MODELING THE WOOD PROPERTIES OF LOBLOLLY PINE (*PINUS TAEDA* L.) GROWING IN SOUTHERN UNITED STATES

by

FINTO ANTONY

(Under the direction of Richard F. Daniels)

Abstract

This dissertation is divided in to two broad sections. In the first section, models are developed that explain specific gravity (SG) variation within planted loblolly pine (*Pinus taeda* L.) from the southern United States. A three segmented quadratic model and a semiparametric model were proposed to explain the longitudinal variation of SG within-trees. Based on both models, regional variation in mean disk SG was observed. SG was highest for trees from the South Atlantic and Gulf Coastal Plain and lowest for trees from the Upper Coastal, Hilly Coastal, North Atlantic Coastal Plain and Piedmont. Maps explaining the regional variation in SG at specified heights within a tree were also developed based on the semiparametric model. A multivariate model system for disk SG and moisture content (MC) was also proposed. The proposed model system took account of the contemporaneous correlation between the two properties and was utilized to improve the prediction of one property given that information on the other property was available. Two subject specific prediction approaches commonly used in forestry (Generalized Algebraic Difference Approach - GADA and Nonlinear Mixed Models - NLMM), were evaluated empirically using a subset of disk SG data. The NLMM approach was found to perform better than GADA in terms of root mean square error (RMSE), and mean absolute residual (MAR).

In the second section the effects of midrotation fertilization on various growth and wood properties were evaluated and the influence of fertilization on latewood SG was modeled. The effects of midrotation fertilization on growth and wood properties of loblolly pine in thinned and unthinned stands were explored. It was observed that both growth and wood property responses were higher in magnitude in the stand that received midrotation fertilization following thinning. Finally, the response of latewood SG following midrotation fertilization in a thinned stand followed a consistent pattern with the rate of nitrogen applied and was modeled successfully.

Index Words: fertilization,generalized algebraic difference approach (GADA), mixed models, moisture content, multivariate, *Pinus taeda*, simultaneous models, specific gravity, thinning, within-tree variation, wood property maps.

MODELING THE WOOD PROPERTIES OF LOBLOLLY PINE (*PINUS TAEDA* L.) GROWING IN SOUTHERN UNITED STATES

by

FINTO ANTONY

B. S., Kerala Agricultural University, India, 2001

M. S., University of Georgia, 2006

A Thesis Submitted to the Graduate Faculty of

The University of Georgia in Partial Fulfillment

of the

Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2010

©2010

Finto Antony

All Rights Reserved

MODELING THE WOOD PROPERTIES OF LOBLOLLY PINE (*PINUS TAEDA* L.)

GROWING IN SOUTHERN UNITED STATES

by

FINTO ANTONY

Approved:

Major Professor: Richard F. Daniels

Committee: Laurence R. Schimleck

Bruce E. Borders

Chris J. Cieszewski

Daniel B. Hall

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2010

Dedication

To my parents Antony Inchodikaran and Philomina Antony, brother Linto Antony and to my wife Assa Finto

"A person who never made a mistake never tried anything new"

Albert Einstein

Acknowledgements

I express my deep sense of gratitude and indebtedness to Dr. Richard F. Daniels, for his constant inspiration and evaluation without which this endeavor would not have been fruitful. I am also deeply indebted to Dr. Laurie Schimleck (Thanks!!) and Mr. Alexander Clark for their valuable advice and help given to me during the course of my study. I express my deep sense of gratitude to Dr. Dan Hall, Lewis Jordan and Dr. Brock Stewart for their valuable suggestions during the statistical analysis of the data. I am also indebted to Dr. Chris J. Cieszewski and Dr. Bruce Borders for their comments and serving as members in my committee. I thank many of my friends for their whole hearted assistance at times of great need without which this study would not have borne fruits. Finally, I would like to express my love and appreciation to my parents, Mr. Antony and Mrs. Philomina and my wife Assa.

Table of Contents

Chapter 1

Introduction and Review of Literature

1.1 Introduction

Pines are major plantation tree species in the southern United States (US), occupying an area of approximately 13 million ha and carrying a timber volume of about 680 million cubic meters. Pine plantation area is expected to increase in the future with a projected expansion of 67 % (22 million ha) by 2040 (Wear and Greis 2002). According to Wear and Greis (2002), the southern US is a major supplier to both the US (58 %) and world timber markets (16 %). The important role that the southern US plays in national and international timber supply was made possible by improved productivity and marked reduction in rotation length of plantation pine during recent decades (Fox et al. 2007b) .

Today the southern US produces more wood than any other region in the world (Prestemon and Abt 2002). However, declining forest area and use of conventional management practices have resulted in a gradual decline in wood fiber production from this region. Typically the growth rate of pine in the southern US is about 5 to 6 tons per acre per year (Allen et al. 2005; Borders and Bailey 2001) which is substantially lower than growth rates obtained in other pine growing regions around the world, such as Argentina, Australia, Brazil, Chile, New Zealand and South Africa (Borders and Bailey 2001). Various field trials (Amateis 2000; Borders and Bailey 2001) have shown that the current productivity of southern US pine plantations is well below what potentially could be achieved and could be substantially improved by widespread adoption of intensive silvicultural practices in pine management.

Intensive silvicultural practices have an important role to play in improving the growth of southern pines (Allen et al. 2005). For example, growth rates of 25 tons per ha per year or more are achievable with some of the newly available plantation management techniques (Stanturf et al. 2003).Various research organizations and cooperatives, such as the Plantation Management Research Cooperative (PMRC), the Consortium for Accelerated Pine Production Studies (CAPPS), and the Forest Nutrition Cooperative (FNC) etc. have used extensive field trials over the last few decades to compare growth gains of intensive management with conventional management. Large gains in growth of pine plantations in the southeastern US have been reported from the use of improved seedlings (Stanturf et al. 2003), vegetation control (Miller et al. 1991; Pienaar and Shiver 1993), thinning (Amateis et al. 1996; Haywood 2005; Thomas et al. 1995) and fertilization (Dickens et al. 2003; Fox et al. 2007a).

In terms of growth and yield improvement, these practices are promising. However, growers and buyers are concerned with the quality of wood produced from such fast grown plantations. A thorough understanding of natural variation of different wood properties, within trees (radially and longitudinally), from tree-to-tree within a stand, from stand-to-stand within a region and from region-to-region for the range of loblolly pine in the southern US, will provide the basic framework for addressing these concerns. Detailed knowledge regarding wood quality responses following different silvicultural practices will add more information to this basic outline. Modeling natural variation in different wood properties which are important from a quality perspective and their responses following different silvicultural practices will provide the growers and buyers tools to predict changes in wood quality resulting from any silvicultural practice. The Wood Quality Consortium is a research cooperative created at the University of Georgia, established in 2000 to address the following three objectives: 1) characterize the variability in wood properties of southern pines; 2) determine the effects of intensive silviculture on wood properties; and 3) develop models to predict the wood properties. The following section of this thesis presents an overview of natural variation in important wood properties of loblolly pine with special emphasis on wood specific gravity (SG). In addition the effect of different silvicultural practices on wood properties will be reviewed.

1.2 Loblolly Pine

Loblolly pine (*Pinus taeda* L.) is an evergreen tree belonging to the family Pinaceae, one of several so-called southern pines native to the US. Loblolly pine typically reaches heights of 30 to 35 m and diameters of 0.5 to 1.5 m. It is a major commercial tree species in the southeastern US, occupying an area of around 11.7 million ha which represents more than half of the standing pine volume in the region (Burns and Honkala 1990). The species is also called old-field pine because it frequently colonizes abandoned fields. Loblolly pine has been planted widely in the southern US and is the most common species used in intensive plantation silviculture.

The natural range of loblolly pine extends throughout the southeast US extending across14 states from New Jersey to Texas. The species grows successfully in the Piedmont, upper Coastal Plain and lower Coastal Plain, the three predominant physiographic regions of the South. Owing to its good growth across a wide range of sites, responsiveness to silvicultural practices, along with the suitability of wood for a variety of uses, loblolly pine is preferred over the other southern hard pines and has become the key commercial species in the region. Loblolly pine is a principal source of raw material for the pulp and paper industry in the US and its wood is also desirable for the production of lumber and composite wood products.

1.3 Wood quality and properties

Wood quality is a cumulative expression of anatomical, physical and mechanical properties of a piece of wood on a particular end product (Punches 2004). Thus the definition of wood quality is complex and multifaceted, depending either on the properties of the product or on the manufacturing process. For example, quality is assessed by strength and stiffness where the wood is intended for structural applications and is assessed by fiber length and proportion of cellulose and lignin where the wood is intended for pulp and paper production. Wood quality encompasses the ability of a product to satisfy the needs of the end user. Some of the most important predictors of wood quality are specific gravity (SG), microfibril angle (MFA), stiffness (modulus of elasticity, MOE) and strength (modulus of rupture, MOR).

According to Megraw (1985) "Of all the parameters practical to measure, SG is recognized as the most useful index to predict the physical behavior of wood". SG describes the amount of woody material in a given volume of wood. Theoretically, it is the ratio of the density of wood to the density of water at 4^{o}C (Megraw 1985). SG is considered as an important wood property because of its strong correlation with the strength of solid wood products, as well as the yield and quality of pulp produced (Panshin and de Zeeuw 1980). SG serves as a measure of wood quality in many wood related studies. An increase in SG of 0.02 units will result in a 22.7 kg increase in dry pulp per ton of round wood (Mitchell 1964) and increases of 31.15 and 3516 kg/cm^2 MOR and MOE respectively (Wahlgren and Schumann 1975). SG varies with age of the tree and growing conditions and can be altered by silvicultural treatments and by genetic manipulation. The wood produced within an annual ring can be divided into earlywood (produced during the early growing season, spring) and latewood (produced during the late growing season). The anatomical properties of wood such as the ratio of earlywood to latewood in the annual ring, fiber wall thickness, fiber length and fiber numerical density ultimately determine SG.

Microfibril angle (MFA) is defined as the angle made by cellulose microfibrils in the S_2 layer of the cell wall relative to the longitudinal axis of the cell (Megraw 1985). MFA has a strong influence on stiffness, strength and dimensional stability of wood and is an important determinant of the quality of sawn timber (MacDonald and Hubert 2002). MFA is inversely correlated with the SG, MOE, MOR and tangential shrinkage and positively correlated with the longitudinal shrinkage.

MOE and MOR are the two widely accepted wood property measures in the solid wood industry. MOE describes the stiffness of a material and is expressed as the ratio between stress and strain. MOR indicates the strength of a material defined as its load carrying capacity. Identifying and defining all these properties are essential for describing the quality of wood produced from a tree. Ultimately, these properties will determine how the wood is used.

1.4 Wood property variation with special emphasis on specific gravity

Wood is a ‘heterogeneous’ product and a versatile raw material suitable for a variety of uses. Large variations in wood properties have been reported in loblolly pine. Variation across growing regions, between stands within a region, trees within stands, within a tree and within annual rings has been identified and reported in loblolly pine. A comprehensive account of wood property variations and its potential causes in different hardwood and softwood tree can be found in Zobel and vanBuijtenen (1989). Marked variation in SG has been reported for loblolly pine across geographical regions by Zobel and Talbert (1984), Tassissa and Burkhart (1998b), Clark and Daniels (2002) and Jordan et al. (2008). According to them, SG was significantly higher in trees from the Coastal Plain compared to the trees from the inland regions. Jordan et al. (2008)

reported higher whole-core average SG (of 0.488) for trees from Atlantic Coastal Plain compared to other regions (Gulf Coastal, Hilly Coastal, Piedmont and Upper Coastal) which averaged 0.455. The reason for Coastal Plain trees having higher SG is their higher latewood percentage, which has been attributed to increased moisture availability during the summer months (Clark and Daniels 2002; Jordan et al. 2008).

Considerable wood property variation within trees has also been reported in loblolly pine. Based on the variation in different wood properties, the wood formed within a tree is divided into two zones: juvenile wood and mature wood. juvenile wood is the wood formed near the center of the tree and has low SG, and short tracheids with large MFA's (Larson et al. 2001). Larson et al. (2001) noted that juvenile wood is formed in the vicinity of live crown, so that there is a core of juvenile wood formed at the center of the tree from stump to tip. Zobel (1972) reported a SG range of 0.36 to 0.45 for juvenile wood and 0.42 to 0.64 for mature wood in loblolly pine. SG also exhibits considerable radial and longitudinal variation. According to Pearson and Gilmore (1971), both juvenile wood and mature wood SG of loblolly pine decreased from stump (juvenile wood SG = 0.474, mature wood SG = 0.525 at ~1 m) to tip (juvenile wood SG = 0.409, mature wood SG = 0.439 at ~13 m) of the tree. Phillips (2002) and He (2004) also reported a decrease in SG from stump-to- tip in loblolly pine. Radially, SG increases from pith-to-bark (Daniels et al. 2002; Megraw 1985; Tasissa and Burkhart 1998a).

Large variation in SG was observed within an annual ring produced within a tree, earlywood and latewood variation (Megraw 1985). Earlywood is characterized by tracheids with thin-walls and large lumen diameter. Latewood, on the other hand, comprised of tracheids with narrow lumen and thick walls. The SG of latewood is higher than the earlywood. In loblolly pine, it was observed that the earlywood SG decreases outwards from pith-to-bark and latewood

SG increases from pith-to-bark of a tree (McMillian 1968). The decrease in earlywood SG is due to an increase in radial diameter of tracheids from pith-to-bark while the wall thickness remains constant. On the other hand, the increase in latewood SG is because of the increase in wall thickness from pith-to-bark while the radial diameter remains constant.

1.5 Silvicultural effects on wood properties

Wood is a byproduct of series of biological processes (growth processes). Silvicultural operations which change any of these biological/growth processes impart changes in the wood properties and thus the quality of wood produced. Various silvicultural practices have been identified as having a positive influence on the growth and yield of southern pines. However improved growth owing to different silvicultural practices may result in an increase in juvenile wood formation and reduced product quality. A summary of the influences of different silvicultural practices on wood properties follows.

1.5.1 Planting density

Planting densities for southern pine plantations typically range from 741 to 2717 trees per ha (TPH) with an average of 1730 TPH. The decision regarding initial planting density depends on the management objective of the land owner. Manipulating stand density is an important silvicultural tool controlling seedling establishment, rate of growth and stem quality of the tree. Clark and Saucier (1989) found that initial spacing did not alter the age of transition from juvenile-to-mature wood in a study based on loblolly pine from the Piedmont of Georgia and slash pine from the Coastal Plain of Georgia, planted at initial spacing of 1.8 x 1.8, 2.4 x 2.4, 3 x 3 and 3.7 x 3.7 m. Wider spacing will enhance diameter growth (Sharma et al. 2002), crown width and increase the diameter of the juvenile core (Clark and Saucier 1989). Clark et al. (1994) studied the effect of initial spacing (1.8 x 1.8, 2.4 x 2.4, 3 x 3 and 3.7 x 3. 7 m) and thinning (to residual basal areas of 14, 18, 23, and 27.5 m^2/ha at age 18 and at 5 year age intervals to age 38) on strength and volume of lumber produced. Stands at 1.8 x 1.8 m spacing thinned to <23 m^2/ha produced more than 60 % No. 2 lumber compared to stands in 3.7 x 3.7 m spacing thinned to the same basal area. Stands spaced at 1.8 x 1.8 and 2.4 x 2.4 m and thinned to 14 m^2/ha produced lumber with less juvenile wood (32-34 %) compared to lumber produced from stands planted at wider spacing (3 x 3 and 3.7 x 3.7 m) and thinned to 23-27.5 m^2/ha (42-49 %). McAlister et al. (1997) examined the effect of initial spacing on strength and stiffness of lumber from 40 year old slash pine planted at 1.8 x 2.4, 2.4 x 2.4, 3 x 3 and 4.6 x 4.6 m. The modulus of rupture of the No.2 grade 0.6 x 1.2 m lumber produced from trees in 4.6 x 4.6 m spacing was significantly lower (25 % lower) than lumber from 1.8 x 2.4 m spacing, but no difference was found among 1.8 x 2.4, 2.4 x 2.4 and 3 x 3 m. Larson et al. (2001) proposed that planting at wide spacing such as 3.7 x 3. 7 m supports uninterrupted crown development and results in the production of a higher proportion of juvenile wood with low specific gravity, short tracheids and large MFA's, and in some cases an abnormal amount of compression wood and extractives.

1.5.2 Site preparation and competition control

Site preparation and competing vegetation control started with the objective of clearing and making the cutover sites to have conditions similar to old-fields. Efforts to recreate old-field conditions and to control competing vegetation led to the development of mechanical and chemical site preparation practices in southern plantation forestry (Fox et al. 2007b). In the lower Coastal Plain, bedding was used to alleviate problems associated with high water tables. Clark and Edwards (1999) studied the effect of six site preparation treatments (1. clear cut only, 2. chain saw removal of residuals, 3. shear and chop, 4. shear, chop and herbicide, 5. shear, root rake, burn and disk, and 6. shear, root rake, burn, disk, fertilize and herbicide) on growth and wood properties of loblolly pine growing in the Piedmont region of the southern US. A significant increase in basal area growth was observed with increased intensity of site preparation treatments, but no effect was observed on length of juvenility and average core specific gravity. However, the juvenile core diameter increased with increased intensity of treatments.

Clark et al. (2006) examined the effect of herbaceous and woody competition control (no weed control, woody control, herbaceous control and woody + herb control) on earlywood and latewood ring SG and latewood proportion on wood samples collected from 13 sites in Southeast US. They observed that the woody plus herbaceous weed control treatment increased growth in all locations and did not alter earlywood and latewood SG, or percent latewood. However, the woody plus herbaceous weed control treatment increased juvenile wood diameter by 19 % and thus resulted in a 10 % decrease in proportion of latewood and a 3 % reduction in specific gravity. Mora (2003) studied the effect of site preparation, early age fertilization and weed control on wood properties of loblolly pine and found a 29-33 % increase in volume from the intensive treatments, however wood properties were not significantly different from the control. The age of transition from juvenile wood to mature wood was not found to be affected by early application of silvicultural treatments, but the demarcation point of juvenile-to-mature wood was changed from site-to-site and from treatment-to-treatment (Mora et al. 2007).

1.5.3 Thinning

Thinning is a practice of removing selected trees from a stand to augment the growth of residual trees. It is an effective practice to earn revenue, generating intermediate cash flow from a plantation. Even though thinning cannot change the total yield significantly compared to an unthinned stand, it can change the product class distribution; with a thinned stand producing a higher proportion of more valuable peeler and saw logs compared to an unthinned stand (Amateis et al. 1996).

Inconsistent results have been reported on the effect of thinning on SG of loblolly pine (Megraw 1985). Burton and Shoulders (1974) reported that wood SG was unaffected following heavy thinning in loblolly pine at the age of 27, while an increase in SG was observed by Smith (1968) and Jackson (1968) following thinning in loblolly pine. Based on a loblolly pine thinning study established throughout southeastern US, Tasissa and Burkhart (1998b) found that both light and heavy thinning (where 30 and 50 % of the basal area removed respectively) did not produce a significant change in ring SG. According to Larson (1969), following thinning trees may behave like an open grown tree with reduced SG in the lower bole. However, reduced competition and increased soil moisture availability following thinning can increase latewood production in the summer and thus overall ring SG. Physiological age of the tree and stand attributes have a large influence on wood property changes following thinning.

1.5.4 Fertilization

The effect of fertilization on wood properties can be explained on the basis of the quantity and quality of wood produced. It is very difficult to generalize the influence of fertilization on wood properties because of the large number of extraneous factors involved in the response process. A mixed response in wood properties has been observed owing to fertilization, which can be related to variation with site characteristics, climatic conditions, age of the tree, initial SG and tracheid length at the time of fertilization.

One of the earliest fertilization trials (Posey 1964) in loblolly pine found a 16 % reduction in SG and a 12 % reduction in tracheid length in the annual ring immediately following treatment (nitrogen fertilizer applied as a single dose at ages 12 and 16 years). Posey (1964) also found that the wood properties of younger trees were more responsive to fertilization than that of older trees. Choong et al. (1970) reported no significant difference in SG following fertilization of loblolly pine at age 8. On the other hand, the quality of wood produced following midrotation fertilization was decreased significantly in loblolly pine which received higher rate of fertilizer (336 kg/ha or more). Based on a study conducted on samples collected from mid-rotation fertilization trials (fertilized with three levels of N: 112, 224 and 336 kg/ha and a control with 28 kg/ha of P), Antony et al. (2009) observed a significant reduction in four year average ring SG and latewood SG following the application of 336 kg/ha of N and 28 kg/ha of P. A study conducted on slash pine and loblolly pine reported similar pattern of response in SG following mid-rotation fertilization (Love-Myers et al. 2009). According to Larson et al. (2001), application of fertilizer decreases latewood percent and thus a reduced SG in young trees, whereas in older trees the response can mainly be attributed to reduced thickness of latewood cells, resulted in lower latewood SG.

Fertilizer application in combination with other silvicultural practices has been found to have a strong influence on the properties of the wood produced. A study (Clark et al. 2004) revealed that an annual application of nitrogen fertilizer through age 12 along with vegetation control led to a 62 % increase in the diameter of the juvenile core, a 6 to 10 % decrease in weighted stem SG, and a 30 to 33 % reduction in toughness compared to untreated trees. In addition, a significant drop in juvenile wood strength (9 to 10 %) and mature wood strength (4 to 7 %) was observed in trees receiving annual nitrogen fertilization in combination with vegetation control compared to untreated trees.

1.6 Modeling wood properties

Predicting variation in the SG of loblolly pine is important for making improved management decisions. Models are needed to explain the variation in properties (for example SG variation) within a tree, from stand-to-stand and from region-to-region and to explain wood property responses following intensive management practices. Predictive models help the forester to make the best use of available resources and to plan for efficient product segregation and utilization.

Models are available to explain regional and within tree variation in SG. Clark and Daniels (2002) developed a model which can predict the SG of loblolly pine growing in different physiographic regions of southeastern US. They developed a linear model for average stand weighted cross-section SG as a function of latitude, longitude, site index, age and logarithm of stand age. Phillips (2002) proposed nonlinear equations to explain the within-tree variation in disk SG and moisture content of loblolly pine growing in different parts of Georgia. Phillips (2002) developed separate models for disk SG which are a function of relative height and diameter outside bark respectively. She also incorporated physiographic differences and site differences in disk SG by expressing the parameters in the base models as appropriate functions of region and site. The models were derived using the algebraic difference approach (ADA), a special case of GADA, proposed by Bailey and Clutter (1974), and the models were fitted using the stochastic parameter estimation technique (Cieszewski et al. 2000) to account for the measurement error in predictor variables.

He (2004) proposed a nonlinear mixed model to explain the within tree variation in disk SG of loblolly pine growing in Georgia as a function of relative height. The model incorporated tree-to-tree variability and within tree variability by appropriately selecting model parameters as random and fixed effect. The proposed model was expanded to take account of the variation in disk SG among physiographic regions and site classes by suitable covariates to the final model in addition to relative height.

Several attempts have been made to model the ring-by-ring variation in SG as a function of physiological age, stand characteristics and latewood percentage etc. Tassissa and Burkhart (1998b) used a linear model to predict ring SG using physiological age, relative height, percent latewood within each ring, latewood width, ring width and competition index as explanatory variables. Daniels et al. (2002) proposed a three parameter logistic function to describe latewood SG changes in loblolly pine from pith to bark. They incorporated the variation in latewood SG with height by expressing the parameters in the logistic model as a function of height, and thus developed a three dimensional model capable of explaining within tree variation, both radially and longitudinally, in latewood SG. He (2004) used a three parameter logistic function to model ring SG of planted loblolly pine in southeastern US. The proposed final model was a three level nonlinear mixed model with ring number, stem taper and relative height as explanatory variables. Regional differences and stand attributes were also integrated into the final model. He (2004) also proposed a linear mixed model to explain percent latewood within rings as a function of ring number, stem taper and relative height. Jordan et al. (2008) modeled the relationship between SG and time using semiparametric regression penalized smoothing splines. Smoothing splines are curves that are formed by joining together several low order polynomials at specified locations known as knots. They used a penalized spline with a quadratic basis for examining regional differences in ring SG.

