area levels just after the third measurement while some plots were left unthinned. An average thinned (removed) basal area was 6.94 m^2ha^{-1} with a range of 0.69 – 19.36 m^2ha^{-1}. A variable that exhibits simple thinning effect ‘THINHA’ = (Thinned basal area per hectare/ (years since thinning) was formulated assuming that thinning effect decreases over the time. The mean and standard deviation (SD) of all variables from the first measurement to the last (sixth) measurement are shown in Table 1. The data consisted of total 14,028 observations, and the summary statistics of the variables used in modeling height prediction and crown ratio estimation are presented in Table 2. The H_i ranged from 3.048 m to 38.100 m with a mean of 20.433 m and SD of 6.222 m and *CR* ranged from 0.055 to 0.80 with a mean of 0.373 and SD of 0.094 (Table 2). The data used for modeling height-dbh relationship and *CR* estimation is shown in Figure 1a and 1b.

Height prediction model with dominant height

To predict individual tree height, Lynch and Murphy (1995) developed a compatible height prediction model which was used in the shortleaf pine growth prediction system described in Lynch et al. (1999). Equation (Eq.) 1 below was used to predict an individual shortleaf pine tree height. Budhathoki et al. (2008) accounted for competition effects of other trees on individual tree growth by including the variable *BAH* (Eq. 2).

$$(H_i - c) = b_0 (H_D - c)^{b_1} \exp(b_2 D_i^{b_3}) \tag{1}$$

$$(H_i - c) = b_0 (H_D - c)^{b_1} \exp(b_2 D_i^{b_3} + b_4 BAH) \tag{2}$$

where H_D = average plot dominant and codominant height; c = the breast height (1.371 m) at which dbh is measured; and b_0, b_1 b_2 ,b_3 and b_4 are parameters to be estimated;

The existing models need to be tested for thinning effect since post-thinning measurements are now available. The simple effect of thinning "*THINHA*" on height prediction model (Eq. 1) was found not significant in the model. Therefore, we tested inclusion of covariates that represents density competition and the relative position of a tree. Both covariates, *QMD* and *RAQD* provided the substantial improvement in fit statistics and showed identical performance, but we preferred *QMD* in (Eq. 3) to avoid correlated covariates in a model, which will be discussed below. A variable that is a function of an inverse of *QMD*, ratio of the number of trees per hectare (*TPH*) to *BAH* (*TPH*/*BAH*) was also used by Sharma and Parton (2007) in the extended model of the

Chapman-Richards function for boreal tree species in Ontario, Canada. The model used by Sharma and Parton (2007) was also modified by using *QMD* variable, instead TPH/BAH in equation (4) for testing on our shortleaf pine dataset.

$$(H_i - c) = b_0 (H_D - c)^{b_1} \exp(b_2 D_i^{b_3} QMD) \tag{3}$$

$$(H_i - c) = b_0 (H_D)^{b_1} (1 - \exp^{-b_2 QMD^{b_3} D_i})^{b_4} \tag{4}$$

After testing several models using OLS, Eq. (3) was selected for fitting with the mixed- effect model approach because it provided root mean square error (*RMSE*) extremely close to the smallest *RMSE* from Eq. (4), but contains only four parameters, making it somewhat simpler in form than Eq. (4). All parameters were tested for possible inclusion of plot-level random effects assuming the same plot effect holds for all remeasurements of the same plot. A plot level random effect associated with the b_0 (asymptotic height), b_1 (slope), and b_2 (curvature) were significant, but due to sizeable differences in Akaike information criteria (*AIC*) value, the random effect associated with the b_3 the parameter multiplicative to D_i in Eq. (3) was selected. This resulted in the mixed-effects model Eq. (5)

$$(H_{ij} - c) = b_0 (H_{Dj} - c)^{b_1} \exp\{b_2 D_{ij}^{(b_3 + u_j)} QMD_j\} + \varepsilon_{ij} \tag{5}$$

where *i* represents attribute of an individual tree and *j* represent attribute of the plot; b_0, b_1, b_2, and b_3 are fixed effects parameters; u_j = random effect associated with parameter b_3 and specific to plot *j* ; and ε_{ij} = within plot error (random error for tree *i* in plot *j*).

Since each tree from the same plot was repeatedly measured for six times, it is possible that autocorrelation within individual observations and non-constant error variance exists. Therefore, to resolve the issues of heteroscedasticity and autocorrelation, we modeled a power variance function and first order autoregressive AR (1) structure, and a combination of both in the best nonlinear and mixed-effects models (Pinheiro and Bates, 2000, pg. 391; Cryer and Chan, 2008, pg. 66). AR (1) structure was selected because the data relatively few lags (5) and AR (1 requires the estimation of only one parameter. The best selected nonlinear models (Eq. 3) and mixed-effects model (Eq. 5) were referred as "base model" forms. Hereafter, the extended model forms with additional power variance function and AR (1) were termed "extended nonlinear models" (ENMs) and "extended mixed-effects models" (EMEMs). The assumption of autocorrelation within an individual tree was considered for ENMs, but this assumption was not compatible while modeling mixed-effects models. This may be due to the differences in the hierarchy of the group at which the errors are correlated, and random effects are associated. So, we assumed autocorrelation within plot for mixed-effects models. The resulted ENMs are as:

Model 3 + power variance function (6)

$$\mathrm{var}(\varepsilon_{ij}) = \sigma^2 \, |\upsilon_{ij}|^{2\delta}$$

error variance (ε_{ij}) was modeled with one covariate. υ_{ij} is covariate and δ is power parameter. H_D was selected as covariate for modeling heterogenouses errors because of smaller *AIC* value and large Likelihood Ratio (*LR*) statistics.

Model 3 + AR (1) (7)

AR (1) models the correlated errors (ε_{it}) as:

$$\varepsilon_{it} = \phi_1 \varepsilon_{it-1} + \omega_{it} \text{ and } \omega_{it} \sim iid \text{ N}(0, \sigma_{\omega}^2)$$

where i is the individual tree and t is measurement (lag) and φ is the autocorrelation between lags. The correlation "rho" (ρ) between residuals of an observation pair declines exponentially with the number of periods (k) apart i.e. $\rho = \phi^k$.

Model 3 + power variance function + AR (1) (8)

The resulted EMEMs are as:

Model 5+ power variance function (9)

Model 5 + AR (1) (10)

Mode 5 + power variance function + AR (1) (11)

Height prediction model without dominant height

A variety of height prediction models with only single independent variable 'dbh' are commonly used in practice when dominant height is not readily available (in Table 3, Eq. 12-19). These eight equations (Eq. 20-27) were fitted to predict the shortleaf pine tree height and tested for modification by including *QMD*, which is typically available from forest inventory data. Other stand-level variables including *TPH*, *BAH*, and *RAQD* were also tested. We found inclusion of *QMD* into these modified "local" models (Eq. 20-27) substantially improved model performance. But, these models were neither tested as mixed-effects model nor as extended models.

The individual tree crown ratio model used by Lynch et al. (1999) for shortleaf pine was also modified and tested. This crown ratio function was developed by Dyer and Burkhart (1987) for planted loblolly pine tree data and also used by Hynynen (1995) for Scots pine stands. The base Eq. (28) used to predict the current individual tree shortleaf pine crown ratio together with an alternative Eq. (29) are shown below;

$$CR_i = 1 - \exp\left[-\left(b_0 + \frac{b_1}{PAG} \right) \left(\frac{D_i}{H_i} \right)^{b_2} \right] \qquad (28)$$

$$CR_i = 1 - \exp\left[-\left(b_0 + \frac{b_1}{H_i} \right) \left(\frac{D_i}{H_{Di}} \right)^{b_2} \right] \qquad (29)$$

where CR_i is the crown ratio of tree i; b_0, b_1 and b_2 are parameters to be estimated; and other variables are as defined above.

The effect of thinning was not included in crown ratio estimation models because though the effect was found significant in the Eq. (28), it didn't markedly reduce the mean square error. Eq. (28) was modified to Eq. (29) by replacing *PAG* with H_D. Both equations were further modified to Eq. (30) and Eq. (31) by adding *BAH*. The Eq. (30) was then modified by using *RSI* in Eq. (32). Relative spacing index (*RSI*) was calculated as $= \left(\sqrt{10{,}000 / TPH} \right) / H_D$.

$$CR_i = 1 - \exp\left[-\left(b_0 + \frac{b_1}{PAG} + BAH^{b_2} \right) \left(\frac{D_i}{H_i} \right)^{b_3} \right] \qquad (30)$$

$$CR_i = 1 - \exp\left[-\left(b_0 + \frac{b_1}{H_i} \right) \left(BAH^{b_2} + \frac{D_i}{H_D} \right)^{b_3} \right] \tag{31}$$

$$CR_i = 1 - \exp\left[-\left(b_0 + \frac{b_1}{H_i} \right) \left(RSI^{b_2} + \frac{D_i}{H_D} \right)^{b_3} \right] \tag{32}$$

The logistic function approach to crown ratio estimation was also tested. The following model (Eq. 33) proved to be a better alternative to the exponential model approach given above.

$$CR_i = \left[1 + \exp\left(-\left(b_0 + b_1 H_i + b_2 \frac{D_i}{H_D} + b_3 BAH \right) \right) \right]^{-1} \tag{33}$$

The mixed-effects approach was also used to fit crown ratio estimation models. Based on performance, Eq. (31) and (32) were modeled with the mixed-effects approach. It was found that the random effect associated with the parameter b_0 for both models as shown below performed better than the other alternatives:

$$CR_{ij} = 1 - \exp\left[-\left((b_0 + \mu_j) + \frac{b_1}{H_i} \right) \left(BAH_j^{b_2} + \frac{D_i}{H_{D_j}} \right)^{b_3} \right] + \varepsilon_{ij} \tag{34}$$

$$CR_{ij} = 1 - \exp\left[-\left((b_0 + \mu_j) + \frac{b_1}{H_i} \right) \left(RSI_j^{b_2} + \frac{D_i}{H_{D_j}} \right)^{b_3} \right] + \varepsilon_{ij} \tag{35}$$

where CR_{ij} is the crown ratio of tree *i* in plot *j*.

The crown ratio estimation model with better fit index and *AIC* values: Eq. (26) nonlinear model; and Eq. (35) mixed-effects model were selected as base models for modeling heterogeneous errors and autocorrelation structures. The resulted ENMs are Eq. (36-38); and EMEMs models are Eq. (39-41).

Model 6 + power variance function (36)

Indiviudal dbh was found a better performing covariate than alternatives in reducing *AIC* value for modeling heterogenous errors.

Model 26 + AR (1) (37)

Mode 26 + power variance function + AR (1) (38)

Model 35 + power variance function (39)

Model 35 + AR (1) (40)

Mode 35 + power variance function + AR (1) (41)

Statistical analysis

All nonlinear models, mixed-effects models and extended form of models were fitted in R (R Development Core team 2012) using the “nls”, “nlme” and “gnls” functions respectively (Phinheiro et al., 2014). Models were compared using the *Fit index*, *RMSE* and *AIC*. Likelihood ratio (*LR*) statistics was also used to compare the ENMs with base nonlinear model (Eq. 3) and, EMEMs with base mixed-effects model (Eq. 5). The *Fit index* for nonlinear models was calculated based on the Eq. (42) and *RMSE* was calculated based on Eq. (43). *Fit indices* and *RMSEs* for mixed models and the extended models were calculated based on the actual height or crown ratio predictions using only parameter estimates of the fixed covariates while setting random effects equal to zero.

$$Fit\ index = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \qquad (42)$$

$$RMSE = \sqrt{\frac{\sum_{i}^{n} (y_i - \hat{y}_i)^2}{n - p}} \qquad (43)$$

where y_i is observed value for i^{th} observation; $\hat{y}_i$ is the predicted value by a model; $\bar{y}_i$ is the mean observed value; p = number of parameter estimated by a model. But for ease in comparison, *Fit index* was interpreted as percentage (multiplied by 100).

The Goldfeld-Quandt test was used to test an assumption of homoscedasticity of the error variance for both height prediction and crown ratio estimation model (Judge et al. 1982, p. 371-372). The dataset was divided into three parts ordered from the smallest to the largest value of the independent variable (dbh), and the middle (1/3rd) of the data were excluded. The Goldfeld-Quandt test compares the ratio of a residual sum of squares of the model from the upper range ($3/8^{ths}$ of total data) observations to the model from lower range ($3/8^{ths}$ of total data) observation. Standardized residual plots were plotted against the fitted (predicted) values, and against dbh values. The standardized residuals were also plotted against the mid-range of the design variables (plot basal area, site index, and plot age) but they are not shown. Standardized residual plots of all height prediction models were similar to each other, and the same was the case for *CR* models. Therefore, only standardized residual plots from the base model and the best model for both height prediction and *CR* estimation are presented.

Results

The patterns of height and crown changes over the time for each measurement can be observed in Figure 2. Some of the changes could be due to thinning from below after the third measurement and removal of many trees due to ice storm damage at the fourth (Figure 2a). Due to the study design large trees in older age classes were present even at the first measurement since the study included a balanced range of age classes at that time. During later measurements trees in the younger age classes grew in height, resulting in an increase in mean height for the study as a whole (Table 2). The mean crown ratio appears to be fairly constant over time, and possibly average increases in total height are balanced by the crown recession over time as might be expected (Figure 2b and Table 2).

Height prediction model with dominant height

Table 4 displays the fit statistics (*RMSE*, *AIC*, and *Fit Index*) of height-dbh relationship models: nonlinear model (Eqs. 1-4); mixed-effects model (Eq. 5); ENMs (Eq. 6-8); and EMEMs (Eq. 9-11). The parameter estimates of all models were significantly different from zero, but estimates of the selected models are shown in Table 5. The fit statistics of the Eq. (3) and Eq. (4) was similar and better than the alternative models (Eq. 1-2). However, the *AIC* value of Eq. (4) was slightly smaller. The smaller *AIC* value might be associated with the likelihood estimation function that involves the number of parameters in a model, i.e. Eq. (3) has 4 parameters; and Eq. (4) has five parameters. Inclusion of *QMD* as stand level covariate in the Eq. (3) showed the reduction of RMSE by 8% compared a model without *QMD* (Eq. 1) and also compared to a model with *BAH* as the covariate (Eq. 2) (Table 5). This suggested that the selection and the position of the stand level covariate also affects the performance of a model (Table 4). Further, the mixed-effects model (Eq. 5) showed similar *Fit index* (95.84) and RMSE (1.27) to Eq. (3) when the random component was assumed to be zero and only the fixed effects were used to make height predictions (Table 4). The SD $\hat{\sigma}(\mu_j)$ =1.19319, of the random component associated with the parameter b_2 was significant (*p-value* < 0.0001) with the confidence interval of [1.208723, 1.80569]. The *AIC* value of the Eq. (5) was lower than of Eq. (3) but this includes random effects parameters in Eq. (5) that are usually not available for prediction unless they can be obtained by calibration.

The *LR* statistics suggested the ENMs and EMEMs were significantly different from their base model form (Eq. 3) and Eq. (5) respectively (Table 4). The ENMs and EMEMs provided similar *RMSE* and *Fit indices*, but Eq. (8) provided smaller *AIC* value and large significant *LR* statistics (Table 4). This suggested that modeling both variance function and autocorrelation structures in a nonlinear model performed better than just modeling variance function (Eq. 6) and other EMEMs (Table 4). In EMEMs, the similar *AIC* value and *LR* statistics indicated that EMEM Eq. (9) could be a better alternative model to Eq. (11) for height prediction but ignoring autocorrelation could lead to an underestimate of standard errors if account is made only for heteroscedasticity.

Both extended model forms (ENMs and EMEMs) showed that relatively small power parameter estimates were needed to stabilize the issue of heteroscedasticity (Eqs.

6, 8, 9, and11) and large autocorrelations within individual tree heights were observed (Eq.7-8) but moderate autocorrelation was observed between the tree heights within plot (Eq. 10-11) (Table 4). The parameter estimates of the Eqs. (5, 8, and11) were very similar (Table 5). Interestingly, the EMEM (Eq. 11) showed greater reduction in the SD of the random effect associated with plot than the mixed-effects model (Eq. 5) (Table 5).

The Goldfeld-Quandt test did not indicate violations of the assumption of homoscedasticity of error variance for any of Eqs. (1-4). For example; Goldfeld-Quandt variance ratio was 0.80 for Eq. (3) which was less than tabulated $F_{(5257, 5257)} = 2.04$ at $\alpha = 0.05$ level. The standardized residuals did not show any systematic pattern to indicate a violation of an assumption of homogeneity of variance. So, the standardized residuals of the better performing models: nonlinear model (Eq. 3), mixed-effects model (Eq. 5), and the ENM (Eq. 8) are shown in Figure 3. The residual distribution pattern was slightly different between nonlinear models and mixed-effects models. The nonlinear model showed some dip near the lower fitted values (Figure 3a) while mixed-effects model showed a more compact distribution of standardized residuals (Figure 3b). The residual distribution of all ENMs was similar as shown in Figure 3c for Eq. (8). Though extended models had better *AIC* values, residual distribution patterns were not different from their base models (Eq. 3 and 5). The standardized residuals plotted against dbh (Figure 3d, 3e, and 3f) also showed similar patterns and trends as discussed for fitted values. Mixed-effects models showed variation at the lower diameter range compared to the nonlinear models and ENMs. The standardized residuals plotted against the range of design variables of shortleaf pine growth study also showed that the median residuals were almost centered to zero with a minimum bias, but these graphs are not shown. .

Height prediction models without dominant height

The parameter estimates of the fundamental local height prediction models with dbh as the only independent variable were significant (*p-value* < 0.0001). The fit indices and *RMSE* for fundamental local models (Eq. 12-19) were very similar but the Eqs. (14), 15), and (17) did not perform as well as alternative models (Table 7). An average *Fit index* and *RMSE* was 73.85 and 3.18 m respectively for the fundamental local models. Fit statistics improved significantly across modified local models when a height prediction model function was multiplied by *QMD* raised to a power (Eq. 20-27) (Table 7). The stand level covariate *QMD* with a parameter in the power position increased values of fit indices and decreased *RMSE* substantially for both two parameters and three parameter dbh-height prediction model functions. On average, model *RMSE* was decreased by 17.31 % (i.e. 0.55 m) and the *Fit index* was increased by 11% (i.e. 8.24). The greatest reduction in *RMSE* (0.64 m) was observed in Eq. (22) that corresponds to base Eq. (14), and the lowest reduction in *RMSE (*0.48 m) was observed in Eq. (23) that corresponds to base Eq. (15). The *AIC* value of each modified local models was reduced by an average of 7.34% (i.e. 5304) compared to the same equation form without *QMD*. Substantial reductions in *AIC* value and *RMSE* for these equation indicates that *QMD* as independent variable plays an important role in minimizing model error compared to models that use the only diameter as an independent variable. For the fundamental local models (Eq. 12-19) and modified local models with *QMD* (Eq. 20-27), the Goldfeld-Quandt test failed to reject homogeneity of variance at $\alpha = 0.05$ of significance because all estimated variance ratios were less than F=2.04.

Crown ratio estimation

The crown ratio estimation model with average dominant height (Eq. 29) instead plot age (Eq. 22) had better fit statistics and a better *AIC* value (Table 7) than other alternatives. Similar improvements in the model performance resulted from adding stand basal area when Eq. (28) was modified to Eq. (30), and Eq. (29) modified to Eq. (31) (Table 7). Crown ratio estimation models with *PAG* and *BAH* (Eq. 30) did not improve model performance as much as the inclusion of H_D (Eq. 29). The crown ratio estimation model that included *RSI* (Eq. 32) had better fit statistics and *AIC* value than all other alternative models. Fitting the logistic function to this dataset suggested that Eq. (33) can be used as an alternative crown ratio estimation model because it had similar fit statistics and a similar *AIC* value compared to other slightly better alternative models. The parameter estimates of the crown ratio estimation models were significant (*p-value* < 0.0001) but only estimates of the selected models are shown (Table 8).

The mixed-effects model for crown ratio estimation (Eq. 35) with RSI had a slightly better *Fit index* and *RMSE* than the model with *BAH* (Eq. 34) though it had a similar *AIC* value (Table 7). However, the fit statistics (*RMSE* and *Fit indices*) of mixed-effects models setting random effects to zero were not smaller than those of the OLS nonlinear models (Eq. 31 and 32) (Table 7). The SD of the error $(\hat{\sigma}(\varepsilon_{ij}))$ of Eq. (35) was slightly smaller (0.01658) with a 95% confidence interval of [0.01478, 0.01859] than of the Eq. (34) but estimate of the SD of the random component $(\hat{\sigma}(\mu_j))$ of Eq. (35) was identical with Eq. (34).

The ENMs (Eq. 36-38), and EMEMs (Eq. 39-41) showed reduced *AIC* values compared to the base models: Eq. (32), and Eq. (35) respectively. The large differences in the *LR* statistics and smaller *AIC* values indicated that ENMs performed better than EMEMs (Table 7). Both ENMs and EMEMs provided negative power parameter estimates (Table 7). However, the ENM with a variance function (Eq. 36) showed *Fit index* better than others, and close to the nonlinear base model (Eq. 32) but not with a greatly reduced *AIC* value (Table 7).

The ENMs with AR (1): Eq. (37), and (38) showed that observed *CR* of an individual tree in repeated measurement were moderately correlated (0.55) but, EMEMs Eq. (40), and (41) showed CR were very weakly correlated (0.12) within a plot (Table 7). Interestingly, ENMs Eq. (38) provided the smallest *AIC* value and comparable *Fit index* with Eq. (35) (Table 7) but had greatly changed parameter estimates (Table 8). It indicated that a nonlinear model with both power variance function and AR (1) structure can be used in the repeated measurements, as a substitute modeling approach to the mixed-effects modeling approach. The parameter estimates of the only selected *CR* estimation models are shown in Table (8).

The Goldfeld-Quandt test also failed to reject homogeneity of variance for the crown ratio estimation models (Eq. 28-32) at the $\alpha = 0.05$ level of significance because the estimated variance ratios were less than F=2.08. For example, the ratios of variance were 1.09 and 1.08 for Eqs. (21) and (32) respectively. The standardized residual plot also indicated homogeneity of variance. The residual distribution patterns were fairly similar within the ENMs and also within the EMEMs with small improvements over the base model. The residual distribution pattern, Figure 4a of nonlinear model (Eq. 32) and Figure 4b of the mixed-effects model (Eq. 35) were similar though the latter was slightly compact and elongated to right, indicating more constant error variance. The residual pattern of the ENM (Eq. 38) showed some errors at the lower fitter values but a more compact error at fitted middle values than other models (Figure 4c). Similarly, the standardized residuals plotted against dbh showed similar trends among the model forms (Figure 4d, 4e, and 4f) and suggested large prediction errors at the lower diameter range. The standardized residuals from Eq. (32) plotted against the design variables also showed that median residuals are almost centered at zero with a minimum bias over the range of design variables, but these graphs are not shown.

Discussion

The inclusion of *QMD* improved the fit statistics of the modified equation for height prediction (Eq. 3) compared to the models presented by Budhathoki et al. (2008) (Eq. 2) and Lynch et al. (1999) (Eq. 1) (Table 4). It was found that the prediction bias of 0.02 m by Eq. (1) was reduced to 0.005 m by Eq. (3) which supported that the conclusion that Eq. (3) has better prediction capability. It was also observed that the stand competition variable '*QMD'* in the Eq. (3) can be substituted by relative measure variable '*RAQD'* to obtain similar parameter estimates, standard error and *RMSE*. The parameter estimates and fit statistics of the mixed-effects model (Eq. 5) were similar to an OLS nonlinear model (Eq. 3). Predictions from a nonlinear mixed model with the random effect parameter set to zero were not as good as the OLS results, as expected (Table 4 and 7). Garber et al. (2009) demonstrated that imputing height from the fixed estimates of OLS nonlinear had less bias in volume prediction (4 m^3ha^{-1}) than a mixed-effects model (20 m^3ha^{-1}) when height was imputed for Douglas fir, as it was discussed by Temesgen et al. (2008) while predicting height for Douglas-fir. The smaller *AIC* value in the mixed-effects model (e.g. Eq. 5) than OLS nonlinear models (e.g. Eq. 1, 2, and 3) might be attributed to the inclusion of the random parameters in the mixed-effects model. However, unless calibration data are available, the random parameters may not be of practical help for most prediction problems. On the other hand, several authors have shown that prediction using mixed models can be attractive when calibration data is available and have indicated, mixed model prediction may not be better than OLS unless calibration data is available (Lynch et al., 2005; Lynch et al., 2012).

The AR (1) structure suggested autocorrelation of residuals within in individual observation was large for height prediction and was moderate for *CR* estimation. The large autocorrelation can be perceived as a wider residual distribution with increasing diameter as in Figure 3 and moderate autocorrelation can be perceived as a circular distribution of residuals as in Figure 4. Ducey (2009) also observed the circular pattern of residual distribution for *CR* modeling of *Pinus strobus*. However, the autocorrelation remediation at the larger lags is not reliable because it is based on fewer residual pairs. The EMEMs did not greatly improve model fits compared to the ENMs for both height-dbh relationship and *CR* estimation. So, the *AIC* values of the EMEMs were not smaller than ENMs (Table 4 and 7). The smaller *AIC* value in the ENMs could be a result of more successful modeling of autocorrelation structure within individual tree than within plot or grouped data as in the mixed model. Because of this, *AIC* value of the EMEMs of both height-dbh relationship and *CR* estimation were not greatly reduced compared to the ENMs (Table 4 and 7. Though *LR* statistics suggested EMEM with AR (1) structure was significantly different from the base mixed-effects model (Table 4 and 7), both models had similar SD of random effect and error variance. Perhaps, the mixed-effects models each group (plots) with its random effect and the assigned AR (1) (within stand) is formally identical to a random effect model that has both among group variance (σ_b^2) and within group variance (σ_w^2) that corresponds to correlation parameters (ρ, σ^2). Therefore, it is possible that ENM with autocorrelation structure and variance function can perform better than the EMEMs.

The small, positive power parameter in both ENMs and EMEMs of height-dbh relationship indicated a small amount of heterogeneous error due to influence of H_D. The negative power parameter associated with dbh could be due to the pattern of the standardized residuals distributed along the lower diameter which showed greater variability. But the fact that large standardized residuals occurred in this study (Figure 3 and 4) may be because the data were from naturally occurring stands as observed in Budhathoki et al. (2008) for shortleaf pine and in Sharma and Patron (2007) for boreal tree species in Ontario, Canada. Height growth in naturally occurring stands is not as uniform as that in plantations, such as described, for example, by Buford (1991) for loblolly pine. It could also partially due to measurement errors since the accuracy of commonly used height measuring devices are likely to be in 0.5 meters if not more, and there can also be variation in the accuracy obtained by individuals in the field measurement crew.

The issues of heteroskedastic residual distribution (non-constant variance) when present can sometimes be substantially resolved, if a weighted OLS nonlinear model is used. After testing different weight functions, the weight 1/dbh, was found beneficial in improving model performance and addressing non constant error variance as used by Haung et al. (1992), Fang and Bailey (1998), Temesgen et al., (2007), and Temesgen et al., (2014). For example, if weight as 1/dbh was used in Eq. (3), the residuals were slightly more compact and constant than shown in Figure 3, but some were still beyond ± 4 at lower diameter range. The standard errors of parameter estimates were not considerably different from those of non-weighted Eq. (3). Because the weighted parameter estimates did not have variances that were substantially different from the unweighted parameter estimates and due to the results of the Goldfeld-Quant test, we used the unweighted parameter estimates. Of course, mathematically the sum of squares between predicted and actual heights cannot be reduced below that obtained with ordinary unweighted OLS by weighing a model that has the same form and the same independent variables.

Modified local models (Eq. 20-27) showed substantially improved and similar model performance except for the Eq. (23) (Table 7). The parameter of *QMD* used in the "power" position was more effective than in the "linear" position. The other variables of stand competition such as *BAH* and *TPH* did not show any improvement but the inclusion of relative measures "*RAQD"* provided RMSE similar to the RMSE with the inclusion of *QMD* in some modified local models (Eq. 17, 20 and 21). However, *RAQD* provided larger RMSE for other modified local models than it was observed with *QMD.* Therefore, the evidence suggests that *QMD* can be a better surrogate covariate in improving model performance when the information on H_D is not available. Temesgen et al. (2007) also suggested that the use of *BAH* and *BAL* in the fundamental local model and that can reduce *RMSE* up to 15%. The height-dbh prediction model has been often modified using H_D to reduce *RMSE* of a model, however in practice H_D may not be easily available for natural stands.

Eq. (28-33) fitted with OLS provided a choice of alternative models for *CR* estimation. The model (Eq. 32) with covariate H_D and *RSI* performed better than a model with *PAG* and *BAH* (Eq. 31), and H_D and *PAG* (Eq. 30) in addressing *CR* variability. This was expected because the distance between the trees reflects the crown competition and also thinning abruptly changes *CR* of an individual tree (Smith et al., 1992; Hynynen, 1995). So, *RSI* has been used a measure of relative competition index in modeling *CR* (Ducey, 2009; Zhao et al., 2012). It was observed that the prediction bias was very small for all *CR* models, but it was negative for Eqs. (28-31), while it was positive for Eq. (32). It suggested that model with *RSI* tended to overestimate than other alternative models. The researcher has also used crown competition factor (*CCF*) as a covariate in association to *BAL* to model *CR* (Hasenauer and Monserud, 1996; Temesgen et al., 2007).