Models to estimate the transition age from juvenile-to-mature wood are also of interest in the proper utilization of wood. Visual inspection of ring-by-ring SG profiles to demarcate the age of transition from juvenile-to-mature wood was proposed by Clark and Saucier (1989). Tassissa and Burkhart (1998a) developed a linear segmented model to determine the transition age from juvenile-to-mature wood in loblolly pine. While, Mora et al. (2007) used a ring SG model to understand the corewood-outerwood transition within loblolly pine. A modified four-parameter logistic function was developed to explain changes in ring SG from pith to bark. The nonlinear mixed model approach was adapted to incorporate site and silvicultural treatment differences and to accommodate the correlation across observations and heteroscedasticity of residuals to the final model.

Based on the above review, there is a need to develop and validate models to explain the natural variation in different wood properties and responses of wood properties following different silvicultural practices. Different modeling methods adapted in the past need to be compared with respect to wood property prediction. This thesis is an effort to address these issues in more detail. Two broad objectives were:

1. to model the within tree variation (longitudinal variation) in SG and,
2. to model the response of SG following midrotation fertilization.

In addition to the above two broad objectives, this dissertation also addressed several other issues related to wood quality. A simultaneous model system for disk SG and moisture content (MC) was proposed. In addition, an empirical comparison of two distinct modeling methodologies used in forestry (Generalized Algebraic Difference Approach and Nonlinear Mixed Models) was conducted using the longitudinal disk SG data.

1.7 References

Allen, H.L., T.R. Fox, and R.G. Campbell. 2005. What is ahead for intensive pine plantation silviculture in the South? Southern Journal of Applied Forestry 29(2):62-69.

Amateis, R.L. 2000. Modeling response to thinning in loblolly pine plantations. Southern Journal of Applied Forestry 24(1):17-22.

Amateis, R.L., P.J. Radtke, and H.E. Burkhart. 1996. Growth and yield of thinned and unthinned plantations. Journal of Forestry 94(12):19-23.

Antony, F., L. Jordan, R.F. Daniels, L.R. Schimleck, A. Clark, and D.B. Hall. 2009. Effect of midrotation fertilization on growth and specific gravity of loblolly pine. Canadian Journal of Forest Research 39(5):928-935.

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

Borders, B.E., and R.L. Bailey. 2001. Loblolly pine-pushing the limits of growth. Southern Journal of Applied Forestry 25(2):69-74.

Burns, R.M., and B.H. Honkala. 1990. Silvics of North America: 1. Conifers; 2. Hardwoods. USDA Forest Service, Agriculture Handbook 654, 877p. Accessible from http://www.na.fs.fed.us/spfo/pubs/silvics_manual/table_of_contents.htm. .

Burton, J.D., and E. Shoulders. 1974. Fast-grown, dense loblolly pine saw logs: A reality. Southern Journal of Applied Forestry 72(10):637-641.

Choong, E.T., B.H. Box, and P.J. Fogg. 1970. Effects of intensive cultural management on growth and certain wood properties of young loblolly pine. Wood and Fiber Science 2(2):105 - 112.

Cieszewski, C.J., W.M. Harrison, and S.W. Martin. 2000. Practical methods for estimating non-biased parameters in self- referencing growth and yield models. Plantation Management Research Cooperative, Athens, GA, Technical Report 2000-7. 11.

Clark, A., B.E. Borders, and R.F. Daniels. 2004. Impact of vegetation control and annual fertilization on properties of loblolly pine wood at age 12. Forest Products Journal 54(12):90-96.

Clark, A., and R.F. Daniels. 2002. Modeling the effect of physiographic region on wood properties of planted loblolly pine in Southeastern United states. P. 54-60 *in* Connection between Forest Resources and Wood Qulaity: Modeling Apporaches and Simulation Software. Fourth Workshop IUFRO Working Party S5.01-04. INRA - Centre de Researches de Nancy, France, Harrison Hot Springs, B.C.

Clark, A., R.F. Daniels, and J.H. Miller. 2006. Effect of controlling herbaceous and woody competing vegetation on wood quality of planted loblolly pine. Forest Products Journal 56(2):40-46.

Clark , A., and M.B. Edwards. 1999. Effect of six site preparation treatments on piedmont loblolly pine wood properties at age 15. P. 316-320 *in* 10th Binniel Southern Silvicutural Research Conference. USDA Forest Service, Shreveport, LA.

Clark, A., and J.R. Saucier. 1989. Influence of initial planting density, geographic location, and species on juvenile wood formation in southern pine. Forest Products Journal 39(7-8):42-48.

Clark, A., J.R. Saucier, V.C. Baldwin, and R.D. Bower. 1994. Effect of initial spacing and thinning on lumber grade, yield and strength of loblolly pine. Forest Products Journal 44(11):14-20.

Daniels, R.F., R. He, A. Clark, and R.A. Souther. 2002. Modeling wood properties of planted Loblolly pine from pith to bark and stump to tip. In: Proceedings of the Fourth Workshop-Connection between Forest Resources and Wood Quality: Modelling Approaches and Simulation Software, September 8-15, British Columbia, Canada, p. 1-4.

Dickens, D.E., D.J. Moorehead, and B. McElvany. 2003. Pine plantation fertilization. Better Crops 87(1):12-15.

Fox, T.R., H.L. Allen, T.J. Albaugh, R. Rubilar, and C.A. Carlson. 2007a. Tree nutrition and forest fertilization of pine plantations in the southern United States. Southern Journal of Applied Forestry 31(1):5-11.

Fox, T.R., E.J. Jokela, and H.L. Allen. 2007b. The development of pine plantation silviculture in the southern United States. Journal of Forestry 105(7):337-347.

Haywood, J.D. 2005. Influence of precommercial thinning and fertilization on total stem volume and lower stem form of loblolly pine. Southern Journal of Applied Forestry 29(4):215-220.

He, R. 2004. Mixed effects modeling of wood properties of loblolly pine in the southeastern United States (Ph.D dissertation), University of Georgia, Athens. 213 p.

Jackson, L.W.R. 1968. Effect of thinning on growth and specific gravity of loblolly and slash pine. Gerogia Forestry Research Council Research Paper No. 50. 6p.

Jordan, L., A. Clark, L.R. Schimleck, D.B. Hall, and R.F. Daniels. 2008. Regional variation in wood specific gravity of planted loblolly pine in the United States. Canadian Journal of Forest Research 38:698-710.

Larson, P.R. 1969. Wood formation and the concept of wood quality. Bulletin No. 74, Yale University, School of Forestry, New Haven, CT.

Larson, P.R., D.E. Kretschmann, A. Clark, and J.G. Isebrands. 2001. Formation and Properties of Juvenile Wood in Southern Pines A Synopsis, USDA Forest Service, General Technical Report FPL-GTR-129 42.

Love-Myers, K.R., A. Clark Iii, L.R. Schimleck, E.J. Jokela, and R.F. Daniels. 2009. Specific gravity responses of slash and loblolly pine following mid-rotation fertilization. Forest Ecology and Management 257(12):2342-2349.

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

McAlister, R.H., A. Clark, and J.R. Saucier. 1997. Effect of initial spacing on mechanical properties of lumber sawn from unthinned slash pineat age 40. Forest Products Journal 47(7/8):107-109.

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

Megraw, R.A. 1985. Wood quality factors in loblolly pine. TAPPI Press, Atlanta, GA.

Miller, J.H., B.R. Zutter, S.M. Zedakker, M.B. Edward, J.D. Haywood, and R.A. Newbold. 1991. A regional study on the influence of woody and herbaceous competition on early loblolly pine growth. Southern Journal of Applied Forestry 15(4):169-179.

Mitchell, H. 1964. Patterns of variation in specific gravity of southern pines and other coniferous species. TAPPI Journal 47:276 - 283.

Mora, C.R. 2003. Effects of early intensive silviculture on wood properties of loblolly pine (MS Thesis), North Carolina State University, Raleigh,North Carolina. 92 p.

Mora, C.R., H.L. Allen, R.F. Daniels, and A. Clark. 2007. Modeling corewood-outerwood transition in loblolly pine using wood specific gravity. Canadian Journal of Forest Research 37(6):999-1011.

Panshin, A.J., and C. de Zeeuw. 1980. Textbook of Wood Technology. McGraw-Hill Book Co., New York. 705 p.

Pearson, R.G., and R.C. Gilmore. 1971. Characterization of the strength of juvenile wood of loblolly pine (*Pinus taeda* L.). Forest Products Journal 21(1):23-31.

Phillips, K.M. 2002. Modeling within tree changes in wood specific gravity and moisture content for loblolly pine in Georgia (M.S. Thesis), University of Georgia, Athens. 95 p.

Pienaar, L.V., and B.D. Shiver. 1993. Early results from an old-field loblolly pine spacing study in the Georgia Piedmont with competition control. Southern Journal of Applied Forestry 17(4):193-196.

Posey, C.E. 1964. The effect of fertilization upon wood properties of loblolly pine. Forest Tree Improvement Program, School of Forestry, North Carolina State University, Technical Report No. 22. 61.

Prestemon, J.P., and R.C. Abt (eds.). 2002. Timber products supply and demand. In *Southern Forest Resource Assessment*, *eds.* D.N. Wear and J.G. Greis 299-325. USDA Forest Service General Technical Report SRS-53. Asheville, NC, Southern Research Station.

Punches, J. 2004. Tree growth, forest management and their implication for wood quality. Pacific Northwest Extension Publication PNW 576. 8p.

Sharma, M., H.E. Burkhart, and R.L. Amateis. 2002. Modeling the effect of density on the growth of loblolly pine trees. Southern Journal of Applied Forestry 26:124-133.

Smith, D.M. 1968. Wood quality of loblolly pine after thinning. USDA Forest Service Research Paper, Forest Product Laborotory, Madison, Wisconsin, Research Paper FPL-89. 14p.

Stanturf, J.A., R.C. Kellison, F.S. Broerman, and S.B. Jones. 2003. Productivity of southern pine plantations - Where are we and how did we get here? Journal of Forestry 101(3):26-31.

Tasissa, G., and H.E. Burkhart. 1998a. Juvenile-mature wood demarcation in loblolly pine trees. Wood and Fiber Science 30(2):119-127.

Tasissa, G., and H.E. Burkhart. 1998b. Modeling thinning effects on ring specific gravity of loblolly pine (Pinus taeda L.). Forest Science 44(2):212-223.

Thomas, C.E., B.R. Parresol, K.H.N. Lê, and R.E. Lohrey. 1995. Biomass and taper for trees in thinned and unthineed longleaf pine plantations. Southern Journal of Applied Forestry 19(1):29-34.

Wahlgren, H.D., and D. Schumann. 1975. Properties of major southren pines: Part I - Wood desnsity survey. USDA Forest Service, Forest Product Laboratory, Madison, Wisconsin, Research Papper FPL-176:57p.

Wear, D.N., and J.G. Greis. 2002. The Southern Forest Resource Assessment: Summary Report. USDA Forest Service, Southern Research Station Report No. 53. 635p

Zobel, B. 1972. Three Rings-Per-Inch, Dense Southern Pine-Can It Be Developed? Journal of Forestry 70:333-336.

Zobel, B.J., and J. Tabert. 1984. Applied forest tree improvement. Wiley, New York. 511 p.

Zobel, B.J., and J.P. vanBuijtenen. 1989. Wood Variation: Its Causes and Control. Springer-Verlag, Berlin. 363 p.

Chapter 2

Modeling the longitudinal variation in wood specific gravity of planted loblolly pine in the United States

2.1 Abstract

Loblolly pine (*Pinus taeda* L.) is the major plantation species grown in the southern United States, producing wood having a multitude of uses including pulp and lumber production. Specific gravity (SG) of the wood is an important property used to measure the quality of wood produced, and it varies regionally and within the tree with height and radius. Disk SG at different height levels was measured from 407 trees representing 135 plantations across loblolly pine's natural range. A three segmented quadratic model and a semiparametric model were proposed to explain the vertical and regional variation in disk SG. Both models were in agreement that a stem can be divided in to three segments based on the vertical variation in disk SG. Based on the fitted models, the mean trend in disk SG of trees from the southern Atlantic and Gulf Coastal Plain was observed to be higher than other physiographical regions (Upper Coastal, Hilly Coastal, northern Atlantic Coastal Plain and Piedmont). The lowest disk SG was observed for trees from the northern Atlantic Coastal Plain. Maps showing the regional variation in disk SG at a specified height were also developed. Maps indicated that the stands in the southern Atlantic and Gulf Coastal Plains have the highest SG at a given height level.

Key Words: Longitudinal variation, specific gravity, juvenile wood

2.2 Introduction

Pine plantations occupy approximately 13 million ha of the southern United States (US), which carries 680 million cubic m of timber, with a projected increase in area of 67 % (22 million ha) by 2040 (Wear and Greis 2002). A 2-fold increase in productivity and 50 % reduction in rotation length of pine plantations during the last few decades turned the South into the wood basket of the US (Fox et al. 2007). Currently, the southern US produces around 58 % of the total wood supply in the United States and 16 % of world's industrial wood supply (Wear and Greis 2002). Loblolly pine (*Pinus taeda* L.) is the most important plantation species in the southern US with more than half of the standing pine volume. Wood from loblolly pine is a principal source of raw material for the pulp and paper industry and is desirable for the production of lumber and composite wood products. The quality of wood produced from a loblolly pine tree is defined by its physical and mechanical properties. Of these, specific gravity (SG) is considered as the most important wood quality indicator because of its strong correlation with the strength of solid wood products, as well as the yield and quality of pulp (Panshin and deZeeuw 1980).

The wood properties of loblolly pine vary considerably across its growing range, between stands within a region, between trees within stands and within the trees. Significant variation in wood properties within trees occurs from pith-to-bark; from stump-to-tip and also within annual rings between earlywood and latewood. Clark and Saucier (1989) divided the radial section of a pine stem in to three zones: core wood and transition wood, which together can be referred to as juvenile wood, and the mature wood. Juvenile wood is the wood formed in the vicinity of the crown forming a core near the center of the stem having low SG, and short tracheids with large microfibril angles (Larson et al. 2001). Zobel (1972) reported an average SG range of 0.36 to 0.45 for juvenile wood and 0.42 to 0.64 for mature wood in loblolly pine.

According to Burden et al. (2004), the concept of juvenile to mature wood progression from pith to bark is inadequate to represent the pattern of variation within a tree and is an oversimplification of the physiological process. They advocated the use of two separate concepts; corewood versus outerwood in the radial direction and juvenile versus mature wood in the longitudinal direction, to explain the within tree variation in wood properties. Based on the proposed classification: juvenile wood occurs in the lower butt log with height < 3 m; transition wood occurs between 3- 5 m in height; and mature wood occurs at heights >5 m.

The longitudinal variation in SG of loblolly pine was reported in several studies. Early studies reported a decrease in SG from stump-to-tip of the loblolly pine tree (Megraw 1985; Zobel and vanBuijtenen 1989). Tasissa and Burkhart (1998b) modeled the within tree variation (stump-to-tip and pith-to-bark) in SG of loblolly pine using a linear function of physiological age, relative height, percent latewood, latewood width and ring width. Phillips (2002) and He (2004) modeled the longitudinal variation in disk SG of loblolly pine using subject specific nonlinear models.

Marked geographical variation in SG has been reported for loblolly pine by Tasissa and Burkhart (1998a), Clark and Daniels (2002) and Jordan et al. (2008). SG was significantly higher in trees from the Coastal Plain compared to the trees from inland areas. The reason for higher SG trees from the Coastal Plain's might be due to the increased latewood production of these trees, which has been attributed to increased moisture availability from frequent summer rainfall in the area (Clark and Daniels 2002; Jordan et al. 2008). Jordan et al. (2008) reported a higher whole-core average SG (of 0.49) for trees from the southern Atlantic Coastal Plain compared to other regions (Gulf Coastal, Hilly Coastal, northern Atlantic, Piedmont and Upper Coastal Plain) which averaged 0.455 using breast height cores collected from trees. They also produced maps showing regional variation in SG at different stand age at breast height. However, the maps showing the regional variation in SG at different height levels within a tree was lacking. It is important to have maps showing the SG variation at different height levels for maximizing product categorization and utilization. The objectives of the present study were to: (1) examine and model the longitudinal variation in disk SG; (2) examine regional variation of disk SG; (3) develop maps depicting the regional variation of disk SG across the southern US.

2.3 Data and Methods

2.3.1 Data

The Wood Quality Consortium at the University of Georgia and the United States Department of Agriculture (USDA) Forest Service Southern Research Station sampled planted loblolly pine trees across their natural range to study the vertical variation in wood SG. Trees were sampled from 135 stands from six physiographic regions across the southeastern US. Regions sampled included: 1- southern Atlantic Coastal Plain, 2- northern Atlantic Coastal Plain, 3- Upper Coastal Plain, 4- Piedmont, 5- Gulf Coastal Plain and 6- Hilly Coastal region. A minimum of 12 plantations from each of the six physiographic regions were sampled. The stands selected for sampling included 20- to 25-year-old loblolly pine plantations planted at 1250 or more trees per hectare and contained 625 trees per hectare or more after thinning. Only stands that were conventionally managed with no fertilization (except phosphorus at planting on phosphorus deficient sites) and no competition control were sampled. Three trees from each stand were felled and cross sectional disks of 3.8 cm thickness were collected from 0.15, 1.37 m and then 1.52 m intervals along the stem up to a diameter of 50 mm outside bark. The disks were sealed in plastic bags and shipped to the USDA Forest Service

Figure 2.1: Plot showing locations of 135 sampled stands ($*$) and a subset of 34 ($\bigcirc$) stands used for reduced knot kriging.

laboratory for physical property analysis. Disk SG based on green volume and oven-dry weight were measured for each disk collected at different heights. A map showing the sampled locations is presented in Figure 2.1. A summary of the stand characteristics along with the number of stands and trees sampled from each region are presented in Table 2.1.

Table 2.1: Mean stand attributes collected from six regions, standard deviation in parenthesis.

Region	# of stands	# of trees	Age	DBH (cm)	Total Ht (m)	Disk SG
southern Atlantic	39	117	22.73 (1.82)	24.07 (4.58)	20.86 (2.50)	0.45 (0.06)
northern Atlantic	7	20	22.46 (1.61)	24.56 (3.74)	18.89 (2.48)	0.41 (0.05)
Upper Coastal	17	51	23.00 (1.46)	24.07 (4.87)	19.39 (3.08)	0.43 (0.05)
Piedmont	26	78	23.08 (2.01)	23.90 (4.54)	18.19 (2.11)	0.42 (0.05)
Gulf Coastal	17	54	23.22 (3.26)	21.16 (3.79)	19.54 (2.58)	0.46 (0.05)
Hilly Coastal	29	87	23.86 (3.58)	23.39 (4.12)	19.59 (2.75)	0.43 (0.05)

2.3.2 Parametric Model

Disk SG follows a nonlinear decreasing trend from stump-to-tip in loblolly pine. Relative height, the ratio of height at any point to the total height of tree, has explained the maximum amount of variation in SG and possesses the property of homogeneous error variance. Relative height was used as a potential variable to explain the change in disk SG from stump to tip in this study. A plot of observed disk SG with relative height for the six physiographic regions in southern US is presented in Figure 2.2. It was observed from Figure 2.2 that disk SG decreases rapidly at the base of the tree (i.e. from the base up to a relative height of ~0.1), decreases at a decreasing rate from ~0.1 to ~0.3 and then decreases at a constant rate above a relative height ~0.3. Large tree-to-tree variation in disk SG profiles was also evident from the plots. Since the rate of change

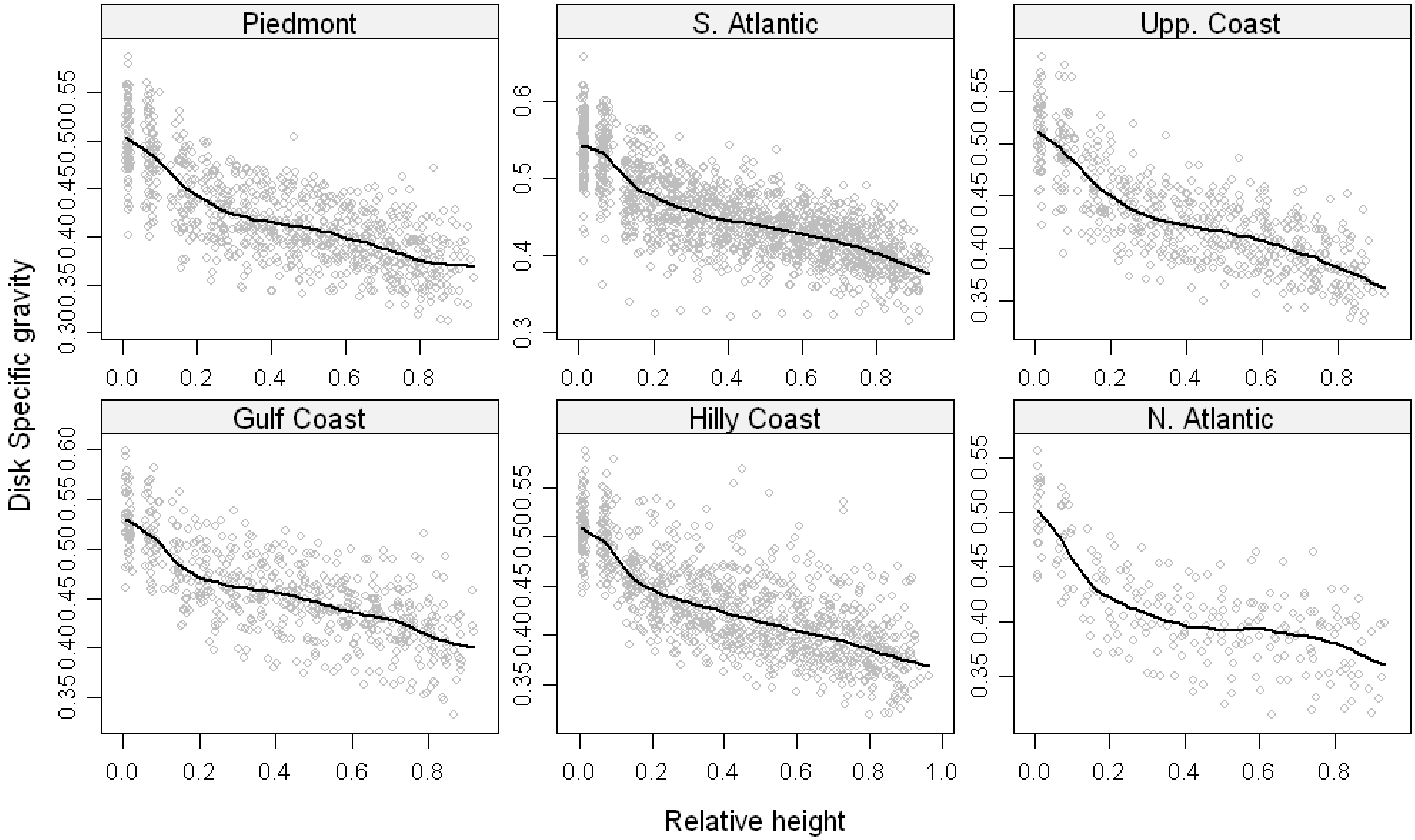

Figure 2.2: Plot of observed disk SG (grey dots) with relative height collected from six physiographic regions with smoothed line from scatter plot smoother in it (solid line).

of SG varies at different parts of the stem (at least two inflection points are present in most of the individual tree profiles), it was difficult to explain the phenomenon using a single function.

A segmented regression model proposed by Gallant and Fuller (1973) was used to explain the change in disk SG with relative height in this study. The general form of the segmented regression model by Gallant and Fuller (1973) can be represented as:

[1] $$y_i = g(x_i) + e_i$$

where, $g(x_i)$ is a sequence of grafted submodels,

$$\begin{aligned} g(x) &= g_1(x,\ \boldsymbol{\beta}_1), \quad 0 \le\ x\ \le \alpha_1 \\ &= g_2(x,\ \boldsymbol{\beta}_2), \quad \alpha_1 <\ x\ \le \alpha_2 \\ &\ \ \vdots \\ &= g_r(x,\ \boldsymbol{\beta}_r), \quad \alpha_{r-1} <\ x\ \le 1 \end{aligned}$$

and $e_i \sim N(0, \sigma^2)$. Each of these submodels is subjected to continuity and smoothness constraints as:

[2] $$\begin{aligned} g_j(\alpha_j, \boldsymbol{\beta}_j) &= g_{j+1}(\alpha_j, \boldsymbol{\beta}_{j+1}), \qquad j = 1, 2, \ldots, k-1 \\ \frac{\partial}{\partial x} g_j(\alpha_j, \boldsymbol{\beta}_j) &= \frac{\partial}{\partial x} g_{j+1}(\alpha_j, \boldsymbol{\beta}_{j+1}), \qquad j = 1, 2, \ldots, k-1 \end{aligned}$$

Following Gallant and Fuller (1973), a segmented model formed after splicing three quadratic submodels was used to explain changes in disk SG with relative height and had the form:

[3] $$g(x, \beta) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 (\alpha_1 - x)_+^2 + \beta_4 (\alpha_2 - x)_+^2$$

where *x* is the relative height *h/H*, *h* is the height above ground, *H* is the total height of the tree, $[\beta_0 \quad \beta_1 \quad \beta_2 \quad \beta_3 \quad \beta_4 \quad \alpha_1 \quad \alpha_2]^T$ are parameters to be estimated, with $[1 > \alpha_1 > \alpha_2 > 0]$.

The $(\alpha_j - x)_+^2$ terms indicates the positive part of the function $\alpha_j - x$ where "+" sets it to zero for those values of x where $\alpha_j - x$ is negative (here $x > \alpha_j$). The above model is equivalent to the standard form of the taper model proposed by Max and Burkhart (1976), which is not constrained to have a value of zero at the tip of the tree. If the knot points $\begin{bmatrix}\alpha_1 & \alpha_2\end{bmatrix}^T$ are known, then the model becomes a simple linear model and the estimates of $\begin{bmatrix}\beta_0 & \beta_1 & \beta_2 & \beta_3 & \beta_4\end{bmatrix}^T$ can obtained through an ordinary least squares solution. However, if the knot points are unknown a solution for the parameters can be estimated using a nonlinear least square procedure. In this study, we are proceeding under the assumption that the knot points are unknown and need to be estimated from the data.

Since the data follows a hierarchical structure by design (stands and trees within stands), a nonlinear mixed model was used to account for the heterogeneity between stands and trees within stands. Let y_{ijk} represent the k^{th} disk SG measurement from the j^{th} tree in the i^{th} stand; the nonlinear mixed model can be represented as:

[4] $$y_{ijk} = \beta_{0ij} + \beta_{1ij}x_{ijk} + \beta_{2ij}x_{ijk}^2 + \beta_{3ij}\left(\alpha_{1ij} - x_{ijk}\right)_+^2 + \beta_{4ij}\left(\alpha_{2ij} - x_{ijk}\right)_+^2 + \varepsilon_{ijk}$$

Following Vonesh and Chinchilli (1997), the mixed effect parameter β_{ij} in the model can be represented as:

[5] $$\beta_{ij} = A_{ij}\beta + B_{ij,1}b_i + B_{ij,2}b_{ij}$$

where

$A_{ij} = B_{ij,1} = B_{ij,2} = I_7$

$\beta = \begin{bmatrix}\beta_0 & \beta_1 & \beta_2 & \beta_3 & \beta_4 & \alpha_1 & \alpha_2\end{bmatrix}^T$

$b_i = \begin{bmatrix}b_i^{(0)} & b_i^{(1)} & b_i^{(2)} & b_i^{(3)} & b_i^{(4)} & b_i^{(5)} & b_i^{(6)}\end{bmatrix}^T$

$b_{ij} = \begin{bmatrix}b_{ij}^{(0)} & b_{ij}^{(1)} & b_{ij}^{(2)} & b_{ij}^{(3)} & b_{ij}^{(4)} & b_{ij}^{(5)} & b_{ij}^{(6)}\end{bmatrix}^T$

where b_i and b_{ij} are the stand and tree level random effects; $B_{ij,1}$ and $B_{ij,2}$ are the associated random effect design matrices; and A_{ij} and β are the fixed effect design matrix and parameter vector respectively. I_7 is a 7 x 7 identity matrix with all the diagonal elements equal to 1.

The random effects and the within-tree error term were assumed to be distributed normally as $b_i \sim N\left(0, \Psi_1\right)$; $b_{ij} \sim N\left(0, \Psi_2\right)$ and $\varepsilon_{ijk} \sim N\left(0, \sigma^2 \Lambda_{ijk}\right)$ and are independent of each other. Here Ψ_1 and Ψ_2 are variance-covariance matrices representing different levels of stand and tree random effects. A full model with random effects associated with all the parameters in the model is considered first by assuming a diagonal variance-covariance matrix structure for random effects and an independent structure for within tree error. These assumptions were relaxed in the later stages of fitting by assuming different variance-covariance structures for the random effects. Several reduced models were also fitted by dropping the random effect terms associated with the parameters. The best of these models was selected by comparing the fitted models using Akaike's Information Criteria (AIC).

The next step in the model building process was to incorporate any covariates, here the region effect, into appropriate parameters in the model. As we had six distinct physiographical regions in the study, we assumed different fixed effect parameters for each region with the southern Atlantic Coastal Plain as the reference region with all other regions having their own parameters which are deviations from the reference (effect version of parameters). After assuming all the parameters in the model as region specific, the fixed effect design matrix and parameter vector for l^{th} parameter in $\boldsymbol{\beta}$ can be represented as:

$$A_l = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\beta_l = [\beta_{l1} \quad \beta_{l2} \quad \beta_{l3} \quad \beta_{l4} \quad \beta_{l5} \quad \beta_{l6}]$$

After formulating the appropriate mean model and the random effect structure, the independent matrix structure associated with the within-tree error was relaxed. This was done to enable us to explain the heteroscedasticity in the data and serial correlation across measurements successfully. Different variance functions usually used in growth modeling such as the power model $\left(Var(\varepsilon_{ijkl}) = \sigma^2 |\upsilon_{ijkl}|^{2\delta}\right)$, the exponential model $\left(Var(\varepsilon_{ijkl}) = \sigma^2 e^{2\delta\upsilon_{ijkl}}\right)$ and the constant power model $\left(Var(\varepsilon_{ijkl}) = \sigma^2 \left(\delta_1 + |\upsilon_{ijkl}|^{\delta_2}\right)^2\right)$ were used to define any non-constant variance within the data. The autoregressive models (AR(p)), moving-average models (MA(q)) and autoregressive with moving average models (ARMA(p,q)) were used with the data to account for dependence across repeated measurements within each tree. AIC criterion was used for checking significant changes in performance of the models. The nonlinear mixed models were implemented using the nlme package available in R (Pinheiro et al. 2009).

2.3.3 Semiparametric model

A more flexible approach to explain the nonlinear trend in disk SG with relative height is by semiparametric regression. Semiparametric regression can model nonlinear relationships, here the change in disk SG with relative height, without having any parametric restriction. The advantage is that these models can be formulated in a linear mixed model frame work (Ngo and

Wand 2004), allowing the use of estimation and inferential tools available in mixed model methodology.

Let y_{ij} represents the disk SG observed at the j^{th} disk of i^{th} tree. A simple model form to explain disk SG with relative height is:

[6] $$y_{ij} = f\left(x_{ij}\right) + \varepsilon_{ij};\ \text{where } \varepsilon_{ij} \sim N\left(0,\ \sigma_{\varepsilon}^{2}\right)$$

where f is a smooth function describing the trend in disk SG with relative height. We utilized penalized smoothing splines, curves that are formed by splicing low-order polynomials at known knot locations, to model the change in disk SG with relative height. A truncated quadratic basis was used to model the function $f\left(x_{ij}\right)$. The model [6] can be represented as:

[7] $$y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 x_{ij}^2 + \sum_{\kappa=1}^{K} u_k \left(x_{ij} - \kappa_\kappa\right)_+^2 + \varepsilon_{ij}$$

where $u_k \sim N\left(0,\ \sigma_u^2\right)$. Here, $\kappa_1, \ldots\ldots, \kappa_\kappa$ are distinct knot locations within the range of x_{ij}'s and $\left(x - \kappa_\kappa\right)_+$ is the positive function where "+" sets it to zero for those values of x_{ij} where $x_{ij} - \kappa_\kappa$ is negative (here $x_{ij} < \kappa_\kappa$). According to Ruppert et al. (2003), a reasonable choice for selecting knots is that there should be 4-5 unique data points between two knots, with 35 knots as the maximum number of allowable knots. They proposed a simple method for knot selection such that knot $\kappa_k = \left(\frac{k+1}{K+2}\right)$ th sample location of the unique x_{ij}'s, $k = 1, \ldots\ldots, K$, where $K = \max\left(5, \min\left(\frac{1}{4} \times \text{number of unique } x_i\text{'s, } 35\right)\right)$. Use of the default knot selection procedure resulted in selecting 35 knots in this study, the maximum allowable knots based on the above procedure. Evenly spaced knots were also recommended and practiced in fitting the semiparametric regression (Jordan et al. 2008; Ruppert et al. 2003). Here, we used 8 evenly spaced knots at an interval of 0.1 between the minimum and maximum of relative height from the available data.

An estimate of $[\boldsymbol{\beta},\ \mathbf{u}]$ can be obtained by formulating the model [7] as a linear mixed model as follows:

[8] $$\mathrm{Y} = \mathrm{X}\boldsymbol{\beta} + \mathrm{Z}\mathbf{u} + \varepsilon$$

Here,

$$\mathbf{Y} = \begin{bmatrix} y_{11} \\ . \\ . \\ y_{1m_1} \\ . \\ y_{nm_n} \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{11}^2 \\ . & . & . \\ . & . & . \\ 1 & x_{1m_1} & x_{1m_1}^2 \\ . & . & . \\ 1 & x_{nm_n} & x_{nm_n}^2 \end{bmatrix}$$

$$\mathbf{Z} = \begin{bmatrix} (x_{11}-\kappa_1)_+^2 & . & . & . & . & (x_{11}-\kappa_\kappa)_+^2 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ (x_{1m_n}-\kappa_1)_+^2 & . & . & . & . & (x_{1m_n}-\kappa_\kappa)_+^2 \\ . & . & . & . & . & . \\ (x_{nm_n}-\kappa_1)_+^2 & . & . & . & . & (x_{nm_n}-\kappa_\kappa)_+^2 \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} u_1 & . & . & u_\kappa \end{bmatrix}$$

The maximum likelihood estimate (MLE) of $\hat{\boldsymbol{\beta}}$ and an empirical best linear unbiased predictor for $\hat{\mathbf{u}}$ can be obtained by fitting the above model form in any standard mixed model software (e.g. lme in S-plus and R, PROC MIXED in SAS). The smoothness of the curve is controlled by the parameter $\lambda = \sigma_\varepsilon^2 / \sigma_u^2$, which is calculated automatically using the restricted MLE's of σ_u^2 and σ_ε^2.

One of the major objectives of this study was to understand the regional variation in the mean trend of disk SG with relative height. The addition of the interaction term in model [7] was used to examine regional differences. Model [7] with an interaction term can be represented as:

[9]
$$\begin{aligned} y_{ij} = \; & \beta_0 + \beta_1 x_{ij} + \beta_2 x_{ij}^2 + \sum_{\kappa=1}^{K} u_k \left(x_{ij} - \kappa_\kappa \right)_+^2 \\ & + \sum_{\ell=2}^{L} z_{i\ell} \left(\gamma_{0\ell} + \gamma_{1\ell} x_{ij} + \gamma_{2\ell} x_{ij}^2 \right) \\ & + \sum_{\ell=2}^{L} z_{i\ell} \left(\sum_{\kappa=1}^{K} \upsilon_\kappa^\ell \left(x_{ij} - \kappa_\kappa \right)_+^2 \right) + \varepsilon_{ij} \end{aligned}$$

Here, $z_{i\ell}$ =1 if $z_{il} = \ell$ and 0 other wise for ℓ = 2, …L (L = 6); $\upsilon_\kappa^\ell \sim N\left(0, \; \sigma_{\upsilon\ell}^2\right)$. The parameters $\left[\beta_0 \quad \beta_1 \quad \beta_2 \quad u_\kappa\right]$ in [9] represent the southern Atlantic Coastal Plain (ℓ = 1) and extra terms represent deviation of other regions from the mean trend of the southern Atlantic Coastal Plain.

In order to account for the heterogeneity between stands and trees within stands from the design, we used random stand (b_i) and tree effects (b_{ij}) in the model. Let y_{ijk} represent SG of the k^{th} disk in the j^{th} tree in the i^{th} stand, the model [9] with random stand and tree effects can be represented as:

$$
\begin{aligned}
y_{ijk} = \; & \beta_0 + \beta_1 x_{ijk} + \beta_2 x_{ijk}^2 + \sum_{\kappa=1}^{K} u_k \left(x_{ijk} - \kappa_\kappa\right)_+^2 \\
& + \sum_{\ell=2}^{L} z_{ij\ell} \left(\gamma_{0\ell} + \gamma_{1\ell} x_{ijk} + \gamma_{2\ell} x_{ijk}^2\right) \\
& + \sum_{\ell=2}^{L} z_{ij\ell} \left(\sum_{\kappa=1}^{K} \upsilon_\kappa^\ell \left(x_{ijk} - \kappa_\kappa\right)_+^2\right) + b_i + b_{ij} + \varepsilon_{ijk}
\end{aligned}
\tag{10}
$$

where $b_i \sim N\left(0,\ \sigma_{b_i}^2\right)$ and $b_{ij} \sim N\left(0,\ \sigma_{b_{ij}}^2\right)$.

It is interesting to know the rate at which SG changes along the length of a loblolly pine stem. The derivative of model [10] could potentially be used to explore the rate of change of disk SG with height. This can answer questions such as: How fast does SG change along a stem? Does it approach a plateau? At what height does the rate of change in SG approach a plateau? Differentiating model [10] with respect to x_{ijk} gives

$$
\begin{aligned}
\frac{\partial y_{ijk}}{\partial x_{ijk}} = \; & \beta_1 + 2\beta_2 x_{ijk} + \sum_{\kappa=1}^{K} 2u_k \left(x_{ijk} - \kappa_\kappa\right)_+ \\
& + \sum_{\ell=2}^{L} z_{ij\ell} \left(\gamma_{1\ell} + 2\gamma_{2\ell} x_{ijk}\right) \\
& + \sum_{\ell=2}^{L} z_{ij\ell} \left(\sum_{\kappa=1}^{K} 2\upsilon_\kappa^\ell \left(x_{ijk} - \kappa_\kappa\right)_+\right)
\end{aligned}
\tag{11}
$$

A solution to the mixed model equation can be utilized to get predicted values and standard errors from models [10] and [11]. For more details of model formulation, fitting and prediction using this procedure, readers are referred to Ruppert et al. (2003). All the models in the above sections were fitted using S-plus software (Pinheiro and Bates 2000).

2.3.4 Specific gravity Maps

Maps explaining the variation in whole disk SG across the geographical range of loblolly pine at a given height are useful for making decisions in product categorization and utilization. The spatial variation in a particular entity is usually explained using a method known as kriging

(Cressie 1993), that has been widely applied in geostatistics. Kriging is an interpolation method which predicts the value of a variable (here disk SG) at an unknown spatial point using the spatial covariance information calculated from the available data. Since SG data in this study were collected across space (latitude and longitude) and covariate height, it is important to model the variation in SG across height and space simultaneously. Since the variation in SG across height was highly nonlinear, a geoadditive approach proposed by Kamman and Wand (2003) was used in this study. The geoadditive approach is a combination of geostatistical and additive models and accounts for the nonlinear covariate effect (here tree height) under the assumption of additivity (Kamman and Wand 2003). These models can be implemented using the mixed-model frame work.

The covariate in the present study was relative height of a tree and the geographical locations are represented by latitude and longitude of a stand from which SG was measured. Following Kamman and Wand (2003) and Ruppert et al (2003), the geoadditive model can be formulated as follows. The additive model component for explaining the change in disk SG with continuous variable relative height is given as

[12] $$y_{ij} = \beta_0 + f\left(Rh_{ij}\right) + \varepsilon_{ij}$$

where y_{ij} is the SG measurement from j^{th} disk in i^{th} stand and f is a smoothing function of relative height Rh_{ij}. The model [12] is equivalent to model [7] with a truncated quadratic basis. The random intercept for stand variable was omitted from model [12].

Given the data of form $\left(\mathbf{x}_{ij,}\ y_{ij}\right)$, where y_{ij} is a scalar and $\mathbf{x}_{ij} \in \Re^2$ represents geographical locations, a simple universal Kriging model with linear covariate is

[13] $$y_{ij} = \beta_0 + \beta_1^T \mathbf{x}_{ij} + S\left(\mathbf{x}_{ij}\right) + \varepsilon_{ij}$$

where $\left\{S\left(\mathbf{x}_{ij}\right):\mathbf{x}\in\Re^2\right\}$ is a stationary mean zero stochastic process. Prediction to a new location $\mathbf{x}_0\in\Re^2$ within the sampling space is done by substituting the estimates of $\hat{\beta}_0$ and $\hat{\beta}_1$ and an empirical best linear predictor $\hat{S}\left(\mathbf{x}_0\right)$ for a known covariance structure for *S* in to model [13]. The geographical component was fitted as a linear mixed model by using a bivariate thin plate spline to a geographic location (Ruppert et al. 2003). The covariance for *S* is assumed to be isotropic, i.e. the covariance between two stands which are $\|\mathbf{h}\|$ units apart is the same regardless of direction and the location of the stand.

The final geoadditive model can be obtained by merging models [12] and [13] as:

[14] $$y_{ij}=\beta_0+f\left(Rh_{ij}\right)+\beta_1^T\mathbf{x}_{ij}+S\left(\mathbf{x}_{ij}\right)+\varepsilon_{ij}$$

which can be expressed as a linear mixed model as

[15] $$\mathrm{Y} = \mathrm{X}\boldsymbol{\beta} + \mathrm{Z}\mathbf{u} + \varepsilon$$

where Y is the vector of response (here SG), $\mathrm{X} = \begin{bmatrix}1 & Rh_{ij} & Rh_{ij}^2 & \mathbf{x}_{ij}\end{bmatrix}$ and Z corresponds to the basis functions for *f* and *S*. The additive component in the model allows us to appropriately explain the nonlinear trend in SG with relative height. The geographical component in the model was fitted using reduced knot kriging, where $\{\kappa_1,........,\kappa_\kappa\}$ are subset of knots selected from sample space $\mathbf{x}_{ij}\in\Re^2$. The knots were selected using the space filling algorithm discussed by Kamman and Wand (2003) and Ruppert et al. (2003). Readers are referred to Ruppert et al. (2003) and Kamman and Wand (2003) for more details of geoadditive model formulation, fitting and prediction. Maps were produced by fitting the geoadditive model to the data. The model [14] was fitted using the SemiPar library in R (Wand et al. 2005).

2.4 Results

2.4.1 Parametric model

The model with stand and tree level random effects on parameters β_0, β_1, β_2 was selected as the best random effects model (AIC = -24447.77). After identifying the mixed effect parameters, all the parameters (except the knot parameters α_1 and α_2) were allowed to vary from region to region (AIC = -24538.6). The heteroskedasticity in residuals was accounted for by using a power variance function with fitted values as covariates (AIC = -24609.94). The correlation across repeated measurement taken from each tree was best represented using an ARMA (1, 1) model (AIC = -25007.74).

The difference between mean trends in disk SG among regions was addressed using a likelihood ratio tests (LRT's) by dropping appropriate region specific parameters from the full model fitted above. We also allowed the knot parameters α_1 and α_2 to vary from region to region at this stage. The final model (AIC = -25057.16) was developed by 'stepwise' procedure where a series of LRT's were conducted between the full model and the reduced model by dropping the non-significant parameters from the full model. The parameter estimates from the final fitted model are presented in Table 2.2. All the parameters in the final model were significant at the 0.05 level, except the parameter $\beta_{4,\ \text{northern Atlantic}}$ (LRT didn't favor dropping this parameter from the final model).

Table 2.2: Estimated Parameter from quadratic – quadratic – quadratic model

Parameter	Estimate	Std Error	t-value	p-value
$\beta_{0,\text{ Intercept}}$	0.4678	0.0044	106.98	0.0000
$\beta_{0,\text{ northern Atlantic}}$	-0.0477	0.0096	-4.96	0.0000
$\beta_{0,\text{ Upper Coastal}}$	-0.0197	0.0055	-3.57	0.0004
$\beta_{0,\text{ Piedmont}}$	-0.0251	0.0047	-5.30	0.0000
$\beta_{0,\text{ Hilly Coastal}}$	-0.0135	0.0046	-2.95	0.0031
$\beta_{1,\text{ Intercept}}$	-0.0437	0.0121	-3.60	0.0003
$\beta_{1,\text{ Gulf Coastal}}$	0.0694	0.0147	4.72	0.0000
$\beta_{2,\text{ Intercept}}$	-0.0493	0.0105	-4.70	0.0000
$\beta_{2,\text{ northern Atlantic}}$	0.0344	0.0122	2.82	0.0048
$\beta_{2,\text{ Gulf Coastal}}$	-0.0636	0.0174	-3.65	0.0003
$\beta_{3,\text{ Intercept}}$	1.2199	0.0946	12.89	0.0000
$\beta_{3,\text{ Upper Coastal}}$	-0.5382	0.1514	-3.56	0.0004
$\beta_{3,\text{ Piedmont}}$	-0.1169	0.0423	-2.77	0.0057
$\beta_{4,\text{ Intercept}}$	-5.4250	1.6637	-3.26	0.0011
$\beta_{4,\text{ northern Atlantic}}$	2.4207	1.2575	1.92	0.0543
$\alpha_{1,\text{ Intercept}}$	0.2878	0.0095	30.35	0.0000
$\alpha_{1,\text{ Upper Coastal}}$	0.0512	0.0257	1.99	0.0464
$\alpha_{1,\text{ Hilly Coastal}}$	-0.0171	0.0049	-3.45	0.0006
$\alpha_{2,\text{ Intercept}}$	0.0800	0.0127	6.31	0.0000
$\alpha_{2,\text{ Upper Coastal}}$	-0.0219	0.0107	-2.05	0.0403
Random parameters				
$\sigma_{b_{0,i}}$	0.0187			
$\sigma_{b_{3,i}}$	0.0234			
$\sigma_{b_{03,i}}$	-0.0003			
$\sigma_{b_{0,ij}}$	0.0151			
σ	0.0362			
Heteroskedasticity and autocorrelation parameters				
ϕ	0.8737			
θ	-0.3424			
δ	0.5122			

Based on the final model, estimates of β_0 from the southern Atlantic and Gulf Coastal Plain were not significantly different. The estimated β_0 parameters from other regions were found to be significantly different from these two regions. The estimate of the β_0 parameter was highest for the southern Atlantic and Gulf Coastal Plain (0.4678) and lowest for the northern Atlantic Coastal Plain (0.4201). The estimate of the β_0 parameter for the other three regions was between these two groups (Piedmont = 0.4427; Upper Coastal Plain = 0.4481; Hilly Coastal Plain = 0.4543). The estimate of the β_1 parameter was not significantly different for all regions except the Gulf Coastal Plain. Similarly, the estimated β_2 parameter was not significantly different for all regions except the northern Atlantic and Gulf Coastal Plain. The estimated β_3 parameter from the Upper Coastal Plain and Piedmont was significantly different from all other regions. The β_4 parameter was significantly different for the northern Atlantic Plain compared to all other regions. The first knot from tip of the tree, α_1, was estimated to be at 0.2878 for all regions except the Upper Coastal Plain (0.3390) and Hilly Coastal Plain (0.2707). The estimate of the second knot parameter from the tip of the tree, α_2 was at a relative height of 0.08 for all regions except Upper Coastal Plain, where the estimate was at a relative height of 0.0581.