The small amount of variation explained (44%) by Eq. (32) in *CR* estimation may be due to the inherent variability of naturally-occurring forests compared to plantations. (Figure 4a, 4b, and 4c). The proportion of variation explained by the *CR* models of Dyer and Burkhart (1987) for loblolly pine plantations was greater (60%) than the proportion of variation in *CR* explained by the natural stand models of Hasenauer and Monserud (1996) in Austrian natural forest stands (49%-17%). The significance of Eq. (29) and Eq. (31) is that the model can be applied when *PAG* is not available. Hynynen, (1995) also found H_D as an important independent covariate in *CR* modeling of Scot pine stands. The addition of *BAH* as a linear term to the ratio of *D* and H_D (Eq. 25) improved fit statistics more than combining it linearly to individual tree height. Smith et al. (1992) and Monserud and Sterba (1996) indicated that using *BAH* as an independent variable can improve prediction of an individual tree crown length. This leads to the expectation that *BAH* could also improve the fit of *CR* models.

The logistic function model for *CR* estimation (Eq. 29) performed similarly to the model of Eq. (31). The logistic function restricts predicted crown ratio bounds within a 0-1 interval. As dbh-height models, the parameter estimates of OLS nonlinear models for *CR* estimation were also better than mixed-effects models (Table 7). The mixed-effects models (Eq. 27-28) of *CR* estimation also performed more poorly than OLS nonlinear *CR* estimation Eqs. (31 and 32), but the *AIC* values were smaller for mixed-effects models (Table 5).

Conclusions

The modified height-dbh relationship (Eq. 3) and crown ratio relationship model (Eq. 32) provided better accuracy than existing models for estimating the height and crown ratio of natural even-aged stands of shortleaf pine. The inclusion of *QMD* as a measure of stand competition rather than *BAH* as an independent variable helped to improve the understanding of the height-dbh relationship. Also, an inclusion of *QMD* improved the precision of height prediction model forms that do not utilize dominant height as a covariate. The *RSI* and H_D enhanced the relationship of crown ratio with height and dbh compared to using *PAG* in *CR* estimations model. It was found that inclusion of *RSI* instead of made small but definite improvements in *CR* estimation and also that the logistic function can be used as a comparable choice to an alternative nonlinear model for *CR* prediction.

By alleviating heterogeneous error at stand level and adjusting autocorrelation at the individual tree level in repeated measurements, a quality nonlinear model with minimum information loss can be obtained. The parameter estimates of such model are preferred as an alternative to the mixed-effects modeling approach in predicting missing height and crown ratios. In repeated measurements, the autocorrelation within in individual observation is larger while predicting the height of that individual tree than estimating its crown ratio. The small, but positive weight of power parameter remediated heterogeneous errors and improved model performance for height prediction. Mixed models provided similar fit statistics when predictions were based only on the fixed effects parameters compared to those of nonlinear models fitted by OLS. However, the mixed-effects models may provide improved predictions when calibration data are available.

Parameter estimates of the ENM, Eq. (8) for height prediction and Eq. (38) for CR estimation can be incorporated in the Shortleaf Pine Stand Simulator (Huebschmann et al. 1988) which can be used to develop information for practical forest management decision making i.e. estimation of the total stand volume and biomass production, for naturally occurring even-aged shortleaf pine forests. The relationships between dbh, height and crown ratio could have important implications in inventories for biomass and carbon estimation of natural stands of shortleaf pine in the southern US. To the extent that these formulations are a novel approach in the forestry literature, they could be considered for application in other forest types in addition to well-known existing equation forms.

Acknowledgements

This research was supported by Department of Natural Resource Ecology and Management) at Oklahoma State University (OSU) and USDA Forest Service Southern Research Station (USFS). This article has been approved for publication by the Oklahoma Agricultural Experiment Station and supported by project OKL0-2843. The authors would like to thank all crew members who contributed in collecting data since permanent plot establishment (1985-87) to the current measurement (2012-2014).

References

Arcangeli, C., Klopf, M., Hale, S.E., Jenkins, T.A.R. and Hasenauer, H. 2014 The uniform height curve method for height–diameter modeling: an application to Sitka spruce in Britain. Forestry, 87, 177-186.

Avery, T.E. and Burkhart, H.E. 2002. Forest Measurements, 5th ed. McGraw-Hill, NY. 456p.

Bates, D.M. and Watts, D.G. 1980 Relative Curvature Measures of Nonlinearity. Journal of the Royal Statistical Society. Series B (Methodological), 42, 1-25.

Budhathoki, C.B., Lynch, T.B. and Guldin, J.M. 2008 A Mixed-Effects Model for the dbh–Height Relationship of Shortleaf Pine (*Pinus echinata* Mill.). South. J. Appl. For. 32, 5-11.

Buford, M.A. 1991 Performance of four yield models for predicting stand dynamics of a 30-year-old loblolly pine (*Pinus taeda* L.) spacing study. For. Ecol. Manage. 46, 23-38.

Burkhart, H.E. and Strub, M.R. 1974 A model for simulation of planted loblolly pine stands. In: Growth models for tree and stand simulation. Royal College of Forestry Stockholm, p. 379.

Burkhart, H.E. and Tomé, M. 2012 Modeling forest trees and stands. Springer Science & Business Media.

Calama, R. and Montero, G. 2004 Interregional nonlinear height–diameter model with random coefficients for stone pine in Spain. Can. J. For. Res. 34, 150-163.

Cryer, J.D. and Chan, K.-S. 2008 Time series analysis: with applications in R. Springerverlag New York.

Ducey, M.J. 2009 Predicting Crown Size and Shape from Simple Stand Variables. J. Sustain. For. 28, 5-21.

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

Fang, Z. and Bailey, R.L. 1998 Height–diameter models for tropical forests on Hainan Island in southern China. For. Ecol. Manage. 110, 315-327.

Garber, S.M., Temesgen, H., Monleon, V.J. and Hann, D.W. 2009 Effects of height imputation strategies on stand volume estimation. Can. J. For. Res. 39, 681-690.

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

Holdaway, M.R. 1986 Modeling Tree Crown Ratio. The Forestry Chronicle, 62, 451-455.

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

Huebschmann, M.M., Lynch, T.B. and Murphy, P.A. 1998 Shortleaf pine stand simulator: an even-aged natural shortleaf pine growth-and-yield model: user's manual. Oklahoma Agricultural Experiment Station, Division of Agricultural Sciences and Natural Resources, Oklahoma State University.

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

Jiang, L., Brooks, J.R. and Hobbs, G.R. 2007 Using Crown Ratio in Yellow-Poplar Compatible Taper and Volume Equations. North. J. Appl. For. 24, 271-275.

Jiang, L.-c. and Liu, R.-l. 2011 Segmented taper equations with crown ratio and stand density for Dahurian Larch (*Larix gmelinii*) in Northeastern China. J. For. Res. 22, 347-352.

Judge, G.G., Hill, R.C., Griffiths, W., Lutkepohl, H. and Lee, T.-C. 1988 Introduction to the Theory and Practice of Econometrics.

Lappi, J. 1991 Calibration of Height and Volume Equations with Random Parameters. For. Sci. 37, 781-801.

Lappi, J. 1997 A Longitudinal Analysis of Height/Diameter Curves. For. Sci. 43, 555-570.

Leites, L.P., Robinson, A.P. and Crookston, N.L. 2009 Accuracy and equivalence testing of crown ratio models and assessment of their impact on diameter growth and basal area increment predictions of two variants of the Forest Vegetation Simulator. Can. J. For. Res. 39, 655-665.

Lynch, T.B., Budhathoki, C. and Wittwer, R.F. 2012 Relationships between Height, Diameter, and Crown for Eastern Cottonwood (*Populus Deltoides*) in a Great Plains Riparian Ecosystem. West. J. Appl. For. 27, 176-186.

Lynch, T.B., Hitch, K.L., Huebschmann, M.M. and Murphy, P.A. 1999 An Individual-

Tree Growth and Yield Prediction System for Even-Aged Natural Shortleaf Pine Forests. South. J. Appl. For. 23, 203-211.

Lynch, T.B., Holley, A.G. and Stevenson, D.J. 2005 A Random-Parameter Height-Dbh Model for Cherrybark Oak. South. J. Appl. For. 29, 22-26.

Lynch, T.B. and Murphy, P.A. 1995 A Compatible Height Prediction and Projection System for Individual Trees in Natural, Even-Aged Shortleaf Pine Stands. For. Sci. 41, 194-209.

Meyer, H.A. 1940 A Mathematical Expression for Height Curves. J. For. 38, 415-420.

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

Peng, C., Zhang, L. and Liu, J. 2001 Developing and Validating Nonlinear Height–Diameter Models for Major Tree Species of Ontario's Boreal Forests. North. J. Appl. For., 18, 87-94.

Pinheiro, J. and Bates, D. 2000 Mixed-effects models in S and S-PLUS. Springer Science & Business Media.

Ratkowsky, D.A. and Giles, D.E. 1990 Handbook of nonlinear regression models. Marcel Dekker New York.

R Core Team 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0

Richards, F.J. 1959 A Flexible Growth Function for Empirical Use. J. Exp. Bot. 10, 290-301.

Schmidt, M., Kiviste, A. and von Gadow, K. 2011 A spatially explicit height–diameter model for Scots pine in Estonia. Eur J Forest Res, 130, 303-315.

Sharma, M. and Parton, J. 2007 Height–diameter equations for boreal tree species in Ontario using a mixed-effects modeling approach. For. Ecol. Manage. 249, 187-198.

Sharma, M. and Yin Zhang, S. 2004 Height–Diameter Models Using Stand Characteristics for *Pinus banksiana* and *Picea mariana*. Scand J. For. Res. 19, 442-451.

Sharma, R.P. and Breidenbach, J. 2015 Modeling height-diameter relationships for Norway spruce, Scots pine, and downy birch using Norwegian national forest inventory data. Forest Sci Technol. 11, 44-53.

Smith, W.R., Farrar Jr, R.M., Murphy, P.A., Yeiser, J.L., Meldahl, R.S. and Kush, J.S. 1992 Crown and basal area relationships of open-grown southern pines for modeling competition and growth. Can. J. For. Res. 22, 341-347.

Stage, A.R. 1975 Prediction of height increment for models of forest growth.

Tahvanainen, T. and Forss, E. 2008 Individual tree models for the crown biomass distribution of Scots pine, Norway spruce and birch in Finland. For. Ecol. Manage. 255, 455-467.

Temesgen, H., Hann, D.W. and Monleon, V.J. 2007 Regional Height-Diameter Equations for Major Tree Species of Southwest Oregon. West. J. Appl. For. 22, 213-219.

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

Temesgen, H., Monleon, V.J. and Hann, D.W. 2008 Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. Can. J. For. Res. 38, 553-565.

Temesgen, H., Zhang, C.H. and Zhao, X.H. 2014 Modelling tree height–diameter relationships in multi-species and multi-layered forests: A large observational study from Northeast China. For. Ecol. Manage. 316, 78-89.

Valenti, M.A. and Cao, Q.V. 1986 Use of crown ratio to improve loblolly pine taper equations. Can. J. For. Res. 16, 1141-1145.

VanderSchaaf, C.L. 2014 Mixed-effects height–diameter models for ten conifers in the inland Northwest, USA. Southern Forests: a Journal of Forest Science, 76, 1-9.

Wykoff, W.R., Crookston, N.L., Stage, A.R. and Forest, I. 1982 User's guide to the stand prognosis model.

Zhao, D., Kane, M., and Borders, B. E. 2012 Crown Ratio and Relative Spacing Relationships for Loblolly Pine Plantations. Open Journal of Forestry, *2*(03), 101.

Table II-1. Descriptive statistics (mean with SD in parentheses) of stand level and tree variables recorded for six times measurement of naturally occurring even-aged shortleaf pine stand.

Variables	Measurements 1st (n= 2682)	2nd (n = 3017)	3rd (n = 3215)	4th (n = 1750)	5th (n = 1677)	6th (n =1687)
CR	0.373 (0.094)	0.365 (0.096)	0.372 (0.096)	0.376 (0.089)	0.374 (0.092)	0.364 (0.084)
D (cm)	28.991 (10.974)	26.426 (10.388)	28.333 (10.5)	31.195 (10.003)	33.276 (10.33)	35.021 (10.68)
H (m)	19.864 (6.222)	18.684 (6.34)	19.894 (6.087)	20.777 (5.396)	21.698 (5.215)	22.877 (5.221)
H_D *(m)*	20.433 (5.596)	19.364 (5.704)	20.625 (5.356)	21.17 (4.83)	22.131 (4.56)	23.293 (4.477)
PAG (yrs)	62.465 (22.114)	56.585 (20.207)	61.537 (20.258)	66.486 (20.742)	72.251 (20.737)	77.828 (20.756)
BAH (m^2ha^{-1})	24.191 (8.653)	23.135 (7.566)	26.156 (8.352)	23.609 (8.693)	26.027 (9.147)	28.063 (9.565)
QMD(cm)	28.997 (9.594)	26.442 (8.942)	28.468 (8.816)	31.118 (8.644)	33.256 (8.861)	35.036 (9.107)
RSI	0.276 (0.101)	0.269 (0.092)	0.256 (0.091)	0.293 (0.112)	0.284 (0.109)	0.274 (0.107)
RAQD	1.063 (0.323)	1.073 (0.347)	1.078 (0.359)	1.043 (0.269)	1.044 (0.268)	1.045 (0.269)
BAHG (m^2ha^{-1}) yrs^{-1}	0.458 (0.274)	0.491 (0.296)	0.505 (0.303)	0.409 (0.241)	0.41 (0.23)	0.404 (0.214)

n = total number of observations*; CR* **=** Crown ratio**;** *D* = diameter at breast height (cm)**;** *H* **=** individual tree height; H_D = average plot dominant and co-dominant height (meters); *PAG* = plot age (yrs); *BAH* = stand basal area per hectare (m^2ha^{-1}), *QMD* = quadratic mean diameter (m); *RSI* = relative spacing index; *RAQD* = ratio of *QMD* to *D*; and *BHAG* = ratio of *BAH* to *PAG*.

Table II- 2. Summary statistics of the variables used to model height prediction and crown estimation of the naturally occurring even-aged shortleaf pine stand (N = 14028).

Variables	Mean	SD	Minimum	Maximum
CR	0.373	0.094	0.055	0.800
D (cm)	28.991	10.974	2.794	67.564
H (m)	19.864	6.222	3.048	38.100
H_D *(m)*	20.433	5.596	6.706	36.019
PAG (yrs)	62.465	22.114	18.000	119.000
BAH (m^2ha^{-1})	24.191	8.653	2.035	48.684
QMD(cm)	28.997	9.594	7.887	58.258
RSI	0.276	0.101	0.126	0.841
RAQD	1.063	0.323	0.386	7.012
BAHG $(m^2ha^{-1})\ yrs^{-1}$	0.458	0.274	0.025	1.338

Table II-3. Fundamental local models for height prediction with dbh only, and with modified height prediction model with quadratic mean diameter (*QMD*).

Eq.	Common height model	Eq.	Modified Model	Source
12	$H_i - c = b_0(1-e^{-(b1Di)})$	20	$H_i - c = b_0(1-e^{-(b1Di)})QMD^{b2}$	Meyer 1940
13	$H_i - c = b_0(1-e^{-b1Di})^{b2}$	21	$H_i - c = b_0(1-e^{-b1Di})^{b2}QMD^{b3}$	Richard 1959
14	$H_i - c = b_0e^{b1/Di}$	22	$H_i - c = b_0e^{b1/Di}QMD^{b2}$	Burkhart and Strub 1974
15	$H_i - c = b_0D_i^{b1}$	23	$H_i - c = b_0D_i^{b1}QMD^{b2}$	Stage 1975
16	$H_i - c = b_0D_i(b_1+D_i)$	24	$H_i - c = b_0D_i(b_1+D_i)QMD^{b2}$	Bates and Watts 1980
17	$H_i - c = e^{b0 + (b1/(Di+1))}$	25	$H_i - c = e^{b0 + (b1/(Di+1))}QMD^{b2}$	Wykoff et al. 1982
18	$H_i - c = b_0e^{b1/(Di + b2)}$	26	$H_i - c = b_0e^{b1/(Di + b2)}QMD^{b3}$	Ratkowsky 1990
19	$H_i - c = (D_i /(b_1 + b_2 D_i))^{b3}$	27	$H_i - c = (D_i /(b_1 + b_2 D_i))^{b3} QMD^{b4}$	Schmidt et al. 2011; Sharma and Briedenbach 2015

Table II-4. Fit statistics (*RMSE, Fit index* (%), and *AIC*), power parameter (δ), and autocorrelation (φ) and likelihood ratio statistics (*LR*) of the different form of nonlinear height-dbh relationship. OLS nonlinear (Eq. 1-4); mixed-effects model (Eq.5); extended nonlinear models (ENM (Eq. 6-8); and extended mixed-effects models (EMEM) (Eq. 9-11).

Parameters	*RMSE*	*Fit Index (%)*	*AIC*	δ	φ	*LR*
Eq.(1)	1.38	95.11	48776			
Eq.(2)	1.37	95.16	48624			
Eq.(3)	1.27	95.85	46453			
Eq.(4)	1.26	95.92	46230			
Eq.(5)	1.27	95.84	45157			
Eq.(6)	1.27	95.85	46203	0.323		251.9
Eq.(7)	1.27	95.85	40305		0.803	9912.6
Eq.(8)	1.27	95.84	36067	0.398	0.810	10389.9
Eq.(9)	1.27	95.83	44790	0.383		369.2
Eq.(10)	1.27	95.83	45140		0.037	19.1
Eq.(11)	1.27	95.83	44771	0.3836	0.039	390.5

Table II-5. Parameter estimates and standard error in the parenthesis of the selected height prediction models. OLS nonlinear (Eq. 1-3); mixed-effects model (Eq. 5); ENM (Eq. 8) and EMEM (Eq. 11).

Parameters	Eq.(1)	Eq.(3)	Eq.(5)	Eq.(8)	Eq.(11)
b_0	2.00052 (0.02780)	1.37232 (0.01111)	1.41103 (0.01583)	1.42302 (0.01659)	1.39325 (0.01473)
b_1	0.82570 (0.00346)	0.94119 (0.00271)	0.93335 (0.00369)	0.93107 (0.00398)	0.93888 (0.00355)
b_2	-9.34600 (0.41130)	-1.58186 (0.06269)	-1.36397 (0.05693)	-1.38491 (0.07173)	-1.25512 (0.04758)
b_3	-1.16510 (0.02090)	-1.69434 (0.01684)	-1.64451 (0.01739)	-1.64367 (0.02327)	-1.61206 (0.01638)
b_4			1.41103 (0.01583)	1.42302 (0.01659)	1.39325 (0.01473)
$\hat{\sigma}(\mu_j)$			1.19456		0.38801
$\hat{\sigma}(\varepsilon_{ij})$			0.03067	0.40902	0.03061

Table II-6. Parameter estimates and fit statistics (*RMSE, Fit index*, and *AIC)* for height prediction models with dbh (Eq. 9-19), and along with quadratic mean diameter (Eq. 20-27).

Eq. (No.)	Parameters estimates				*RMSE*	*Fit index (%)*	*AIC*
	b_0	b_1	b_2	b_3			
12	41.2084	0.0215			3.16	74.27	72061
13	38.8776	0.0247	1.0551		3.16	74.28	72058
14	37.3973	-18.5592			3.25	72.64	72921
15	1.6609	0.7209			3.19	73.76	72336
16	67.0409	72.9095			3.16	74.24	72076
17	3.6529	-20.2320			3.23	73.05	72708
18	51.5498	-39.2751	10.7639		3.16	74.28	72055
19	1.4364	0.0324	1.1802		3.16	74.28	72055
20	3.5866	0.0716	0.5393		2.62	82.24	66863
21	3.5567	0.0741	1.0387	0.5405	2.62	82.24	66864
22	4.1877	-8.2663	0.5365		2.62	82.26	66847
23	1.0548	0.3493	0.5055		2.71	81.09	67738
24	4.48945	13.7148	0.5188		2.63	82.08	66983
25	1.4838	-9.2260	0.5279		2.62	82.26	66845
26	4.2827	-8.6416	0.4166	0.5327	2.62	82.26	66847
27	0.7076	0.8722	10.6304	0.5329	2.62	82.26	66847

Table II-7. Fit statistics (*RMSE, Fit index* (%), and *AIC*), variance function (δ), and autocorrelation (φ) and likelihood ratio statistics (*LR*) of the different form of crown ratio (*CR*) estimation model. OLS nonlinear (Eq. 28-31); logistic model (Eq. 32) mixed-effects models (Eq. 34-35); extended nonlinear models (ENM) (Eq. 36-38); and extended mixed-effects models (EMEM) (Eq. 39-41).

Parameters	*RMSE*	*Fit index (%)*	*AIC*	δ	φ	*LR*
Eq.(28)	0.07225	41.3	-33919			
Eq.(29)	0.07169	42.17	-34128			
Eq.(30)	0.07141	42.56	-34223			
Eq.(31)	0.07071	43.67	-34496			
Eq.(32)	0.07052	44.07	-34588			
Eq.(33)	0.07106	43.18	-34374			
Eq.(34)	0.07117	42.99	-35467			
Eq.(35)	0.07088	43.44	-35475			
Eq.(36)	0.07053	44.02	-34978	-0.246		390.74
Eq.(37)	0.07073	43.69	-37878		0.548	3291.8
Eq.(38)	0.07095	43.35	-38397	-0.364	0.55	3636
Eq.(39)	0.07089	43.44	-35726	-0.282		498.2
Eq.(40)	0.07081	43.57	-35703		0.129	230
Eq.(41)	0.07079	43.6	-35467	-0.277	0.127	720.7

Table II-8. Parameter estimates standard error in the parenthesis of the selected *CR* estimation models. OLS nonlinear (Eq. 28, 31, 32); mixed-effects model (Eq. 35); ENM (Eq. 36, 38) and EMEM (Eq. 41).

Parameters	Eq. (28)	Eq. (31)	Eq.(32)	Eq.(35)	Eq. (36)	Eq. (38)	Eq.(41)
b_0	0.26736 (0.00217)	0.17518 (0.00477)	0.16527 (0.00293)	0.149905 (0.00352)	0.16644 (0.00499)	0.26438 (0.00375)	0.15291 (0.00403)
b_1	3.24279 (0.08715)	1.23511 (0.04487)	1.19436 (0.03247)	0.971799 (0.0377)	1.258 (0.05096)	1.56297 (0.054)	1.02018 (0.04693)
b_2	0.98527 (0.0105)	-0.2342 (0.02597)	0.40621 (0.03076)	0.152942 (0.03284)	0.44583 (0.02861)	2.38424 (0.14721)	0.19512 (0.03973)
b_3		1.0726 (0.02635)	1.07876 (0.02026)	1.18193 (0.02442)	1.06131 (0.02732)	0.74909 (0.01253)	1.14689 (0.02515)
b_4				0.06721	0.16644 (0.00499)	0.26438 (0.00375)	0.16404
$\hat{\sigma}(\mu_j)$				0.01658			0.01654
$\hat{\sigma}(\varepsilon_{ij})$				0.06721	0.15588	0.19222	0.01654

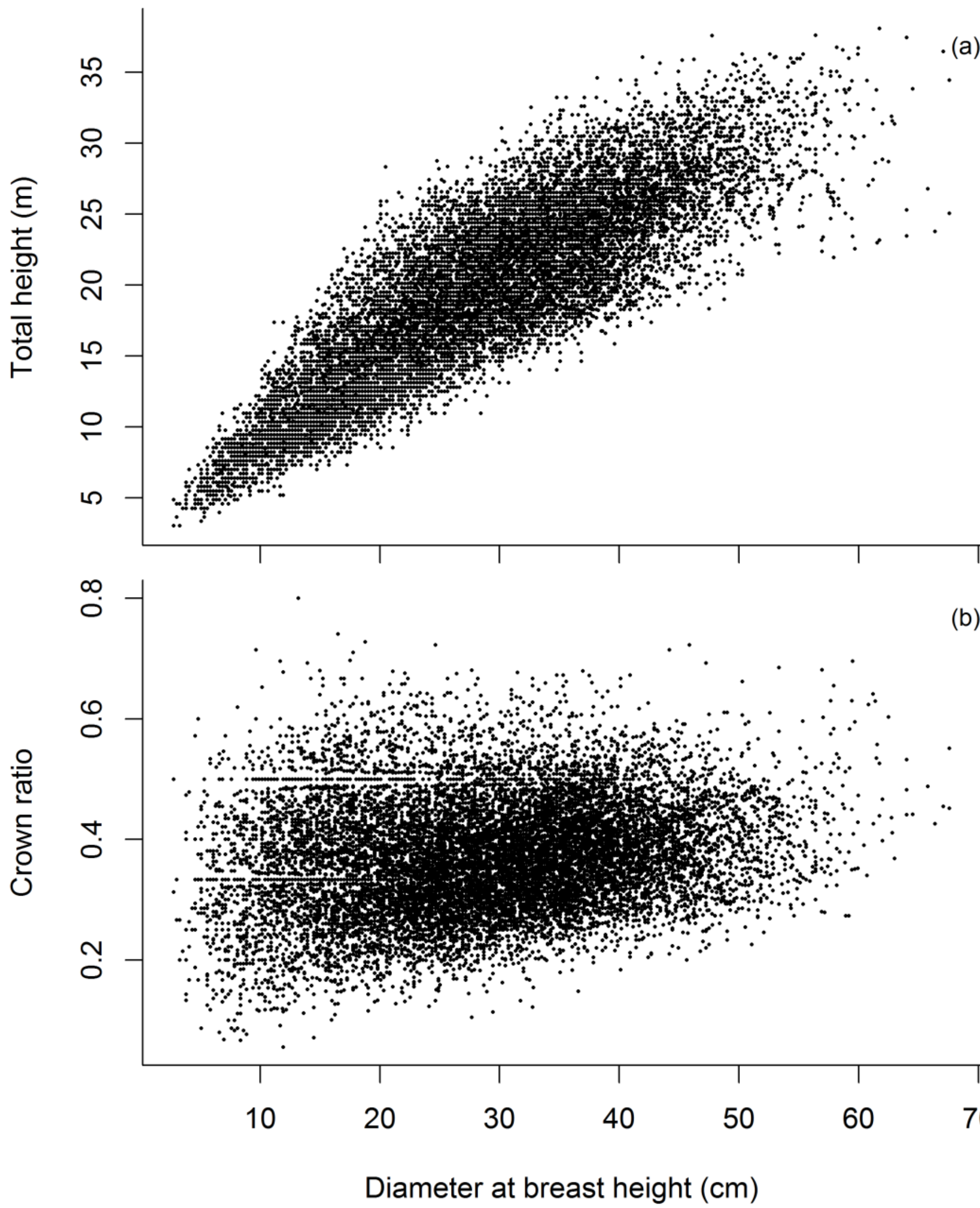

Fig. II- 1 Scatter plot of whole data showing the distribution of height (a) and crown ratio (b) of shortleaf pine along the diameter range.