Plots of mean predicted disk SG is presented in Figure 2.3. Based on the three segmented quadratic model, the mean disk SG trends of trees from the southern Atlantic and Gulf Coastal Plain were higher than all other regions with the mean trend of Gulf Coastal Plain above the southern Atlantic Plain. The mean trend in disk SG was lowest for trees from the northern Atlantic Coastal Plain. Mean disk SG trend of the other regions fell between these two limits with the Hilly Coastal Plain having the highest SG's followed by the Upper Coastal Plain and then the Piedmont. It was also observed that the mean trend in disk SG of trees from the northern

Atlantic Coastal Plain merged with the mean SG trend of the Hilly Coastal, Upper Coastal and Piedmont above a relative height of 0.8.

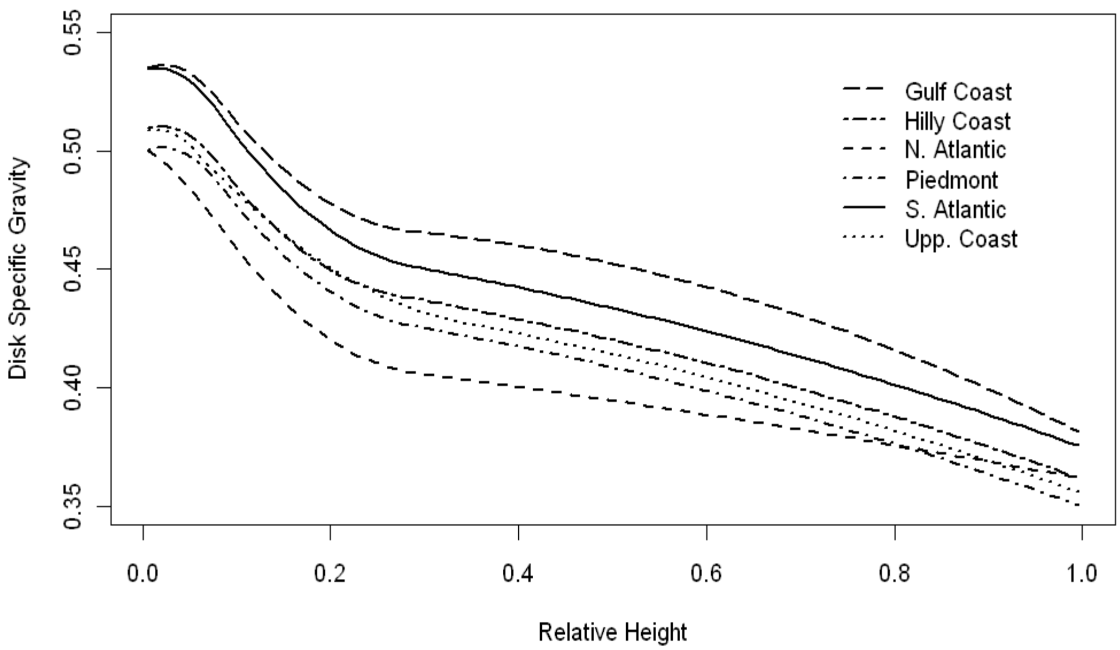

Figure 2.3: Predicted disk SG for six regions from the three segmented (two knots) parametric model

2.4.2 Semiparametric model

The nonlinear trend in disk SG with tree height and the regional variation in mean trend was explained more thoroughly using a semiparametric model. A model with common smoothing parameter for all regions $\left(\sigma_u^2 = \sigma_{\upsilon\ell}^2\right)$ was favored and fitted based on preliminary analysis. Based on the fitted model, disk SG follows a decreasing trend with relative height. A test of regional variation on the mean trend of disk SG was addressed by using a LRT test by fitting the full model (Eq. 10 with assumed common smoothing parameters for all regions) and a reduced model with $\mathrm{H}_0: \gamma_{p\ell} = 0$ where p=0, 1, 2 and ℓ = 2, 3, 4, 5 or 6. Based on the LRT, significant differences between regions was found with a test statistic of 103.94 (p-value <0.0001) which follows an asymptotic $\chi^2_{\mathrm{df}=15}$.

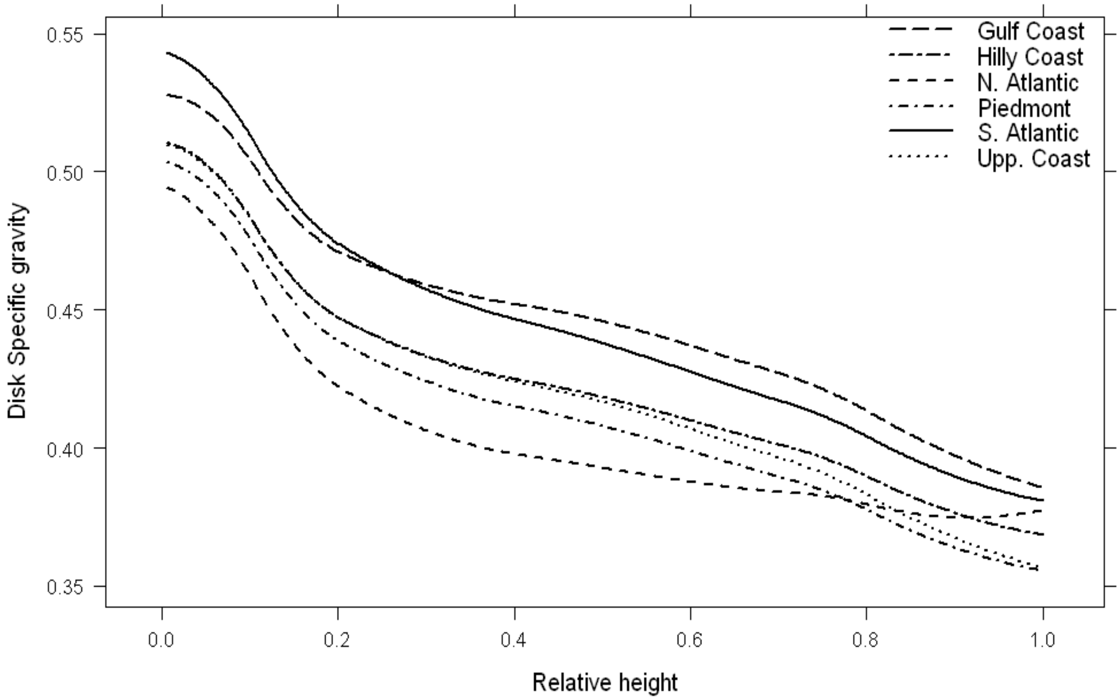

Figure 2.4: Predicted disk SG for six regions from semiparametric model (8 known knots)

A plot of predicted SG from the model is presented in Figure 2.4. Mean trends of SG for trees from southern Atlantic and Gulf Coastal Plain was higher than all other regions. It was observed from the mean plot of these two regions that at the base of the tree the mean trend of disk SG for the southern Atlantic Plain was above the Gulf Coastal Plain up to a relative height ~0.25, at relative heights >0.25 the trend was reversed. The predicted disk SG of trees from the northern Atlantic Coastal Plain was the lowest. The predicted SG of other regions again fell between these two groups with the Hilly Coastal Plain having the highest predicted disk SG's,

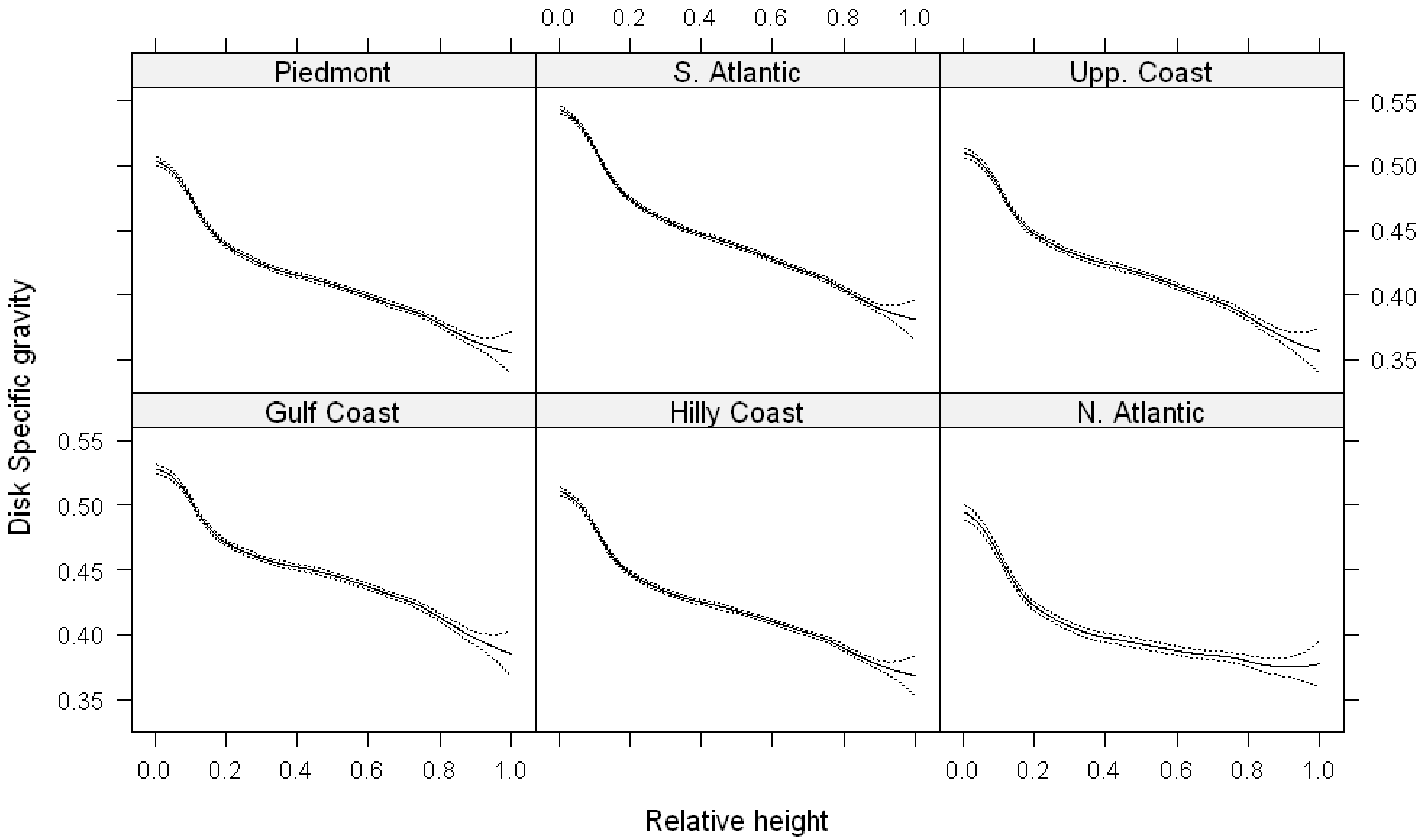

Figure 2.5: Predicted plot with 95 % individual prediction band from semiparametric model (8 known knots) by region.

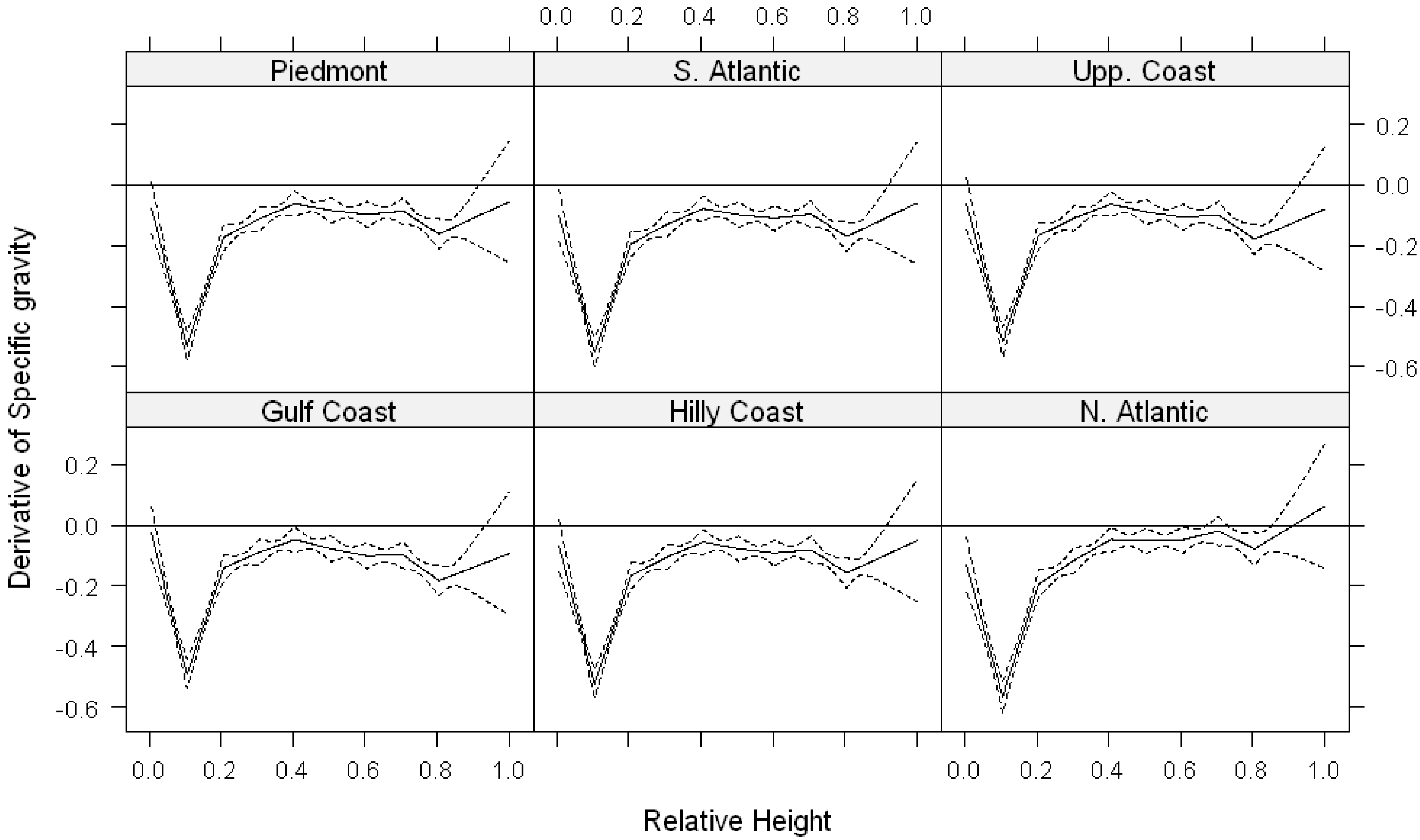

Figure 2.6: Derivative plot from semiparametric model (8 known knots) by region.

followed by the Upper Coastal Plain and the Piedmont. The predicted SG of trees from northern Atlantic Coastal Plain again merged with the mean SG trend of the Hilly Coastal, Upper Coastal Plain and Piedmont above a relative height of 0.8. A plot of the predicted disk SG with 95 % prediction intervals is presented in Figure 2.5. It was observed that the variability around the predictions was very narrow. A plot of the derivatives of mean predicted disk SG along with 95 % point wise confidence bands is presented in Figure 2.6. Based on the plot, SG decreases very rapidly near the base of the tree to a relative height of ~0.1, then decreases at a decreasing rate from a relative height of ~0.1 – ~0.3, and then decreased at a constant rate above a relative height ~0.3 to the top of the tree.

2.4.3 Specific gravity maps

After fitting the geoadditive model (Eq. 14; Figure 2.7), it was observed that the stand average disk SG followed a similar pattern as described based on semiparametric model (Figure 2.6).

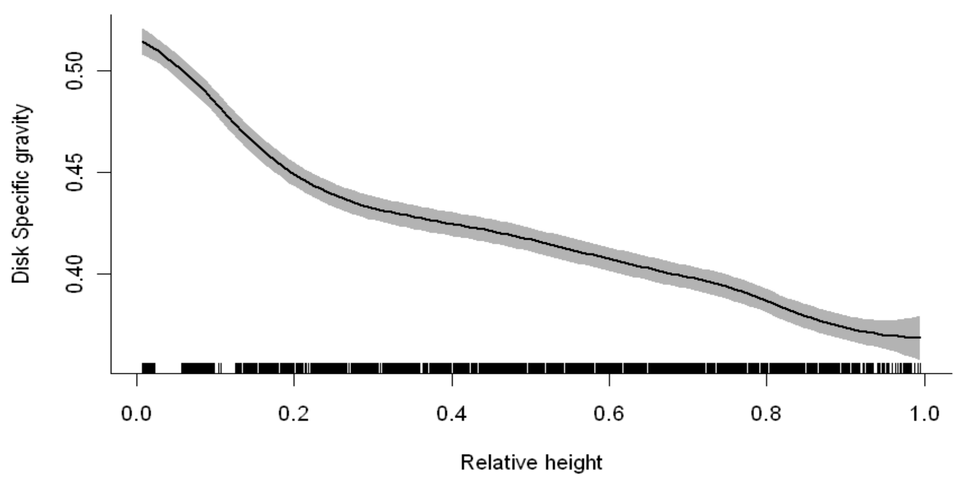

Figure 2.7: Plot showing the effect of relative height on whole stand disk SG, with 95 % variability bar from geoadditive model fitting.

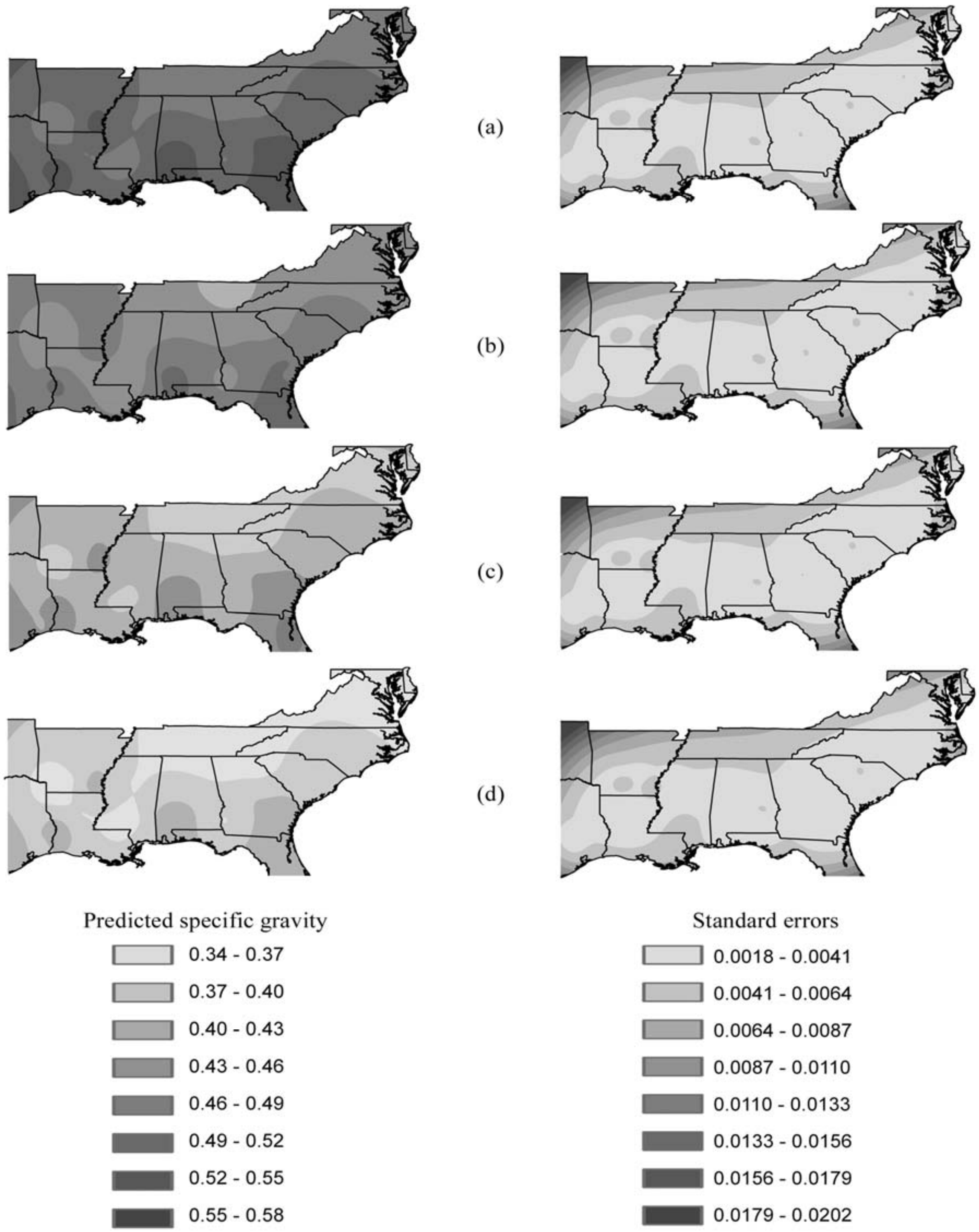

Figure 2.8: Maps showing the predicted disk SG and standard error using the geoadditive model at specific relative height: (a) Relative height = 0.05; (b) Relative height = 0.15; (c) Relative height = 0.5; (d) Relative height = 0.8.

Maps showing the geographical variation in whole stand disk SG along with the standard error of predictions at specific relative heights (0.05, 0.15, 0.5, and 0.8) are presented in Figure 2.8. The maps were made under the assumption that all the stands sampled are of same age (average age of ~23). The primary reason for making such an assumption is that the stands sampled were comes from a narrow range of age (Table 1). Lack of any disk SG trend with stand age further supports this assumption (based on plot, not presented here). Based on the maps a decreasing trend in disk SG was observed from south to north and from east to west. Whole stand SG was higher near the Coastal Plain with high SG bands in the southern Atlantic and Gulf Coastal Plains. Disk SG was high in the southern part of Georgia, the southwest of Alabama and the western edge of Texas. The Upper Coastal Plain, Hilly Coastal Plain and Piedmont formed a band of lower SG wood, while the lowest SG wood was from the northern Atlantic Plain and parts of the Piedmont. Areas with low sampling intensity, such as Tennessee, the northern part of Arkansas, Alabama, Virginia, and southern parts of Mississippi and Louisiana can be identified from the large standard errors of prediction.

2.5 Discussion

Disk SG of loblolly pine trees decreases in a nonlinear fashion with tree height. Both parametric and semiparametric approaches were used to explain the longitudinal and regional variation in disk SG along the stem. A geoadditive approach was used to describe the regional variation in disk SG at a specific disk height and significant regional variation in mean SG trend was observed. Generally, mean SG trends were higher for trees from the southern Atlantic and Gulf Coastal Plain and lowest for trees from the northern Atlantic Coastal Plain with trees from the Hilly Coastal, Upper Coastal Plain and Piedmont, between these extremes. Both parametric and semiparametric modeling approaches agreed and resulted in similar conclusions.

Our study suggests that the stem of loblolly pine can be divided into three zones based on the longitudinal variation of disk SG. Based on the derivative plots from the semiparametric model (Figure 2.6) for all regions, mean SG decreased rapidly from the base of the tree to a relative height ~0.1; SG then decreased at a decreasing rate between relative heights of ~0.1 - ~0.3; for relative heights >~0.3 SG decreases at constant rate. The result from semiparametric model supports the proposed parametric model where a stem is represented as three segments with each segment represented by a quadratic function of relative height with two knot points which are unknown and estimated from the data. Based on the three segment parametric model, the first change in curve shapes occurred approximately at a relative height of 0.08. The second change in curve shape of mean disk SG was around a relative height of 0.29. These findings agree with the three segmented classification of the stems of loblolly pine proposed by Burdon et al. (2004).

The mean trend of disk SG was highest for the southern Atlantic and Gulf Coastal Plain. The overall mean SG observed for these two regions was 0.46, which was higher than the mean disk SG observed for the other regions (0.42) (Table 2.1). Both parametric and semiparametric models support this conclusion with higher mean SG curves of trees from southern Atlantic and Gulf Coastal Plains compared to all other regions (Figure 2.3 and 2.4). The high SG of trees from southern Atlantic and Gulf Coastal Plain might be attributed to two major reasons: (1) reduced length of core wood formation and proportion of core wood formed in these two regions compared to other inland regions (Clark and Daniels 2002; Clark and Saucier 1989; Jordan et al. 2008); (2) high latewood percent in the rings of trees growing in these regions (~40 %) compared to other regions (~35 %). The proportion of latewood formed is highly correlated with summer precipitation, mean annual temperature and number of growing days. The trees growing in the southern Atlantic and Gulf Coastal Plains, on average, receive more summer precipitation, have a higher mean annual temperature and more growing days than the other regions (Clark and Daniels 2002).

Maps of mean stand disk SG showed similar trends of regional variation in SG as described based on the parametric and semiparametric models. A decreasing trend in disk SG was present from south to north and from east to west. The disk SG maps, depending on the specified height, divided the loblolly pine growing range into three major regions. A high SG band which mainly included parts of the southern Atlantic and Gulf Coastal Plain, a medium SG band which included the northern parts of the southern Atlantic and Gulf Coastal Plain, parts of the Upper Coastal and Hilly Coastal Plain, and a low SG band which included the Piedmont, Hilly Coastal and northern Atlantic Coastal Plains. The above findings agrees with the earlier results by Clark and Daniels (2002) and Jordan et al. (2008) where they found decrease in SG with increase in latitude and increase in SG with increase in longitude based on the ring by ring data collected from breast height of trees.

Significant longitudinal and regional variation in SG was observed in loblolly pine. For forest product industries an understanding of both longitudinal and regional variation in SG is important as it allows raw material segregation and optimization of manufacturing processes. SG is an important wood quality index and highly correlated with the strength and stiffness of wood and determines the pulp yield and quality. An increase in SG of 0.02 units will result in a 22.7 kg increase in dry pulp per ton of round wood (Mitchell 1964) and/or 31.15 and 3516 kg/cm^2 increase in modulus of rupture and modulus of elasticity (Wahlgren and Schumann 1975). Hence the strength of lumber or yield of pulp from a tree harvested from the southern Atlantic and Gulf

Coastal Plains will generally be greater than trees harvested from other regions at an equivalent age.

2.6 Acknowledgements

Thanks to the Wood Quality Consortium at the University of Georgia for providing the funding and data for this study. Authors are indebted to USDA Forest Service, Athens, GA for their invaluable help in collecting the data.

2.7 References

Burdon, R.D., R.P. Kibblewhite, C.F.W. John, A.M. Robert, E. Robert, and J.C. David. 2004. Juvenile versus mature wood: A new concept, orthogonal to corewood versus outerwood, with special reference to Pinus radiata and P. taeda. Forest Science 50:399-415.

Clark, A., and R.F. Daniels. 2002. Modeling the effect of physiographic region on wood properties of planted loblolly pine in Southeastern United states. P. 54-60 *in* Connection between Forest Resources and Wood Qulaity: Modeling Apporaches and Simulation Software. Fourth Workshop IUFRO Working Party S5.01-04. INRA - Centre de Researches de Nancy, France, Harrison Hot Springs, B.C.

Clark, A., and J.R. Saucier. 1989. Influence of initial planting density, geographic location, and species on juvenile wood formation in southern pine. Forest Products Journal 39(7-8):42-48.

Cressie, N. 1993. Statistics for spatial data. Wiley, New York. 887 p.

Fox, T.R., E.J. Jokela, and H.L. Allen. 2007. The development of pine plantation silviculture in the southern United States. Journal of Forestry 105(7):337-347.

Gallant, A.R., and W.A. Fuller. 1973. Fitting segmented polynomial regression models whose join points have to be estimated. Journal of the American Statistical Association 68(341):144-147.

He, R. 2004. Mixed effects modeling of wood properties of loblolly pine in the southeastern United States (Ph.D dissertation), University of Georgia, Athens. 213 p.

Jordan, L., A. Clark, L.R. Schimleck, D.B. Hall, and R.F. Daniels. 2008. Regional variation in wood specific gravity of planted loblolly pine in the United States. Canadian Journal of Forest Research 38:698-710.

Kamman, E.E., and M.P. Wand. 2003. Geoadditive Models. Applied Statistics 52:1-18.

Larson, P.R., D.E. Kretschmann, A. Clark, and J.G. Isebrands. 2001. Formation and Properties of Juvenile Wood in Southern Pines A Synopsis, USDA Forest Service, General Technical Report FPL-GTR-129 42.

Max, T.A., and H.E. Burkhart. 1976. Segmented polynomial regression applied to taper equations. Forest Science 22:283-289.

Megraw, R.A. 1985. Wood quality factors in loblolly pine. TAPPI Press, Atlanta, GA.

Mitchell, H. 1964. Patterns of variation in specific gravity of southern pines and other coniferous species. TAPPI Journal 47:276 - 283.

Ngo, L., and M.P. Wand. 2004. Smoothing with mixed model software. Journal of Statistical Software 9(1):1-54.

Panshin, A.J., and C. deZeeuw. 1980. *Textbook of Wood Technology*. McGraw-Hill Book Co., New York.

Phillips, K.M. 2002. Modeling within tree changes in wood specific gravity and moisture content for loblolly pine in Georgia (M.S. Thesis), University of Georgia, Athens. 95 p.

Pinheiro, C.J., and D.M. Bates. 2000. Mixed effects models in S and S-plus. Springer, New York. 523 p.

Pinheiro, C.J., M.D. Bates, S. DebRoy, D. Sarkar, and R.C. team. 2009. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3. 1-92.

Ruppert, D., M.P. Wand, and R.J. Carrol. 2003. Semiparametric Regression. Cambridge University Press, New York. 385 p.

Tasissa, G., and H.E. Burkhart. 1998a. Juvenile-mature wood demarcation in loblolly pine trees. Wood and Fiber Science 30(2):119-127.

Tasissa, G., and H.E. Burkhart. 1998b. Modeling thinning effects on ring specific gravity of loblolly pine (Pinus taeda L.). Forest Science 44(2):212-223.

Vonesh, E.F., and V.M. Chinchilli. 1997. Linear and Nonlinear Models for the Analysis of Repeated Measurements. Marcel Dekker, New York, NY.

Wahlgren, H.D., and D. Schumann. 1975. Properties of major southren pines: Part I - Wood desnsity survey. USDA Forest Service, Forest Product Laboratory, Madison, Wisconsin, Research Papper FPL-176:57p.

Wand, M.P., B. Coull, J. French, B. Ganguli, E.E. Kamman, J. Staudenmeier, and A. Zanobetti. 2005. SemiPar 1.0 R Package [online]. Available from cran.r-project.org.

Wear, D.N., and J.G. Greis. 2002. The Southern Forest Resource Assessment: Summary Report. USDA Forest Service, Southern Research Station Report No. 53. 635p

Zobel, B. 1972. Three Rings-Per-Inch, Dense Southern Pine-Can It Be Developed? Journal of Forestry 70:333-336.

Zobel, B.J., and J.P. vanBuijtenen. 1989. Wood Variation: Its Causes and Control. Springer-Verlag, Berlin. 363 p.

Chapter 3

A multivariate mixed model system for wood specific gravity and moisture content of planted loblolly pine stands in the southern United States

3.1 Abstract

Specific gravity (SG) and moisture content (MC) both have a strong influence on the quantity and quality of wood fiber in a given volume. SG and MC are highly negatively correlated: high SG is associated with low MC and vice-versa. We proposed a multivariate mixed model system to model the two properties simultaneously. Disk SG and MC at different height levels were measured from three trees in 135 stands across the natural range of loblolly pine (*Pinus taeda* L.). Stand average disk SG and MC were used for simultaneous modeling of the SG-MC system. Regional variation in mean trend of the properties was incorporated in the model. Contemporaneous correlation between the two components in the model was accounted for by defining within stand error structure appropriately. The predictive performance of the multivariate model relative to univariate models for SG and MC was evaluated assuming that one variable was available to predict the other variable. Compared to univariate models for disk SG and MC predictions based on the multivariate model were improved by approximately 29 and 26 % in root mean square prediction error (RMSEP) respectively after taking account of the contemporaneous correlation between the two properties.

Key words: multivariate, multilevel, nonlinear mixed model, systems of equations, wood properties, specific gravity

3.2 Introduction

A forest is a complex dynamic system with inter-related individual components. Foresters commonly rely on simultaneous modeling systems to explain such inter-dependent systems. One familiar example of such a system to forest biometricians is simultaneous modeling of dominant height, basal area, trees per hectare and volume (Borders 1989; Fang et al. 2001; Hall and Clutter 2004). Two main reasons for the popularity of simultaneous modeling systems in forestry are: 1) compatibility requirement of individual components in the system (Clutter 1963); 2) contemporaneous correlation of error among individual components in the system.

Specific gravity (SG) and moisture content (MC) both have a strong influence on the quantity and quality of wood. SG describes the mass of woody material present in a given volume of wood. It is a unit-less measure and expressed as the ratio of wood basic density (oven dry weight divided by green volume) with the density of water at 4^{o}C (Megraw 1985). SG is considered an important wood property because of its strong correlation with the strength of solid wood products, as well as the yield and quality of pulp produced (Panshin and deZeeuw 1980). Generally the moisture content of wood is expressed as a percentage of the oven dry weight of wood. Moisture content influences the physical and mechanical properties of wood, resistance to biological deterioration and dimensional stability (Haygreen and Bowyer 1996).

SG and MC vary considerably within loblolly pine (*Pinus taeda* L.) trees. SG follows a decreasing trend with tree height (He 2004; Megraw 1985; Phillips 2002; Zobel and Blair 1976), while MC increases with height (Koch 1972; Phillips 2002). It has been reported that these two variables are highly negatively correlated with high SG associated with low MC and vice-versa (Koch 1972; Zobel and Blair 1976). The primary factor controlling the longitudinal variation in disk SG and MC in a loblolly pine tree is the proportion of juvenile wood (Zobel and Blair 1976;

Zobel and vanBuijtenen 1989). In general, the proportion of juvenile wood is higher towards the top of a tree than at the base and juvenile wood has lower SG and higher MC than mature wood.

The objective of this study was to model the longitudinal variation in disk SG and MC as a simultaneous multivariate mixed model system. We will show how contemporaneous correlation between these two variables (disk SG and MC) can be potentially utilized to improve the prediction of disk SG or MC for loblolly pine at any height.

3.3 Data

The Wood Quality Consortium at the University of Georgia and the United States Department of Agriculture (USDA) Forest Service Southern Research station sampled planted loblolly pine across its natural range to study the longitudinal variation in wood SG and MC. Trees were sampled from 135 stands from six physiographic regions across the southeastern United States. Regions sampled included: 1- southern Atlantic Coastal Plain (**R1**), 2- northern Atlantic Coastal Plain (**R2**), 3- Upper Coastal Plain (**R3**), 4- Piedmont (**R4**), 5- Gulf Coastal Plain (**R5**) and 6- Hilly Coastal Plain (**R6**). A minimum of 12 plantations from each of the six physiographic regions were sampled. The stands selected for sampling included 20- to 25-year-old loblolly pine plantations planted at 1250 or more trees per hectare and having 625 trees per hectare or more after thinning. Only stands that were conventionally managed with no fertilization (except phosphorus at planting on phosphorus deficient sites) and no competition control were sampled. Three trees from each stand were felled and cross sectional disks of 3.8 cm thickness were collected from 0.15, 1.37 m and then 1.52 m intervals along the stem up to a diameter of 50 mm outside bark. The disks were sealed in plastic bags and shipped to the USDA Forest Service laboratory for physical property analysis. Disk SG (based on green volume and oven-dry weight) and disk MC (based on green and oven-dry weights) were determined for each sampling height. Stand averages (at each height) for disk SG

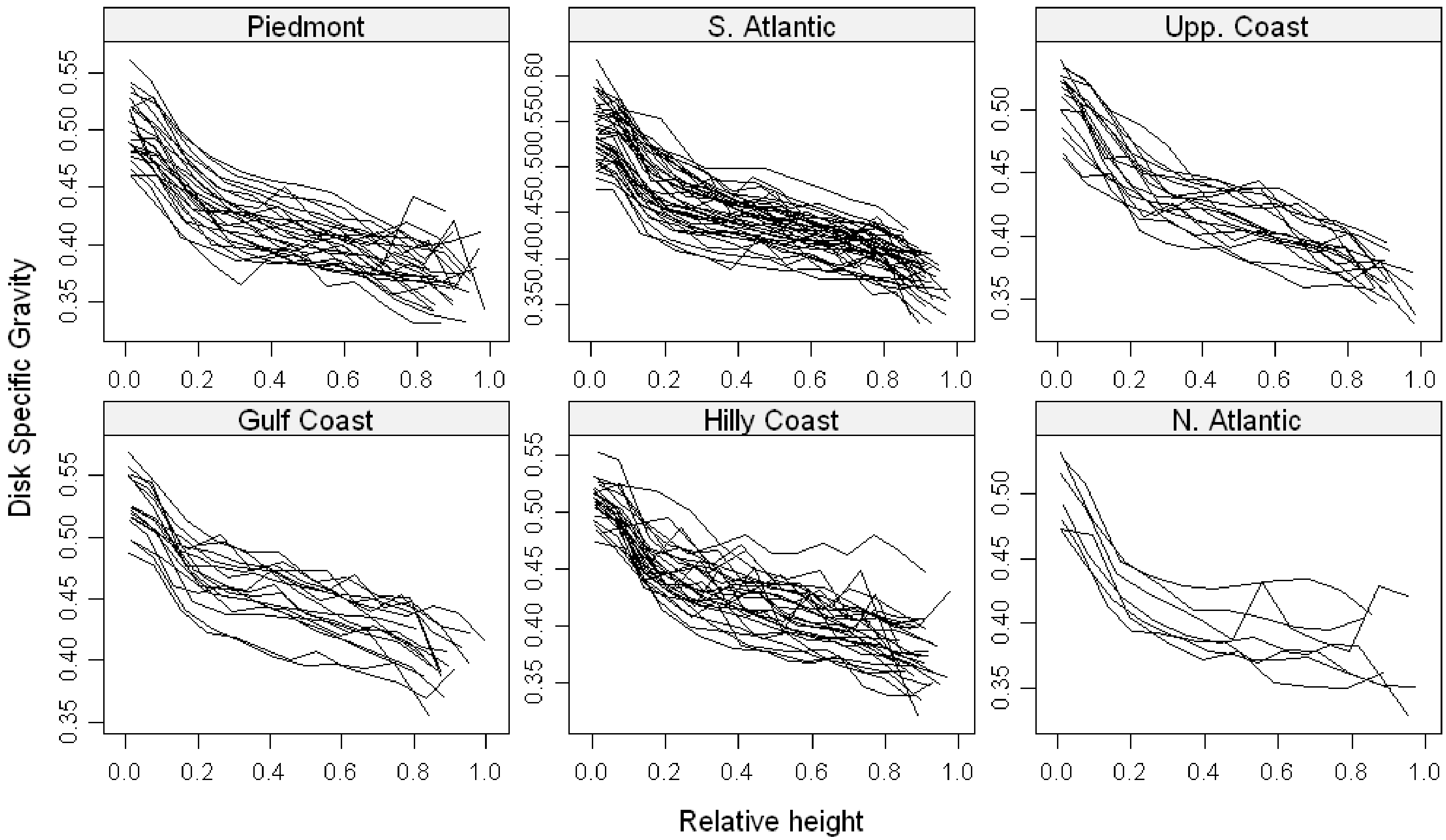

Figure 3.1: Plot showing observed stand level disk specific gravity with relative height by region.

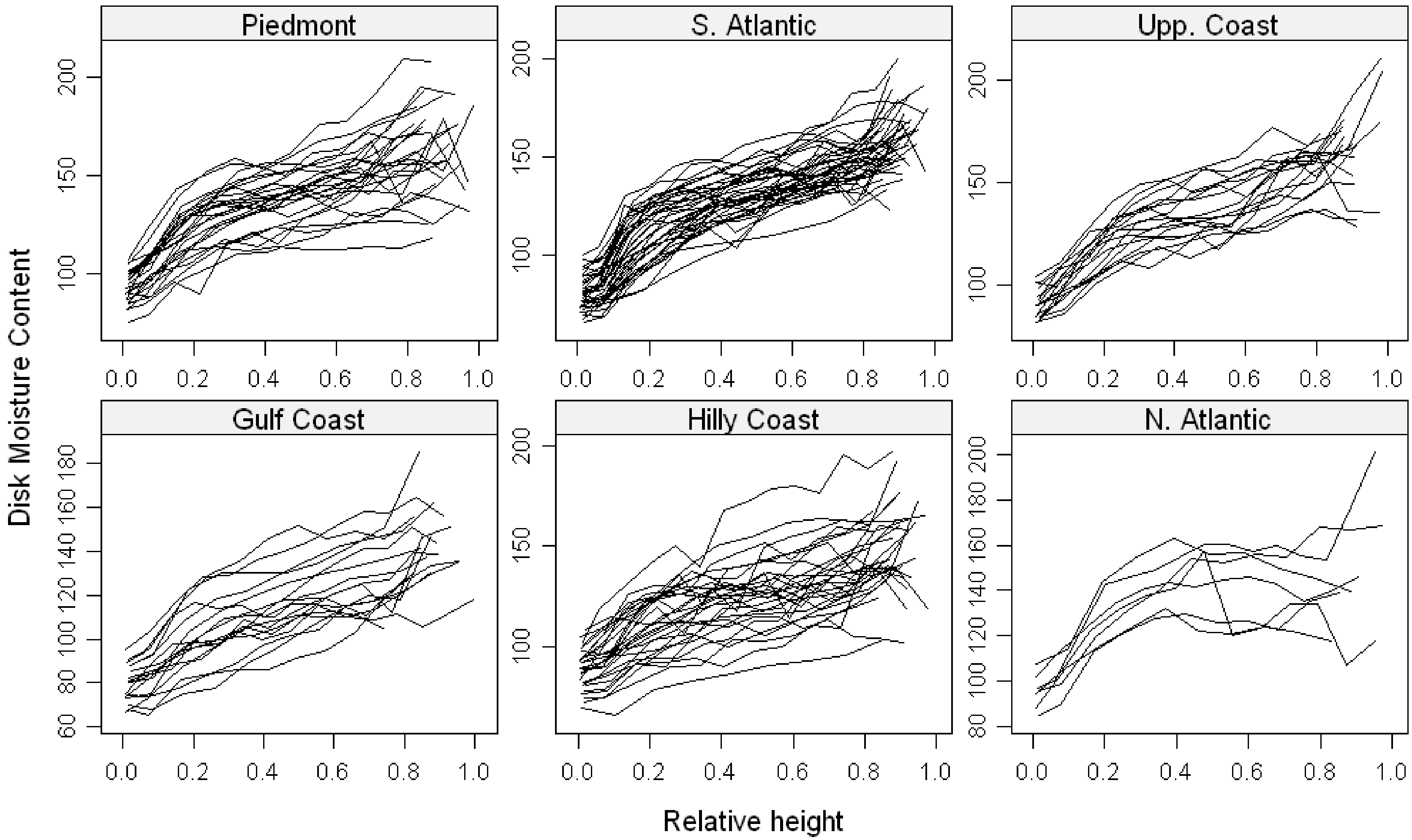

Figure 3.2: Plot showing observed stand level disk moisture content with relative height by region.

and MC were calculated using the three trees sampled per stand. A summary of average stand characteristics for each region is presented in Table 3.1. Plots of stand average disk SG and MC with relative height are presented in Figures 3.1 and 3.2.

Table 3.1: Average stand attributes by region, standard deviations are in parenthesis.

Region	Age	DBH (cm)	Total Ht (m)	Disk SG	Disk MC
southern Atlantic	22.74 (1.82)	24.13 (3.41)	20.86 (2.15)	0.4489 (0.052)	126.48 (26.28)
northern Atlantic	22.33 (1.72)	24.63 (3.08)	18.72 (2.26)	0.4085 (0.044)	133.78 (22.37)
Upper Coastal	23.00 (1.46)	24.25 (3.24)	19.38 (2.66)	0.4254 (0.046)	133.25 (25.34)
Piedmont	23.08 (2.03)	23.97 (2.49)	18.19 (1.69)	0.4206 (0.046)	134.85 (25.71)
Gulf Coastal	23.21 (3.43)	21.08 (2.47)	19.38 (2.36)	0.4553 (0.042)	112.29 (23.10)
Hilly Coastal	23.86 (3.60)	23.40 (3.27)	19.55 (2.60)	0.4288 (0.046)	122.15 (24.18)

3.4 Model development

Two response components are considered in this simultaneous model system, disk SG and MC measured at the same heights for 3 trees in a stand. The basic models adopted for these two components are

[1]
$$SG = f_1(x,\ \beta) = \beta_{0,1} + \beta_{1,1}x + \beta_{2,1}x^2 + \beta_{3,1}(\alpha_{1,1} - x)_+^2 + \beta_{4,1}(\alpha_{2,1} - x)_+^2 + \varepsilon_{SG}$$

[2]
$$MC = f_2(x,\ \beta) = \beta_{0,2} + \beta_{1,2}x + \beta_{2,2}(\alpha_{1,2} - x)_+^2 + \varepsilon_{MC}$$

where *SG* = disk SG; *MC* = disk MC; *x* = relative height *h/H*, *h* is the average height above ground and *H* is the average total height of the stand calculated from the three sampled trees; $[\beta_{0,1} \quad \beta_{1,1} \quad \beta_{2,1} \quad \beta_{3,1} \quad \beta_{4,1} \quad \beta_{0,2} \quad \beta_{1,2} \quad \beta_{2,2} \quad \alpha_{1,1} \quad \alpha_{2,1} \quad \alpha_{1,2}]^T$ are parameters to be estimated, with knot parameters $[1 > \alpha_{1,1} > \alpha_{2,1} > 0]$ and $[1 > \alpha_{1,2} > 0]$; ε_{SG} and ε_{MC} are error terms for disk SG and MC respectively. The $\left(\alpha_j - x\right)_+^2$ terms indicates the positive part of the function $\alpha_j - x$ where "+" sets it to zero for those values of x where $\alpha_j - x$ is negative (here $x > \alpha_j$). The basic model form for disk SG is equivalent to the standard form of the taper model proposed by Max and Burkhart (1976), which is not constrained to have a value of zero at the tip of the tree.