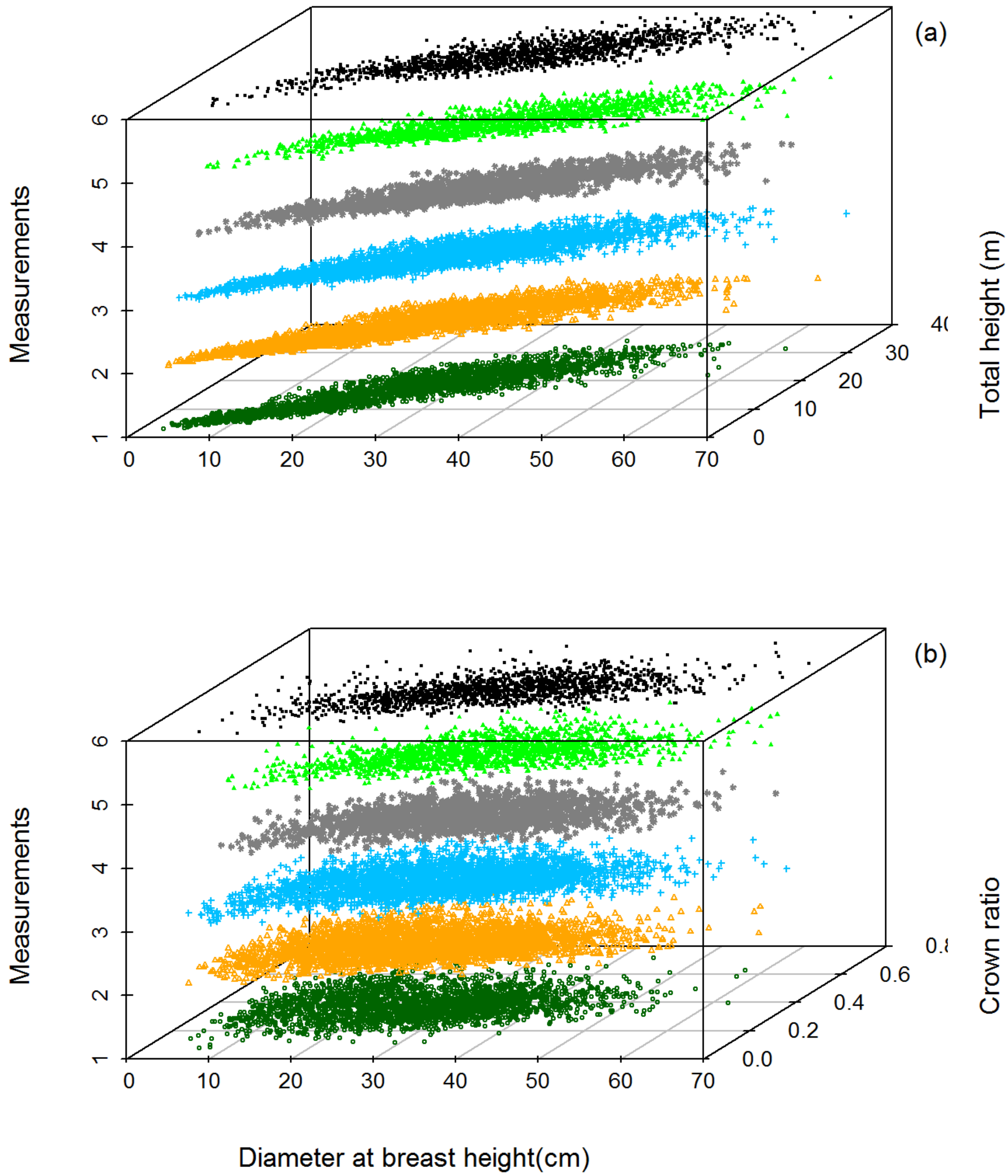

Fig. II-2 3D Scatter plot of each measurement the distribution of height (a) and crown ratio (b) of shortleaf pine along the diameter range. X-axis dbh, y-axis: total height, crown ratio, and z-axis-measurement time.

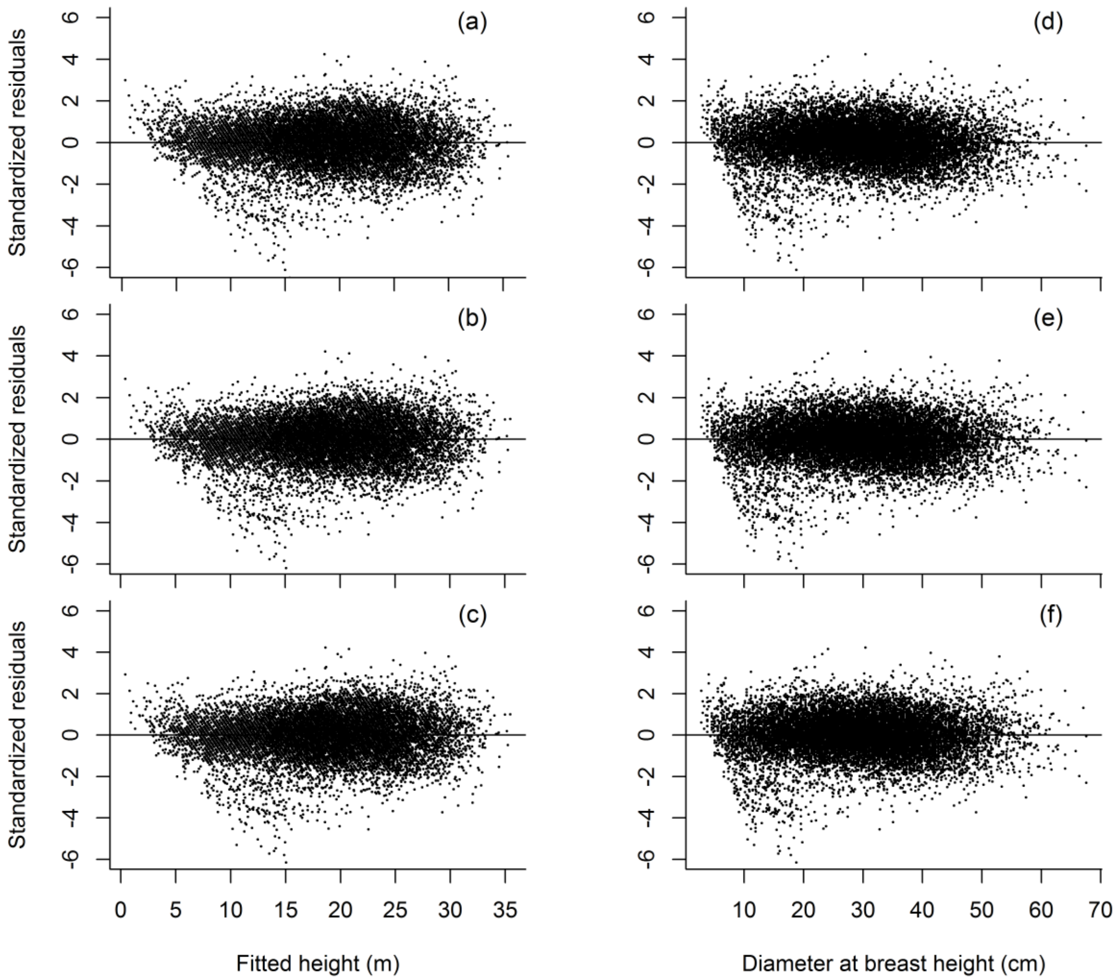

Fig. II-3 Scatter plot of standardized residuals vs. fitted values (left panel) and vs. dbh (right panel) for the total height prediction models: a) OLS nonlinear (Eq. 3); b) mixed-effects model (Eq. 5); and c) extended nonlinear model with AR (1) and power variance function (Eq. 8).

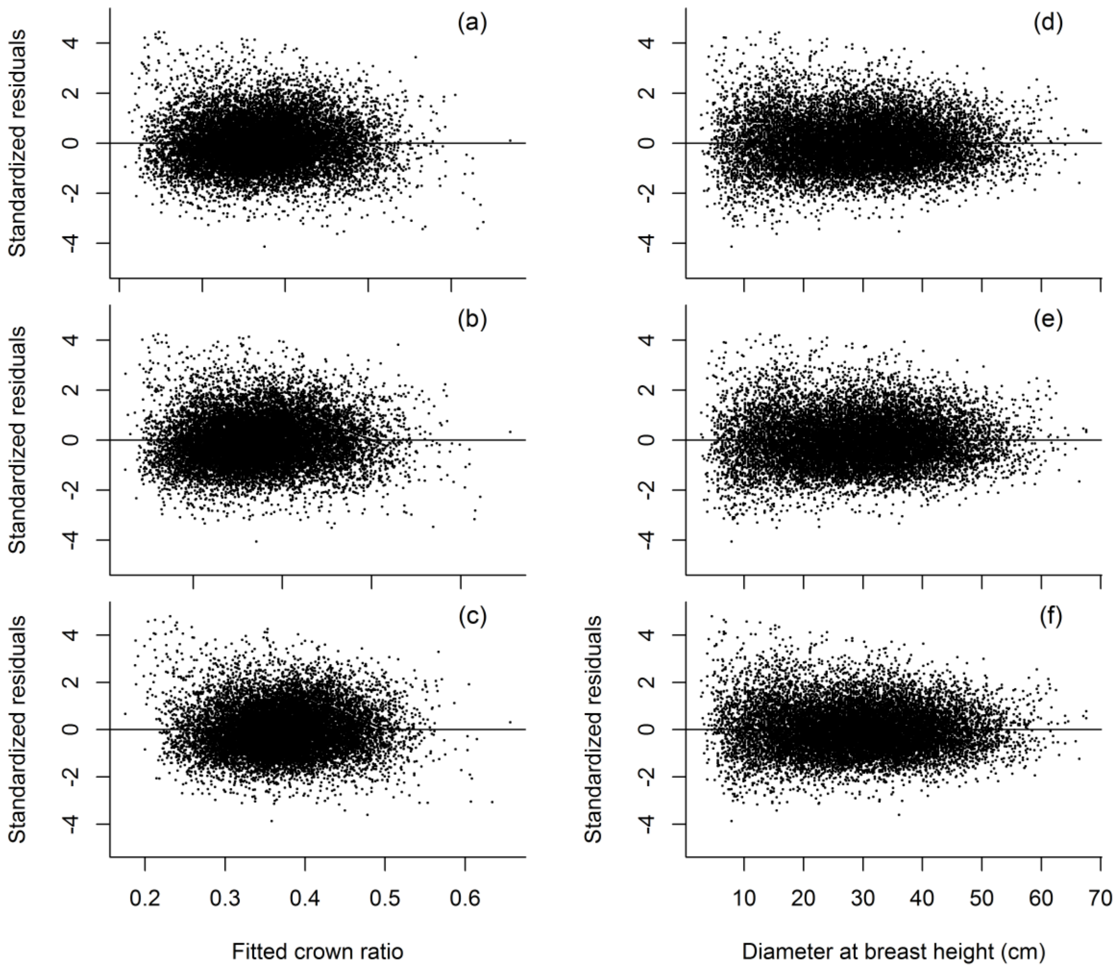

Fig. II-4 Scatter plot of standardized residuals vs. fitted values (left panel) and vs. dbh (right panel) for the crown ratio estimation models: a) OLS nonlinear (Eq. 32); b) mixed-effects model (Eq. 35); and c) extended nonlinear model with AR (1) and power variance function (Eq. 38).

CHAPTER III

MORTALITY MODELS FOR INDIVIDUAL SHORTLEAF PINE TREES BASED ON 25 YEARS OF REMEASUREMENTS DATA

Abstract

Mortality prediction models for individual trees were developed for naturally occurring even-aged shortleaf pine (*Pinus echinata* Mill.*)* stand. Data was collected from 208 permanently established plots located in the Ozark and Ouachita National Forests in Oklahoma and Arkansas, USA. The plots were re-measured six times over the period of 25 years. Re-measured data suggested that annual mortality rates were greater than 5% at the lower diameter (dbh) range (< 15 cm) and less than 1% for greater dbh ranges. Two models, with and without the effect of thinning, were fitted using the logistic function with a binary response (0 = alive, 1 = dead). Variables that measure competition indices with age: including the ratio of basal area per hectare to plot or stand age (*BAHG*), the ratio of individual dbh to stand age (*DAG*)), and without age: quadratic mean diameter (*QMD*) were found significant in predicting the probability of mortality. The logistic model estimated directly using a binary distribution ("binary logistic") provided smaller standard errors than the traditional iterative re-weighted regression, despite providing similar parameter estimates. Fitting a logistic model with thinning effect ("binary thinned"), with the addition of thinning variable (*THINHA*) showed inferior performance compared to the binary logistic model in chi-square (χ^2) test based on mortality in mid dbh classes, though it had more area under curve (AUC) value. The model fitted using nonlinear mixed-effects approach provided better Akaike information criteria (AIC) than binary logistic and binary thinned models but performed poorly in a χ^2 test.

Keywords**:** binary logistic, mixed-effects model, mortality rate, even-aged, thinning, competition indices

Introduction

Accurate and reliable growth and yield information is an important support for better forest management due to the prediction of future stand characteristics. It is important to obtain good estimates of probability of tree survival/mortality because it is a major factor in influencing forest growth and yield (Lynch et al. 1999; Yang et al. 2003) which also determines the changes in stand structures. Individual tree survival models are used to predict the survival of individual trees in a stand. Tree sizes, stand density, species composition, site quality, and competition are major factors that influence tree mortality (Peet and Christensen 1987; Temesgen and Mitchell 2005; Zhao et al. 2006; Cao and Strub 2008). Based on the nature of a mortality or survival model; predicting either probability or number of trees, the selected variables for a model differ. In either case, the stand level attributes are more appealing as candidate variables in a model in conjunction with the tree level attribute (height, diameter) because it measures the effect of competition experienced by a tree (Lynch et al. 1999; Zhao 2006; Crecente-Campo et al. 2010). Variables that measure competition (e.g. basal area of large trees) and relative position of a tree in a stand are also often found to be important covariates in mortality models (Monserud and Sterba 1996; Temesgen and Mitchell 2005; Cao and Strub 2008).

Annual survival equations predict the probability that a tree survives the following year. It can be challenging to obtain estimates of annual survival probability from data that were repeatedly measured at intervals longer than one year (Monserud 1976; Monserud and Sterba 1999; Cao 2000). However, most individual tree models require equations that predict the probability of tree survival on an annual basis. Logistic regression model is the most commonly used model to estimate the survival/mortality of individual trees (Monserud 1976; Cao 2000; Yao et. al. 2001; Zhao et al. 2004; Zhao et al. 2006; Cao and Strub 2008; Crecente-Campo et al. 2010; Groom et al. 2012). It has the flexibility to produce survival curves and estimate hazard rates for censored data (Efron 1988). Raising the logistic function to the power of the number of years in the measurement interval has often been used to generalize the logistic function for application to unequal measurement intervals (Yao et al. 2001). After the model is fitted in this way, the power can be set to 1 to predict the annual probability of survival/mortality. This procedure helps to overcome the problem of unequal measurement intervals (Yang et al. 2003; Cao and Strub 2008; Crecente-Campo et al. 2010).

Data for building mortality or survival models require repeated measurement of the permanent plots or several years of data collection. The mixed-effects modeling approach accounts for such structured and non-independent data. It includes plot-specific random effects to account for the heterogeneity that may occur due to the clustering of trees within a plot. Such statistical models that include both fixed and random effects of parameters which are associated nonlinearly to the response variable in the model have widely been used (Lynch et al. 2005; Lynch et al. 2012; Groom et al. 2012). The mixed-effects approach often helps to account for temporal and spatial correlation in the model. According to Lappi (2006) mixed models perform better when items are grouped within datasets. Grouped datasets may contain longitudinal or repeated measurements or can be defined as multilevel or block designs (Pinheiro and Bates 2000). However, mixed effects are limited to continuous static variables (e.g. height, dbh) but not often to a mortality; a binary, rare, and dynamic response (Rose et al. 2006; Groom et al. 2012).

Shortleaf pine is one of four major southern pines and an important species in southern US forests. It is economically important as a timber producing species as well an important component of wildlife habitat for species such as the red-cockaded woodpecker. Huebschmann et al. (1998) developed Shortleaf Pine Stand Simulator (SLPSS) for even-aged natural shortleaf pine forests. A survival model is one of the major components of the SLPPS that includes a prediction equation for the probability of tree survival based on repeatedly measured plots permanently located in the Ozark and Ouachita National Forests with diverse ages, site qualities, and densities. Other important components of SLPSS are a basal area growth model for individual trees (Hitch 1994; Lynch et al. 1999) and system of equations for height prediction and projection for shortleaf pine trees in even-aged natural stands (Lynch and Murphy 1995).

This study consists of a large data set repeatedly measured over a 25-year period and is an important source of information regarding the survival of individual trees of naturally occurring even-aged shortleaf pine forest system of Arkansas and Oklahoma. A previously published survival model by Lynch et al. (1999) analyzed only the first measurement period. Additionally, some plots were thinned after the third measurement. This provides an opportunity further to understand the influence of thinning on individual tree mortality. Thinning operations change the diameter growth rate of individual trees, but thinning from below in young stands may accelerate mortality rate if the residual stands experience thinning shock in combinations of insect or disease damage (Bailey et al. 1985). Post-thinning, a survival model, may not perform well for lower diameter classes (Bravo-Oviedo et al. 2006). On the other hand, it is possible that after thinning the survival probability of an individual tree might increase due to reduced competition for space and nutrients (Zhao et al. 2004).

The length of the study period (over 25 years) provided an opportunity to investigate mortality nature/trend, the potential influence of treatment and a variable corresponding to age in a mortality model covering wide ranges of tree size and age distribution. Considering the significance of the database and shortleaf pine in the forests of the southern US, it is important to enhance our understanding of shortleaf mortality models. This study is focused on developing a mortality model to predict the probability of mortality of an individual tree accounting for the treatment effects, e.g., thinning, and improvement our understanding of the significance of a variable in predicting mortality using re-measured data over the period of 25 years. Specifically, this study compares two different mortality models; one which includes a variable indicating the effect of thinning and another that does not. The prediction ability of the selected mortality model using parameter estimates obtained from the logistic regression with a Bernoulli distribution (this is termed a "binary" model in this paper), through iteratively reweighted nonlinear regression and mixed-effects modeling approach with a Bernoulli distribution will be compared. In addition to this, we will enhance our understanding of annual mortality rate (*AMR*) over the study period, and assess the influence of the selected variable in predicting the probability of individual tree mortality.

Material and methods

Data

The study plots are located in the Ozark and Ouachita National Forests in western Arkansas and southeastern Oklahoma. Until 1985, the major sources of data on the growth and yield of naturally occurring shortleaf pine forests were from fully stocked plots or unmanaged stands (Lynch et al. 1999). Because of this, during the period of 1985-1987, the Department of Forestry (now part of the Department of Natural Resource Ecology and Management) at Oklahoma State University (OSU) and USDA Forest Service Southern Research Station (USFS) at Monticello, Arkansas collaboratively established growth and yield plots in even-aged natural shortleaf pine stands. These plots were designed to represent a range of ages, densities and site qualities (for the detailed description see Murphy 1988; Lynch et al. 1999). Plots were thinned to specified residual densities at establishment and hardwood understory trees were removed using a chemical herbicide.

Plots have been re-measured in every 4 to 6 years, and the last (sixth) measurement occurred during the period from 2012 to 2014. Six measurements provided five measurement periods. The survival status of each tree was recorded at each measurement. Variables including diameter at breast height (*dbh*) (cm), tree height (*HT*) (m), and height to base of live crown (m) were recorded for each tree on each of the measurement plots. Each tree was classified either as dominant, co-dominant, intermediate, or suppressed. The sample plots were circular with the radius of 17.4 m (57.2 ft) and 0.0809 ha (0.2 ac). The measurement plots are surrounded by a buffer strip of 10 m (33 ft) wide that received the same silvicultural treatments at the establishment as the measurement plot. The total sample consisted of 208 plots. OSU-USFS established 183 plots of these, and 25 plots were from a thinning study established by Frank Freese during 1963-1964. In 1988, the Freese study plots were balanced to be consistent with the basal area levels of OSU-USFS shortleaf pine growth study (for the detailed description see Murphy 1988; Lynch et al. 1999).

An ice storm impacted the study area just before the fourth measurement. In this study, the 101 ice damaged plots from the third and subsequent measurements period were excluded. Many plots were re-thinned to their original basal area levels shortly after the third measurement while a portion were left unthinned. The thinned basal area (m2ha-1) was deducted from the estimated basal area per hectare of the third measurement because this better reflects the competitive pressure experienced by the trees on the plot during the measurement interval subsequent to thinning. The average thinned basal area was 6.94 m^2ha^{-1} with a range of 0.69 – 19.36 m^2ha^{-1}. The total numbers of trees available for mortality analysis at the beginning of each of the five measurements were 8288, 8078, 4027, 2470 and 2355. The summary of the tree level and stand level attributes associated with the dead trees for each measurement period are shown in Table 1.

Model Development

Logistic regression model

In our data, the response variable '*y*' represents mortality for an individual shortleaf pine tree, and has values of 1 for trees that died in between the measurement interval or 0 for trees that survived during that time with probabilities p_i and $1-p_i$

respectively. Since y_i (an individual tree) is an independent Bernoulli random variable with parameter $E\{y\} = p$, the simple logistic regression form was used.

$$E\{y_i\} = p_i = \frac{e^{\beta_0+\beta_1 x_1+......+\beta_n x_n}}{1+e^{\beta_0+\beta_1 x_1+......+\beta_n x_n}}$$

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_1 + ... + \beta_n x_n \qquad (1)$$

Hamilton and Edwards (1976) modified Eq. (1) to Eq. (2) which he used to fit the parameters for a logistic model to estimate the probability of mortality.

$$p = \left[1+\exp\left\{-\left(\beta_0 + \beta_1 x_1 + ... + \beta_n x_n\right)\right\}\right]^{-1} \qquad (2)$$

where p is the probability of annual mortality, $x_1, x_2, \ldots, x_n$ are set of n predictors, $\beta_0, \beta_1, \ldots, \beta_n$ are regression parameters to be estimated and exp is the base of the natural logarithm. Later, many researchers set p to the annual probability of survival (e.g., Monserud (1976)). Eq. (2) provides the probability of mortality rather than the probability of survival. The probability of survival would be $1-p$ because, by definition, p is bounded by 0 and 1.

Variable selection

For preliminary screening of variables, SAS/ PROC LOGISTIC procedure was used to select the best set of predictor variables at α = 0.05 level of significance. It is a typical logistic regression model to fit data with dichotomous outcomes by the method of maximum likelihood (SAS Institute Inc. 2007; Allison 2001). Additionally, selected variables were also tested for their ideal performance to predict either mortality or survival when response label of the binary variable is interchanged. The attributes of dead trees at the tree and stand level shown in Table 1 were used as candidate variable in the model building process. The ratio of stand basal area per hectare to plot age (*BAHG*), the ratio of dbh to plot age (*DAG*), and quadratic mean diameter (*QMD*) were found highly significant with *p-value* < 0.0001, and more promising than other variables. Summary statistics of the selected variables are given in Table 2. These variables measure relative stand level competition with age, hierarchical position of an individual tree with age, and stand level competition respectively. For most plots, only a subsample of trees was selected for measurement of total height and crown length. Therefore, height prediction and crown ratio estimation models for shortleaf pine of Saud et al. (2015 submitted) were used to predict the missing measurements for testing independent variables that use height and/or crown ratio.

In the model building process, an interaction terms were also included. Though the interaction terms were highly significant (*p-value* < 0.0001), the performance of the model with interaction terms was not satisfactory in terms chi-square (χ^2) goodness of fit test value. Therefore, models with interaction term are not discussed here. To represent the possible effects of thinning at stand level in the model (*THMD*) the following variable was created: *THINHA* = (*Thinned basal area per hectare/ (years since thinning*)). Because thinned basal area is divided by the number of years since thinning, it causes the effects of thinning to be reduced with time. In the thinning model, all above mentioned variables were found significant along with *THINHA* and these variables were used for with and without interaction term in alternative models.

Binary logistic regression model and iteratively re-weighted regression model

Data from re-measured plots include multi-year intervals in which tree survival or death is observed. Let '*t*', a time period in years, be such an interval over which to observe tree dead or survived. Hamilton and Edwards (1976), Monserud (1976), and Monserud and Sterba (1999) described a model (Eq. 2) to estimate the probability of survival assuming that survival time follows uniform distribution over the growth interval. The model of Eq. (2) was modified to obtain Eq. (3), 'binary logistic regression model', to predict mortality probability rather than survival probability where the response variable is mortality. This formulation is more consistent with biomedical studies (right censored data), and in related work we want to compare these methods to methods from the biomedical literature. Flewelling and Monserud (2002) stated that mortality is not a Markov process because a tree can only die once. However, survival is a Markov process. Therefore "this property requires that all algebra be mediated in terms of survival, not mortality" (Flewelling and Monserud 2002). Because of this, we modeled mortality (with tree death = 1) as 1-*p* (tree survives to the end of the measurement period).

$$p_j^t = 1 - \left[1 + \exp\left\{-\left(\beta_0 + \beta_1 BAHG_j + \beta_2 DAG_j + \beta_3 QMD_j\right)\right\}\right]^{-t} \quad (3)$$

where p_j^t is the probability that tree *j* died during the measurement interval of *t* years, and the β_i are parameters to be estimated.

The parameter estimates obtained from the Eq. (3) are identical to the estimates obtained when survival was modeled (live tree =1) using Eq. (2). Hence the Eq. (3) modeled mortality but provides parameter estimates of survival. Eq. (3) was used to predict the mortality probability in different time intervals. When the interval *t* is zero or at the beginning of growth period the probability that the tree survives is *1* which means the tree is definitely alive. Conversely, as $t \rightarrow +\infty$, the probability of survival decreases and eventually approaches zero (Yao et al. 2001). The model with a thinning effect includes the variable *THINHA* in Eq. (4).

$$p_j^t = 1 - \left[1 + \exp\left\{-\left(\beta_0 + \beta_1 BAHG_j + \beta_2 DAG_j + \beta_3 QMD_j + \beta_4 THINHA_j\right)\right\}\right]^{-t} \quad (4)$$

where, the notation are as described above for Eq.(3).

Eq. (3) and Eq. (4) were used in what we term here as “binary logistic model” and “binary thinned” by specifying response as binary with probability of p^t. The NLMIXED procedure in SAS (SAS Institute Inc. 2007) was used to fit those models where no random effect was specified.

PROC NLIN is the major SAS procedure for nonlinear (or curvilinear) regression analysis (SAS Institute Inc. 2007). It fits nonlinear regression models and estimates the parameters by nonlinear least squares or weighted nonlinear least squares (SAS Institute Inc. 2007). Since mortality is a binary or Bernoulli random variable, it has where p^t is the probability of mortality during the period of *t* years. The inverse of the variance i.e. *1*/ (p^t

($1-p^t$)) was used as a weight while fitting the nonlinear regression. McCullagh and Nedler (1989) recommended iteratively re-weighted regression as an effective procedure to find the maximum likelihood estimates for the mortality model. Both Eq. (3) and (4) were used for the iteratively reweighted nonlinear regression model and the resulted in Eq. (5) and Eq. (6).

Eq. (3) + Weight Eq. (5)

Eq. (4) + Weight Eq. (6)

Mixed- effects model

Mixed-effects models were also fitted to develop a model for the probability of mortality for shortleaf pine. The NLMIXED procedure in SAS (SAS Institute Inc. 2007) was used to fit the nonlinear mixed-effect models by specifying the response as a binary variable. We investigated random effects associated with plot, additive to the intercept parameter, as well as to the parameters associated with all variables of the model. The best results were obtained by adding random effect to the parameter associated with variable *QMD* for without thinning effect "binary mixed" model (Eq. 7).

$$p_{jk}^{t} = 1 - \left[1 + \exp\left\{-\left(\beta_0 + \beta_1 SBAG_{jk} + \beta_2 DAG_{jk} + (\beta_3 + \mu_k) QMD_{jk}\right)\right\}\right]^{-t} + \varepsilon_{jk} \quad (7)$$

where β_0, β_1, β_2, β_3, and β_4, are parameters to be estimated for tree j on plot k, u_k is random effect associated with plot k normally distributed with mean 0 and variance σ^2_u

and ε_{jk} is error term with mean zero. The other alternative binary mixed models are mentioned as; Eq. (8-10).

Eq. (7) with random effect at intercept: $(\beta_0 + \mu_k)$ Eq. (8)

Eq. (7) with random effect at BAHG: $(\beta_1 + \mu_k)BAHG_{jk}$ Eq. (9)

Eq. (7) with random effect at *DAG*: $(\beta_2 + \mu_k)DAG_{jk}$ Eq. (10)

In the mixed-effects model with thinning effect "binary mixed thinned", the random effect of the plot associated with variable THINHA (Eq. 11) was found better over the alternative binary mixed thinned models.

$$p_{jk}^{t} = 1 - \left[1 + \exp\left\{-\left(\beta_0 + \beta_1 BAHG_{jk} + \beta_2 DAG_{jk} + \beta_3 QMD_{jk} + (\beta_4 + \mu_k)THINHA_{jk}\right)\right\}\right]^{-t} + \varepsilon_{jk} \qquad (11)$$

where notation are as described above. And the alternative models are shown as; Eq. (12-15).

Eq. (11) with random effect at intercept: $(\beta_0 + \mu_k)$ Eq. (12)

Eq. (11) with random effect at *BAHG*: $(\beta_1 + \mu_k)BAHG_{jk}$ Eq. (13)

Eq. (11) with random effect at *DAG*: $(\beta_2 + \mu_k)DAG_{jk}$ Eq. (14)

Eq. (11) with random effect at *QMD*: $(\beta_3 + \mu_k)DAG_{jk}$ Eq. (15)

Mortality rates and variable influence in mortality

We estimated annual mortality rate of shortleaf pine for each measurement period and overall the period of 25 years of re-measurement. The annual rate of tree mortality was estimated using Eq. (16), a negative compound interest formula (Hamilton & Edwards 1976).

$$AMR = 1 - \left\{ \left(\frac{N_t}{N_0} \right)^{1/t} \right\} \qquad (16)$$

where *AMR* = annual mortality rate; N_t = number of trees survived at re-measurement time (t = 4 - 6 years); and N_0 = number of trees at previous measurement).

To examine the trend of tree mortality along the tree size, *AMR* were plotted against the dbh classes having widths of five centimeters. The dbh classes were labelled by the midpoint; for example: 2.5 cm = ≤ 5 cm; 7.5 cm = >5 to ≤ 10 cm; ….. ; and 42.5 = > 45 cm. The predicted probability of mortality by the selected model was plotted against the variables used in the model at their mid-point ranges to compare the distribution of predicted probability of mortality across the variables. The ranges of mid-points were similar for the both variables *BAHG* and *DAG*. It was of 0.20 for all classes, except the uppermost class; 1.1 = >1.0. The ranges used for mid-points for *QMD* were: ≤ 15; >15 to ≤ 23; >23 to ≤ 31; >32 to ≤ 39; >39 to ≤ 47; and >47 cm. It was used for *THINHA* as: ≤ 1; >1 to ≤ 5; > 5 to ≤10; >10 to ≤15; and > 15 m^2ha^{-1}. Similarly, the influence of a variable in annual survival probability was examined by plotting the predicted probability of annual survival by a model against the selected variables. The value of the selected variable ranged beyond the data set while all other variables were held constant at their mean values for the data set.

Model evaluation and accuracy

The Akaike information criterion (AIC) was used to select the best of the alternative models (Pinheiro and Bates 2000). The AIC and log-likelihood were estimated as given below (Eq.17 and 18) where *d* is the number of parameters in the model and *l* is the log-likelihood:

$$AIC = -2\log(l) + 2d \quad (17)$$

$$l = \sum_{i=1}^{n}\left[y_i \log_e p_i + (1 - y_i)\log_e(1 - p_i) \right]$$

Once the fitted response function was obtained, a chi-square goodness-of-fit test was conducted to check the fitness of the response function. The χ^2 test is often used to evaluate the appropriateness of the model that has a binary response variable (Neter and Maynes 1970). After obtaining parameter estimates, comparisons between observed and predicted numbers of live and dead trees were made by using mid-dbh classes. The models having the lowest AIC values were evaluated using the χ^2 goodness-of-fit test (Eq. 18), based on these diameter classes. The χ^2 value by mid-diameter classes was calculated as:

$$\chi^2 = \sum_{j=1}^{c}\sum_{k=0}^{1}\frac{\left(O_{jk} - E_{jk}\right)^2}{E_{jk}} \quad (18)$$

where O_{j1} and O_{j0} are an observed number of mortality and survival trees, respectively in diameter class *j*. Similarly, E_{j1} and E_{j0} are numbers of dead trees and surviving trees respectively in diameter class *j* expected from the model and *c* is the number of diameter classes.

Model prediction accuracy was evaluated using the principal of the Receiver Operating Characteristic (ROC) curve (Metz, 1978). ROC analysis was conducted in R suing ROCR package (R Core Team 2012; Sing et al. 2005). The ROC curve is often used to measure the accuracy of a logistic model form, but it is less commonly used practice in forestry (Crecente-Campo et al. 2010; Hein and Weiskittel 2010; Groom et al. 2012). It provides a measure of model discrimination by showing the area under the ROC curve (AUC) but shows the tradeoff between sensitivity and specificity. Sensitivity is the true positive rate, the rate of probability to be predicted positive given that a subject or an individual is positive. Specificity is the false positive rate, the rate of the probability to be predicted positive given that a subject or an individual is negative. The false positive rate is also denoted as 1-specificity.

Results

Mortality

The mortality of shortleaf pine was 13.67% of initial total sample population (8288 trees), from the plot establishment year 1985-85 till the last sixth measurement period, with an average of 3.73 % periodic mortality rate (4-6 years). The proportion of tree mortality, 2.44, 6.21, 6.63, 4.70 and 1.95% for the consecutive five measurement periods (Table 1) was significantly different ($df = 4, \chi^2 = 215.91, p < 0.001$).

The reversed J-shaped of *AMR* curve along the mid-dbh class of all measurement periods and of overall period of 25 years indicate that *AMR* was high for the smallest trees and declined rapidly as dbh increase and remain stable for large dbh class (Fig. 1). However, the *AMR* curve of the fifth measurement period was not typically reversed J-shaped because of lower *AMR* at the lowest mid-dbh class (12.5 cm). Overall the whole period of 25 years of remeasurement data suggested that an average *AMR* for the trees of lower mid-dbh class (7.5 cm) was very high (> 10%), but it was less (on an average of 1%) for trees belonging to medium-sized (17.5 cm) to large-sized dbh class (Fig. 1).

Annual mortality rate based on measurement period was 1.19, 5.68, 2.34, 1.98 and 1.26 %, but it was 3.54% over the period of 25 years. It was higher for all mid-dbh classes during the third measurement period than other periods (Fig. 1). High *AMR* was observed for the smallest mid-dbh class (2.5 cm) till the second measurement period, but it was not recorded in the later measurement periods because of shift of dbh in the upper mid-dbh class (Fig. 1). However in the third measurement period, *AMR* was 100% for the lower 7.5 cm mid-dbh class because of a single observation.

A high frequency of dead trees were observed in the lower dbh range in each measurement period, but in the second measurement period it was greatest at the lowest dbh distribution (Table 1), which could be due to the effects of competition as density increased. But the stand level competition in the third measurement was reduced due to thinning. In the second measurement period, 290 dead trees were recorded in 7.5 cm class and 79.3% of that mortality occurred in the young plot age classes (27-30 years) with densities greater than 2,200 trees per hectare.

Binary logistic regression model and iteratively re-weighted regression model

The performance of binary logistic (Eq. 3) model and re-weighted regression model (Eq. 5) was identical (Table 3). It was observed because the highly significant ($p <$ 0.0001) parameter estimates were all similar for both models (Table 4). This was as expected since these are simply two different ways of solving the same likelihood function. However, standard errors and the confidence intervals were slightly different (Table 4). The standard errors for the parameter estimates of binary logistic model (Eq. 3) were smaller by an average of 33% making the confidence interval narrower than that of re-weighted regression model (Eq. 5).

As mentioned above, parameter estimates of binary thinned model (Eq. 4) and the re-weighted regression models with thinning effect (Eq. 6) were also similar and highly significant ($p < 0.0001$) (Table 4). The standard error for the parameter estimates of the binary model with thinning effect (Eq. 4) were also found almost 37% smaller than of the iteratively re-weighted regression model form with thinning effect (Eq. 6). Though both approaches provided similar fit statistics (Table 3) because fit statistics involves parameter estimates of a model, the binary logistic model approach provided smaller standard errors for parameter estimates. Therefore, the binary model was preferred over the iteratively re-weighted regression estimates.

Parameter Estimation Using a Nonlinear Mixed Model

In the binary mixed model, the random effect of plot associated with variable *QMD* (Eq. 7) was found better than other alternatives (Eq. 8-10) when fit statistics were compared (Table 3). All parameter estimates of the binary mixed model (Eq. 7) were highly significant ($p < 0.0001$) and are shown in Table (4). In binary mixed model, the random effect of plot associated with intercept (Eq. 8) and with the parameter of *BAHG* (Eq. 9) had comparable AIC value with Eq. (7), but the AIC value of the Eq. (10) was quite large (Table 3). Eq. (7) provided better χ^2 test value for prediction, so it was selected as the best binary mixed model. The variance component (σ^2_{uk} = 0.0024) associated with the parameter b_3 of *QMD* was significant with a 95% confidence interval of [0.0017, 0.0033]. It was found that the parameter b_1 of *BAHG* was highly positively correlated (0.82) with b_3 of *QMD*. The correlation and covariance matrix of the binary mixed model (Eq. 7) is shown in Table (5).

In binary mixed thinned models, all parameter estimates were highly significant except the random effect of plot associated with the parameter (b_2) of DAG (Eq. 14) (Table 3). The AIC value of the binary mixed thinned model, Eq. (11) was similar to the random effect of plot associated with the parameter of *BAHG*, (Eq. 13) (Table 3). Likewise, the AIC value with random of plot effect associated with the parameter of intercept (Eq. 12) and *QMD* (Eq. 15) was similar but smaller compared to former models (Table 3). However, the random effect of plot associated with the parameter of, *THINHA* (Eq. 11) provided better χ^2 test value for prediction than the models with lower AIC values (Table 3). Therefore Eq. (11) was selected as the best binary mixed thinned model and parameter estimates are shown in Table 4. The variance component (σ^2_u = 0.1518) associated with parameter b_4 of *THINHA* was significant with a confidence interval of [0.0977, 0.2060]. It was also found that the parameter b_1 of *BAHG* was positively correlated (0.85) with b_3 of *QMD*. The correlation and covariance matrix of the binary mixed thinned model (Eq. 11) is shown in Table (5).

Mortality prediction

The chi-square test statistic has two components, one based on observed minus expected mortality, and the other based on observed minus expected survival. However, individual survival and mortality components were sometimes of a more modest magnitude. The estimated χ^2 goodness of fit test value for both models (Eq. 3 and 4) was less than tabulated ($\chi^2_{(9, 0.05)}$ =16.91) for survival prediction in mid-dbh class but it was not for mortality prediction (Table 3). The total χ^2 test value was 24.62 for the binary logistic model (Eq. 3) with χ^2 = 1.57 for survival prediction, and χ^2 = 23.05 for mortality prediction in mid-dbh class. Similarly, the total χ^2 test value was 33.02 for binary thinned model (Eq. 4) with χ^2 = 2.01 for survival prediction, and χ^2 = 30.01 for mortality prediction in mid-dbh class (Table 3).

The χ^2 test value was larger for the both mixed-effects model approaches of binary mixed, and binary mixed thinned model than for the binary models (Table 3). In the binary mixed model, Eq. (7) had a total χ^2 test value of 52.35, with χ^2= 3.31 for survival prediction and χ^2 = 49.04 for mortality prediction for dbh mid class (Table 3). In the binary mixed thinned, Eq. (11) had a total χ^2 test value of 61.60, with χ^2= 2.54 for survival prediction and χ^2 = 59.06 for mortality prediction for dbh mid class. However, the total χ^2 test value was substantially larger for the alternative binary mixed thinned models (Table 3). For example; the total χ^2 test value was 88.34 and 255.54 for the Eq. (13) and Eq. (14) respectively.

Both binary logistic model (Eq. 3) and binary thinned model (Eq. 4) showed a better fit of the predicted mortality curve with observed mortality than the mixed-effects models (Fig. 2a). The percentage difference in observed and predicted mortality for each mid-dbh class showed that binary logistic model provided marginally better fit in predicting mortality than binary thinned model (Fig. 2b; Table 6). In mixed-effects model approach, binary mixed model, Eq. (7) appeared to have greater percentage differences in mortality prediction than binary mixed thinned, (Eq. 11) (Fig. 2b). All significant models (Eq. 3, 4, 7 and 11) under predicted mortality for dbh mid classes at 2.5 cm, and 7.5 cm (Fig. 2b; Table 6). The total number of predicted the mortality trees was nine trees greater than observed dead trees for the binary logistic model and eight trees greater than observed for binary thinned model (Table 6). However, the total number of predicted tree mortality was 161 trees fewer than observed dead trees for the binary mixed model but it was 15 trees greater by binary mixed thinned model (Table 6).

Model accuracy

The reasonable AUC value of all models indicated a good ability to discriminate the mortality risk for an individual tree of even-aged shortleaf pine growing in a natural forest stands (Table 3). AUC showed that a model while assigning 55% of true positive rate (true dead trees) would assign 10 % of false positive rate (Fig. 3). Examination of the ROC curve showed AUC for the selected models with thinning effect were better than models without thinning effect (Table 3 and Fig. 3). It was observed that mixed-effects models with comparable AIC value (Eq. 8 and 13) with the selected mixed-effects models (Eq. 7 and 11) had comparable AUC (Table 3). However, in the mixed thinned model, Eq. (13) showed better AUC than the selected model, Eq. (11). Table 3 and Fig. 3 indicated that the binary mixed model (Eq. 7) was slightly better than the binary logistic model. However, the binary thinned model was slightly better according to the AUC criterion than the binary mixed thinned model (Eq. 11) and was the best model by a small margin according to the AUC criterion (Table 3).

Influence of variables for predicting mortality

The relative influence of variables used for mortality prediction was similar in all of the models tested both those with a thinning effect and those without a thinning effect. Fig. 4 illustrates the behavior of mortality predictions for classes of several of the important variables used in modeling. The predicted probabilities were from the binary logistic model (Eq. 3) except *THINHA* from the binary thinned model (Eq. 4). The probability of mortality increased with the increasing values of stand level competition; *BAHG, QMD* and *THINHA* (Fig. 4a, 4c, and 4d.), while the probability of mortality decreased with increasing relative hierarchical position of individual level tree variables with age, *DAG* (Fig. 4b).

The change in the risk of mortality was very small for modestly increasing values of *BAHG, QMD, and THINHA*. The predicted mortality over the measurement length period showed that the mortality risk was nearly 5% for the median of the population of trees belonging to the uppermost mid-class for *BAHG* (Fig. 4a), *QMD* (Fig. 4c), and *THINHA* (Fig. 4d). The influence of the thinning variable in mortality could be possible due to effects of logging. Similarly, the mortality risk was 5% for the median of the population of trees belonging to lower mid class of *DAG* (Fig. 4b). The influences on predicting mortality by each variable could be greater for some individual trees than shown in Fig. 4 because some of the predicted mortalities were beyond the upper whisker of the boxplot for all variables.

Figure 5 shows the graphs of predicted probability of annual survival on the change in a selected individual variable while all other variables were held constant at their mean values for the dataset. The reference of all variables for Fig. 5 is as discussed above for Fig. 4. It appeared that *DAG* has the greatest impact on annual survival probability over the range plotted (Fig. 5a). Although *BAHG*, *QMD,* and *THINHA* showed a very small change in survival probability for increasing values, *BAHG* showed (Fig. 5b) relatively smaller decreasing annual survival probability compared to the other two (Fig. 5c and 5d). The ratio of diameter to stand plot age greater than 0.3 (cm $yr^{-1)}$ showed stable and better survival probability for an individual tree (Fig. 5a). The increased value of the variable *BAHG* (> 1.4 m^2 ha^{-1}/yrs), and *QMD* (> 50 cm) are likely to increase the stand level competition and influence survival of a suppressed tree by 5% (Fig. 5b and 5c).

Discussion

Mortality is one of the most critical and fragile components of growth and yield because changes in forest dynamics over time cause major increases or decreases in mortality rate. Also, mortality can have a large effect on measured productivity since a dead tree counts as negative growth. Since growth and mortality are known to alter with time a predicting variable that corresponds to the age of a stand was a rational choice for mortality. The selected variables not only measure relative stand level competition and hierarchal position of an individual with age but also stand level competition that reflects the site productivity. Additionally, the selected variables (*BAHG, DAG*, and *QMD*) can be easily constructed with from dbh and ages measurements that are commonly available data for even-aged stands. Ideally, these variables were also highly significant in predicting survival probability when the binary response label was interchanged from “1” to “0” for dead trees.

A variety of different variables for explaining either mortality or survival has been used by Lynch et al. (1999), Monserud and Sterba (1999), Cao (2000), Zhao et al. (2004), Zhao et al. (2006), Cao and Strub (2008), and Crecente-Campo et al (2010). We also tested a variety of candidate variables (Table 1), not all of which were discussed above, but several were discarded due to their poor performance in a model. There were sets of models provided by stepwise logistic regression which had better AUC and AIC value in the binary logistic model form than the models finally selected, but they performed poorly in the chi-square goodness of fit test (e.g. Eq. (12, 15) in Table 3). Examples of these included binary logistic model with variables a) *CRT*, *RAQD*, *BAHA,* and *DAG*, and b) *HDR*, *BAHA*, *QMD,* and *DAG*.

The selected variable *QMD* was very sensitive compared to other variables in estimating the probability of mortality of an individual tree because it could have either an increasing or decreasing impact depending on the competitive environment as well as other aspects of stand dynamics. Such effect was observed in a simple logistic model with *QMD* and *BAHG* that showed an increase in *QMD* has the effect on mortality (odds ratio estimate =1.05), while in a model with *QMD* and *DAG* showed the null effect of *QMD* on mortality (odds ratio estimate = 0.99). So, it was inferred that this sensitive variable is important for model performance.

The fact that the χ^2 test value of binary logistic model was better than the binary thinned model suggested that deducting thinned basal area (m^2ha^{-1}) or the residual stand basal area ($m^2 ha^{-1}$) could show the effect of reduced stand level competition rather than using a simple variable to represent the thinning effect in the model. It could be due to the variability in *AMR* curve level at lower dbh range (≤ 12.5 cm) and *AMR* curve below than 1% at upper dbh ranges (≥ 17.5) for all measurement periods (Fig. 1) indicated that thinning or reduced stand basal area affected only the *AMR* of trees at lower dbh ranges. The observed *AMR* curve at the higher level for the third measurement was the effect of the shorter remeasurement interval i.e. length of four years, recorded for the majority (2/3rd) of dead trees. Using Eq. (16); for the fixed proportion of dead trees; shorter the interval, larger the *AMR*; and longer the interval, smaller the *AMR*.

Annual mortality rates over the period of 25 years for mid-dbh class showed a reserved J shaped curve (Fig. 1). However, the mortality curve (Fig. 2a and 2b) over the period is not either U shaped or reversed J shaped curve, which are often presumed to be the main alternative relationships for mortality trends (Coomes 2003; Woodall et al. 2005; Westphal 2006). This was as a consequence of the fact that the highest mortality occurred at lower mid-dbh classes (Table 6) during 2nd and the 3rd measurement period (Fig. 1). Mortality was almost 9.2% for multi-year measurement intervals from the 7.5 and 12.5 cm mid-dbh classes though subsequently increasing dbh classes had a decreased the mortality rate. A similar pattern of high mortality rate (7%) at lower dbh class (5-15 cm) was observed at ten years after planting shortleaf pine (Dipesh et al. 2014). This tends to confirm that it is possible for high spikes in mortality to occur in lower dbh classes (Fig. 2a, and 2b). If the mortality from the lower mid-dbh class is ignored or merged into classes of wider intervals (10 cm) then Fig. 2a and 2b could display reversed a J-shaped curve. Another factor that may influence the mortality trends in this study is the fact that at the beginning of the study there were many “young-aged” plots containing large numbers of small trees per hectare. As these plots approached competition-induced mortality densities, significant numbers of mortality trees in mid-range dbh classes probably occurred.

In alternative binary mixed thinned models, the random effect associated with intercept, and *QMD* had smaller and similar AIC values. However, their performance in the χ^2 test was poorer than with a random effect associated with *THINHA* (Eq. 6) (Table 3). It appeared from the χ^2 test results that the mixed-effects modeling approach was not satisfactory for a natural stand with longer re-measurement intervals. This approach may be a better fit for survival prediction with shorter re-measurement intervals with plantation tree species as in Rose et al. (2006). A similar study in which the mixed-effects model did not provide compelling evidence of better fit was reported recently by Groom et al. (2012) for predicting Douglas-fir mortality. Groom et al. (2012) found that AUC of the mixed-effects model with only fixed parameter estimates was not different from binary logistic regression model. Unlike AIC value, all tested binary models and mixed-effects models of both with and without thinning effect also did not show clearly a distinguishable pattern in AUC (Table 3). It can be observed that except for a few mixed-effects models, the binary logistic and binary thinned models had a better-discriminating ability. However, the discriminating ability can be optimized by increasing default cutoff relative to the standard point of 0.5 (Hein and Weiskittel 2010).

There can be an apparent conflict between using the AIC criterion and the χ^2 goodness of fit test used to evaluate model performance. The mixed model appeared to be better than the binary model using the AIC criterion but did not perform well in the χ^2

goodness of fit test. Part of the reason may be that we evaluated probabilities for the mixed model using the fixed parameter estimates only to conduct the chi-square test. However, the AIC evaluation, of course, includes the random-effects parameters. In typical applications of the model, the random effect for a particular forest would not be known. In other mixed model applications to forest resources (for example; dbh-height relationship, e.g. Lynch et al. 2005; Lynch et al. 2012) one often may be able to calibrate a mixed model by estimating random parameters with locally-obtained data, but this would usually be very difficult for applications of a mortality model because re-measured plots would be required, but are typically not available. Furthermore, there is some evidence in the literature of forest resources that indicates that OLS models without mixed-effects may be preferred to the use of mixed-effects models if calibration using locally-available data cannot be applied (Monleon 2003; Lynch et al. 2005; Lynch et al. 2012). However, these latter studies did not address mortality models specifically.

The high χ^2 test values for binary models suggest that the model is predicting a significantly different number of dead trees than were observed for one or more mid-dbh classes (Table 6). The estimated χ^2 test value for mortality prediction in this binary logistic model (Eq.3) and binary thinned model (Eq. 4) could be marginally reduced by increasing the dbh class width from five cm to ten cm. For example; the reduced total χ^2 test value for Eq. (3) was 21.5 and for Eq. (4) it was 29.89. The χ^2 test value cannot always be made small enough to fail to reject the mortality model for all tree species even with a large dbh intervals (10 cm) (Monserud and Sterba 1988; Temesgen and Mitchell 2005). The large χ^2 test value obtained here was possibly due to the variability in measured variables of natural shortleaf pine stands in the dataset used to develop these mortality prediction models.

Conclusions

An independent variable that provides variation in the point estimates of their associated parameters either quite small or very large in a combination of the other variables in a model might be helpful to predict better mortality probability of an individual tree. It was found that measures of relative stand level competition (*BAHG*), hierarchal position (*DAG*) with age and stand level competition (*QMD)* were highly influential variables for predicting mortality. Ideally the same independent variables would be important whether the dependent variable in the probability prediction model was survival or mortality. The binary model performed better than an iteratively reweighted nonlinear regression model with the same independent variables because the binary estimation technique produced smaller standard errors for parameter estimates though both estimation methods provide similar parameter estimates. The mixed-effects model approach for predicting mortality produced a significant model with smaller AIC value than the binary model but did not produce lower χ^2 test values than the binary models. Therefore, the binary model appears to be better for predicting the probability of mortality than the mixed-effects model.

A model with thinning effects (Eq. 4) could be used where thinning effects are present although the effect was not very strong in the model. The inclusion of the thinning variable in the binary model did not improve the χ^2 the test of goodness of fit although the thinning variable was significant and associated with increasing risk of mortality. This may be due to logging damage. The binary logistic model (Eq. 3) can be used to predict the annual mortality or survival rate of individual trees of even-aged shortleaf pine forests. Therefore, the binary logistic regression model performed better than alternative models in this study and could be tested for use in a shortleaf pine growth simulation model such as the Shortleaf Pine Stand Simulator (SLPSS) (Huebschmann et al. 1998), a distance-independent individual tree growth simulator for naturally-occurring shortleaf pine.

Acknowledgements

This research was supported by Department of Natural Resource Ecology and Management at Oklahoma State University (OSU) and USDA Forest Service Southern Research Station (UAFS) at Monticello, Arkansas. The authors would like to thank all crew members who contributed in collecting data since permanent plot establishment (1985-87) to the current measurement (2012-2014).

References

Allison, P.D. 2001. Logistic Regression using the SAS System: theory and applications, 2nd ed. SAS® publishing, books by users. press. 288p.

Bailey, R.L., Borders, B.E., Ware, K.D., and Jones, E.P. 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.

Bravo-Oviedo, A., Sterba, H., del Río M., and Bravo, F. 2006. Competition-induced mortality for Mediterranean *Pinus pinaster* Ait. and *P. sylvestris* L. For Ecol Manag 222(1):88-98

Cao, V. 2000. Prediction of annual diameter growth and survival for individual trees from periodic measurements. For. Sci. 46(1):127-131.

Cao, Q, and Strub, M. 2008. Evaluation of four methods to estimate parameters of an annual tree survival and diameter growth model. For. Sci. 54(6):617-624

Coomes, D.A., Duncan, R.P., Allen, R.B., and Truscott, J. 2003. Disturbances prevent stem size-density distributions in natural forests from following scaling relationships. Ecol. Lett. 6(11): 980-989.

Crecente-Campo, F., Soares, P., Tomé, M., and Diéguez-Aranda, U. 2010. Modelling annual individual-tree growth and mortality of Scots pine with data obtained at irregular measurement intervals and containing missing observations. For. Ecol. Manage. 260(11):1965-1974.

Dipesh, KC, Will, R.E., Lynch, T.B., Heinemann, R., and Holeman, R. 2014. Comparison of loblolly, shortleaf, and pitch x loblolly pine plantation in Oklahoma. For. Sci. 61(3):540-547.

Efron, B. 1988. Logistic regression, survival analysis, and the Kaplan-Meier curve. J. Amer. Statist. Assoc. 83(402):414-425.

Flewelling, J.W. and Monserud, R.A. 2002. Comparing methods for modeling tree mortality. USDA For. Serv. Proc. RMRS-P-25:168-177.

Groom, J.D., Hann, D.W. and Temesgen, H. 2012. Evaluation of mixed-effects models for predicting Douglas-fir mortality. For. Ecol. Manage. 276:139-145.

Hamilton, D.A., and Edwards, B.M. 1976. Modeling the probability of individual tree mortality. USDA For Ser Res Pap INT (USA).

Hein, S., and Weiskittel, A.R. 2010. Cutpoint analysis for models with binary outcomes: a case study on branch mortality. Eur. J. For. Res. 129(4):585-590.

Hitch, K.L. 1994. A distance-independent individual tree basal area growth model for natural, even-aged stands of shortleaf pine in eastern Oklahoma and western Arkansas. M.Sc. thesis. Oklahoma State Univ., Stillwater, OK. 78p.

Huebschmann, M.M., Lynch, T.B., and Murphy, P.A. 1998. Shortleaf pine stand simulator: An even-aged natural shortleaf pine growth and yield model user's manual. Res. Rep. P-967, Oklahoma State Univ. Agric. Exp. Stn. 25p.

Lappi. J. 2006. A multivariate, non-parametric stem-curve prediction method. Can. J. For. Res. 36(4):291-301.

Lynch. T.B., and Murphy, P.A. 1995. A compatible height prediction and projection system for individual trees in natural, even-aged shortleaf pine stands. For. Sci. 41(1):194-209.

Lynch, T.B., Hitch, K.L., Huebschmann, M.M., and Murphy, P.A. 1999. An individual-tree growth and yield prediction system for even-aged natural shortleaf pine forests. Sout. J. Appl. For. 23(4):203-211

Lynch, T.B., Holley, A.G., and Stevenson, D.J. 2005. A Random-Parameter Height-Dbh Model for Cherrybark Oak. South. J. Appl. For. 29(1):22-26.

Lynch, T.B., Budhathoki, C., Wittwer, and R.F. 2012. Relationships between Height, Diameter, and Crown for Eastern Cottonwood (*Populus Deltoides*) in a Great Plains Riparian Ecosystem. West. J. Appl. For. 27(4):176-186.

McCullagh, P., and Nelder, J.A. 1989. Generalized Linear Models. 2nd Ed. Chapman and Hall, New York. 511p.

Metz, C.E. 1978. Basic principles of ROC analysis. In *Seminars in nuclear medicine* 8(4):283-298 WB Saunders.

Monleon, V.J. 2003 A hierarchical linear model for tree height prediction. P 2865-2869 in *Proc. of the 2003 Meeting of the American Statistical Association, Section on Statistics and the Environment, 3–7 August 2003, San Francisco, California*. The American Statistical Association, Alexandria, VA.

Monserud, R.A. 1976. Simulation of forest tree mortality. For Sci 22(4):438-444.

Monserud, R.A, and Sterba, H. 1999. Modeling individual tree mortality for Austrian forest species. For. Ecol. Manage. 113(2):109-123.

Murphy, P.A. 1988. Study plan and progress report: Growth relationships for even-aged stands of shortleaf pine. South. For. Exp. Station. For. Sci. Lab., Monticello, AR. 22 p.

Neter, J., Maynes, E.S. 1970. On the appropriateness of the correlation coefficient with a 0, 1 dependent variable. J. Amer. Statist. Assoc. 65(330):501-509.

Peet. R. K., and Christensen, N.L. 1987. Competition and tree death. BioScience 37(8):586-595.

Pinheiro, J.C., and Bates, D.M. 2000. Mixed-Effects Models in S and S-Plus. Springer-Verlag, New York. 528p.

R Core Team 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rose, C.E., Hall, D.B., Shiver, B.D., Clutter, M.L., and Borders, B. 2006. A multilevel approach to individual tree survival prediction. For. Sci. 52(1):31-43.

Saud, P., Lynch, T.B., Anup, KC., and Guldin, J.M. 2015. Models for estimating height and crown ratio of Shortleaf pine based on 25 years of remeasurement data. (Submitted for publication).

SAS Institute, Inc. 2007. SAS online help documentation. SAS Institute, Inc. Cary, NC. Available online at http://support.sas.com/documentation/onlinedoc/guide/; last accessed Sep. 3, 2014.

Sing, T., Sander, O., Beerenwinkel, N., and Lengauer, T. 2005. ROCR: visualizing classifier performance in R. *Bioinformatics*, 21(20):3940-3941.

Temesgen, H., and Mitchell, S.J. 2005. An individual-tree mortality model for complex stands of southeastern British Columbia. West. J. Appl. For. 20(2):101-109.

Westphal, C., Tremer, N., Oheimb, G.V., Hansen, J., Gadow, K.V., and Härdtle, W. 2006. Is the reverse J-shaped diameter distribution universally applicable in European virgin beech forests? For. Ecol. Manage. 223(1):75-83.

Woodall, C.W., Grambsch, P.L., and Thomas, W. 2005. Applying survival analysis to a large-scale forest inventory for assessment of tree mortality in Minnesota. Ecol Model 189(1):199-208.