In order to account for stand-to-stand variability in the data, we used a nonlinear mixed effect model (NLMM). Let y_{ijk} represent the k^{th} response (k = 1, 2) variable measured at j^{th} relative height from i^{th} stand; the univariate nonlinear mixed model for each property can be represented as

[3] $$y_{ij1} = \theta_{0,i1} + \theta_{1,i1}x_{ij} + \theta_{2,i1}x_{ij}^2 + \theta_{3,i1}\left(\alpha_{1,i1} - x_{ij}\right)_+^2 + \theta_{4,i1}\left(\alpha_{2,i1} - x_{ij}\right)_+^2 + \varepsilon_{ij1}$$

[4] $$y_{ij2} = \theta_{0,i2} + \theta_{1,i2}x_{ij} + \theta_{2,i2}\left(\alpha_{1,i2} - x_{ij}\right)_+^2 + \varepsilon_{ij2}$$

The mixed effect parameter θ_{ijk} in the above models takes the form

[5] $$\theta_{ik} = A_{ik}\beta_k + B_{ik}b_{i,k}$$

where $b_{i,k}$ is the i^{th} stand level random effect vector specific to the k^{th} response variable with $b_{i,k} \sim N\left(0, \Psi_k\right)$; B_{ik} is the associated random effect design matrix; A_{ik} is the fixed effect design matrix and β_k is the fixed effect parameter vector specific to the k^{th} response variable.

In order to develop the bivariate model, we first fitted the univariate stand level NLMM's model for disk SG (Eq. 3) and MC (Eq. 4) separately. Initially we assumed all the parameters in the univariate models were mixed. Final specification of mixed effect parameters in the univariate models were decided based on Akaike's Information Criteria (AIC), a model selection criterion used for NLMM's. Parameters $\beta_{0,i1}$, $\beta_{1,i1}$, $\beta_{2,i1}$, $\beta_{0,i2}$ and $\beta_{1,i2}$ were selected as mixed, with random stand level intercepts in these parameters. The regional variation in mean trend for both properties was incorporated by appropriate fixed effect specification (fixed effect design matrix) for all parameters, except the knot parameters, in both univariate modes. The knot parameters were assumed as common for all regions for both properties. Since we had six distinct physiographical regions in the study, we assumed different fixed effect parameters for each region with the southern Atlantic Coastal Plain as the reference region with all other regions having their own parameters which are deviations from the reference region. The final fixed effect specifications for each parameter were identified using univariate models for each property and likelihood ratio test between full model and reduced model. The fixed effect specifications corresponds to all parameters used in the bivariate model are presented in Table 3.2.

Table 3.2: Fixed effect specifications.

Parameter	Fixed effect specification
$\beta_{0,1}$	1+ R2 + R3 + R4 + R6
$\beta_{1,1}$	1
$\beta_{2,1}$	1 + R2
$\beta_{3,1}$	1 + R3 + R4 + R5 + R6
$\beta_{4,1}$	1
$\alpha_{1,1}$	1
$\alpha_{2,1}$	1
$\beta_{0,2}$	1 + R2 + R3 + R4 + R5
$\beta_{1,2}$	1 + R2 + R6
$\beta_{2,2}$	1 + R2 + R5 + R6
$\alpha_{1,2}$	1

The variance-covariance structure for $\text{var}\left(\text{b}_{i,\ 1}\right)$ and $\text{var}\left(\text{b}_{i,\ 2}\right)$ in the univariate models were selected based on the model selection criteria (AIC and Bayesian information criterion (BIC)). We selected a general positive definite form of variance-covariance structure for disk SG and a diagonal form of variance-covariance structure for disk MC. The model information criteria and log likelihood values for the final selected univariate models, called **SG1** and **MC1** respectively for each response, are presented in Table 3.3.

For fitting the bivariate model, the univariate model equations for two responses were stacked together and can be represented as

[6] $$\mathbf{y}_{ij} = f\left(\mathbf{x}_{ij},\ \boldsymbol{\theta}_i\right) + \boldsymbol{\varepsilon}_{ij}$$

where $\mathbf{y}_{ij} = \left(y_{ij1},\ y_{ij2}\right)$. To take account of the correlation between responses measured from the same stand at the same height level, we assumed the within stand variance-covariance matrix as $\boldsymbol{\varepsilon}_{ij} \overset{i.i.d}{\sim} N\left(0,\ \sigma^2\boldsymbol{\Lambda}\right)$ where $\boldsymbol{\varepsilon}_{ij} = \left(\varepsilon_{ij1},\ \varepsilon_{ij2}\right)$. Following Eq. 5, after stacking the fixed effect and random effect vectors and design matrices for two response variables, we can write $\boldsymbol{\theta}_i = \left(\theta_{i1},\ \theta_{i2}\right)$ as

[7] $$\theta_i = A_i\beta + B_ib_i$$

where $A_i = \text{diag}\left(A_{i1},\ A_{i2}\right)$; $B_i = \text{diag}\left(B_{i1},\ B_{i2}\right)$; $\beta = \left(\beta_1^T,\ \beta_2^T\right)^T$; $b_i = \left(b_{i,1}^T,\ b_{i,2}^T\right)^T$ and we assumed that $b_i \overset{i.i.d}{\sim} N\left(0,\ \boldsymbol{\psi}\right)$.

All the models were fitted using the nlme package in R, version 2.9.1 (Pinheiro et al. 2009). Initially the two univariate models (Eq. [3] and [4]) were simultaneously fitted, referred to as **SGMC1**, with a positive definite form of variance-covariance structure for disk SG, a diagonal form of variance-covariance structure for disk MC and unique variance parameter estimate for each response variable. Here, a block-diagonal form was used to define the random effect structure of two responses as follows

$$\boldsymbol{\psi} = \begin{pmatrix} \operatorname{var}\left(b_{i,1}\right) & \mathbf{0} \\ \mathbf{0} & \operatorname{var}\left(b_{i,2}\right) \end{pmatrix}$$

The advantage of multivariate fitting over univariate fitting is that we can incorporate correlation among errors and random effects associated with different response variables in the model by specifying different forms of $\boldsymbol{\Lambda}$ and $\boldsymbol{\psi}$ (Fang et al. 2001; Hall and Clutter 2004). The contemporaneous correlation between responses was incorporated by relaxing the form of $\boldsymbol{\Lambda}$ from an identity matrix to a symmetric positive definite matrix (referred to as **SGMC2**). We also allowed for correlation among random effects associated with the two models. The final best fitted model (referred to as **SGMC3**) is represented as follows

$$\begin{aligned} y_{ij1} = {} & \left(\beta_{0\ell,\,1} + b_{0i,1}\right) + \left(\beta_{1\ell,\,1} + b_{1i,1}\right)x + \left(\beta_{2\ell,\,1} + b_{2i,1}\right)x^2 + \beta_{3\ell,\,1}\left(\alpha_{1,\,1} - x\right)_+^2 \\ & + \beta_{4\ell,\,1}\left(\alpha_{2,\,1} - x\right)_+^2 + \varepsilon_{ij1} \end{aligned}$$

$$y_{ij2} = \left(\beta_{0\ell,\,2} + b_{0i,2}\right) + \left(\beta_{1\ell,\,2} + b_{1i,2}\right)x + \beta_{2\ell,\,2}\left(\alpha_{1,\,2} - x\right)_+^2 + \varepsilon_{ij2}$$

[8] $$\begin{pmatrix} b_{0i,1} \\ b_{1i,1} \\ b_{2i,1} \\ b_{0i,2} \\ b_{1i,2} \end{pmatrix} \sim N\left(\mathbf{0}, \boldsymbol{\Psi}\right) \qquad \text{where} \qquad \boldsymbol{\Psi} = \begin{pmatrix} \varphi_{00,1} & \varphi_{01,1} & \varphi_{02,1} & \varphi_{00,12} & 0 \\ & \varphi_{11,1} & \varphi_{12,1} & \varphi_{10,12} & 0 \\ & & \varphi_{22,1} & \varphi_{20,12} & 0 \\ & & & \varphi_{00,2} & 0 \\ & & & & \varphi_{11,2} \end{pmatrix}$$

$$\left(\varepsilon_{ij1} \quad \varepsilon_{ij2}\right)^T \Big| \mathbf{b}_i \sim N\left(\mathbf{0}, \mathbf{R}_i\right),$$

where

$$\mathbf{R}_i = \sigma^2 \mathbf{G}_i^{1/2}(\delta)\boldsymbol{\Gamma}(\rho)\mathbf{G}_i^{1/2}(\delta)$$

$$\mathbf{G}_i(\delta) = diag\left(\mathbf{I}_1 \quad \delta^2 \mathbf{I}_2\right)$$

$$\boldsymbol{\Gamma}(\rho) = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

In [8] the fixed effect $\beta_{\ell,\, k(k=1,2)}$ indicates parameter β specific to ℓ^{th} region specified in Table 3.2 for response variable SG (k=1) and for response variable MC (k=2). The random effect $b_{i,k(k=1,2)}$ indicates the random effect parameter specific to the i^{th} stand for response variable SG (k=1) and for response variable MC (k=2).

The model information criteria (AIC and BIC) and log likelihood values from simultaneous fitting of the models (**SGMC1**, **SGMC2** and **SGMC3**) are presented in Table 3.3. The log likelihood and information criteria from SGMC1 were equal to the sum of log likelihood

Table 3.3: Model information criteria and log likelihoods for univariate and bivariate models.

Model	**AIC**	**BIC**	**Log likelihood**
SG1	-9744.01	-9619.34	4895.01
MC1	11942.86	12029.59	-5955.43
SGMC1	2198.85	2437.28	-1060.42
SGMC2	1210.49	1455.04	-565.25
SGMC3	1106.06	1368.95	-510.03

and information criteria from univariate fitting SG1 and MC1. Incorporation of contemporaneous correlation into the model (SGMC2) significantly improved the model fitting criteria. The final model SGMC3 found to have a significant improvement in model information criteria over

SGMC2. The estimated fixed effect parameter from the final simultaneous model is presented in Table 3.4. The estimated random effect variance-covariance matrix $\boldsymbol{\Psi}$ is

$$\boldsymbol{\Psi} = \begin{pmatrix} 0.00074 & -0.00100 & 0.00058 & -0.23368 & 0 \\ & 0.00335 & -0.00257 & 0.32946 & 0 \\ & & 0.00241 & -0.20459 & 0 \\ & & & 121.92 & 0 \\ & & & & 284.19 \end{pmatrix}$$

and the within stand residual parameters are $\delta = 638.68$ and $\rho = -0.779$.

3.5 Prediction

Our primary objective of developing a simultaneous system is to make predictions. The reported advantage of using a multivariate method over univariate method is its improvement in predictive performance (Fang et al. 2001; Hall and Clutter 2004). The information on contemporaneous correlation among response variables can be potentially utilized to improve the prediction of a variable at a particular measurement occasion (here at a particular stand height level) given that the observed value of other response variables at the specified measurement occasion. For example in the proposed multivariate system, information of disk SG at any specific height can be utilized to improve the prediction of disk MC at that height. Similarly, observed disk MC at any specific stand height can be utilized to improve the prediction of disk SG at that height.

There are several situations where we can utilize a multivariate model to make predictions. Fang et al. (2001) dealt with several such prediction scenarios based on their height-basal area-volume simultaneous mixed model system. In the present study, we are primarily interested in prediction from a multivariate model system where observations on one of the correlated response variables are available. For example, we may want to predict disk MC for a stand at different heights when measurements of disk SG are available. To this extent, we can

Table 3.4: The estimated parameters for the fixed effects from the simultaneous model system.

Parameter	**Estimate**	**Std. Error**	**t-value**	**p-value**
$\beta_{0,1}$	0.4805	0.0037	130.34	<0.0000
$\beta_{0,1.R2}$	-0.0543	0.0087	-6.27	<0.0000
$\beta_{0,1.R3}$	-0.0248	0.0053	-4.68	<0.0000
$\beta_{0,1.R4}$	-0.0309	0.0045	-6.81	<0.0000
$\beta_{0,1.R6}$	-0.0215	0.0037	-5.84	<0.0000
$\beta_{1,1}$	-0.0706	0.0090	-7.85	<0.0000
$\beta_{2,1}$	-0.0258	0.0075	-3.45	0.0006
$\beta_{2,1.R2}$	0.0408	0.0098	4.18	<0.0000
$\beta_{3,1}$	0.9980	0.0985	10.14	<0.0000
$\beta_{3,1.R3}$	-0.1052	0.0473	-2.23	0.0261
$\beta_{3,1.R4}$	-0.1486	0.0422	-3.52	0.0004
$\beta_{3,1.R5}$	-0.1892	0.0584	-3.24	0.0012
$\beta_{3,1.R6}$	-0.1784	0.0489	-3.65	0.0003
$\beta_{4,1}$	-2.2348	1.1540	-1.94	0.0529
$\alpha_{1,1}$	0.2914	0.0112	26.05	<0.0000
$\alpha_{2,1}$	0.0849	0.0249	3.42	0.0006
$\beta_{0,2}$	102.2878	1.4256	71.75	<0.0000
$\beta_{0,2.R2}$	31.3519	4.6761	6.70	<0.0000
$\beta_{0,2.R3}$	9.1917	2.7728	3.32	0.0009
$\beta_{0,2.R4}$	13.7670	2.3638	5.82	<0.0000
$\beta_{0,2.R5}$	-13.5634	2.3857	-5.69	<0.0000
$\beta_{1,2}$	57.0750	2.0681	27.60	<0.0000
$\beta_{1,2.R2}$	-43.4926	7.6022	-5.72	<0.0000
$\beta_{1,2.R6}$	-8.2443	4.0510	-2.04	0.0419
$\beta_{2,2}$	-255.2719	26.2173	-9.74	<0.0000
$\beta_{2,2.R2}$	-142.2978	32.4391	-4.39	<0.0000
$\beta_{2,2.R5}$	133.4103	26.6937	5.00	<0.0000
$\beta_{2,2.R6}$	104.9782	21.6648	4.85	<0.0000
$\alpha_{1,2}$	0.3250	0.0162	20.02	<0.0000

utilize a predictor proposed by Hall and Clutter (2004) for NLMM's which is based on a linear mixed model (LMM) approximation of NLMM. The proposed predictor is analogous to the empirical best linear unbiased predictor (BLUP) of LMM. It is supposed to perform better than the plug-in-predictor proposed for NLMM by Pinheiro and Bates (2000). The following on the derivation of a predictor was extracted from Hall and Clutter (2004). Generically a NLMM can be represent as

[9] $$\mathbf{y} = \mathbf{f}\left(\beta, \mathrm{b}, \mathbf{A}, \mathbf{B}\right) + \varepsilon$$

where β is p x 1 vector of fixed effect parameters and $\mathbf{A}$ is a corresponding fixed effect design matrix; b is q x 1 vector of random effect parameters and $\mathbf{B}$ is a corresponding random effect design matrix; and ε is N x 1 vector of error term with $\varepsilon \overset{i.i.d.}{\sim} N\left(0,\ \sigma^2\Lambda\right)$. Taking first-order Taylor series linearization of Eq. [9] around the estimates of $\left(\beta, \mathrm{b}\right) = \left(\hat{\beta}, \hat{\mathrm{b}}\right)$ gives

[10] $$\mathbf{y} \approx \mathbf{f}\left(\hat{\beta}, \hat{\mathrm{b}}, \mathbf{A}, \mathbf{B}\right) + \tilde{\mathbf{A}}\left(\beta - \hat{\beta}\right) + \tilde{\mathbf{B}}\left(\mathrm{b} - \hat{\mathrm{b}}\right) + \varepsilon$$

where

$$\tilde{\mathbf{A}} = \left.\frac{\partial \mathbf{f}\left(\beta, \mathrm{b}, \mathbf{A}, \mathbf{B}\right)}{\partial \beta}\right|_{\beta=\hat{\beta}, \mathrm{b}=\hat{\mathrm{b}}}, \quad \tilde{\mathbf{B}} = \left.\frac{\partial \mathbf{f}\left(\beta, \mathrm{b}, \mathbf{A}, \mathbf{B}\right)}{\partial \mathrm{b}}\right|_{\beta=\hat{\beta}, \mathrm{b}=\hat{\mathrm{b}}}$$

Now the Eq. 10 can be represented as a LMM on $\mathbf{z} = \mathbf{y} - \mathbf{f}\left(\hat{\beta}, \hat{\mathrm{b}}, \mathbf{A}, \mathbf{B}\right) + \tilde{\mathbf{A}}\hat{\beta} + \tilde{\mathbf{B}}\hat{\mathrm{b}}$ as follows

[11] $$\mathbf{z} = \tilde{\mathbf{A}}\beta + \tilde{\mathbf{B}}\mathrm{b} + \varepsilon$$

Let us decompose the response vector $\mathbf{y} = \left(\mathbf{y}_s^T,\ \mathbf{y}_h^T\right)$, where $\mathbf{y}_s$ represents the observed component and $\mathbf{y}_h$ represents the unobserved component. Accordingly, all other model quantities can be divided as

$$\mathbf{z} = \begin{pmatrix} \mathbf{z}_s \\ \mathbf{z}_h \end{pmatrix}; \qquad \tilde{\mathbf{A}} = \begin{pmatrix} \tilde{\mathbf{A}}_s \\ \tilde{\mathbf{A}}_h \end{pmatrix}; \qquad \tilde{\mathbf{B}} = \begin{pmatrix} \tilde{\mathbf{B}}_s \\ \tilde{\mathbf{B}}_h \end{pmatrix};$$

$$\mathbf{f}\left(\hat{\beta}, \hat{\mathrm{b}}, \mathbf{A}, \mathbf{B}\right) = \begin{pmatrix} \mathbf{f}_s\left(\hat{\beta}, \hat{\mathrm{b}}, \mathbf{A}_s, \mathbf{B}_s\right) \\ \mathbf{f}_h\left(\hat{\beta}, \hat{\mathrm{b}}, \mathbf{A}_h, \mathbf{B}_h\right) \end{pmatrix}$$

Then based on LMM [11], the empirical BLUP of $\mathbf{z}_h$ based on $\mathbf{z}_s$ is given as

[12]
$$\mathbf{z}_h = \tilde{\mathbf{A}}_h\hat{\beta} + \tilde{\mathbf{V}}_{hs}\tilde{\mathbf{V}}_{ss}^{-1}\left(\mathbf{z}_s - \tilde{\mathbf{A}}_s\hat{\beta}\right)$$

where $\tilde{\mathbf{V}} = \tilde{\mathbf{B}}\,\hat{\mathrm{var}}\left(\mathrm{b}\right)\tilde{\mathbf{B}}^{\mathbf{T}} + \hat{\mathrm{var}}\left(\varepsilon\right)$, the variance-covariance matrix of $\mathbf{z}$ based on LMM approximation [11], which can be decomposed into

$$\tilde{\mathbf{V}} = \begin{pmatrix} \tilde{\mathbf{V}}_{ss} & \tilde{\mathbf{V}}_{sh} \\ \tilde{\mathbf{V}}_{hs} & \tilde{\mathbf{V}}_{hh} \end{pmatrix}$$

By rearranging [12] using the relation between $\mathbf{z}$ and $\mathbf{y}$, we will get our predictor for $\mathbf{y}_h$ as

[13]
$$\begin{aligned} \hat{\mathbf{y}}_h = {} & \mathbf{f}_h\left(\hat{\beta}, \hat{\mathrm{b}}, \mathbf{A}_h, \mathbf{B}_h\right) - \tilde{\mathbf{B}}_h\hat{\mathrm{b}} \\ & + \tilde{\mathbf{V}}_{hs}\tilde{\mathbf{V}}_{ss}^{-1}\left\{\mathbf{y}_s - \mathbf{f}_s\left(\hat{\beta}, \hat{\mathrm{b}}, \mathbf{A}_s, \mathbf{B}_s\right) + \tilde{\mathbf{B}}_s\hat{\mathrm{b}}\right\} \end{aligned}$$

When $\mathrm{cov}\left(\varepsilon_s, \varepsilon_h\right) \neq 0$, the predictor specified in Eq. [13] takes account of this dependence through $\tilde{\mathbf{V}}_{hs}$. However when $\mathrm{cov}\left(\varepsilon_s, \varepsilon_h\right) = 0$, $\mathbf{y}_h$ and $\mathbf{y}_s$ are correlated only through the shared random effects and is best approximated by the plug-in-predictor $\hat{\mathbf{y}}_h = \mathbf{f}_h\left(\hat{\beta}, \hat{\mathrm{b}}, \mathbf{A}_h, \mathbf{B}_h\right)$. Since we are interested in predicting the value of one response variable using data where another response variable is available or measured at the same height from the same stand, we expect that the predictor [13] performs better than the plug-in-predictor.

In order to evaluate the predictive performance of the fitted multivariate model, we randomly selected data from 25 stands. We created a new data set with data from the 25 selected

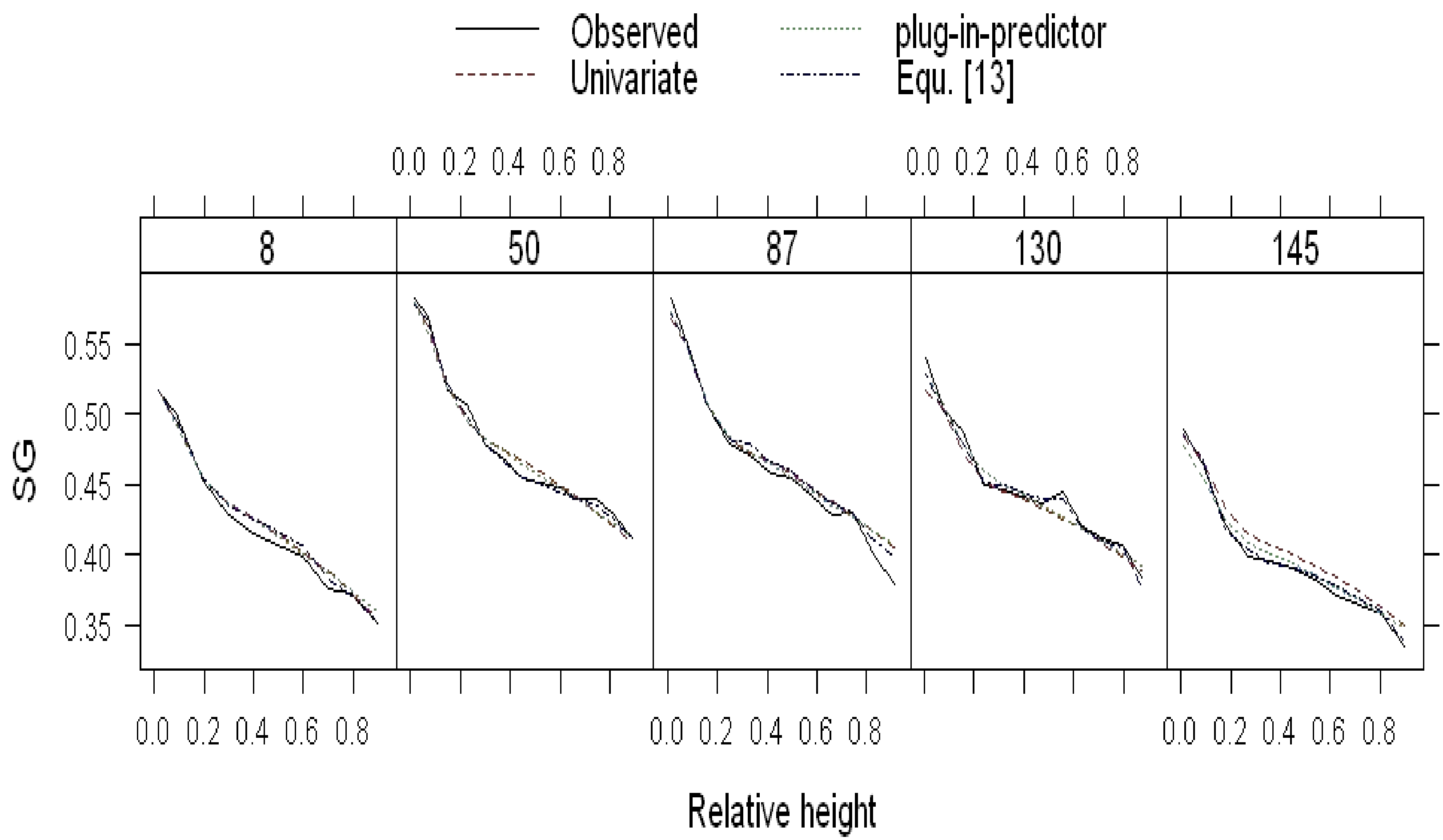

Figure 3.3: Plot showing observed stand level disk specific gravity with relative height along with predictions of specific gravity using different predictors for 5 selected stands.

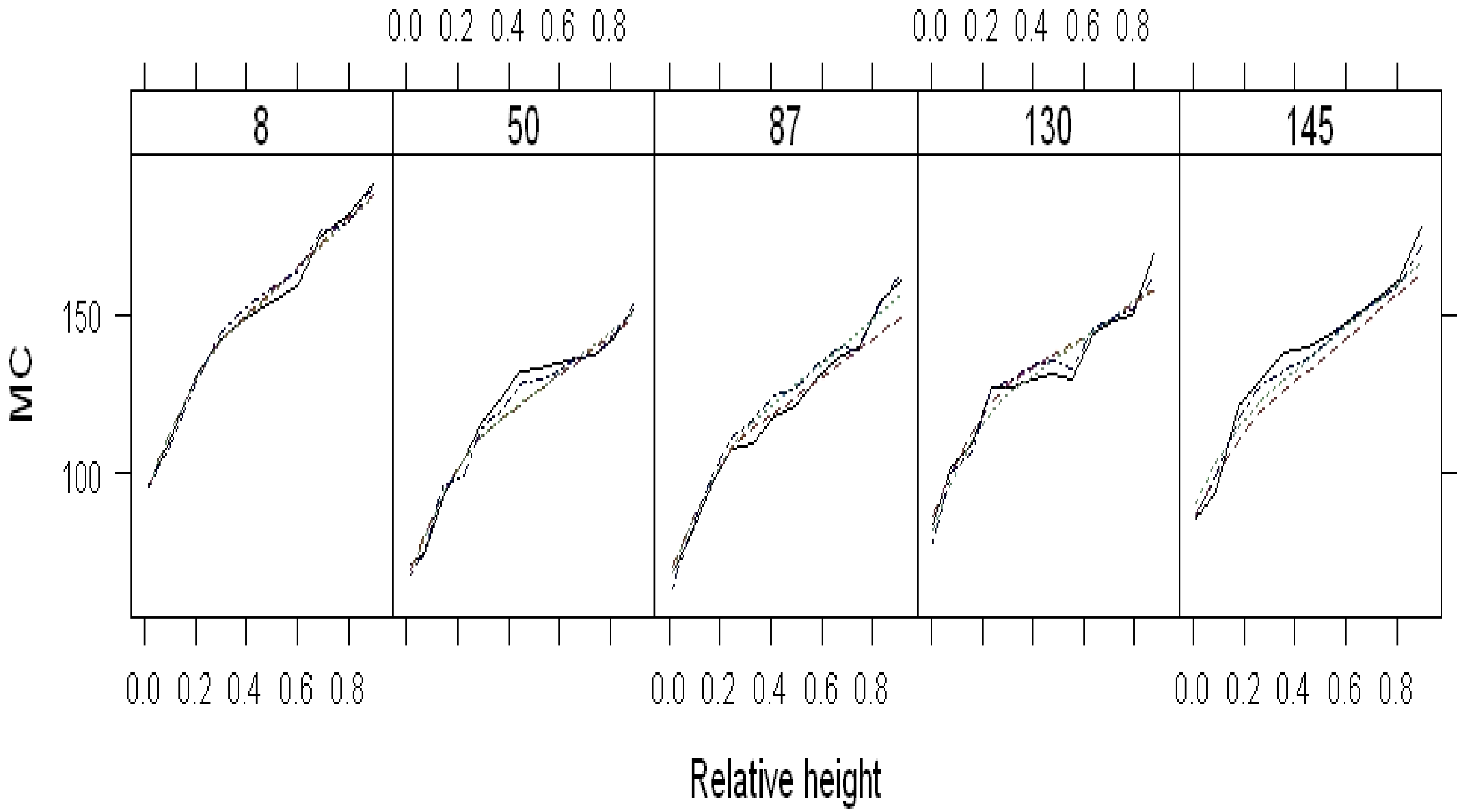

Figure 3.4: Plot showing observed stand level disk moisture content with relative height along with predictions of moisture content using different predictors for 5 selected stands. Legend same as described in Figure 3.3.

stands excluded (apart from data measured at relative heights equivalent to heights of 1.37 m and 13.7 m to get the estimate of random effect while fitting) and refitted the final model SGMC3 to this new data. We made predictions based on [13] for both disk SG and MC for the selected 25 stands that were not used for model fitting. Disk SG was predicted for the 25 excluded stands assuming that disk MC measurements were available for all heights and stands. The same assumption was made for disk SG when disk MC was predicted for the excluded stands.

Plots showing the univariate plug-in-prediction, multivariate plug-in-prediction and multivariate improved prediction (based on Eq. [13]) of disk SG and MC for 5 stands randomly selected from the excluded 25 are presented in Figure 3.3 and 3.4. We can see from the figures that additional information for one response variable significantly improved the prediction of the other response variable using Eq. [13] compared to the plug-in-predictors. The curves are closer to their observed values for both disk SG and MC using the Eq. [13] predictor. Table 3.5, presents the root mean square prediction error (RMSPE) for the three prediction methods based on predictions of SG and MC for trees from the 25 excluded stands. Prediction from multivariate approaches, both plug-in-predictor and Eq. [13], was considerably better than those of the univariate approach. Prediction based in Eq. [13] were improved by 29 (SG) and 26 % (MC) (Table 3.5).

Table 3.5: Root mean square prediction error (RMSPE) from univariate approach, multivariate plug-in-prediction and multivariate prediction using Eq. [13].

Response	**univariate**	**plug-in-prediction**	**Eq. [13]**
SG	0.0118	0.0112	0.0083
MC	7.70	6.86	5.71

3.6 Discussion

Nonlinear mixed models are an important tool for modeling and predicting growth and wood quality attributes in forestry (Fang 1999; Hall and Bailey 2001; Jordan et al. 2008; Jordan et al. 2006). Univariate mixed models were commonly used in forestry to model different growth and wood properties. Compared to conventional methods univariate mixed models provide improved predictions because of their ability to capture different levels of variability within the data, e.g. variability from stand-to-stand, plot-to-plot and tree-to-tree (Fang et al. 2001) through random effects in the models. In addition to variability observed at different levels of the data, individual components (properties) measured from a forest are usually inter-dependent. The simultaneous modeling technique can take account of the inter-dependency in a system through random effects and the inter-dependency among different components in the system through contemporaneous correlation.

In this article, we proposed a multivariate simultaneous mixed model for stand average disk SG and MC at different tree heights. We observed a high correlation (-0.78) between two components in our system. The inverse relation between SG and MC was identified by Koch (1972), Zobel and Blair (1976) and Zobel and van Buijtenen (1989). Various explanations have been proposed for the inverse relation between SG and MC within trees such as the amount of heartwood, the presence of extractives and the proportion of juvenile wood. According to Zobel and Blair (1976), the dominant factor controlling SG and MC variation within a loblolly pine tree is the proportion of juvenile wood and the proportion of juvenile wood increases longitudinally from stump-to-tip of loblolly pine trees.

The advantage of multivariate simultaneous systems is their improvement in prediction in one component given the other components in the system (Fang et al. 2001; Hall and Clutter

2004). Based on this study, we found a significant improvement in prediction for both properties, approximately 29 and 26 % reduction in RMSPE for both disk SG and MC respectively, based on the simultaneous system after taking account of the contemporaneous correlation between the components. The multivariate plug-in-predictor improved by 5 and 11 % in RMSPE compared to univariate approach for both disk SG and MC respectively. This clearly indicates the potential of multivariate model fitting over univariate approach. Operationally, the proposed system can be used to improve the prediction of stand disk SG at different height levels using the measured disk MC using non-destructive sampling methods.

3.7 Acknowledgements

Thanks are due to the Wood Quality Consortium, University of Georgia for providing the data for this study. Thanks also to the USDA Forest Service Southern Research Station, Athens, GA for preparing the samples and providing the lab facilities.

3.8 References

Borders, B.E. 1989. System of equation in forest stand modeling. Forest Science 35:548-556.

Clutter, J.L. 1963. Compatible growth and yield models for loblolly pine. Forest Science 9(3):354-370.

Fang, Z. 1999. A simultaneous system of linear and nonlinear mixed effects models for forest growth and yield prediction (Ph.D dissertation), University of Georgia, Athens. 177 p.

Fang, Z., R.L. Bailey, and B.D. Shiver. 2001. A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects. Forest Science 47(4):550-562.

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

Hall, D.B., and M. Clutter. 2004. Multivariate multilevel nonlinear mixed effects models for timber yield predictions. Biometrics 60(1):16-24.

Haygreen, J.G., and J.L. Bowyer. 1996. *Forest Products and Wood Science: An Introduction. Third Edition*. Ames: Iowa State University Press, Iowa. pp. 80-102.

He, R. 2004. Mixed effects modeling of wood properties of loblolly pine in the southeastern United States (Ph.D dissertation), University of Georgia, Athens. 213 p.

Jordan, L., A. Clark, L.R. Schimleck, D.B. Hall, and R.F. Daniels. 2008. Regional variation in wood specific gravity of planted loblolly pine in the United States. Canadian Journal of Forest Research 38:698-710.

Jordan, L., R. He, D.B. Hall, A. Clark, and R.F. Daniels. 2006. Variation in loblolly pine cross-sectional microfibril angle with tree height and physiographic region Wood and Fiber Science 38(3):390-398.

Koch, P. 1972. *Utilization of the Southern Pines*. USDA Forest Service, Southern Forest Experiment Station, Agriculture Handbook No. 420. 235-260.

Max, T.A., and H.E. Burkhart. 1976. Segmented polynomial regression applied to taper equations. Forest Science 22:283-289.

Megraw, R.A. 1985. Wood quality factors in loblolly pine. TAPPI Press, Atlanta, GA.

Panshin, A.J., and C. deZeeuw. 1980. *Textbook of Wood Technology*. McGraw-Hill Book Co., New York.705 p.

Phillips, K.M. 2002. Modeling within tree changes in wood specific gravity and moisture content for loblolly pine in Georgia (M.S. Thesis), University of Georgia, Athens. 95 p.

Pinheiro, C.J., and D.M. Bates. 2000. Mixed effects models in S and S-plus. Springer, New York. 523 p.

Pinheiro, C.J., M.D. Bates, S. DebRoy, D. Sarkar, and R.C. team. 2009. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3. 1-92.

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

Zobel, B.J., and J.P. vanBuijtenen. 1989. Wood Variation: Its Causes and Control. Springer-Verlag, Berlin. 363 p.

Chapter 4

Generalized Algebraic Difference Approach and Nonlinear Mixed Effect Models – A comparison using disk specific gravity data from loblolly pine

4.1 Abstract

The Generalized Algebraic Difference Approach (GADA) and nonlinear mixed models (NLMM) are two modeling tools used to deal with longitudinal or re-measurement data collected in forestry. The present study was an attempt to evaluate the predictive performance of these two approaches using disk specific gravity (SG) data collected from stump-to-tip of 81 loblolly pine trees. Two base models were proposed and for each, their GADA and NLMM forms were used to fit to the data. The M-fold cross validation technique was used to assess predictive performance of the two approaches where SG data from one tree was retained to validate predictions made using models based on the remaining trees. The predictive performance of models was assessed by using 1, 2, 3, 4 or 5 observed SG-height pairs as prior information available for the estimation of subject specific effects. Root mean square error (RMSE), mean absolute residual (MAR) and mean residual (MR) were used to evaluate the performance of the models. It was found that the NLMM consistently performed better than the GADA methodology in terms of both RMSE and MAR. A 3-19 % improvement in RMSE and MAR was observed for the NLMM approach compared to the GADA approach with improvement varying with the number of prior observations used to estimate the subject specific effect. RMSE and MAR for GADA and NLMM decreased considerably as the number of data points used for estimating the subject specific parameter increased from 1 to 5.

Key words: Nonlinear mixed effect models, generalize algebraic difference approach, prediction.

4.2 Introduction

The southern region of the United States (US) supplies approximately 58 % of the total wood used in the USA and 16 % of all wood supplied to the world timber market (Wear and Greis 2002). Southern pines contribute the majority of wood supply from this region and occupy an area of approximately 13 million hectares (Fox et al. 2007). Understanding the growth of these stands is of critical importance for foresters. Mathematical models for accurate estimation of growth and yield of even-aged pine plantations have received considerable attention in recent decades (Calegario et al. 2005; Jordan et al. 2005; Zhang and Borders 2004; Zhang et al. 2002). Apart from growth and yield estimation, the forest product industries are also concerned about the quality of the wood produced from these plantations, yet modeling of wood property variation has received little attention.

Forests are dynamic systems with data being collected continuously over time to make inferences about the system and also to make predictions of the future. Since the measurements are often taken from permanent sample plots or from individual trees repeatedly over a given time interval (longitudinal measurements), the data from an individual plot/tree will be correlated i.e. multiple measurements taken from a single stand or plot or individual will not be independent (Berhe 2009; Lindstrom and Bates 1990). The generalized algebraic difference approach (GADA) and nonlinear mixed models (NLMM) are two common methods used by foresters to deal with longitudinal data.

The GADA methodology was first illustrated and used in the forestry literature by Cieszewski and Bailey (2000). Conceptually, GADA expands a two dimensional functional relation between the response *y*, to a longitudinal explanatory variable *t* (usually in forestry data collected at different ages or across height as in this study), to a three dimensional functional relation between *y*, *t* and a new variable χ, which is defined as "the growth intensity factor" by Cieszewki and Bailey (2000). According to Cieszewki and Bailey (2000), χ is ***a variable or variables*** which are "continuous, monotonic and relevant to the modeled dynamics" and used to describe the change in curve shape with site/subject productivity. A simple notational illustration of GADA is as follows

[1]
$$y = f\left(\overbrace{\{\theta_1,\ldots., \theta_j, \theta_k,\ldots., \theta_p\}}^{\boldsymbol{\theta}}, t\right) + \varepsilon$$

where the response *y* is measured longitudinally across *t* points and is expressed as a function of the parameter vectors $\boldsymbol{\theta}$ and t; ε is the error term with $\varepsilon \sim N\left(0, \sigma^2\right)$. If we assume two of the parameters θ_j and θ_k in the base function (Eq. 1) as site/subject specific, i.e. these parameters vary from site-to-site (subject-to-subject). Following Cieszewski and Bailey (2000), each of these site/subject specific parameters can be represented as functions of new parameter vectors, $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$, and the continuous unobserved theoretical variable χ as $\theta_j = g_j(\boldsymbol{\gamma}, \chi)$ and $\theta_k = g_k(\boldsymbol{\beta}, \chi)$. After substituting the relation between site variable χ with site specific parameter to [1] we have

[2]
$$y = f\left(\theta_1,\ldots.,\theta_{j-1}, g_j(\boldsymbol{\gamma}, \chi), g_k(\boldsymbol{\beta}, \chi),\theta_{k+1}\ldots., \theta_p, t\right) + \varepsilon$$

The new model [2] has *p*-2 site/subject independent parameters plus the new parameters introduced in the model through parameter vectors $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ (from functions used to define the relation between site/subject productivity and the site/subject specific parameters θ_j and θ_k) with two independent variables, the observed variable t and the unobserved variable χ. Since χ is an unobserved variable, a solution for χ using initial conditions $(t_0,\ y_0)$ is usually substituted into Eq. [2] (Cieszewski 2002; Cieszewski and Bailey 2000; Cieszewski and Strub 2008). GADA methodology is not restricted to two subject specific parameters as illustrated above and can extend to any desired number of subject specific parameters from a base model. The algebraic difference approach (ADA) proposed by Bailey and Clutter (1974) is a simplified version of GADA where instead of multiple parameters, one parameter is allowed to vary with respect to site/subject productivity χ. Both ADA and GADA have been widely applied in forestry literature for the last few decades (Bailey and Clutter 1974; Borders et al. 1984; Cieszewski 2003; Cieszewski and Strub 2008).

NLMM have long been used in the statistical modeling of repeated correlated measurements in variety of scientific fields, including forestry (Berhe 2009; Calegario et al. 2005; Davidian and Giltinan 2003; Fang 1999; Hall and Bailey 2001; Hall and Clutter 2004). The general framework for NLMM can be represented following Lindstrom and Bates (1990) and Davidian and Gallant (1995). Let y_{ij} be the j^{th} response variable measured from i^{th} subject/site at time t_{ij} then:

[3] $$y_{ij} = f\left(\boldsymbol{\theta}_i,\ t_{ij}\right) + \varepsilon_{ij}$$

where *f* is a real-value, differentiable function of a subject-specific parameter vector $\boldsymbol{\theta}_i$ and a covariate vector t_{ij}, and $\varepsilon_{ij} \sim N\left(0,\ \sigma^2\right)$. The parameter vector $\boldsymbol{\theta}_i$ has the form

[4] $$\boldsymbol{\theta}_i = \mathbf{A}_i\boldsymbol{\beta} + \mathbf{B}_i\mathbf{b}_i$$

where $\boldsymbol{\beta}$ is a ($p \times 1$) vector of fixed effects, $\mathbf{b}_i$ is a ($q \times 1$) vector of random effects associated with the i^{th} series, and $\mathbf{A}_i$ and $\mathbf{B}_i$ are the fixed effect and random effect design matrices, $\mathbf{b}_i$'s are site/subject specific random effects or site/subject specific deviations from the fixed effect with $\mathbf{b}_i \sim N(0, \mathbf{\Psi})$. Similar to χ (the unobserved variable representing site productivity) in GADA, the random effect explains the unexplained variation in the curve shapes through the deviation in parameters from site-to-site (subject-to-subject).

Both GADA and NLMM have been used to model specific gravity (SG) for loblolly pine (*Pinus taeda* L.). Phillips (2002) used ADA models to explain SG changes within trees as a function of relative height. Jordan et al. (2006) developed a self referencing function using ADA methodology to estimate SG changes from stump-to-tip in slash pine (*Pinus elliottii* Engelmann) and successfully incorporated the model into a system of equations to estimate the biomass of trees. NLMMs were also used to model the within tree changes of disk SG. He (2004) developed a NLMM version of the base model used by Phillips (2002) to model disk SG changes within trees.

Foresters rely on statistical models/equations to make predictions of the future which allows them to evaluate future conditions of the forest resource and manage the resource successfully. In the present context, predicting disk SG at different heights within a tree, given prior measurements from the same tree or similar tree, enables foresters to optimize product categorization and utilization of wood and thus maximize profit from forest products. Both GADA and NLMM have been used to make predictions in forestry, however, which method performs better is debatable. The aim of this study is to compare the predictions of disk SG given previous observed data with models constructed using the GADA and NLMM methodologies.

4.3 Materials and Methods

4.3.1 Data

Data from 81 loblolly pine trees sampled from 27 stands across the Atlantic Coastal Plain of the southern USA were utilized. The stands selected were 20- to 25-year-old loblolly pine planted at 1250 or more trees per hectare and had 625 or more trees per hectare or more after thinning. Unfertilized stands (except possibly P at planting) with no competition control were sampled. Three trees from each stand were felled and 3.8 cm thick disks were collected from 0.15 m, 1.37 m, 3.05 m and at 1.52 m intervals along the stem to a diameter outside bark of 5.08 cm. The disks were sealed in plastic zip-lock bags and shipped to the USDA Forest Service laboratory in Athens, GA for SG analysis. Disk SG based on green volume and oven dry weight were determined for each disk. The original data were truncated such that all data collected below 1.37 m (4.5 feet) height were removed. A plot of the disk SG data by relative height is presented in Figure 4.1.

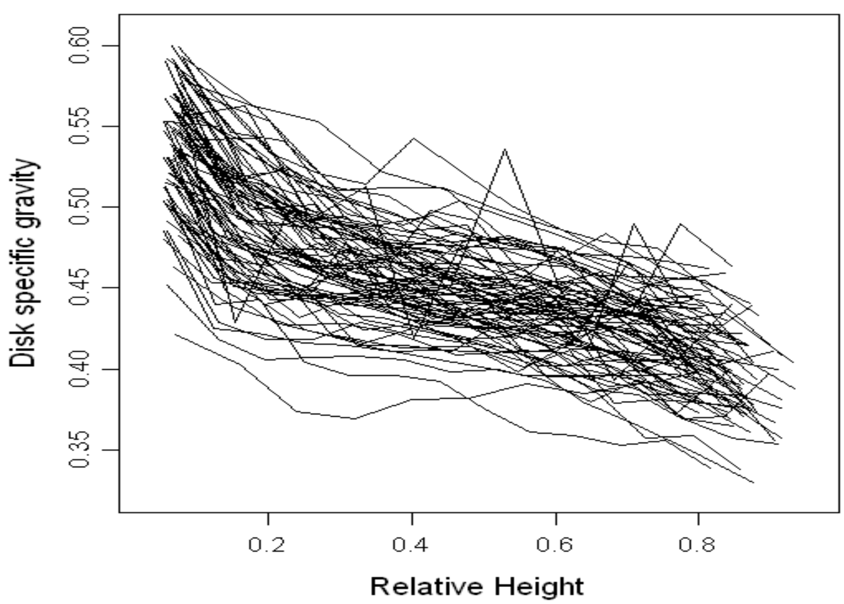

Figure 4.1: Plot of the data used in the study

4.3.2 Model derivation

Two base models, each with three parameters, were used in this study to illustrate the predictive performance of GADA and NLMM methodology. The nonlinear functions used were of the following base form

[1a] $$\mathbf{Y}_i = \beta_0 \exp\left(\frac{\beta_1}{\mathbf{x}_i + \beta_2}\right) + \varepsilon_i$$

[1b] $$\mathbf{Y}_i = e^{\beta_0}\left(\mathbf{x}_i^{-\beta_1} - 1\right)^{\beta_2} + \varepsilon_i$$

$$\varepsilon_i \sim N\left(0,\ \sigma^2\right)$$

$$i = 1, 2, ...M$$

The scripts *a* and *b* along with the equation number (Eq. [1a] and [1b]), here after, are used to represents different formulations of both models. Let $\mathbf{Y}_i = \left(y_{i1}, y_{i2}, ..., y_{in_i}\right)$ = the n_i-dimensional vector of the i^{th} tree's SG measurements; $\mathbf{x}_i = h_i / H_i$ is the relative height vector of the i^{th} tree, where h_i is the vector of above ground heights for the i^{th} tree having a total height of H_i; $\boldsymbol{\beta} = \left(\beta_0\ \beta_1\ \beta_2\right)^T$ is a vector of parameters to be estimated from the data. The equations [1a] and [1b] are referred to as the nonlinear least squares (NLS) form of the models and represent the mean trend in the data ignoring any variation in subject specific curves. Since NLS models 1a and 1b do not utilize any available information for making subject specific prediction, the equation for predicting SG of the k^{th} disk from the j^{th} tree from both models [1a] and [1b] will be

[2a] $$\hat{y}_{jk} = \hat{\beta}_0 \exp\left(\frac{\hat{\beta}_1}{x_{jk} + \hat{\beta}_2}\right)$$

[2b] $$\hat{y}_{jk} = e^{\hat{\beta}_0}\left(x_{jk}^{-\hat{\beta}_1} - 1\right)^{\hat{\beta}_2}$$

The Eq. [1b] with site/subject specific parameters β_{0i} in it can be represented as:

[3b] $$\mathbf{Y}_i = e^{\beta_{0i}}\left(\mathbf{x}_i^{-\beta_1} - 1\right)^{\beta_2} + \varepsilon_i$$

Let $\chi_i = \beta_{0i}$, the parameter representing the intensity of the modeled process (Cieszewski and Bailey 2000), then the ADA version of the above model [3b] is