Yang, Y., Titus, S.J., and Huang, S. 2003. Modeling individual tree mortality for white spruce in Alberta. Ecol. Model. 163:209-222.

Yao, X., Titus, S.J., and Macdonald, S.E. 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(2):283-291.

Zhao, D., Borders, B., and Wilson, M. 2004. Individual-tree diameter growth and mortality models for bottomland mixed-species hardwood stands in the lower Mississippi alluvial valley. For. Ecol. Manage. 199(2):307-322.

Zhao, D., Borders, B., Wilson, M., Rathbun, S.L. 2006. Modeling neighborhood effects on the growth and survival of individual trees in a natural temperate species-rich forest. Ecol. Model. 196(1):90-102.

Tables

Table III-1 Summary of the tree and stand level attributes of dead trees of five measurement periods.

Variables	Measurement periods				
	1st (n= 202)	2nd (n = 502)	3rd (n = 267)	4th (n = 116)	5th (n = 46)
	Mean (Median)	Mean (Median)	Mean (Median)	Mean (Median)	Mean (Median)
Tree level					
dbh (cm)	14.35(10.67)	10.83(7.87)	25.54(26.42)	20.47(18.16)	20.74(18.48)
BA (m^2)	0.27(0.1)	0.16(0.05)	0.62(0.59)	0.43(0.28)	0.43(0.29)
CRT	0.35(0.34)	0.3(0.3)	0.35(0.33)	0.31(0.29)	0.31(0.3)
HT (m)	11.34(9.51)	9.95(8.13)	19.55(21.37)	16.69(15.13)	16.86(16)
RAQD	1.46(1.4)	1.68(1.62)	1.22(1.11)	1.33(1.28)	1.37(1.3)
HDR (cm)	1.28(1.25)	1.59(1.54)	0.88(0.81)	1.06(0.98)	1.08(1.06)
Stand Level					
BAH (m^2ha^{-1})	23.54(26.51)	30.95(31.73)	20.75(20.72)	32.08(32.23)	38.62(46.21)
QMD(cm)	17.33(15.31)	15.49(12.77)	28.71(31.1)	24.7(20.45)	25.19(23)
HD (m)	13.9(13.65)	13.8(12.19)	20.6(22.86)	18.77(17.23)	19.44(19.12)
PAG (yr)	38(31)	36(29)	63(73)	52 (41)	52(47)
BAHG (m^2ha^{-1})/yrs	0.75(0.87)	0.98(1)	0.37(0.34)	0.72(0.75)	0.78(0.98)
DAG(cm yrs^{-1})	0.36(0.33)	0.29(0.26)	0.42(0.41)	0.39(0.37)	0.4(0.35)
SIND (m)	17(16.14)	16.32(15.49)	18.02(18.59)	17.47(17.79)	17.82(18.33)
TPH	2087(1440)	2511(2459)	429(321)	947(803)	873(1112)

Note: Diameter at breast height (*dbh*); basal area per hectare (*BAH*); individual tree basal area (*BA*); crown ratio (*CRT*); quadratic mean diameter (*QMD*); ratio of *QMD* to *dbh* (*RAQD*); stand or plot age (*PAG*); ratio of *BAHA* to *PAG* (*BAHG*); ratio of dbh to *PAG* (*DAG*); individual tree height (*HT*); average dominant and codominant height (*HD*); ratio of *HD* to dbh (*HDR*); site index (*SIND*); and trees per hectare (*TPH*).

Table III-2 Statistic summary of 25218 shortleaf pine trees from natural, even-aged stands used to fit mortality regression models.

Variables	Mean	Std. dev.	Minimum	Maximum
BAHG (m^2ha^{-1})/yrs	0.593	0.319	0.041	1.280
DAG (cm yrs^{-1})	0.470	0.138	0.052	1.362
QMD (cm)	22.945	9.503	7.887	56.878
THINHA (m^2ha^{-1})	1.241	3.110	0.000	19.363

Note: *THINHA* = Thinned basal area per hectare/ years since thinning.

Table II-3 Model evaluation of the binary logistic model, re-weighted, and mixed-effects model forms of with and without thinning effect using the fit statistic (AIC, χ^2, and AUC).

Model form	Fit statistics				
	AIC	χ^2 (Mortality)	χ^2 (Survival)	AUC	σ^2_{uk}
Unthinned					
Binary Eq. (2)	7983	24.63	1.57	76.2	
Weighted Eq. (3)	7983	24.63	1.57	76.2	
Mixed					
Eq. (7)	7498	49.04	3.31	76.53	0.003
Eq. (8)	7511	98.02	9.15	76.73	1.628
Eq. (9)	7591	1554.42	55.98	72.7	8.902
Eq. (10)	8177	510.01	59.63	67.78	2.115
Thinned					
Binary Eq. (4)	7878	33	2	77.85	
Weighted Eq. (6)	7878	33	2	77.85	
Mixed					
Eq. (11)	7438	59.01	2.55	77.63	0.152
Eq. (12)	7264	958.23	478.81	75.63	3.138
Eq. (13)	7435	79.39	8.95	78.49	4.067
Eq. (14)[#]	18770	3256.93	847.04	70.62	$-1.1*10^{-14}$
Eq. (15)	7281	217.93	37.62	77.18	0.0005

Note: σ^2_{uk} = variance of the random effect of the corresponding mixed-effects model.
[#] The parameter estimates of b_0 was significant with *p-value* = 0.0212 and random effect associated plot was not significant (*p-value* = 1).

Table III-4 Parameter estimates (standard errors) of the selected binary logistic model, re-weighted regression model, and mixed-effects model form of with and without thinning effect for predicting the probability of mortality of an individual tree of naturally occurring shortleaf pine.

Coefficient	Without thinning effect			With thinning effect		
	Logistic Eq. (3)	Weighted Eq. (5)	Mixed (Eq. 7)	Logistic (Eq. 4)	Weighted (Eq. 6)	Mixed (Eq. 11)
b_0	4.5930	4.5929	5.6797	4.8955	4.8955	5.1883
	(0.2484)	(0.3727)	(0.5047)	(0.2587)	(0.4107)	(0.2900)
b_1	-2.4309	-2.4308	-3.2635	-2.7012	-2.7012	-2.9844
	(0.1654)	(0.2465)	(0.3646)	(0.1731)	(0.2732)	(0.1904)
b_2	8.1132	8.1130	9.6223	8.6290	8.6290	8.2498
	(0.2840)	(0.4238)	(0.3484)	(0.2933)	(-0.4637)	(0.3047)
b_3	-0.0720	-0.0720	-0.1108	-0.0806	-0.0806	-0.0744
	(0.0063)	(0.0094)	(0.0128)	(0.0065)	(-0.1016)	(0.0075)
b_4				-0.1016	-0.1016	-0.1442
				(0.0085)	(0.1370)	(0.045)

Note: All parameter estimates had *p-value* <0.0001, and parameters b_1 b_2, b_3, and b_4 are associated with the variable *BAHG, DAG, QMD* and *THINHA* respectively.

Table III-5 Matrix of correlations and covariance for mixed-effects approach model; binary mixed (Eq.7) and binary mixed thinned (Eq. 11).The values above diagonal are correlation and those below are covariance.

Model	Coefficient	b_0	b_1	b_2	b_3	b_4	σ^2_{uk}
Eq. (5)	b_0		-0.93630	-0.17130	-0.88010	-	0.26050
	b_1	-0.17230		0.02587	0.81940	-	-0.25760
	b_2	-0.03012	0.00329		-0.15600	-	0.15440
	b_3	-0.00569	0.00383	-0.00070		-	-0.21830
	σ^2_{uk}	0.00005	-0.00004	0.00002	0.00000	-	
Eq. (6)	b_0		-0.93300	-0.19210	-0.85240	-0.00134	0.04218
	b_1	-0.05154		-0.06830	0.85360	0.05581	-0.06152
	b_2	-0.01698	-0.00396		-0.25750	-0.06837	0.01811
	b_3	-0.00184	0.00121	-0.00058		-0.05726	-0.02270
	b_4	-0.00002	0.00048	-0.00094	-0.00002		0.24210
	σ^2_{uk}	0.00034	-0.00032	0.00015	0.00000	0.00030	

Table III- 6 Observed and predicted survival and mortality by five-centimeter classes from a nonlinear approach of binary logistic model fit, and a mixed-effect model fit for with and without thinning effect.

Mid-dbh class (cm)	Observed Survival	Predicted Survival				Observed Mortality	Predicted Mortality			
		Eq.3	Eq.4	Eq.7	Eq.11		Eq.3	Eq.4	Eq.7	Eq.11
2.5	158	168	167	167	165	82	72	73	73	75
7.5	2358	2387	2395	2427	2380	373	344	336	304	351
12.5	4018	3985	3975	4028	3951	192	225	235	182	259
17.5	4057	4048	4042	4079	4030	114	123	129	92	141
22.5	4098	4080	4081	4106	4086	97	115	114	90	109
27.5	3431	3431	3434	3449	3447	103	103	100	85	87
32.5	2648	2661	2664	2671	2676	86	73	70	63	58
37.5	1819	1832	1834	1836	1843	61	48	46	44	37
42.5	887	881	881	880	885	17	23	23	24	19
47.5	611	604	604	602	607	8	15	15	17	12

Note: Eq. (3) without thinning effect; Eq.(4) with thinning effect; Eq.(5) binary mixed model with random effect associated with variable *QMD*; and Eq. (6) binary mixed thinned model with random effect associated with variable *THINHA*.

Figures

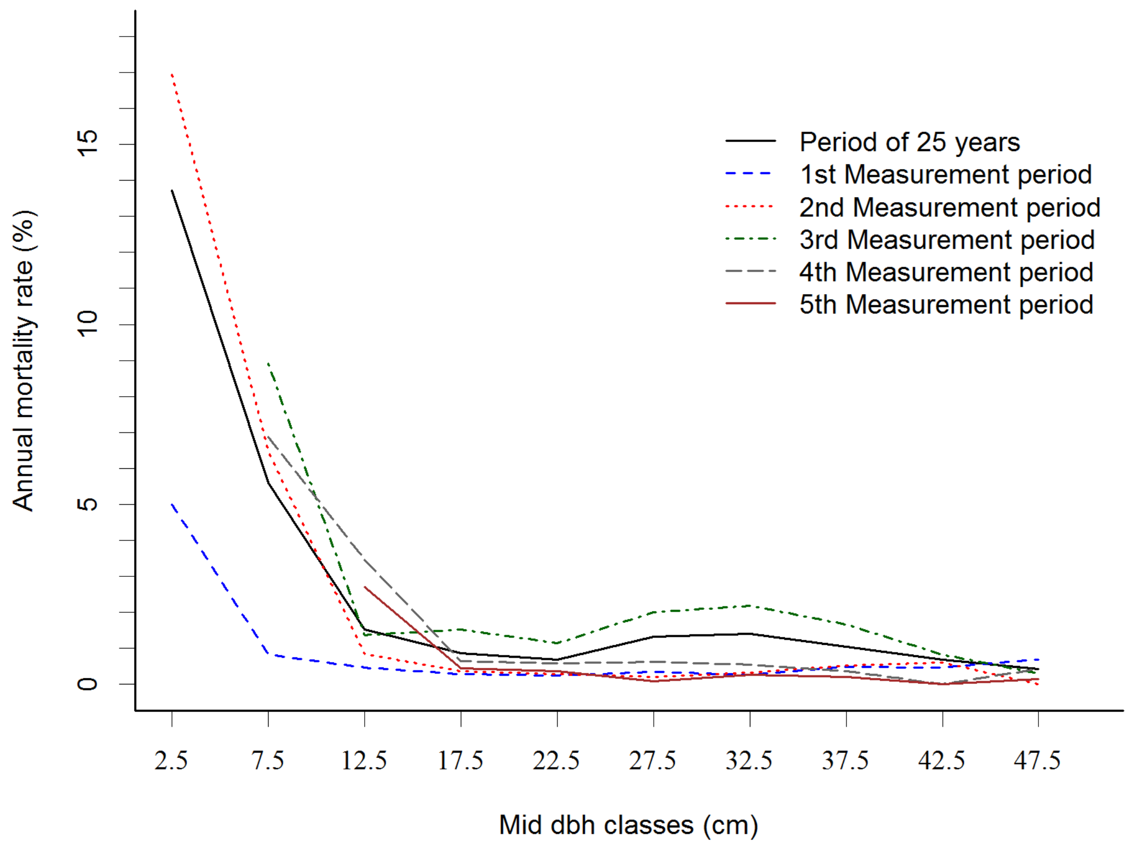

Fig. III- 1 The annual mortality rate of naturally occurring even-aged shortleaf pine based on mid-dbh classes for over the period of 25 years and for each successive measurement periods.

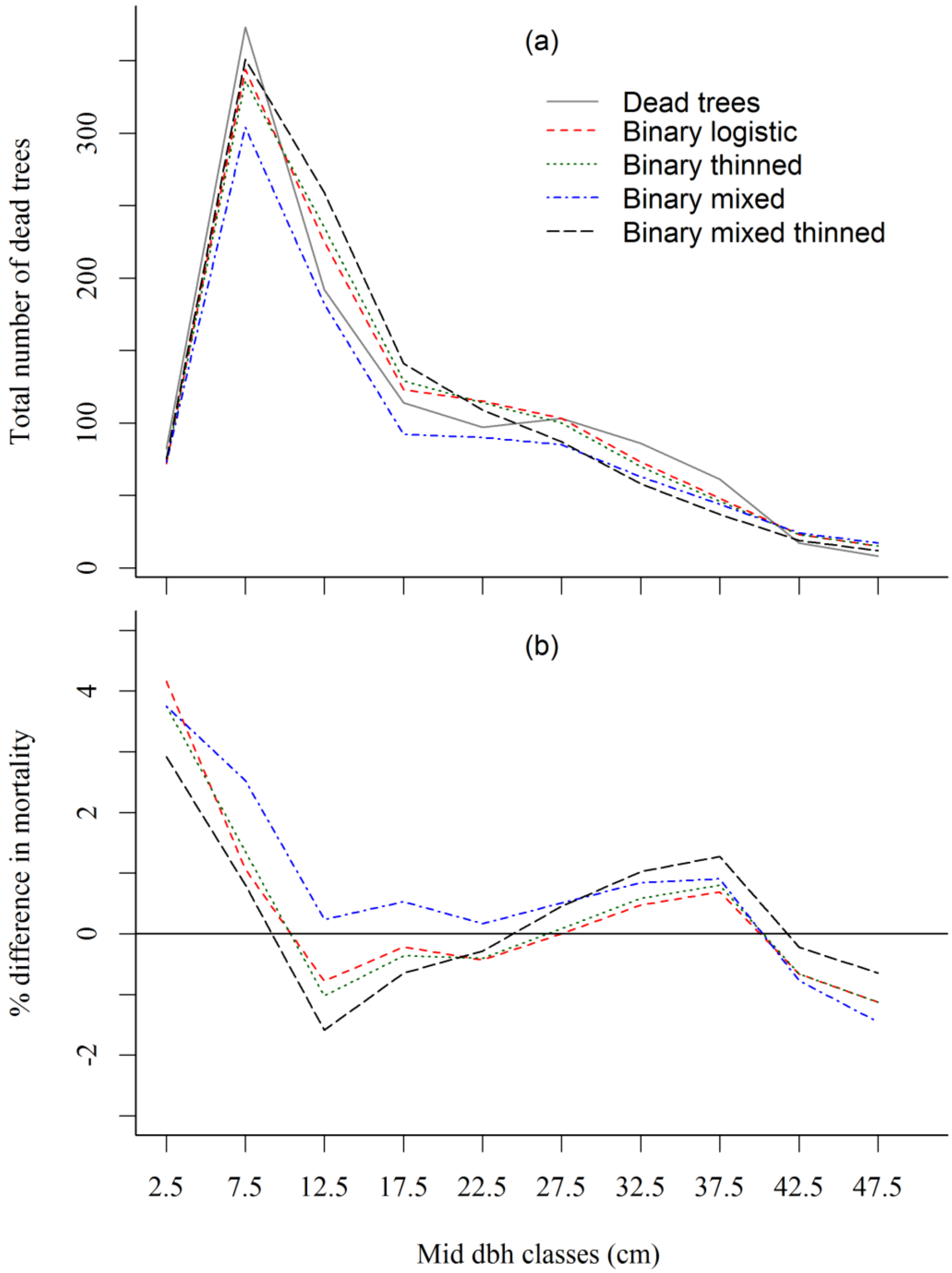

Fig. III-2 Predicted number dead trees for each diameter class (a) by four different best models (b) percentage difference of observed and predicted number dead trees for each mid diameter class by the binary model fit and mixed-effects model fit for with and without thinning effect.

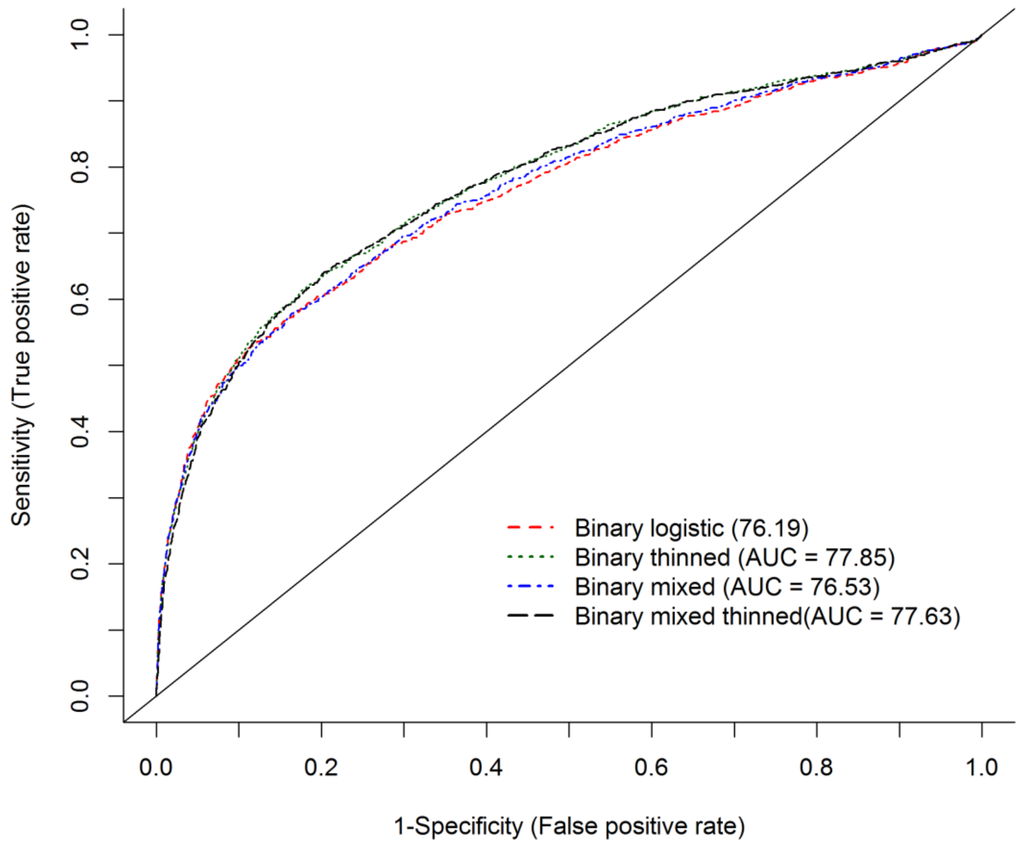

Fig. III- 3 ROC curve showing the area under the curve (AUC) for binary models and mixed-effects models for mortality prediction of an even aged natural stand of shortleaf pine.

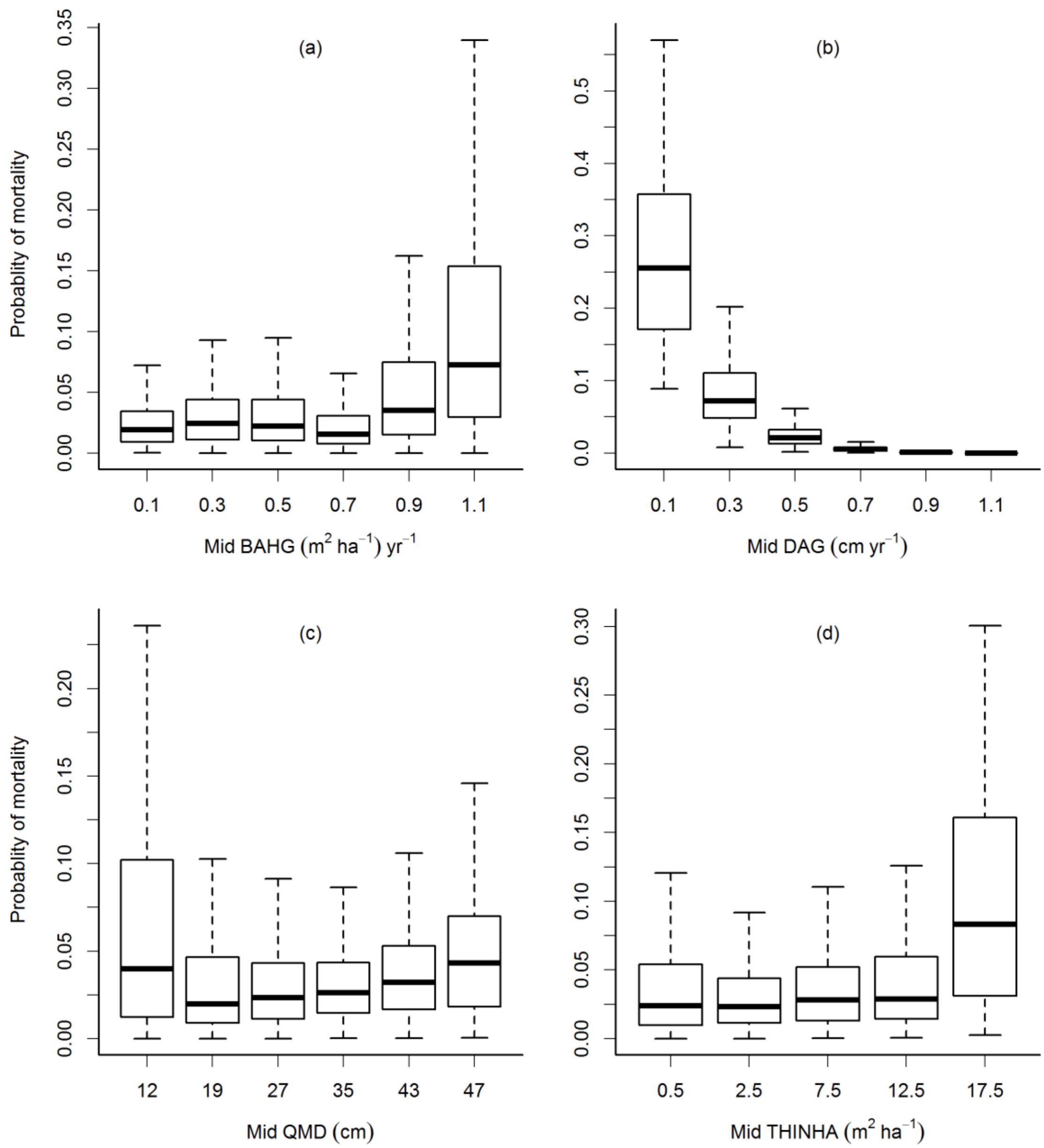

Fig. III- 4 Box plot showing the predicted probability of mortality against the variables used in the model at their mid-class ranges. The boxplot for variable '*THINHA* 'was plotted from the parameter estimates of binary thinned (Eq. 4) while other boxplots were plotted from the parameter estimates of the binary logistic model (Eq. 3).

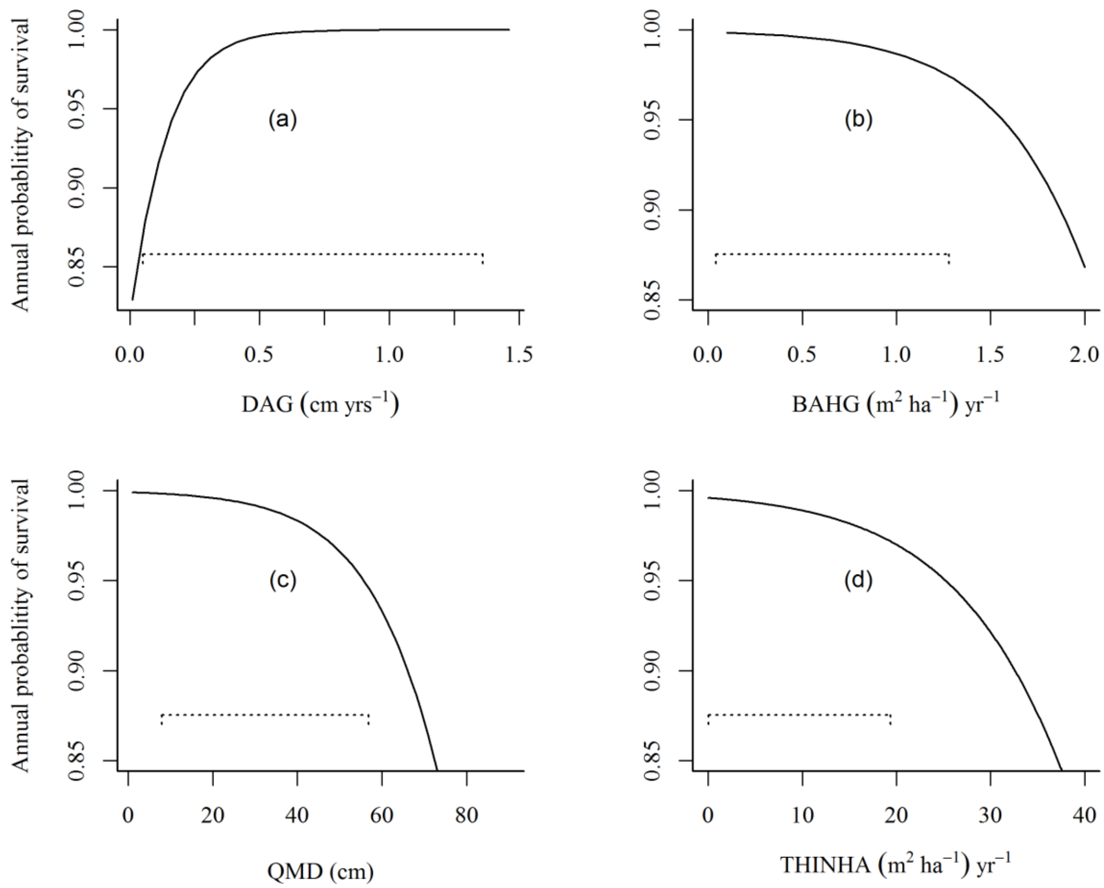

Fig. III- 5 Influence of a variable in predicting annual survival probability while fixing all other independent variables as constant equal to their data means. The plot for *THINHA* was plotted using parameter estimates of binary thinned (Eq. 4) while other plots were plotted using parameter estimates of binary logistic (Eq. 3).The dotted line indicates the data ranges of those respective variables.

CHAPTER IV

CLIMATE-BASED GROWTH MODELS FOR ANNUAL BASAL AREA GROWTH OF INDIVIDUAL SHORTLEAF PINE TREE BASED ON 25 YEARS OF REMEASUREMENT DATA

Abstract

To understand the effect of climate on annual basal area growth rate, we modified an existing annual basal area growth model for individual shortleaf pine by incorporating growing season (April-September) climate variables. The data for this study were 208 permanently established 208 plots, which were measured six times (1985-87 to 2012-14) at intervals of 4-6 years in naturally occurring even aged shortleaf pine (*Pinus echinata* Mill) stands in western Arkansas and eastern Oklahoma.. Various nonlinear growth models with and without climate variables were formulated. Although growth models were fitted with: mixed effects approach, first-order autoregressive (AR (1)) structure, and employing a power variance function had better Akaike information criteria (AIC) values than ordinary least squares (OLS) models, a growth model fitted using OLS had a better fit index. The climate-based growth models with deviation in seasonal climate variables and also in monthly climate variables from a decade-long (1980-2014) mean climate response during the growing season over the for each plot reduced the residual standard error (RSE) by 1.3% and 2.3% respectively compared to a growth model without climate variables. The deviation in maximum temperature and precipitation amount were more effective in an exponential factor multiplied by a base growth model rather than when included in a linear manner and convincingly addressed climate change scenarios. The simulation showed that increase in mean maximum air temperature by 0.5 ^{0}C with mean monthly precipitation amounts could easily reduce the annual basal area growth rate of an individual tree in the long term. Due to variation in climate, young trees experienced more stress in growth rate and the cumulative effects on growth were smaller for older trees over the course of a 50 year period of change in climate.

Keywords annual basal area growth, climate-based growth model, climate modifier, seasonal climate variable, monthly climate variable, random forests.

Introduction

An individual tree basal area growth or diameter increment model is a core component of an individual tree forest growth model. These models are formulated using individual tree and stand-level covariates such as initial diameter, basal area, stocking level, and measures of stand-level competition (Andreassen and Tomter, 2003; Budhathoki et al., 2008; Subedi and Sharma, 2011). Growth models that do not account the spatial arrangement i.e. distance independent growth models are common in practice and such models can be applied to the range of a species (Lynch et al., 1999; Andreassen and Tomter, 2003; Lhokta and Loewenstein, 2011; Pokharel and Dech, 2012). Prediction equations for annual growth have also usually been fitted using ordinary least squares (OLS) methods (Budathoki et al. 2008; Subedi and Sharma, 2011) that assume the repeated growth observations are independent. However, when the data structure includes repeated measurements, temporal autocorrelation exists that may influence the prediction of individual basal area growth rate. To overcome such issues from repeated observation, power variance functions, autocorrelation structures, and mixed-effects approaches have sometimes been included in modeling efforts (Budhathoki et al., 2008; Uzoh and Oliver, 2008; Lhokta and Loewenstein, 2011; Subedi and Sharma, 2011; Pokharel and Dech, 2012).

Site productivity and climate often account for significant variations in model-based predictions of tree growth rate (Pokharel and Froese, 2009; Pokharel and Dech, 2012; Subedi and Sharma, 2013; Jiang et al., 2014, Manso et al., 2015). The annual growth rate or radial increment over a period can be readily studied, using tree-ring studies (dendrochronology). Variations in tree annual growth rate are often explained by climate variables such as summer temperature, precipitation, and drought index of the growing seasons (Graumlich et al., 1993; Biondi, 2000). These climate variables can be significantly correlated with radial growth anomaly patterns. For example; average summer temperature was strongly correlated with tree ring density at high northern latitudes (Briffa et al., 1998) and extensions of temperature-induced drought stress during the grown seasons affected the radial growth (Barber et al., 2000; Wilmking et al., 2004). Therefore, these climate variables have often been regressed against annual growth rates to develop climate-based growth models (Biondi, 2000; Jump et al., 2006; Chhin et al. 2008; Martín-Benito, 2008; Duschesne, 2012; Foster et al., 2015). These models can help to explain the variations in annual growth responses that can be attributed to the relative precipitation and temperature for the period of study (Yeh and Wensel, 2000; Way and Oren, 2010).

Diameter increment responses may vary among tree species when exposed to climatic stressors (Callaway et al., 1994; Hanson and Waltzing., 2000). Insight into the growth-climate relationship can obtain through the study of the largest diameter trees in a forest stand (Chhin, 2008). However, these large trees have highly variable growth rates (Hanson and Weltzin, 2000; Carrer and Urbinati, 2004). The growth rates of mature trees are also affected by the severe or prolonged drought stress or warmer temperatures (Hanson and Weltzin, 2000; Lloyd and Fastie, 2002; Jump et al., 2006; Martìn-Benito et al., 2008). These factors can also induce tree mortality. Consequently, the effects of an increased growing season length, are dictated by regional conditions and site-specific limiting factors to productivity in ways similar to the effects of variation in precipitation and temperature (Graumlich, 1993; Hanson and Weltzin, 2000; Yeh and Wensel, 2000;

Boisvenue and Running, 2006). Therefore, monitoring and evaluating the response of trees to changes in climate is useful for understanding the limiting factors of forest growth on the particular sites (Boisvenue and Running, 2006; Moore et al., 2006; Park et al. 2013).

Shortleaf pine (Pinus echinata Mill.) grows in more than 22 states and is second in volume among the southern pines of USA only to loblolly pine (*Pinus taeda* L.) but its abundance is declining due to replacement by plantations of other commercially viable southern pine species (Lawson, 1990). It is more drought tolerant than loblolly pine and is found to grow well where annual precipitation averages between 100 cm to 150, and average annual temperature ranges from 90 C to 210 C (USDI Geological Survey, 1970). Studies focused on growth and yield of naturally grown shortleaf pine include Murphy et al. (1992), Lynch et al. (1991), Lynch et al. (1999), and Budhathoki et al. (2008). These studies are based on permanent plots established in even-aged naturally-occurring shortleaf pine forests in western Arkansas (Ozark and Ouachita National Forest) and southeastern Oklahoma (Ouachita National Forest). Previous studies have fitted an annual basal area growth models for individual shortleaf pine trees using OLS (Lynch et al.,1999) or with a mixed-effects modeling approach (Budhathoki et al., 2008; Lhokta and Loewenstein, 2011). The model fitted by Lhokta and Loewenstein (2011) was for uneven-aged shortleaf pine stands while the models developed by Lynch et al. (1999) and Budhathoki et al. (2008) were for even-aged forests. None of the latter studies included climate variables in the diameter or basal area growth model. Interestingly, Saud et al. (2015) presented a climate-based growth model for naturally occurring even-aged shortleaf pine with a smaller dataset (129 plots) than that used in the current study in which there were only four measurement times. The result demonstrated potentiality for improving an existing growth model (Lynch et al., 1999) by adding climate variables. However, the current study is based on a larger dataset and includes an additional fifth and sixth measurement time. Climate-based growth models are often based on tree ring data that use only climate variables; however, a few studies have extended typical growth models also to include climate variables for species other than shortleaf pine (Subedi and Sharma, 2013; Manso et al., 2015, Saud et al., 2015). These studies have used climate variables as linearly added terms in existing growth models. It is evident that addition of a significant climate variables in these models can improve model prediction by a small amount. However, these studies did not investigate whether the climate variables used as linearly added term or as multiplicative (modifier) functions in a growth model performs better. Therefore, the objective of this study was to develop a climate-based growth model with potential climate modifier function that can enhance annual basal area growth response of an individual tree using remeasured permanent plots of shortleaf pine that have been measured six times during a period of 25 years. The other specific objectives included updating and evaluating the existing annual basal area growth model with different modeling approaches. An additional specific aim was to examine the influence of the climate-based growth model under different climate change scenarios.

Materials and methods

Tree Data

Data for this study were obtained from 208 permanent plots located in even-aged natural stands of shortleaf pine in the Ozark and Ouachita National Forests in western

Arkansas and southeastern Oklahoma. These permanent plots were collaboratively established by the Department of Forestry (now part of the Department of Natural Resource Ecology and Management) at Oklahoma State University (OSU) and USDA Forest Service Southern Research Station (USFS) at Monticello, Arkansas. These research plots were established in 1985-87 to monitor growth and yield performance of managed naturally occurring shortleaf pine stands because, until 1985 the major sources of data were from fully stocked plots or unmanaged shortleaf pine stands. For additional details concerning plot establishment, see Lynch (1991), Lynch et al. (1999), and Budthathoki et al. (2008).

The permanent plots were circular and 0.081 ha (0.2 ac) in size with a radius of 16.06 m. At an interval of 4 to 6 years, each plot was remeasured, and the last (sixth) measurement made during the period from 2012 to 2104. Therefore, the data contain six measurements and five growth (measurement) periods. At the time of plot establishment, the minimum diameter at breast height (dbh) recorded was 2.78 cm and the maximum was 61.98 cm. Similarly, the minimum plot age recorded was 18 years, and the maximum was 93 years. Diameter for the individual trees on each plot were recorded at each measurement (Table 1). However, tree height (HT) in m and height to base of live crown (CL) in m were recorded for a subsample of trees from each plot selected to represent the range of tree diameters in the plot. Using the ring count method, ages of dominant and codominant trees were averaged to obtain the plot or stand age. Most of the plots were re-thinned to their original basal area levels at establishment plot time just after the third measurement while a portion were left unthinned. Therefore, the thinned basal area was deducted from the stand basal area of the third measurement. Individual trees recorded as not suitable for growth calculation (Maverick) and having ice damage were not included in the growth model development process. The summary statistics of tree level and stand level variables for the five growth periods that were used in the model development process are shown in Table 2.

Climate data

Climate data (air temperatures and precipitation) were accessed using the global position system (GPS) locations (Latitude and Longitude) of each plot as long-term data from 1980- 2014. Climate data were obtained with 4 km resolution (grids) from the Parameter-elevation Relationships on Independent Slopes Model (PRISM) climatic group based at Oregon State University (PRISM Climate Group 2014). The length of the active growing season of shortleaf pine was assumed from the month of April to September. This length of growing season corresponds to the six months that are normal for calculating the standardized precipitation index (SPI). SPI is believed to be a surrogate for the effect of the both temperature and precipitation (Guttman, 1999). The six months SPI indicates seasonal to medium-term trends in precipitation by comparing the precipitation for that period with the same six months period over the range of the data (historical data). SPI is considered as more sensitive and an alternative index to the Palmer drought index but it requires long-term (20-30 years) records of monthly precipitation so that the precipitation data can be normalized with a Weibull transformation (Guttman, 1999; Beguería et al., 2013). The SPI index indicates the wet and dry events using scales. The scale ranges from ≤ -2.0 for extremely dry, ≤ 1.5 for moderately dry, $\leq$ -1.0 for dry, < 1.0 for neutral, < 1.5 wet, < 2.0 moderately wet, and > 2 for extremely wet.

Initially, seven candidate variables were derived at the growth interval level from climate variables i.e. air temperature, precipitation, and SPI, for each plot during the growing season over the growth period. The derived variables were: an average monthly maximum temperature (MTMAX), an average monthly mean temperature (MTMEAN), and an average monthly minimum temperature (MTMIN), average monthly precipitation (MPPT), average monthly SPI (MPSI), average total precipitation (APPT), and average total SPI (ASPI). Additionally, to understand the influence of climate of the specific month of the growing season during the growth period, we created monthly climate variables (24 variables) with an averages values of the climate variables for the respective months of the growth period. Hereafter, the first seven climate variables were termed "seasonal climate variables" and the latter were termed "monthly climate variables". The monthly average values of the seasonal climate variables during the growing season over the length of 34 years are shown in Fig. (1a) and (1b), and the values of the monthly climate variables during the growing season are shown in Fig. (1c) and (1d). Over the period of 34 years, the climate response during the growing season were: average monthly precipitation was 111.36 ± 25.13 cm (mean ± standard deviation (SD)) with a range of 37.14 to 209 cm; average monthly maximum temperature was 28.95 ^{0}C ± 1.20 ^{0}C with a range of 25.84 to 33.43 ^{0}C; average monthly mean temperature 22.63 ± 0.88 ^{0}C with a range of 20.18 to 26.25 ^{0}C; and average monthly minimum temperature 16.29 ± 0.85 with range of 14.00 to 19.55 ^{0}C.

Models and statistical analysis

The annual basal area growth model of Lynch (1999) is based on parameter estimates using ordinary least squares (OLS). Because only one growth period was used for the Lynch (1999) analysis, it was not necessary to consider autocorrelation in time. OLS parameters estimate along with alternative mixed-effects estimates were given by Budhathoki et al. (2008), but the effects of autocorrelation on growth periods (two were used) was not considered. Repeated measurements give rise to a hierarchical pattern i.e. tree within stands and correlation between individual observations within the stand (sampling unit). To account for correlation among trees measured on the same plot, the temporal correlation within a series of observations on each tree within stands as well as possibly heteroscedastic variation among stands, the techniques of power variance function (Budhathoki et al., 2008), correlation structures (Diéguez -Aranda et al., 2006), mixed effects modeling (Lynch et al., 2005; Budhathoki et al., 2008; Lynch et al., 2012; Manso et al., 2015) and or combination of these techniques (Subedi and Sharma, 2013) are widely used. Therefore, a nonlinear annual basal growth model (Eq. 1) developed by Lynch et al. (1999) was modified by Budhathoki et al. (2008) to accommodate plot-specific random coefficients. Annual basal area growth was estimated as the ratio of the total basal area growth or increment during a growth period over the measurement interval. A variety of models developed using techniques as mentioned above were evaluated using the goodness of fit statistics including: Akaike Information Criteria (AIC), residual standard error /root mean square error (RSE), likelihood ratio test, and fit index (pseudo R^2). The base annual basal area growth model (Model 1) is:

$$y_{ij} = \frac{\beta_1 B_{ij}^{\beta_2} - \left(\beta_1 B_{ij} / B_{\max}^{1-\beta_2}\right)}{1 + \exp\left(\beta_3 + \beta_4 B_{si} + \beta_5 A_i + \beta_6 R_{ij} + \beta_7 B_{ij}\right)} + \varepsilon_{ij} \quad (1)$$

where y_{ij} = average annual basal area growth ($m^2 year^{-1}$) of a tree *j* in the plot *i*, B_{ij} = mid basal area (m^2) of tree *j* in plot *i*, B_{max} = 0.6566828736 (m^2) (the maximum expected basal area for a shortleaf pine tree). B_{si} = stand basal area per hectare ($m^2\ ha^{-1}$), A_i = stand age (years) of plot *i*, R_{ij} is the ratio of a tree *j* diameter to the quadratic mean plot *i* diameter $\beta_1, \beta_2, \ldots, \beta_7$ are fixed parameters estimates, and ε_{ij} is within-plot error, residual for tree *j* in the plot *i* $\varepsilon_{ij} \sim N(0,\sigma^2)$.

The existing base equation was fitted as a generalized nonlinear regression model. The gnls function available in nlme package of R (R Core Team, 2012) allowed us to test the heteroscedasticity of error variance and correlated errors (Pinheiro et al., 2015). Model 1 was modified to evaluate the model heteroscedasticity of errors with weighted regression using a power variance function to account for heterogeneous error variance (Pinheiro and Bates, 2002 pg. 391). The modification resulted in Model 2:

Model 1 + power of variance function (2)

$$\mathrm{var}(\varepsilon_{ij}) = \sigma^2 \,|\, \upsilon_{ij} \,|^{2\delta};$$

where error variance (var(ε_{ij})) was modeled with one covariate using variance function. υ_{ij} is the covariate and δ is a power parameter. Tree basal area was selected as the covariate for modeling error variance of the basal area growth model.

The first order of autoregressive process (AR (1)) was used to investigate the correlation structure within an individual tree growth errors with Model 1. The AR (1) model assumes the series is stationary i.e. variance and covariances are independent of time lag and autocorrelation at the second lag is the square of the lag-1 autocorrelation (Cryer and Chan, 2008 pg. 66). The resulting Model 3 with autocorrelation structure is:

$$y_{ijk} = \frac{\beta_1 B_{ijk}^{\beta_2} - \left(\beta_1 B_{ijk} / B_{\max}^{1-\beta_2}\right)}{1+\exp\left(\beta_3 + \beta_4 B_{sik} + \beta_5 A_{ik} + \beta_6 R_{ijk} + \beta_7 B_{ijk}\right)} + \varepsilon_{ijk} \quad (3)$$

where k = 1, 2, 3, 4, 5; corresponding to the growth (time) periods; ε_{ijk} = error term of a tree j in the plot i at the growth period k *;* the error follow the AR(1) process:

$$\varepsilon_{ik} = \phi e_{ik-1} + \omega_{ik}$$

where, $\omega_{it} \sim iid\ \mathrm{N}(0,\sigma_\omega^2)$*, i* is the individual tree and k is measurement period (lag) and ‘phi’ (ϕ) is the autocorrelation coefficient , $|\phi|<1$. ε_{ijk} is related to its own past values instead other independent variables. The correlation “rho” (ρ) between residuals declines exponentially with the number of periods (k) apart i.e. $\rho = \phi^k$.

The mixed effect model approach provides maximum likelihood estimates of parameters accounting for possible correlation and heterogeneity in errors within plots. A mixed effect model that include random-effects associated with plots is Model 4, annual basal area growth model in which random-effect (u_7) was associated with the fixed-effect (β_7) is shown below:

$$y_{ij} = \frac{\beta_1 B_{ij}^{\beta_2} - \left(\beta_1 B_{ij} / B_{\max}^{1-\beta_2}\right)}{1 + \exp\left(\beta_3 + \beta_4 B_{si} + \beta_5 A_{ik} + \beta_6 R_{ij} + (\beta_7 + u_{7i}) B_{ij}\right)} + \varepsilon_{ij} \qquad (4)$$

where u_{7i} is a random parameter specific to i^{th} plot associated with mid-tree basal area fixed effect coefficient β_7, and the rest of the terms are described as above for base model. The random effect (u_{7i}) is normally distributed with mean 0 and variance σ^2_u, the error term (ε_{ij}) is normally distributed with mean zero and variance (σ^2_ε), and the covariance of u_{7i} and ε_{ji} is zero. The variance of (u_{7i}) is an indication of the spread of the random coefficients.

Model 4 was further examined by testing addition of a power variance function and which resulted in the Model 5:

Model 4 + power variance function (5)

Model 4 was further modified with a two-level extension that uses growth period (Model 6). In this model, there is a random parameter associated with each combination of plot and growth period. Model 6 can be written as below:

$$y_{ij(k)} = \frac{\beta_1 B_{ij(k)}^{\beta_2} - \left(\beta_1 B_{ij(k)} / B_{\max}^{1-\beta_2}\right)}{1 + \exp\left(\beta_3 + \beta_4 B_{si(k)} + \beta_5 A_{i(k)} + \beta_6 R_{ij(k)} + (\beta_7 + \mu_{7i(k)}) B_{ij(k)}\right)} + \varepsilon_{ij(k)} \qquad (6)$$

Model 6 was further examined with the variance function in a way similar to Model 2, where basal area was used as covariate. This modification resulted in the Model 7:

Model 6 + power variance function. (7)

After evaluating the performance of the models above, all climate variables were included to examine the effects of climate on the growth of individual trees. These annual basal area growth models are termed "climate-based growth models". We found that the model performance was somewhat better using climate variables in the form of a multiplicative modifier to Model 1 instead of using linearly added terms. So, the climate variables in a model were termed "climate modifier". The climate modifier was further modelled with monthly climate variables for the growth period. New variables (that represent the difference between the observed minus mean value of the respective seasonal and monthly climate variables) were also created. The letter "D" beginning a variable name indicates a variable minus the mean of that respective variable, for e.g. DMTMAX = MTMAX – mean of MTMAX. Random forests (RF) in R (Liaw and Wiener, 2002) was used for preliminary selection of the predictive seasonal climate variables. RF reports the variable importance score by the percentage increase in mean square error (%incMSE) computed for no permuted versus randomly permuted predictors (Liaw and Wiener, 2002, Jiang et al., 2014). Variables with higher %incMSE are the most promising predicting variables as linear predictors in terms of reducing the model mean square error. However, this may not be completely reliable for nonlinear models but the variables with higher %incMSE of seasonal climate variables (Fig. 2a) and of differenced monthly climate variables (Fig. 2b), severed as candidate variables in the climate modifier.

The climate-based growth model that includes seasonal climate variables as climate modifier is Model 8 and can be written as:

$$y_{ij} = \frac{\beta_1 B_{ij}^{\beta_2} - \left(\beta_1 B_{ij} / B_{\max}^{1-\beta_2}\right)}{1 + \exp\left(\beta_3 + \beta_4 B_{si} + \beta_5 A_{ik} + \beta_6 R_{ij} + \beta_7 B_{ij}\right)} \text{Climate Modifier} + \varepsilon_{ij} \qquad (8)$$

where the climate modifier is $\left\{\exp\left(\beta_8(\bar{x}_{jk} - \mu_{xj}) + \beta_9(\bar{y}_{jk} - \mu_{yj})\right)\right\}$;

$\bar{x}_{jk}$ = mean value of the climate variable *x* assigned for the plot *j* during the growth period *k*;

$\bar{y}_{jk}$ = mean value of the climate variable y assigned for the plot *j* during the growth period k;

μ_{xj} = is the mean of the variable *x* for the plot j over the study period of 34 years.

μ_{yj} = the mean of the variable y for the plot j over the study period of 34 years.

β_8 ,and β_9 are the parameters to be estimated for the seasonal climatic variables; the rest of the terms are described as above.

The climate modifier of the Model 8 were: DMTMAX = MTMAX - μ_{xj}, and DMPPT = MPPT-μ_{yj}; where *x* = temperature and *y* = precipitation. Monthly climate variables were included in climate modifier to develop the following climate-based growth model (Model 9).

$$\text{Model1}\left\{\exp\left(\beta_8 DTMAX6_{jk}+\beta_9 DPPT9_{jk}\right)\right\}+\varepsilon_{ij} \tag{9}$$

where β_8 and β_9 are parameters to be estimated for monthly climatic variables. $DTMAX6_{jk}$ is the difference between the average maximum temperatures of the month of June of the plot j during the period k minus the mean maximum temperature for the month June for the plot j over the period of 34 years (μ_{xj}). $DPPT9_j$ is average precipitation of the month of September of the plot j during the period k minus the mean precipitation of the month September for the plot j over the period of 34 years (μ_{yj}).

The performance of climate-based growth Model 9 was better than Model 8. We selected Model 9 and the climate-based growth model was also further analyzed using a power variance function, AR (1) structure, and a mixed effects model approach. The resulting models were:

$$\text{Model 2}\left\{\exp\left(\beta_8 DTMAX6_{jk}+\beta_9 DPPT9_{jk}\right)\right\}+\varepsilon_{ij} \tag{10}$$

$$\text{Model 3}\left\{\exp\left(\beta_8 DTMAX6_{jk}+\beta_9 DPPT9_{jk}\right)\right\}+\varepsilon_{ij} \tag{11}$$

$$\text{Model 4}\left\{\exp\left(\beta_8 DTMAX6_{jk}+\beta_9 DPPT9_{jk}\right)\right\}+\varepsilon_{ij} \tag{12}$$

Climate change scenario analysis

The estimated coefficients of the Model 8 were used instead Model 9 to simulate the effect of changing temperature and precipitation on annual basal area growth of an individual tree because we wanted to demonstrate how much a small improvement in a model can effect prediction. Additionally, the model with seasonal climate variable was the more sensible choice for the interval measurement data than using monthly climate variables because the large SD (58.19) associated with estimate of monthly precipitation (PPT9) amount (μ = 118.78) would show high variability that might affect the prediction of annual growth. This indicates such a model might be useful in the study of tree ring data to address the variability (relaxed or stressed) in the late wood growth of an individual tree due to the large variation amounting of precipitation. Individual trees representative of the dataset having stand age of 38-year and 78-year were selected. Both the 38-year old tree (dbh = 20.2 cm) and the 78-year old tree (dbh = 32.9 cm) were grown 50 years with an annual basal area increment and age increment while other variables in the model were held constant. The values of the input variables for the 38 year old tree were: $B_i = 0.0321$ cm^2, $B_{si} = 14.7543$ m^{2}ha^{-1},$R_{ij}= 0.5788$, and for the 78 year old tree were: $B_i = 0.0851$ cm^2, $B_{si} = 11.7767$m^{2}ha^{-1}, $R_{ij} = 0.5825$.

Though the climate change scenario models are available, the increase in temperature and change in precipitation projection differs among the models (Loehle and David, 1996). So, to avoid such inconsistency in projection, we used climate variables values of lower quartile (LQ), mean (μ), upper quartile (UQ) and SD of *MTMAX* and *MTPPT* of the growing season over the years of 1980-2014. The value at LQ, μ and UQ and SD were 28.07, 28.95, 29.77, and 1.21^{0}C for temperature and 93.12, 111.36, 126.39, and 25.13 cm for precipitation. To ascertain the validity of the data value used for climate change simulation, future climate data at the year 2060 were obtained for a sample plot location from Moscow Forestry Sciences Laboratory (2015) using global climate model (GCM) prediction models for different scenarios (A1B, A2, B1, B2): CGCM3_A1B; CGCM3_A2; CGCM3_B1; and HADCM3_B2 (for detail description see Jiang et al., 2014). The projected mean monthly precipitation for the growing season at year 2060 under different scenarios were consistent with the value at either LQ or median value of precipitation during growing season. It was an average of 103.8 ± 32.5, 100.8 ± 24.3, 95.5 ± 24.4 and 89.1 ± 25.7 cm for respective scenarios. However, the projected average maximum temperature at 2060 during the growing season were slightly larger with larger SD than temperature value simulated here. The mean maximum temperature of the growing season was 31.6 ± 4.5, 31.6 ± 4.4, 31.2 ± 4.1 and 32.5 ± 4.9 ^{0}C for the respective scenarios. A high SD associated with the point estimates of monthly temperature was also observed in the climate dataset that we used.

Fifty years of an individual tree's growth was simulated using Model 8 to demonstrate the possible influence of climate variation on annual basal area growth. In simulation process, both climate variables were randomly generated using SD of the data over the period of 34 years and subtracted from the μ of the respective climate variable. The simulation was conducted under four different scenarios having three different levels (cases) as mean responses, (μ, LQ, and UQ) ± SD, of one climate variable while the response of the other climate variable was at mean level (μ ± SD). In simulation scenarios, the mean precipitation amount and temperature degree were also increased to generate more variability in annual growth response. The simulation scenarios consist: 1) temperature with μ ± SD and precipitation amount at three levels; 2) precipitation amount with μ ± SD and degree of temperature at three levels; 3) increased mean amount of precipitation and degree of temperature at three levels; and 4) increased mean degree of temperature and amount of precipitation at three levels. These four different scenarios were simulated for 100,000 runs with three different cases.

Results

Annual basal area growth

The average annual basal growth rate of individual shortleaf pine trees in the dataset was estimated to be 0.001395 m^2 over the period of twenty-five years. The average periodic mean annual basal area growth rate of individual shortleaf pine trees in the data was: 0.00121, 0.00135, 0.00150, 0.00153 and 0.00138 m^2 respectively for the 1^{st}, 2^{nd}, 3^{rd}, 4^{th}, and 5^{th} growth periods (Table 1). The difference in average annual basal area growth rate for individual trees in the dataset between the first and second growth periods was 8.3%. However, the average annual basal area growth rate of an individual tree in the data increased by 15.4% in the third growth period and by 16.9% in the fourth growth period compared to the second growth period. However, the annual basal area growth rate declined by 5.4% in the fifth growth period compared to the second growth period.

Models without climatic variables

Summaries of fit statistics of different annual basal area growth models are shown in Table 3. The parameter estimates of all models (1-7) were highly significant, but other models tested having poorer performance are not shown (Table 4). The residual patterns were very similar for most of the models except that Model 2 which used power variance function displayed a greatly improved, narrow and compact residual distribution (Fig. 2). The standardized residuals for Model 1, 2, 3, and 6 are shown in Fig. 2. A likelihood ratio test for goodness of fit indicated that Model 2 with the power variance function was better than Model 1. In Model 2, the estimate of the power used in the variance function associated with covariate tree basal area was 0.56248 within 95% confidence interval of 0.55321 and 0.57176. When Model 2 and Model 3 with AR (1) structure compared, a likelihood ratio test indicate no significant difference in goodness of fit, although Model 2 had better AIC, BIC, and log-likelihood value than Model 3. But Model 3 with AR (1) performed better than Model 2 in terms of the fit index. A likelihood ratio test showed that the goodness of fit of the mixed effects model with a power variance component (Model 5) was significantly different from a mixed effect model (Model 4). Similarly, a likelihood ratio test indicated that the goodness of fit of the mixed effect model with two level extension that use growth period and power variance (Model 7) was significantly different from Model 6 (Table 3). This latter result was supported by corresponding lower AIC values for the models having the better significantly better goodness of fit.

The residual standard errors for Model 2, Model 5 and Model 7 were noticeably larger than the other growth models in Table 3. The mixed effect model (Model 4) had the smallest residual standard error, but Model 1 and Model 3 had better fit indices than other alternative models (Table 3). The fit indices of the mixed effects models (Model 4, 5, 6, 7) were estimated using the estimated fixed parameter estimates with the random effects set to zero. Among these three models (Model 1, 3, & 6) with better fit indices, Model 3 with AR (1) was preferred because the AIC and BIC values were noticeably smaller than for the other two models and the covariance structure accounted the correlated errors in the model. The estimated autocorrelation coefficient, 'phi' (φ) was 0.4412 for autocorrelation among individual tree measurements with a 95% confidence interval of 0.4296 to 0.4527. AR (1) structures indicate a decreasing empirical autocorrelation function between individual tree measurements for successive lags. The performance of the Model 7 was judged to be better than the other models tested because the fit index was similar to Model 1, and its residual standard error was smaller than for Model 1 or Model 3. The SD of the random effects associated with the growth period effect was small (0.8063) but the SD among the random effects associated with the plot within growth period was almost negligible.

Models with climatic variables

The seasonal climate variables *DMTMAX* and *DMPPT* and the monthly climate variables *DTMAX6* and *DPPT9* were highly significant ($p < 0.0001$) in climate modifiers for Models 8 and 9 respectively. The climate modifier showed that *DMTMAX* of the growth season and *DTMAX6* have a negative effect (negative sign associated with coefficient) and *DMPPT* and *DPPT9* of the growth season have a positive impact (positive sign associated with coefficient) on annual basal area growth for individual trees (Table 4). It indicated that the negative difference in mean maximum temperature and the positive difference in mean precipitation would greatly favor the annual growth rate. Fit statistics (fit index and AIC) and RSEs of the climate-based growth models (Model 8 and 9) were better than the growth model without climate variables (Model 1) (Table 3). The goodness of fit statistics and fit indices for Models 9 was better than Model 8. The improvement in the fit index was 2.56 and 1.45% and the reduction in RSE was 2.30 and 1.31% respectively for Model 9 and Model 8 over the Model 1 without climate variable. The AIC values and the RSEs of the climate-based growth models (Models 10- 12) were better than for the corresponding Models 2- 4 without climate variables (Table 3). The fit indices were improved by 2.53, 2.77 and 4.18% and RESs were lowered by 4.64, 2.44 and 3.31 % for the Models 10-12 compared to Models 2-4 respectively. Though the climate-based growth model with mixed effects (Model 12) had a smaller RSE than other climate-based growth models, it had a poor fit index (Table 3). We preferred Model 9 because of better fit statistics, and Model 11 with AR (1) because of better AIC value and also had a fit index close to the best results from the other models. Therefore, the parameter estimates of the competitive climate-based growth models (Model 8, 9 and 11) are shown in Table 4.

Climate influence on annual growth

The simulation study showed that the percentage of difference (decrease) in annual basal area growth rate between beginning and the end of 50 years for each of the trees of different ages was influenced more strongly by variability in mean monthly maximum degree of temperature (Fig. 4c and 4d) than variability in mean amount of precipitation during the growing season (Fig. 4a and 4b). An annual basal area growth rate of a tree grown using Model 1 (without climate variables) differed by 14.48 and 18.94 % at the end of 50 years for the 38 and 78 year old trees respectively (Fig. 4). In scenarios 1 and 4 (Fig. 4a and 4b; 5a and 5b), the smallest difference in annual basal area growth rate change was in the case (1c) and (4c) while the largest difference was in the case (1a) and (4a) for tree ages 38 and 78 years respectively (Table 5). Similarly, in scenario 2 and 3 (Fig. 4c and 4d; 5c and 5d), the smallest difference was observed in the case (2a) and (3a) and the largest difference in the case (2c) and (3c) for stand ages 38 and 78 years respectively (Table 5).

It was evident that for both stand ages, the maximum temperature at LQ with mean monthly precipitation amount (case 2a) showed better growth (Fig. 4c and 4d). It resulted in higher percentage difference in growth rates (14.84 and 17.17%) at the end of 50 years, when compared to the growth rate with no climate variables included model at the end of the prolonged scenario (Table 5). Similarly, increased mean monthly precipitation with maximum temperature at LQ (case 3a) showed better growth (Fig. 5a and 5b) with higher percentage differences in growth rate than other scenarios (Table 5). It was observed that increased mean temperature by 0.5 ^{0}C could easily affect the prediction of an annual basal area growth rate even when the mean monthly precipitation amount was of upper quartile range (case 4c) (Fig. 5c and 5d). The predicted annual growth rate by climate-based growth model were much less than the predicted growth rate by Model 1 without climate variable (Table 5). In such condition, the estimated total basal area growth at the end of 50 years would be smaller by 5 to 18 % than estimated by Model 1 depending upon the stand age (Table 5). However, the favorable climate scenarios: 2a, 3a and 3b maximized the estimate of the total basal area growth higher by 0.45 to 17.66 % at the end of 50 years (Table 5). In Table 5, the high negative value of differences in the total basal area growth indicated that the growth rate would highly stress in trees of younger stands with an increase in temperature than in trees of old stand age. Similarly, the high positive value of differences in the total basal area growth a tree of old stand age indicated the relaxed growth rate compared to younger stands grown in similar climatic condition. Therefore, it was observed that variability in the annual growth rate within different climate change scenarios was higher for the younger stand, and the old age stand was less affected at the end of the prolonged period of climate change (Fig 4 and 5).

Discussion

Growth Model 2 provided better standardized residual distribution patterns than other models tested but interestingly, the Model 3 with AR (1) performed about as well as Model 1. Theoretically, the estimated correlation (ρ) between the observations at successive lags (k) by Model 3 would be less reliable because it would base on only fewer residual pairs of observations (k = 5). Though we had fewer pairs of observations, the successive lags i.e. time apart are composed of a period of years, not of a single year. So, we believe the estimated correlation at successive lags would be still reliable. We tested a model with both power variance function and AR (1) structure that had a superior AIC value (-333777) but its fit index (62.65%) was inferior to both Models 2 and 3. The mixed effect model with power variance function (Model 5) provided substantially better AIC value than all alternative models, but the point estimate of the variance component (b7) and RSE were not superior. The large variance component and increased RSE of the Model 5 was also reported by Budhathoki et al. (2008). Interestingly, the mixed effect model with a hierarchical level of the random parameter that represents growth period and plots within-growth period (Model 6) performed as well as Model 1. Budhathoki et al. (2008) also fitted Model 6 without covariate “stand age” because of the convergence problem. Though, Model 7 successfully modeled power variance function for heterogeneous within-plot errors, it performed worse than Model 6. It could be that Model 7 did not greatly reduce RSE and the random effect associated with growth periods, as Model 6 did (Table 3). In fact, the performance of the mixed-effects model for prediction is often not substantially better than OLS models unless the mixed-effect models are calibrated or localized by using local data to estimate the random effect (Lynch et al., 2005; Temesgen et al. 2008; Lynch et al., 2012; Pokharel and Dech, 2012; Subedi and Sharma, 2013).

The growth models with power variance function were found to resolve the issue of heteroscedasticity of errors better than other competitive models, as seen in Fig (3) but RSE associated with the models with power variance function was large (Table 3). Mixed effect models with both AR (1) structure and variance power function can be used to reduce heteroscedasticity of errors as shown by Subedi and Sharma (2013) in their development of a diameter growth model with climate variables using tree ring data for a Black spruce and Jack pine trees in boreal Ontario, Canada. But in their study AR (1) structure was used to resolve the temporal autocorrelation among residuals within-stands not within-individual trees, and the power variance function was also used at a stand level to address the problem of heteroscedasticity within the stand. However, fitting a mixed-effects model (Model 4) with AR (1) within-individual tree observation was not possible because of conflict in modelling structures arising from the random effect associated with plot level (stand level), and the autocorrelation structure within individual trees. Interestingly, fitting Model 3 with AR (1) within-stand, the coefficient of autocorrelation among residuals was found slightly smaller (0.3584), but RSE (0.000653) and AIC (-

320365) were slightly bigger than Model 3 with AR (1) within-individual tree. It was also observed that adding climate variables in a model (climate-based growth model) did not significantly change the estimates of power variance, autocorrelation and random effect than a model with fewer predictors. For example the estimates: δ was 0.554928 for Model 10; ϕ was 0.452898 for Model 11; and $\hat{\sigma}_{\mu}$ was 3.672901 for Model 12; were similar in magnitude to the corresponding parameter estimates of Model 2, 3, and 4 respectively.

Previous climate-based growth models have often been formulated by introducing climate variables by adding them linearly to the existing growth model (Wykoff, 1990; Pokharel and Forsee 2009; Subedi and Sharma, 2013; Manso et al., 2015; Saud et al., 2015). However, linearly added climate variables did not perform as well as the multiplicative climate function did in our growth model. This reason could be that most previous climate-based growth models have been formulated either on tree-ring chronologies or on annual remeasurement data, where the climate variables are regressed against the annual growth (Biondi, 2000; Jump et al., 2006; Chhin et al. 2008; Martín-Benito, 2008; Duschesne, 2012; Foester et al., 2015; Subedi and Sharma, 2013). It may be that for growth models that already are linear in form (Subedi and Sharma, 2013; Manso et al., 2015), the linear inclusion of climate variables performs better than was the case with our nonlinear model. Interestingly, the model reported by Saud et al. (2015) was a nonlinear model form with the parameters and used linearly added climate variables in the Model 1.

The performance of climate-based growth models could differ for different tree species. For example, Subedi and Sharma (2013) reported 0.34 % and 1.56% reductions in RSE of annual diameter growth model of Jack pine and Black spruce respectively. However, Manso et al., (2015) reported substantial reduction (15%) in the AIC value rather than a decrease in RSE in annual diameter increment model of beech and oak stands. Saud et al., (2015) reported 0.94% improved the fit index of the climate-based growth model with the addition of climate variables. Subedi and Sharma (2013) used mean temperature, precipitation during the wettest quarter, and total precipitation during the growing season as climatic variables while Manso et al., (2015) used polynomial terms of mean temperature. Saud et al. (2015) used periodic average daily maximum temperature, periodic average daily mean temperature, and periodic mean daily precipitation during the growing season in the model.

The difference in the performance of the climate-based growth models i.e. Models 8 and 9 also demonstrated that the choice of climate variables used for the modelling purpose also greatly affects model improvement. The above contrasting findings suggested that improvements in the climate-based growth model not only depends on the variability in tree growth rate but are also controlled by the nature of climate variables used. For e.g. precipitation and temperature, which are dictated by regional conditions (altitudinal and latitudinal range) and site-specific factors (micro climate) limits forest productivity (Hanson and Weltzin, 2000; Boisvenue and Running, 2006; Park et al., 2013). Because of this, climate variables that are often expected to be influential, including mean temperature, maximum temperature, total annual precipitation (Pokharel and Froese, 2009; Subedi and Sharma, 2012; Manso et al., 2015) did not always perform well while fitting climate-based growth models.

Our results indicated that difference in maximum air temperature (DMTMAX, DTMAX6) and the difference in precipitation amount (DMPPT, DPPT9) were a better choice for providing better fit statistics (AIC, fit index) and reducing the RSE in climate-based growth models. This may be because the air temperature in the warmer areas during the growing season is stable and high that limits the plant productivity by inducing water stress (Llyod and Fastie, 2002; Way and Oren, 2010). However, if there is a drop in high temperature i.e. maximum air temperature, during the growing season it may show a significant effect on plant growth. The selected seasonal climate variables, DMTMAX and DMPPT, could be preferred over the monthly climate variables, DTMAX6 and DPPT9 while modelling climate-based growth model for periodic remeasurement data because the monthly climate variable for the short length of study period would have large variability (SD). These monthly climate variables could be a better choice for formulating the climate-based growth model using tree ring data because large tree ages are used for the study that would increase the total number of observations for the monthly climate variable to reduce variability. Additionally, these monthly climate variables would be useful for addressing early and latewood growth in a tree.

In our climate modifier, the difference of SPIs (DSPI) variables did not perform well to address the effect of climate in annual growth. However, randomForest showed that the linear combination of DSPI could be useful in addressing annual growth rate change (Fig. 2). The difference in monthly climate variables performed better than the difference in seasonal climate variables in improving model performance. Monthly climate variables have also been found to be important predictors of tree growth (Carrer and Urbinati, 2004). Shortleaf pine, which is more drought resistance than other southern pines (Lawson, 1990), had an annual growth rate that responded very well to monthly climate variables, compared to other alternatives. The selected monthly climate variables (DTMAX6 and DPPT9) suggested that moderate temperature and enough rainfall later in the growing season would favor growth. These variables also correspond to the latewood formation period in tree ring. It also indicated that shortleaf pine is facing water stress condition moderately during growth season. Binodi, (2000) indicated that Douglas-Fir in Idaho was grown on an arid site or had a moisture stress during growth period because the annual growth had a negative response to summer temperature and a positive response to late spring/early summer precipitation. Warmer growing season or the extended growing season, either spring warm up, or late fall has favored the growth of drought resistant tree species, for example, Jack Pine (Subedi and Sharma, 2013) and Rocky Mountain juniper (*Juniperus scopulorum*) (Spond et al., 2014). Before the start of summer, growing season precipitation also favored the growth of Lodgepole pine ((*Pinus contorta*) in Alberta, Canada (Chhin et al., 2008).

Tree-ring chronologies provide data that show a stronger correlation between annual growth rate and climate variables, but the correlation magnitude varies with tree species and location (Biondi, 2000; Spond et al., 2013). Interestingly, measurement data exhibited highly significant (*p-value* <0.0001) correlation between annual basal area growth and differences in seasonal climate variables but very weak, both positive and negative correlation. The significant correlation could be due to a large total number of observations, and very weak correlation could be a reason for fewer observed data point (five growth periods). The correlation of annual basal area growth with DMTMIN and DMTMEAN was positive while with DMPPT and DAPPT was negative. However, the correlation between DMTMAX and growth was not significant (*p-value* = 0.9317). The negative correlation with precipitation was due to a chance that annual basal area growth response of more than 2/3rd individual trees were confined within the range of minimum to mean amount of precipitation (Fig. 1b and 1d) during the growing season over a measurement period. Usually, the correlation of annual growth with mean temperature and precipitation during the growth season is positive (Briffa et al., 1998; Barber et al., 2000; Biondi, 2000; Manso et al., 2015). However, sometimes it could be negative in response to temperature and positive to precipitation (Jump et al., 2006; Chinn et al., 2008).

The annual basal area growth rate for individual trees in stands of different age was influenced by climate variables (Fig. 4 and 5). The growth modifier containing climate variables induced a noticeable effect on tree growth compared to the growth predicted by growth equation that did not contain any climate variables (Table 5). The variation in mean maximum temperature highly affected the basal area growth rates in both young and older stand ages than by the change in mean precipitation amount when one of the climate variables was held constant in the simulation (Fig. 4 and Table 5). It could be possible because the magnitude of the coefficient estimate of temperature was large than of precipitation to show an effect of small variation (Table 4). The decreased in the growth rate than Model 1 was associated with increased temperature then increased precipitation during the growing season (Fig. 5 and Table 5). It was evident that increased temperature greatly affected the growth of the younger-aged tree more than the older-aged tree (Table 5). In general, older stand age or larger diameter stands are more responsive to climate variation than younger stands, as reported by Callaway et al.

(1994), Hanson and Weltzin (2002), Carrer and Urbinati (2004), and Chhin (2008) in their studies. However, young stands exposed to increased temperature with varying precipitation amount could have more negative or positive growth responses than older stand ages (Table 5). Negative growth differences from Model 1 represent a stressed response in annual growth rate due to the positive difference in temperature degree, and positive growth differences from Model 1 are a relaxed response in annual growth rate due to the negative difference in temperature degree. However, a slight variation in mean precipitation amount does not greatly affected the growth rate (Fig. 4 and Table 5). It is also suggested that decrease in annual basal area growth rate is associated with the low amount of precipitation and increase in temperature during the growing season (Barber et al., 2000; Jump et al., 2006; Anderegg et al., 2013).

The simulations also indicate that changes in climate variables having a magnitude similar to that of feasible climate change scenarios can result in significant changes in tree growth rates, even though the inclusion of climate variables reduced the RSE by rather modest amounts. For example, simulations indicated very different percentage differences between initial and final tree basal areas after fifty years of growth under different climate scenarios (Table 5). The climate change with maximum temperature and precipitation amount as the mean response of the period of the 34 years of the period showed a moderate difference in the total basal area growth with a range of 11 to 12.58 66 % (Table 5). However, the moderate monthly maximum temperature and increased mean monthly precipitation amount showed indicated an increase of 0.45 % to 17.66 % in simulations for total basal area growth at the end of 50 years of growing length (Table 5). However, increased in mean maximum temperature by 0.5 0C showed greater differences of 4.95-18.8% in the total basal area growth even with high mean monthly precipitation amount (Fig. 5 and Table 5). It is difficult to establish the effects of changing the climate on growth conclusively, but some studies have demonstrated positive impacts on forest productivity if the water was not limiting (Boisvenue and Running, 2006).

Conclusions

Estimation of parameters in a growth model with repeated measurements using OLS violates the assumption of independent errors and may also violate the assumption of non-constant variance. This issue can be addressed by using weighted estimation with variance functions and mixed-effects estimation, or also by applying autoregressive correlation structures. However, these approaches do not always improve actual predictive capability as measured by fit statistics. As a result, there may be tradeoffs between predictive ability and application of remedial measures to enhance adherence to modeling assumptions. For our data, the basal area prediction equations with the overall best performance were either the model with AR (1) covariance structure or the mixed effects model that included growth period effects. However, a mixed-effects model that used a power variance functions provided better statistics of fit (AIC).

Though climate-based growth models showed modest improvements in RSE, this improvement could play a significant role in understanding the dynamics of growth change in response to climate variability. The potential effects on basal area growth of changes in climate variables similar in magnitude to feasible climate change scenarios were demonstrated in 50-year simulations of individual tree growth. The monthly average climate variables used in the growth models can be used as a proxy for seasonal climate variables while establishing the relationship of the annual growth with climate change. It is expected that the level of accuracy and precision of the climate-based growth model would increase if measurement data were reconstructed to annual growth or were obtained from dendrochronology. However, it is assumed that the developed climate-modifier can be easily replicated with the tree ring width data that would be more useful in providing the estimates of future forest growth response. The climate-based growth model will make it easier to forecast the changes in the productivity since the seasonal climate variables are obtained easily and estimated directly. These growth models might also improve the possibility of understanding stand structure changes over the time due to changing climate scenarios.

References

Anderegg, W.R.L., Kane, J.M., and Anderegg, L.D.L. 2013. Consequences of widespread tree mortality triggered by drought and temperature stress. Nat. Clim. Chang. 3(1): 30-36.

Andreassen, K., and Tomter, S.M. 2003. Basal area growth models for individual trees of Norway spruce, Scots pine, birch and other broadleaves in Norway. For. Ecol. Manage. 180(1–3): 11-24.

Barber, V.A., Juday, G.P., and Finney, B.P. 2000. Reduced growth of Alaskan white spruce in the twentieth century from temperature-induced drought stress. Nature 405(6787): 668-673.

Beguería, S., Vicente-Serrano, S.M., and Beguería, M.S. 2013. Package ‘SPEI’.

Biondi, F. 2000. Are Climate-Tree Growth Relationships Changing in North-Central Idaho, U.S.A.? Arct., Antarc., Alp. Res. 32(2): 111-116.

Boisvenue, C., and Running, S.W. 2006. Impacts of climate change on natural forest productivity – evidence since the middle of the 20th century. Glob. Chang. Biol. 12(5): 862-882.

Briffa, K.R., Schweingruber, F.H., Jones, P.D., Osborn, T.J., Shiyatov, S.G., and Vaganov, E.A. 1998. Reduced sensitivity of recent tree-growth to temperature at high northern latitudes. Nature 391(6668): 678-682.

Budhathoki, C.B., Lynch, T.B., and Guldin, J.M. 2008. Nonlinear mixed modeling of basal area growth for shortleaf pine. For. Ecol. Manage. 255(8–9): 3440-3446.

Callaway, R.M., DeLucia, E.H., and Schlesinger, W.H. 1994. Biomass Allocation of Montane and Desert Ponderosa Pine: An Analog for Response to Climate Change. Ecology 75(5): 1474-1481.

Carrer, M., and Urbinati, C. 2004. Age-dependent tree-ring growth responses to climate in Larix decidua and *Pinus cembra*. Ecology 85(3): 730-740.

Chhin, S., Hogg, E.H., Lieffers, V.J., and Huang, S. 2008. Potential effects of climate change on the growth of lodgepole pine across diameter size classes and ecological regions. For. Ecol. Manage. 256(10): 1692-1703.

Cryer, J.D., and Chan, K.-S. 2008. Time series analysis: with applications in R. Springerverlag New York.

Diéguez-Aranda, U., Burkhart, H.E., and Amateis, R.L. 2006. Dynamic Site Model for Loblolly Pine (*Pinus taeda*) Plantations in the United States. For. Sci. 52(3): 262-272.

Duchesne, L., Houle, D., and D'Orangeville, L. 2012. Influence of climate on seasonal patterns of stem increment of balsam fir in a boreal forest of Québec, Canada. Agric. For. Meteorol. 162–163: 108-114.

Foster, T., Schmalzer, P., and Fox, G. 2015. Seasonal climate and its differential impact on growth of co-occurring species. Eur. J. For. Res. 134(3): 497-510.

Graumlich, L.J. 1993. Response of tree growth to climatic variation in the mixed conifer and deciduous forests of the upper Great Lakes region. Can. J. For. Res. 23(2): 133-143.

Guttman, N.B. 1999. Accepting the standardized precipitation index: a calculation algorithm. J. Am. Water Resour. Assoc. 35(2): 311-322.

Hanson, P.J., and Weltzin, J.F. 2000. Drought disturbance from climate change: response of United States forests. Sci. Total Environ. 262(3): 205-220.

Hopkins, W.G., and Hüner, N. P. 1995. Introduction to plant physiology. Vol. 355. New York: Wiley.

Jiang, H., Radtke, P.J., Weiskittel, A.R., Coulston, J.W., and Guertin, P.J. 2014. Climate- and soil-based models of site productivity in eastern US tree species. Can. J. For. Res. 45(3): 325-342.

Jump, A.S., Hunt, J.M., and PeÑUelas, J. 2006. Rapid climate change-related growth decline at the southern range edge of Fagus sylvatica. Glob.Chang. Biol. 12(11): 2163-2174.

Lawson, E.R. 1990. *Pinus echninata* Mill. Shortleaf Pine. *In* Silvics of North America; Volume 1, Conifers: 316-326. Burns RM, Honkala BH, technical coordinators. U.S. Department of Agriculture, Forest Service Handbook 654. Washington, DC.

Lhotka, J.M., and Loewenstein, E.F. 2011. An individual-tree diameter growth model for managed uneven-aged oak-shortleaf pine stands in the Ozark Highlands of Missouri, USA. For. Ecol. Manage. 261(3): 770-778.

Liaw, A., and Wiener, M. 2002. Classification and regression by randomForest. R news 2(3): 18-22.

Lloyd, A., and Fastie, C. 2002. Spatial and Temporal Variability in the Growth and Climate Response of Treeline Trees in Alaska. Clim. Chang 52(4): 481-509.

Loehle, C., and LeBlanc, D. 1996. Model-based assessments of climate change effects on forests: a critical review. Ecol. Model. 90(1): 1-31.

Lynch, T.B., Budathoki, C. and Wittwer, R.F. 2012. Relationship between Height, Diameter, and Crown for Eastern Cottownwood (*Populus Deltoides*) in a Great Plains Riparian Ecosytem. West. J. Appl. For. 27, 176-186.

Lynch, T.B., Hitch, K.L., Huebschmann, M.M., and Murphy, P.A. 1999. An Individual-Tree Growth and Yield Prediction System for Even-Aged Natural Shortleaf Pine Forests. South. J. Appl. For. 23(4): 203-211.

Lynch, T.B., Holley, A.G., and Stevenson, D.J. 2005. A Random-Parameter Height-Dbh Model for Cherrybark Oak. South. J. Appl. For. 29(1): 22-26.

Lynch, T.B., Murphy, P.A., and Lawson, E.R. 1991. Stand volume equations for managed natural even-aged shortleaf pine in eastern Oklahoma and western Arkansas. Research report P-Oklahoma Agricultural Experiment Station.

Manso, R., Morneau, F., Ningre, F., and Fortin, M. 2015. Effect of climate and intra- and inter-specific competition on diameter increment in beech and oak stands. Forestry. doi: 10.1093/forestry/cpv020.

Martín-Benito, D., Cherubini, P., del Río, M., and Cañellas, I. 2008. Growth response to climate and drought in *Pinus nigra* Arn. trees of different crown classes. Trees 22(3): 363-373.

Monserud, R.A., and Sterba, H. 1996. A basal area increment model for individual trees growing in even- and uneven-aged forest stands in Austria. For. Ecol. Manage. 80(1–3): 57-80.

Moore, D.J.P., Aref, S., Ho, R.M., Pippen, J.S., Hamilton, J.G., and De Lucia, E.H. 2006. Annual basal area increment and growth duration of *Pinus taeda* in response to eight years of free-air carbon dioxide enrichment. Glob. Chang. Biol. 12(8): 1367-1377.

Moscow Forestry Sciences Laboratory. 2015. Custom Climate Data Requests. Available from http://forest.moscowfsl.wsu.edu/climate/customData/index.php [accessed 4 Aug 2015]

Murphy, P.A., Lawson, E.R., and Lynch, T.B. 1992. Basal area and volume development of natural even-aged shortleaf pine stands in the Ouachita Mountains. South. J. Appl. For. 16(1): 30-34.

Park, W.A., Allen, C.D., Macalady, A.K., Griffin, D., Woodhouse, C.A., Meko, D.M., Swetnam, T.W., Rauscher, S.A., Seager, R., Grissino-Mayer, H.D., Dean, J.S., Cook, E.R., Gangodagamage, C., Cai, M., and McDowell, N.G. 2013. Temperature as a potent driver of regional forest drought stress and tree mortality. Nat. Clim. Chang. 3(3): 292-297.

Pinheiro, J., and Bates, D. 2002. Mixed-effects models in S and S-PLUS. Springer Science & Business Media. pp. 391.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. and R Core Team 2015. Linear and Nonlinear Mixed Effects Models, R package version 3.1-122.

PRISM Climate Group. 2015. Parameter-elevation regression on independent slopes model, http://www.prism.oregonstate.edu/explorer/ created 14 June, 2015.

Pokharel, B., and Dech, J.P. 2012. Mixed-effects basal area increment models for tree species in the boreal forest of Ontario, Canada using an ecological land classification approach to incorporate site effects. Forestry 85(2): 255-270.

Pokharel, B., and Froese, R.E. 2009. Representing site productivity in the basal area increment model for FVS-Ontario. For. Ecol. Manage. 258(5): 657-666.

R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Saud, P., Lynch, T.B., Wilson, D.S., Stewart, J., Guldin, J.M., Heinemann, B., Holeman, R., Wilson, D., and Anderson, K. 2015. Influence of weather and climate variables on the basal area growth of individual shortleaf pine trees. In Holley, A.G., Connor, K.F., and Haywood, J.D (eds) Proceedings of the 17th Biennial Southern Silvicultural Research Conference, Shrevport, Louisiana, March 5-7, 2013 (pp.406-408), , Southern Research Station, Asheville, NC.

Spond, M.D., van de Gevel, S.L., and Grissino-Mayer, H.D. 2014. Climate-growth relationships for Rocky Mountain juniper (*Juniperus scopulorum* Sarg.) on the volcanic badlands of western New Mexico, USA. Dendrochronologia 32(2): 137-143.

Subedi, N., and Sharma, M. 2011. Individual-tree diameter growth models for black spruce and jack pine plantations in northern Ontario. For. Ecol. Manage. 261(11): 2140-2148.

Subedi, N., and Sharma, M. 2013. Climate-diameter growth relationships of black spruce and jack pine trees in boreal Ontario, Canada. Glob. Chang. Biol. 19(2): 505-516.

Temesgen, H., Monleon, V.J., & Hann, D.W. 2008. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. *Can. J. For. Res. 38*(3), 553-565.

U.S. Department of Interior, Geological Survey. 1970. The National Atlas of the United States of America. U.S. Department of Interior, Geological Survey, Washington, DC. 97 pp.

Uzoh, F.C., and Oliver, W.W. 2008. Individual tree diameter increment model for managed even-aged stands of ponderosa pine throughout the western United States using a multilevel linear mixed effects model. For. Ecol. Manage. 256(3): 438-445.

Way, D.A., and Oren, R. 2010. Differential responses to changes in growth temperature between trees from different functional groups and biomes: a review and synthesis of data. Tree Physiol 30(6): 669-688.

Wilmking, M., Juday, G.P., Barber, V.A., and Zald, H.S.J. 2004. Recent climate warming forces contrasting growth responses of white spruce at treeline in Alaska through temperature thresholds. Glob. Chang. Biol. 10(10): 1724-1736.

Wykoff, W.R. 1990. A Basal Area Increment Model for Individual Conifers in the Northern Rocky Mountains. For. Sci. 36(4): 1077-1104.

Yeh, H.-Y., and Wensel, L.C. 2000. The relationship between tree diameter growth and climate for coniferous species in northern California. . Can. J. For. Res. 30(9): 1463-1471.

Tables

Table IV-1 Mean and standard deviation of tree level and stand variables recorded for six times measurement of the natural stand of shortleaf pine. Values inside bracket are standard deviation and outside are means of the recorded observations of each measurement.

Variables	Measurements					
	1st (N=8290)	2nd (N=8088)	3rd (N=7587)	4th (N=4722)	5th (N=4403)	6th (N=4286)
Dbh (cm)	18.832 (9.891)	20.716 (9.976)	23.176 (10.068)	27.595 (9.798)	29.495 (9.943)	31.202 (10.37)
Basal Area (m^2)	0.036 (0.036)	0.042 (0.038)	0.05 (0.042)	0.067 (0.047)	0.076 (0.05)	0.085 (0.055)
Plot age (years)	41.448 (19.628)	46.286 (19.659)	52.02 (19.804)	59.606 (20.067)	65.181 (20.02)	70.993 (20.092)
Site index (m)	17.449 (2.89)	17.227 (2.982)	17.298 (3.008)	17.544 (3.045)	17.421 (3.022)	17.397 (3.031)
Plot basal area (m^2 ha^{-1})	1.724 (0.542)	2.084 (0.597)	2.406 (0.68)	2.052 (0.678)	2.262 (0.703)	2.444 (0.734)

Table IV-2 Summary of the tree level and stand level variables used in the development of the model. The values inside bracket are standard deviation and outside are means of the recorded observations (N) for the respective five growth periods.

Variables	Growth periods				
	1st (N = 8088)	2nd (N = 7582)	3rd (N = 3862)	4th (N = 3631)	5th (N = 3611)
ABAG	0.0012	0.0014	0.0015	0.0015	0.0014
(m^2/tree)/year	(0.001)	(0.0011)	(0.0012)	(0.0011)	(0.0011)
Mid dbh (cm)	19.8633	22.303	26.2529	28.216	30.1064
	(9.8877)	(9.8844)	(9.5488)	(9.7952)	(10.1687)
Mid Basal area (m^2)	0.0387	0.0467	0.0613	0.0701	0.0793
	(0.0369)	(0.0402)	(0.0438)	(0.0477)	(0.0523)
Mid stand basal	9.4417	11.1523	9.8758	10.8464	11.9934(
area (m^2ha^{-1})	(2.7713)	(3.1506)	(3.2463)	(3.315)	3.5879)
Mid ratio of QMD	0.704	0.6828	0.6788	0.6778	0.6716
to dbh	(0.218)	(0.1993)	(0.1674)	(0.1693)	(0.1653)
Mid stand age	44.29	49.96	56.96	61.98	68.00
(year)	(19.66)	(19.76)	(20.21)	(20.22)	(20.25)

Table IV-3 Summary statistics for fitted basal area growth models without climate variables (total observations = 26774 and the total number of plots = 208).

Model	Residual d.f.	AIC	Log-Likelihood	L. Ratio[a]	Fit Index	$\hat{\sigma}_{T(p)\mu}$	$\hat{\sigma}_{\mu}$	$\hat{\sigma}_{(\varepsilon)}$
1	26767	-316881	158448	-	64.13	-	-	0.0006511
2	26767	-328136	164077	1 vs 2 **	63.17	-	-	0.0034059
3	26767	-321126	160572	-	63.95	-	-	0.0006486
4	26559	-320420	160219	-	61.29	-	3.786194	0.0006007
5	26559	-331432	165726	4 vs 5 **	60.79	-	6.894290	0.0029902
6	25764	-317209	158615	-	64.12	0.806356	2.07×10^{-10}	0.0006475
7	25764	-328426	164224	6 vs 7 **	63.11	1.202574	3.91×10^{-10}	0.0033842
8	26765	-317424	158772		64.85	-	-	0.0006445
9	26774	-317406	158714		64.83	-	-	0.0006447
10	26765	-328881	164452	9 vs 10**	63.79	-	-	0.0033811
11	26765	-322232	161127		62.47	-	-	0.0006419
12	26558	-320819	160420		61.52		3.415525	0.0005962

[a]L.Ratio = Likelihood ratio test; ** Likelihood ratio test significant with *p-value* < 0.0001.

Table IV-4 Parameter estimates and standard error of the fitted annual basal area growth models without climate variables (Mod 1, 3, 5 and 6) and with climate variables (Mod 8, 9 and 11).

Parameter	Mod 1	Mod 3	Mod 5	Mod 6	Mod 8	Mod9
	Estimate (Std err.)	Estimate (Std err.)	Estimate (Std err.)	Estimate (Std err.)	Estimate (Std err.)	Estimate (Std err.)
β_1	0.032751	0.034853	0.028650	0.032910	0.030788	0.037402
	(0.001794)	(0.002425)	(0.001524)	(0.001797)	(0.001682)	(0.001908)
β_2	0.506295	0.531768	0.538360	0.500289	0.514333	0.622298
	(0.015336)	(0.018314)	(0.013301)	(0.015348)	(0.015533)	(0.010852)
β_3	-1.501148	-1.608526	-3.616883	-1.456620	-1.569555	-3.214365
	(0.084148)	(0.114678)	(0.091735)	(0.082768)	(0.082990)	(0.100866)
β_4	0.105153	0.096481	0.153471	0.105266	0.108146	0.145373
	(0.002039)	(0.002585)	(0.002988)	(0.002000)	(0.002077)	(0.002846)
β_5	0.020915	0.021051	0.010896	0.020591	0.020792	0.025758
	(0.000569)	(0.000764)	(0.000801)	(0.000549)	(0.000568)	(0.000689)
β_6	1.330286	1.521592	3.485364	1.335173	1.329651	2.426962
	(0.048986)	(0.064697)	(0.060988)	(0.048441	(0.049061)	(0.049605)
β_7	-10.816504	-10.511766	-4.456743	-10.616484	-11.040174	-11.252718
	(0.300462)	(0.395363)	(0.661425)	(0.465513)	(0.304651)	(0.391013)
β_8	-	-	-	-	-0.599612	-0.430701
					(0.046033)	(0.038073)
β_9	-	-	-	-	0.055932	0.039677
					(0.004855)	(0.003865)

Table IV-5 Percentage difference in the annual basal area growth rate and total basal area of an individual tree at stand age of 38 and 78 years at the end of 50 years under four different scenarios using Model 8 compared to Model 1 without climate.

Percentage (%) difference in growth rate	Cases	Stand Age 38 year				Stand age 78 year			
		Scenario				Scenario			
		1	2	3	4	1	2	3	4
Between beginning and at the end of 50 years	a	18.65	7.89	4.79	21.86	21.84	11.62	8.82	24.69
	b	15.86	15.94	13.49	19.42	19.18	19.62	16.80	22.53
	c	13.42	21.91	20.10	17.33	16.93	26.47	23.08	20.38
At the end of 50 years compared to Model 1	a	-13.00	14.84	21.68	-22.45	-9.90	17.17	23.91	-19.30
	b	-5.40	-5.39	1.07	-15.25	-2.35	-3.21	4.16	-12.12
	c	1.06	-23.01	-17.12	-9.00	4.01	-24.93	-13.73	-5.86
In total basal area at the end of 50 years compared to Model 1	a	-10.54	11.08	16.27	-18.08	-8.04	12.58	17.66	-15.45
	b	-4.56	-4.59	0.45	-12.31	-2.16	-2.82	2.75	-9.76
	c	0.42	-18.57	-13.79	-7.40	2.75	-19.84	-10.99	-4.95

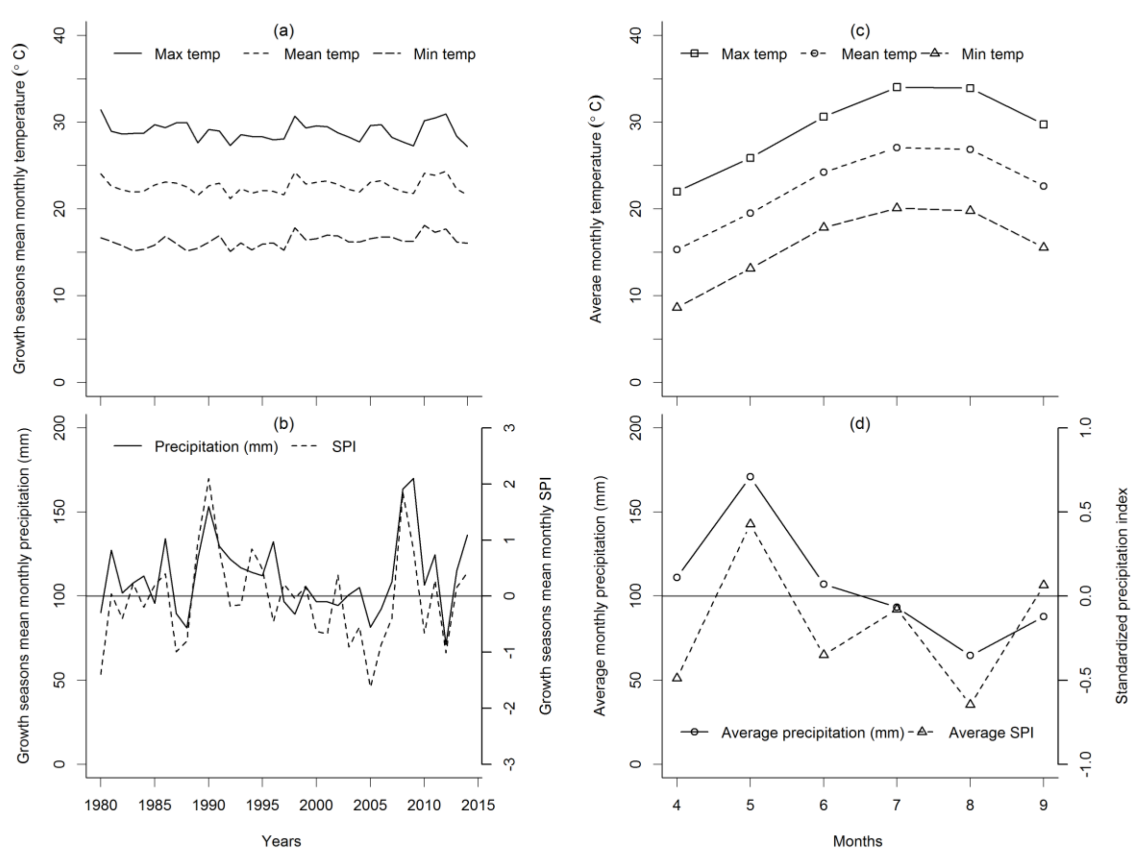

Fig. IV- 1 Distribution of temperature and precipitation during the months of growing season over the length of 1980-2014**. a)** Mean maximum, minimum and mean air temperature during growing season of the year; **b)** Mean total monthly perception and monthly standardized perception index (SPI) during growing season of the year; **c)** An average temperature of the month of growing seasons over the length of the study period (1984-2014); **d)** Average precipitation and standardized precipitation index of the month of the growing seasons over the length of the study period (1984-2014). Months “4-9” represents “April-September”.

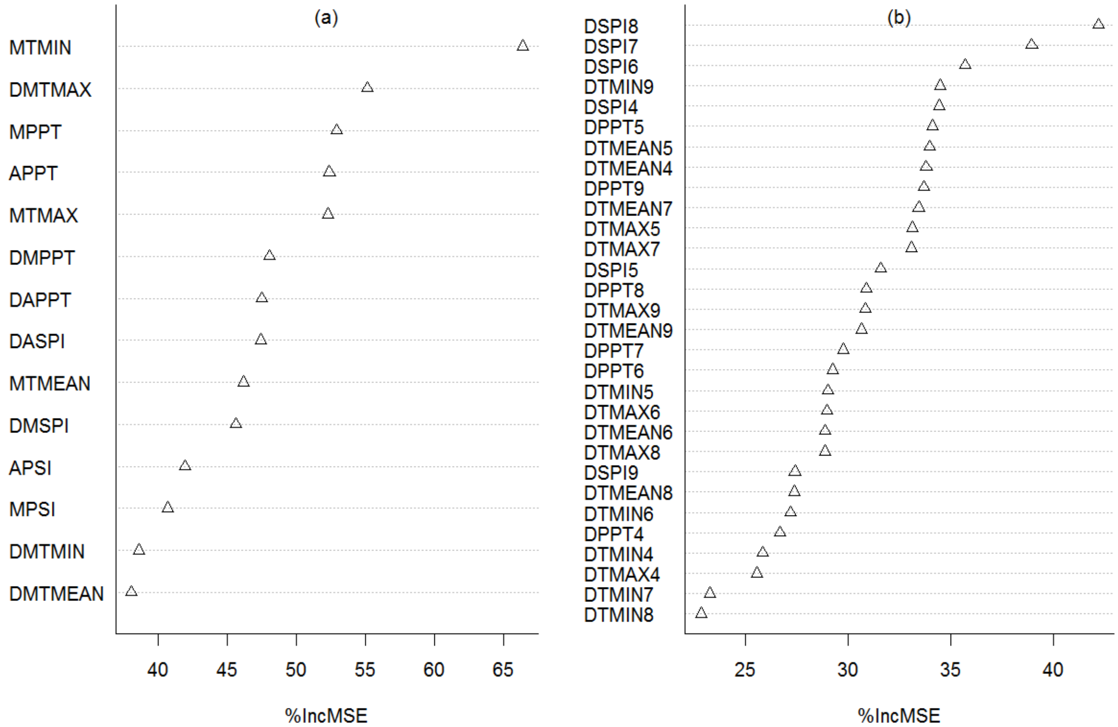

Fig. IV- 2 Random forests plot showing the importance value of each seasonal climate variables and monthly climate variables (from top to bottom in order) when used for predicting annual basal area growth. The variable importance is measured by %IncMSE, which refers to the mean decrease in accuracy (the greater the value, the better). (a) Seasonal climate variables over the growth period, and (b) Monthly climate variables, where number indicate the corresponding month of a calendar year; TMIN = minimum temperature, TMAX = maximum temperature, *TMEAN* = mean temperature and *SPI* = standardized precipitation index. "D" as the initial letter indicates the difference between the estimated value of a climate variable during the growth period and the mean value of the respective climate variable over the period of 34 years.

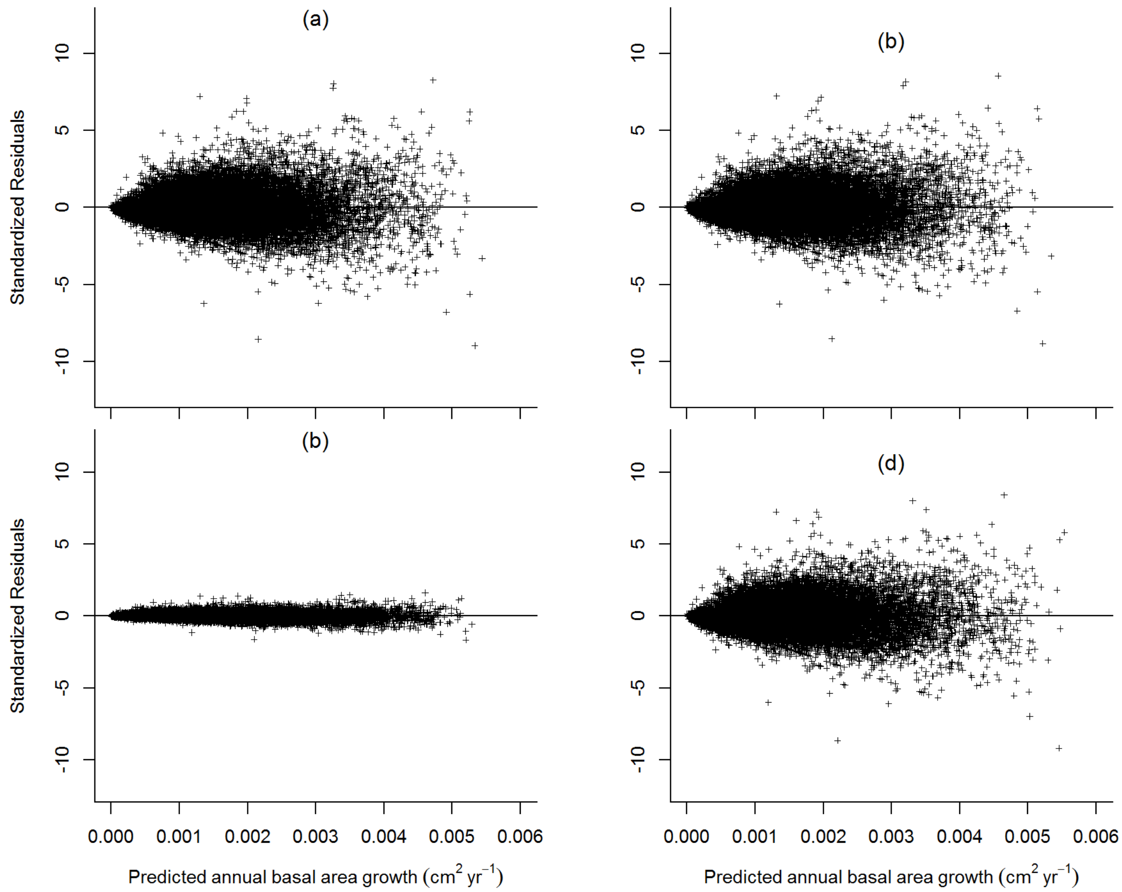

Fig. IV-3 Standardized residuals (observed-predicted) from annual basal area growth models for shortleaf pine: (a) Model 1; (b) Model 2 with power variance function; (c) Model 3 with AR (1); and (c) Model 6 with random-effects for growth period and with plots within growth period.

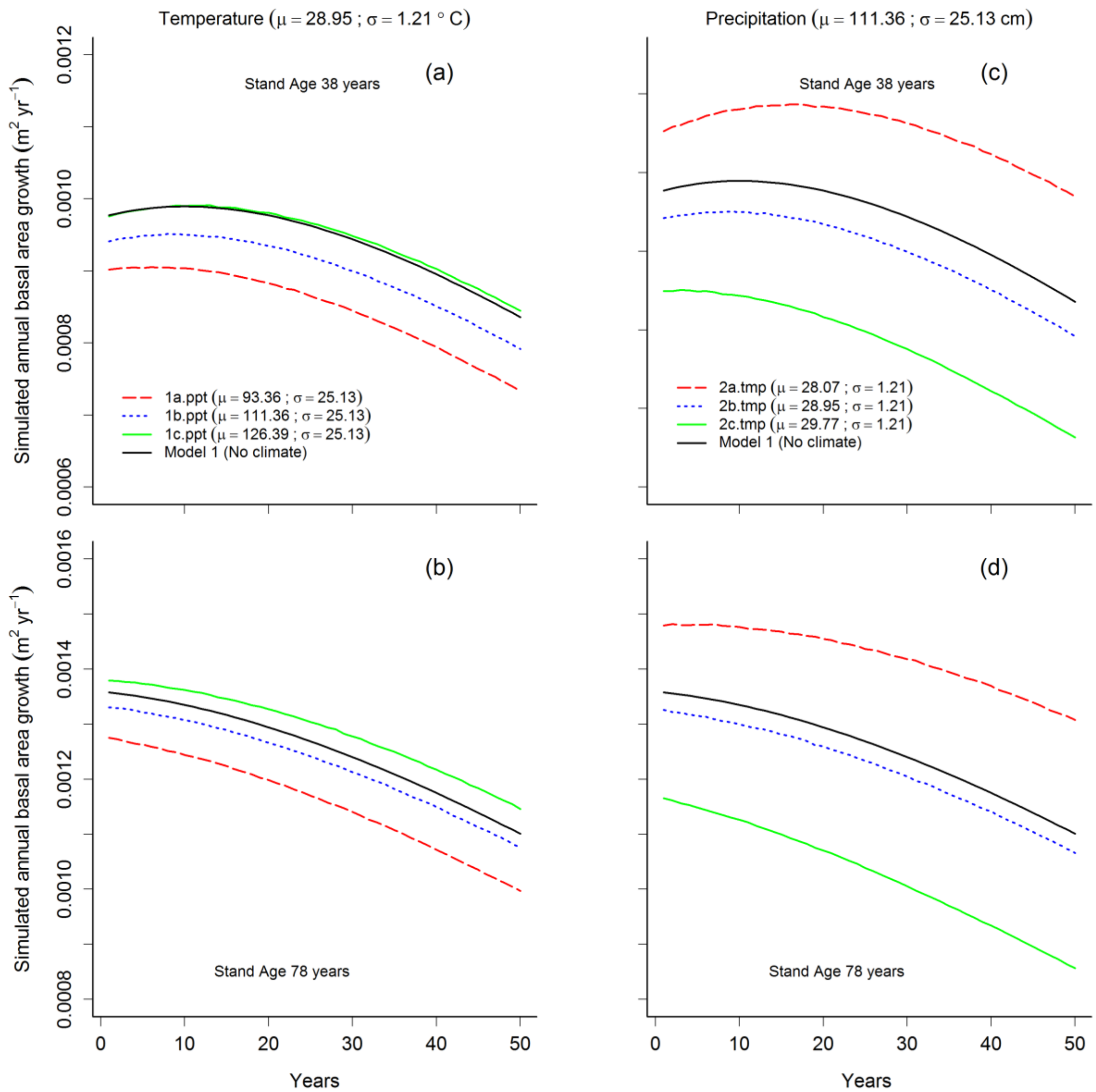

Fig. IV-4 Simulated scenario 1 and 2 for annual basal area growth of an individual tree from stand age of 38 and 78 years with climate-based growth Model 8 and Model 1 (No climate). Figure (a) and (b) (left panel) shows variation in annual basal area growth rate from scenario 1 with three different cases. Similarly, figure (c) and (d) (right panel) show variation in annual basal area growth rate from scenario 2 with three different cases.

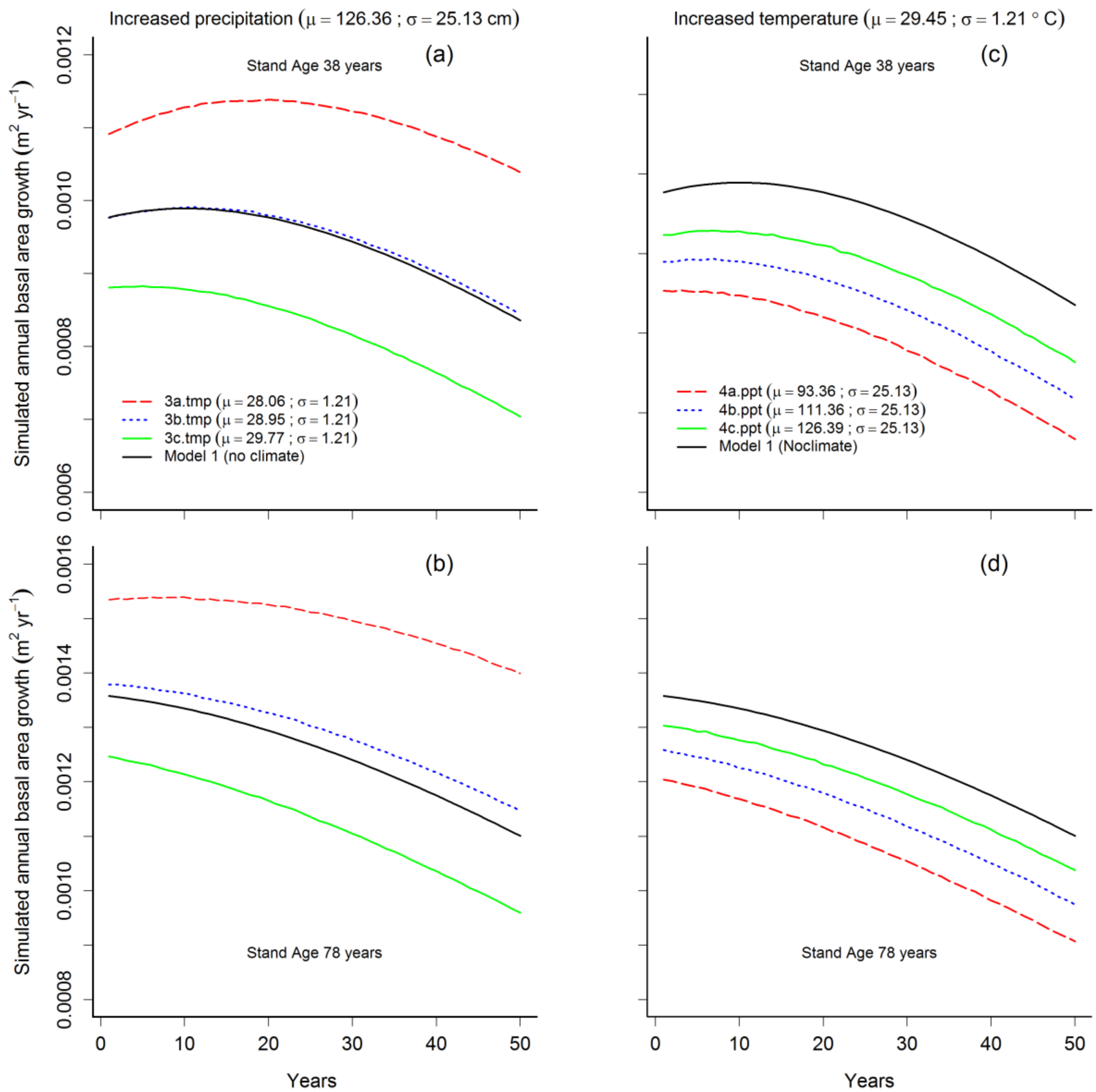

Fig. IV-5 Simulated scenario 3 and 4 for annual basal area growth of an individual tree from stand age of 38 and 78 years with climate-based growth Model 8 and Model 1 (No climate). Figure (a) and b) (left panel) shows variation in annual basal area growth rate from scenario 3 with three different cases. Similarly, figure (c) and (d) (right panel) show variation in annual basal area growth rate from scenario 4 with three different cases.

CHAPTER V

CONCLUSIONS

The modified dbh-height relationship by using stand-level competition, quadratic mean diameter, provided better prediction ability than the existing model. Quadratic mean diameter also enhanced the precision of height prediction model forms that uses the only dbh as an independent variable and do not utilize dominant height as a covariate. Quadratic mean diameter can be easily constructed from dbh data that are available in inventory data. Crown ratio estimation of an individual tree was improved by using the relative spacing index in the existing model. The relative spacing better explained variability in crown ratio estimation than stand level competition (basal area per hectare) and stand age in the model. The logistic function could be the comparable alternative to nonlinear model choice for estimating the crown ratio of an individual tree.

Parameter estimates of the revised model for height prediction and crown ratio estimation can be incorporated in the Shortleaf Pine Stand Simulator (Huebschmann et al. 1988) which can be used to develop information for practical forest management decision making for naturally occurring even-aged shortleaf pine forests. The relationships between dbh, height and crown ratio could have an significant implication in inventories for biomass and carbon estimation of a natural stand of shortleaf pine in the southern USA. To the extent that these formulations are novel in the forestry literature, they could be considered for application in other forest types in addition to well-known existing equation forms.

A substantial amount of the variability in predicting the probability of mortality of an individual tree was explained by the measures of stand level competition and sizes of individual trees with age i.e. BAHG and DAG, and stand level competition (QMD). The binary logistic model and an iteratively reweighted nonlinear regression model with the same independent variables provided similar parameter estimates. However, the binary estimation technique performed better because it provided smaller standard errors for parameter estimates. The binary logistic model fitted with the mixed-effects approach provided significant parameter estimates with smaller AIC values but the mortality predictions from this model evaluated using $\chi 2$ test value were inferior to binary model. The low thinning would not have strong effects in predicting mortality but could increase a risk of morality if it is triggered by logging effect. The binary logistic regression model could be used for assessing the mortality of naturally occurring even-aged shortleaf pine forests in the Southern USA. It also can be tested for use in a shortleaf pine growth simulation such as the Shortleaf Pine Stand Simulator (SLPSS) (Huebschmann et al., 1998), a distance-independent individual tree growth simulator.

The climate-based basal area growth models showed small improvements in fit statistics compared to the basal area growth models that did not include climate variables. However, this relatively small improvement could play a significant role in helping us to understand the dynamics of growth change in response to climate variability. The seasonal climate variables, deviation in average minimum monthly air temperature and average monthly precipitation from the decade-long average climate are helpful in explaining variability in annual basal area growth rate. Similarly, deviation in monthly average climate variables used in the growth models can be used as a proxy for seasonal climate variables in establishing the relationship with annual radial growth. The formulation of climate modifier successfully modelled the influence of deviation in mean climate response in the growth of an individual tree. Such climate-based growth models might also help to improve understanding of possible of stand structure changes over the time due to climate change scenarios.

The fit statistics (RMSE and fit index) of the mixed-effect models of all individual growth models, when predictions were based only on the fixed effects parameters were similar to or inferior to those of the best fitted nonlinear models fitted by OLS, though AIC values were smaller for mixed-effect models. However, the mixed-effect models may provide improved predictions when calibration data are available. Estimation of parameters in individual tree growth models with repeated measurements using OLS violates the assumption of independent errors and may also violate the assumption of non-constant variance. This issue can be addressed by using weighted estimation with variance functions and mixed-effects estimation, or also by applying autoregressive correlation structures. However, these approaches do not always improve actual predictive capability as measured by fit statistics. As a result, there may be tradeoffs between predictive ability and application of remedial measures to enhance adherence to modeling assumptions. For our data, the OLS model for height-dbh relationship and crown ratio estimation with AR (1) structure and power variance function showed better performances in AIC value and prediction ability than other alternatives. Similarly, the annual basal area prediction equations with the overall best performance were either the model with AR (1) structure or the mixed effects model that included growth period effects. However, a mixed-effects model that used a variance power functions provided better statistics (AIC).

VITA

Pradip Saud

Candidate for the Degree of

Doctor of Philosophy

Thesis: MODELLING INDIVIDUAL TREE GROWTH OF NATURALLY OCCURING EVEN-AGED SHORTLEAF PINE (*Pinus echinata* mill.)

Major Field: Natural Resource Ecology and Management

Biographical:

Completed the requirements for the Doctor of Philosophy in Natural Resource Ecology and Management at Oklahoma State University, Stillwater, Oklahoma in December 2015.

Completed the requirements for the Master of Science in Forestry at West Virginia State University, Morgantown, West Virginia, 2011.

Completed the requirements for the Bachelor of Science in Forestry at Tribhuvan University, Kathmandu, Nepal in 2006.

Experience:

Research Assistant (2011-2015), Department of Natural Resource Ecology and Management, Oklahoma State University, Stillwater, Oklahoma.

Research Assistant (2009-2011), Division of Forestry and Natural Resources, West Virginia University, Morgantown, West Virginia.

Program Officer (2007-2009), Program Officer, Green Governance Nepal.

Program Officer (2006-2007), Field Research Assistant, Nepal Agroforestry Foundation, Kathmandu.

Professional Memberships:

Green Governance Nepal, Kathmandu, Nepal.
Nepal Foresters Association.
Society of American Foresters.
Xi Sigma Pi