[4b] $$\mathbf{Y}_i = e^{\chi_i}\left(\mathbf{x}_i^{-\beta_1} - 1\right)^{\beta_2} + \varepsilon_i$$

where χ_i is the subject specific parameter (or null parameter) and β_1 and β_2 represents the global parameter in the model. Generally model [4b] is defined in terms of the expected value parameterization of χ_i given $\left(y_{0i}, x_{0i}\right)$ in site index modeling literature. The solution for χ_i given $\left(y_{0i}, x_{0i}\right)$ from model [4b] is

[5b] $$\chi_i = \ln\left(\frac{y_{0i}}{\left(x_{0i}^{-\beta_1} - 1\right)^{\gamma}}\right)$$

The model [1a] with subject specific parameters β_{1i} and β_{2i} and model [1b] with subject specific parameters β_{0i} and β_{2i} can be represented as:

[6a] $$\mathbf{Y}_i = \beta_0 \exp\left(\frac{\beta_{1i}}{\mathbf{x}_i + \beta_{2i}}\right) + \varepsilon_i$$

[6b] $$\mathbf{Y}_i = e^{\beta_{0i}}\left(\mathbf{x}_i^{-\beta_1} - 1\right)^{\beta_{2i}} + \varepsilon_i$$

Let $\chi_i = \beta_{1i}$ in Eq. [6a], the derived GADA model is

[7a] $$\mathbf{Y}_i = f_i\left(\mathbf{x}_i, \boldsymbol{\beta}, \chi_i\right) = \beta_0 \exp\left(\frac{\chi_i}{\mathbf{x}_i + \gamma_1 + \gamma_2\chi_i}\right) + \varepsilon_i$$

Let $\chi_i = \beta_{0i}$ in Eq. [6b], the derived GADA model is

[7b] $$\mathbf{Y}_i = f_i\left(\mathbf{x}_i,\ \boldsymbol{\beta}, \chi_i\right) = e^{\chi_i}\left(\mathbf{x}_i^{-\beta_1} - 1\right)^{\gamma\ \chi_i} + \varepsilon_i$$

where $\beta_{2i} = \gamma_1 + \gamma_2\chi_i$ in Eq. [7a] and $\beta_{2i} = \gamma\ \chi_i$ in Eq. [7b], a linear relation between subject specific parameters; χ_i 's are subject specific fixed effects known as 'local parameters' in the site index modeling literature; $\boldsymbol{\beta} = \left(\beta_0\ \gamma_1\ \gamma_2\right)^T$ in model [7a] and $\boldsymbol{\beta} = \left(\beta_0\ \gamma\right)^T$ in model [7b] are fixed effect parameters known as 'global parameters' common to all subjects in the data. Both model [7a] and [7b] can also defined in terms of expected value parameterization of χ_i given $\left(y_{0i}, x_0\right)$ as in site index modeling literature. The solution for χ_i given $\left(y_{0i}, x_{0i}\right)$ from both models is:

[8a] $$\chi_i = \frac{\ln\left(\dfrac{y_{0i}}{\beta_0}\right)(x_{0i} + \gamma_1)}{1 - \ln\left(\dfrac{y_{0i}}{\beta_0}\right)\gamma_2}$$

[8b] $$\chi_i = \frac{\ln\left(y_{0i}\right)}{1 + \gamma\ln\left(x_{0i}^{-\beta_1} - 1\right)}$$

where in [5b], [8a] and [8b] y_{0i} is the expected value of response, here disk SG, observed at a fixed height x_{0i} from the i^{th} tree. Usually the solution for χ_i is substituted in to the models [4b], [7a] and [7b] and estimates y_{0i} as subject specific fixed effects. The shape of the curves defined by non-differenced form of the models ([4b] and [7a] and [7b]) and differenced form of models (models derived from substituting solutions [5b], [8a] and [8b] in to model [4b], [7a] and [7b]) were exactly the same. So, we used the non-differenced form of the models for parameter estimation (Wang et al. 2008b). Here, χ_i was allowed to vary from subject to subject and was considered as subject specific fixed effects.

Both ADA and GADA models use prior available information about the subject to make future predictions. Conventionally, ADA and GADA models were devised to make predictions for an individual subject where one prior observation is available. If a single observation is available for the j^{th} subject at x_{0j}, $\left(y_{0j}, x_{0j}\right)$, then the subject specific effect χ_j can be estimated in ADA as

[9b] $$\hat{\chi}_j = \ln\left(\frac{y_{0j}}{\left(x_{0j}{}^{-\beta_1} - 1\right)^{\gamma}}\right)$$

and for the GADA models as

[10a] $$\hat{\chi}_j = \frac{\ln\left(\frac{y_{0j}}{\beta_0}\right)(x_{0j} + \gamma_1)}{1 - \ln\left(\frac{y_{0j}}{\beta_0}\right)\gamma_2}$$

[10b] $$\hat{\chi}_j = \frac{\ln\left(y_{0j}\right)}{1 + \gamma \ln\left(x_{0j}{}^{-\beta_1} - 1\right)}$$

Based on Stewart et al. (2010) and Jordan et al. (2010), predictions from ADA and GADA models were considerably improved by using more than one observation. In cases where more than one observation is available, a least square method proposed by Wang et al (2008a) and Stewart et al. (2010) was followed to estimate the subject specific effect χ_j. Let $\left(\mathbf{y}_{0j}, \mathbf{x}_{0j}\right)$ be a vector of prior observations available for j^{th} subject, then the best solution for χ_j will be the one which minimizes the sum of square error evaluated at $\hat{\boldsymbol{\beta}}$ and given as:

[11] $$\mathrm{SSE}(\chi_j) = \sum_{k=1}^{n}[\mathbf{y}_j - f_j(\hat{\boldsymbol{\beta}}, \chi_j, \mathbf{x}_j)]^2$$

A best linear unbiased estimate for χ_j satisfies the following criteria as

[12] $$0 = \sum_{k=1}^{n}[\mathbf{y}_j - f_j(\hat{\boldsymbol{\beta}}, \chi_j, \mathbf{x}_j)]\frac{\partial f_j(\hat{\boldsymbol{\beta}}, \chi_j, \mathbf{x}_j)}{\partial \chi_j}$$

This is a simple nonlinear least square solution for χ_j. The solution for $\hat{\chi}_j$ can be used to make predictions for the k^{th} disk from the j^{th} tree from ADA (Eq. [4b]) as

[11b] $$\hat{y}_{jk} = f_{jk}\left(x_{jk}, \hat{\boldsymbol{\beta}}, \hat{\chi}_j\right) = \exp\left(\hat{\chi}_j\right)\left(x_{jk}^{-\hat{\beta}_1} - 1\right)^{\hat{\beta}_2}$$

and from GADA models (Eq. [7a] and [7b]) as

[12a] $$\hat{y}_{jk} = f_{jk}\left(x_{jk}, \hat{\boldsymbol{\beta}}, \hat{\chi}_j\right) = \hat{\beta}_0 \exp\left(\frac{\hat{\chi}_j}{x_{jk} + \hat{\gamma}_1 + \hat{\gamma}_2\hat{\chi}_j}\right)$$

[12b] $$\hat{y}_{jk} = f_{jk}\left(x_{jk}, \hat{\boldsymbol{\beta}}, \hat{\chi}_j\right) = \exp\left(\hat{\chi}_j\right)\left(x_{jk}^{-\hat{\beta}_1} - 1\right)^{\hat{\gamma}\,\hat{\chi}_j}$$

A general nonlinear mixed model (NLMM) formulation of Eq. [1a] and [1b] can be written as

[13] $$\mathbf{Y}_i = \mathbf{f}_i\left(\boldsymbol{\theta}_i, \mathbf{x}_i\right) + \varepsilon_i$$

where $\mathbf{f}_i$ is a real valued differential function of mixed effect parameter $\boldsymbol{\theta}_i$ and covariate $\mathbf{x}_i$. The mixed effect parameter $\boldsymbol{\theta}_i$ can be represented as

[14] $$\boldsymbol{\theta}_i = \mathbf{A}_i\boldsymbol{\beta} + \mathbf{B}_i\mathbf{b}_i$$

where,

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

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

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

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

Following the above formulation, the best NLMM's used in this study can be written as

$$\mathbf{Y}_i = \theta_{0i} \exp\left(\frac{\theta_{1i}}{x_i + \theta_{2i}} \right) + \varepsilon_i$$

[15a]

$$\underbrace{\begin{bmatrix} \theta_{0i} \\ \theta_{1i} \\ \theta_{2i} \end{bmatrix}}_{\boldsymbol{\theta}_i} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{A}_i} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{B}_i} \underbrace{\begin{bmatrix} b_{0i} \\ b_{2i} \end{bmatrix}}_{\mathbf{b}_i} = \begin{bmatrix} \beta_0 + b_{0i} \\ \beta_1 \\ \beta_2 + b_{2i} \end{bmatrix}$$

$$\mathbf{b}_i \sim N\left(0, \begin{pmatrix} D_{00} & D_{02} \\ D_{02} & D_{22} \end{pmatrix} \right); \ \varepsilon_i \sim N\left(0, \ \sigma^2\right)$$

$$\mathbf{Y}_i = e^{\theta_{0i}} \left(\mathbf{x}_i^{-\theta_{1i}} - 1 \right)^{\theta_{2i}} + \varepsilon_i$$

[15b]

$$\underbrace{\begin{bmatrix} \theta_{0i} \\ \theta_{1i} \\ \theta_{2i} \end{bmatrix}}_{\boldsymbol{\theta}_i} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{A}_i} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{B}_i} \underbrace{\begin{bmatrix} b_{0i} \\ b_{2i} \end{bmatrix}}_{\mathbf{b}_i} = \begin{bmatrix} \beta_0 + b_{0i} \\ \beta_1 \\ \beta_2 + b_{2i} \end{bmatrix}$$

$$\mathbf{b}_i \sim N\left(0, \begin{pmatrix} D_{00} & 0 \\ 0 & D_{22} \end{pmatrix} \right); \ \varepsilon_i \sim N\left(0, \ \sigma^2\right)$$

An unstructured variance-covariance structure of random effects was preferred as the best mixed model in [15a] and a diagonal variance-covariance structure of random effects was preferred as the best mixed model [15b] based on the model information criteria and log likelihood. Let $\left(\mathbf{y}_{0j}, \mathbf{x}_{0j}\right)$ be a vector of prior information available for j^{th} subject, then the estimated best linear unbiased predictor (EBLUP's) for random effects, $\mathbf{b}_j$, given the maximum likelihood estimate of $\left(\hat{\boldsymbol{\beta}}, \hat{\mathbf{D}}, \hat{\boldsymbol{\sigma}}\right)$ is given as

[16] $$\hat{\mathbf{b}}_j = \hat{\mathbf{D}}\mathbf{Z}_j^T(\mathbf{Z}_j\hat{\mathbf{D}}\mathbf{Z}_j^T + \hat{\sigma}^2\mathbf{I})^{-1}[\mathbf{y}_j - \mathbf{f}_j(\hat{\boldsymbol{\beta}}, \mathbf{x}_j)]$$

where $\mathbf{Z}_j = [\partial\mathbf{f}(\hat{\boldsymbol{\beta}}, \mathbf{x}_j) / \partial\hat{\boldsymbol{\beta}}^T]\mathbf{B}_j$ evaluated at $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}$. The estimated random effects were used to make subject specific predictions for the k^{th} disk from the j^{th} tree:

[17a] $$\hat{y}_{jk} = \left(\hat{\beta}_0 + \hat{b}_{0j}\right)\exp\left(\frac{\hat{\beta}_1}{x_{jk} + \left(\hat{\beta}_2 + \hat{b}_{2j}\right)}\right)$$

[17b] $$\hat{y}_{jk} = e^{\left(\hat{\beta}_0 + \hat{b}_{0j}\right)}\left(x_{jk}^{-\hat{\beta}_1} - 1\right)^{\left(\hat{\beta}_2 + \hat{b}_{2j}\right)}$$

We are proposing four methodologies using two separate models for making predictions of disk SG for a new subject with available information. They are: 1. the NLS method where no subject specific information is used for making predictions and predictions of a new subject was obtained using Eq. [2a] and [2b]; 2. the ADA method where prediction for new subjects were obtained using Eq. [11b]; 3. the GADA method where predictions for new subjects were obtained using Eq. [12a] and [12b]; 4. the NLMM method where predictions for new subjects were obtained using Eq. [17a] and [17b].

4.3.3 Model comparison

A subject-wise *M*-fold cross validation was used to assess the predictive performance of the four methods. The original sample is partitioned into *M* subsamples (*M* = 81 trees). From the *M* trees, data from one tree is retained as the validation data for making predictions using the models and data from the remaining *M* − 1 tree's are used for fitting the models. The process is repeated for each tree such that each of the trees is used once as validation data *(M-folds)*. The *M* results from the folds can then be combined to produce a single estimate.

The predictive performance of the models for each tree retained in the validation data set was evaluated by using 1, 2, 3, 4, or 5 observed height-SG pairs to predict the subject specific effects (subject specific parameter in ADA, GADA and random effect in NLMM). For each subject in the validation set, we used all possible combinations of disk SG–height pairs for a given number of SG–height pairs, say *r* (*r* = 1, 2, 3, 4, 5), in each fold of cross-validation. If the number of pairs used was *r*, then the total number of subject specific effects needing to be estimated for each subject was $n_i!/[r!(n_i - r)!]$. The estimated subject specific effects were then used to make disk SG predictions for all heights of the tree, including those used for estimating the subject specific effect, in the validation data set. The total number of predictions made on an individual tree given *r* pairs, is $N_i = n_i n_i!/[r!(n_i - r)!]$.

The models were compared using three statistics: root mean square error (*RMSE*), mean absolute residual (*MAR*) and mean residual (*MR*), each evaluated at *r* (Loague and Green 1991; Mayer and Butler 1993). These criteria are given below as:

[18] $$RMSE_r = \sqrt{\frac{\sum_{i=1}^{M}\left(y_{N_i} - \hat{y}_{N_i}\right)^2}{\sum_{i=1}^{M} N_i}}$$

[19] $$MAR_r = \frac{\sum_{i=1}^{M}\left|y_{N_i} - \hat{y}_{N_i}\right|}{\sum_{i=1}^{M} N_i}$$

[20] $$MR_r = \frac{\sum_{i=1}^{M}\left(y_{N_i} - \hat{y}_{N_i}\right)}{\sum_{i=1}^{M} N_i}$$

where, y_{N_i} and $\hat{y}_{N_i}$ are the actual and predicted values of disk SG.

4.4 Results

Tables 4.1 and 4.2 summarize the comparison criteria (RMSE, MAR and MR) from the NLS, ADA, GADA and NLMM versions of the two models. Overall both subject specific approaches: GADA and NLMM, performed better than the NLS approach where no subject specific information was used for making future predictions. A 3-19 % improvement in RMSE and MAR was observed for the NLMM approach compared to the ADA and GADA approaches. For model 1a the improvement in RMSE for the NLMM was 8-14 % compared to GADA and it varied depending on the number of data pairs used to estimate the subject specific parameter (Table 4.1). For model 1b the improvement in RMSE for NLMM ranged from 8-15 % (compared to ADA) and 10-19 % (compared to GADA), again with the results varying with the number of data pairs used to estimate the subject specific parameter. The MAR of model 1a for NLMM also showed consistent improvement from GADA (7-11 % improvement) irrespective of the prior information available for making predictions (Table 4.1). Similar to model 1a, the MAR of model 1b was consistently better for NLMM than ADA (3-15 % improvement for NLMM) and GADA (7-18 % improvement for NLMM) (Table 4.2). Based on the comparison criteria for model 1b, ADA did slightly better than the GADA (3-4% better) with lower RMSE and MAR (Table 4.2) irrespective of the prior information used for making predictions (Table 4.2).

In terms of MR, NLMM over predicted disk SG in all cases for both models irrespective of the number of data points used. MR results were mixed for GADA and ADA, in cases where 1 and 2 data pairs (prior information) were used for making predictions, GADA under predicted disk SG using model 1a, while it over predicted disk SG when more than two data pairs were used as prior information. It was observed that the ADA method over predicted disk SG when

Table 4.1: Summary from M-fold cross-validation from Model 1a

		RMSE			MAR			MR	
Total Pairs	**Data points used**	**GADA**	**NLMM**	**%Diff**	**GADA**	**NLMM**	**%Diff**	**GADA**	**NLMM**
10697	1	0.0272	0.0234	-14	0.0188	0.0175	-7	0.00039	-0.00061
58076	2	0.0221	0.0203	-8	0.0160	0.0149	-7	0.00002	-0.00121
195884	3	0.0202	0.0185	-8	0.0149	0.0136	-9	-0.00003	-0.00163
459859	4	0.0192	0.0174	-9	0.0142	0.0127	-10	-0.00005	-0.00191
797087	5	0.0184	0.0166	-10	0.0137	0.0121	-11	-0.00007	-0.00209

For the NLS M-fold cross-validation the predicted data points were 921, RMSE = 0.0318; MAR = 0.0254; MR = 2.19E-05.

Table 4.2: Summary from M-fold cross-validation from Model 1b

		RMSE					MAR					MR		
Total Pairs	**Data points used**	**ADA**	**GADA**	**NLMM**	**%Diff (ADA vs NLMM)**	**%Diff (GADA vs NLMM)**	**ADA**	**GADA**	**NLMM**	**%Diff (ADA vs NLMM)**	**%Diff (GADA vs NLMM)**	**ADA**	**GADA**	**NLMM**
10697	1	0.0245	0.0252	0.0227	-8	-11	0.0172	0.0179	0.0166	-3	-7	-4.23E-05	-2.71E-05	-1.37E-03
58076	2	0.0206	0.0212	0.0192	-8	-10	0.0151	0.0156	0.0139	-8	-11	2.23E-06	3.36E-06	-1.60E-03
195884	3	0.0190	0.0196	0.0173	-10	-13	0.0141	0.0145	0.0124	-12	-14	3.76E-05	7.14E-08	-1.66E-03
459859	4	0.0181	0.0186	0.0160	-13	-16	0.0134	0.0139	0.0115	-14	-17	7.09E-05	-9.45E-06	-1.66E-03
797087	5	0.0173	0.0179	0.0151	-15	-19	0.0129	0.0134	0.0109	-15	-18	1.03E-04	-2.07E-05	-1.64E-03

For the NLS M-fold cross-validation the predicted data points were 921, RMSE = 0.0314; MAR = 0.0250; MR = -1.99E-05.

% Diff = [(NLMM-GADA)/GADA]*100

one prior data pair was used with model 1b, in all other cases (2, 3, 4 and 5 data pairs) it under predicted disk SG. Using model 1b, GADA under predicted when 2 and 3 data pairs were used and over predicted when 1, 4 and 5 data pairs were used.

RMSE and MAR based on the GADA and NLMM approaches decreased considerably as the number of data points used for estimating/predicting subject specific parameters was increased. For models 1a and 1b, RMSE improvements of approximately 30 % were observed using both GADA and NLMM when five data points were used to make predictions instead of a single data point. Improvements in MAR were similar when 5 data points were used rather than one data point; 27 % (model 1a) and 25 % (model 1b) for GADA and 31 % (model 1a) and 34 % (model 1b) for NLMM. The predictive performances of both GADA and NLMM (in terms of RMSE and MAR) were improved considerably when the number of prior observations was increased from one to two. Improvements in RMSE and MAR ranging from 12-19 % were observed for ADA, GADA and NLMM for both model forms when two data points was used for making predictions instead of just one data point. The relative change in RMSE and MAR was marginal (only 4-10 %) when more than two observations were used to make the predictions, i.e. increasing the number of observations from two to three or more.

Improvements in RMSE for NLMM over GADA for both models 1a and 1b were consistent with relative height (Figure 4.2 and 4.3). RMSE of both GADA and NLMM approached each other as the number of observations used for making predictions increased from one to five. Similarly, MAR of NLMM was better than GADA in both model 1a and 1b irrespective of relative height (Figure 4.6 and 4.7). As the number of prior observations used for

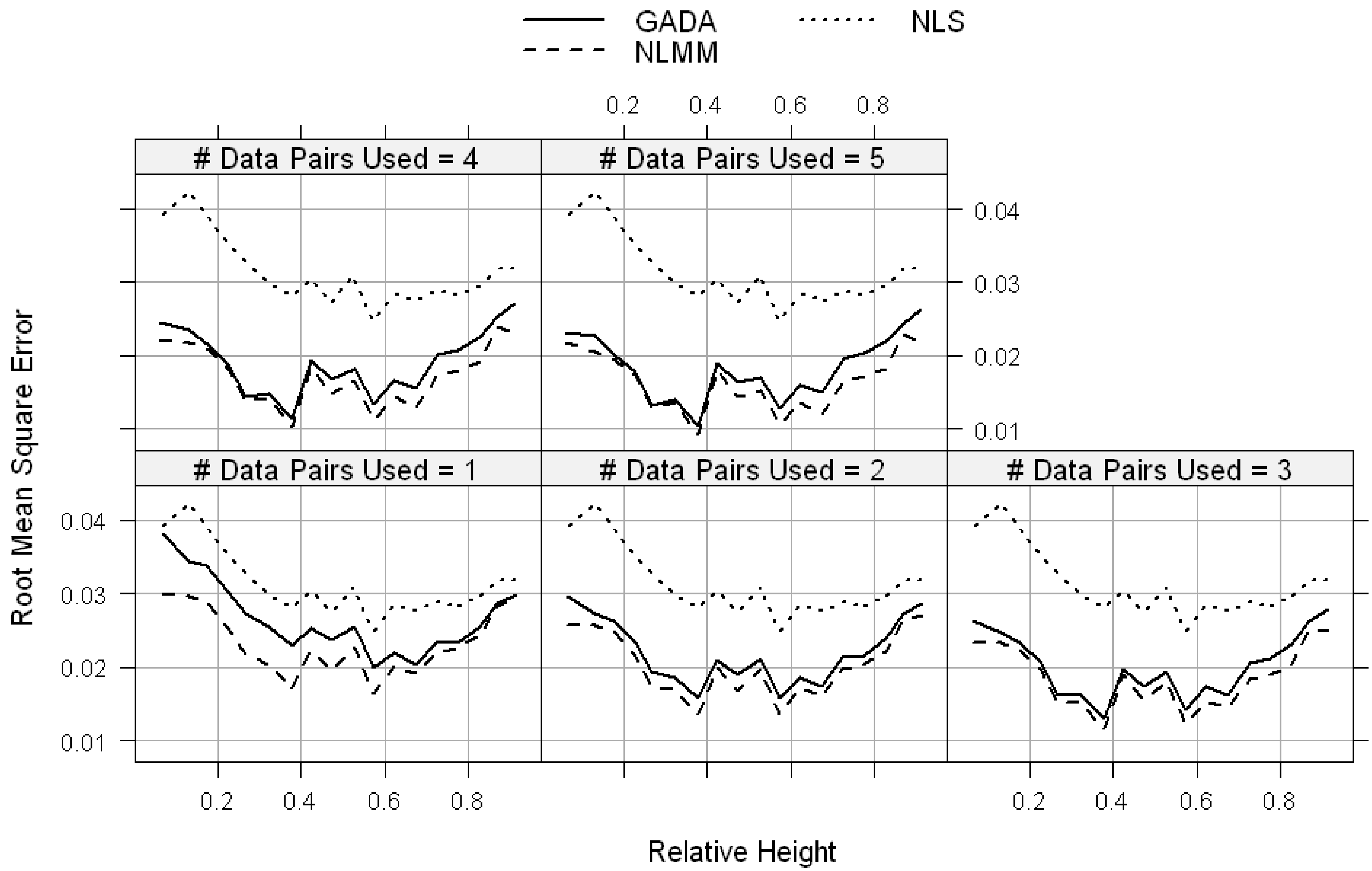

Figure 4.2: Root Mean Square Residual from M-fold cross validation with relative height by method used from model 1a

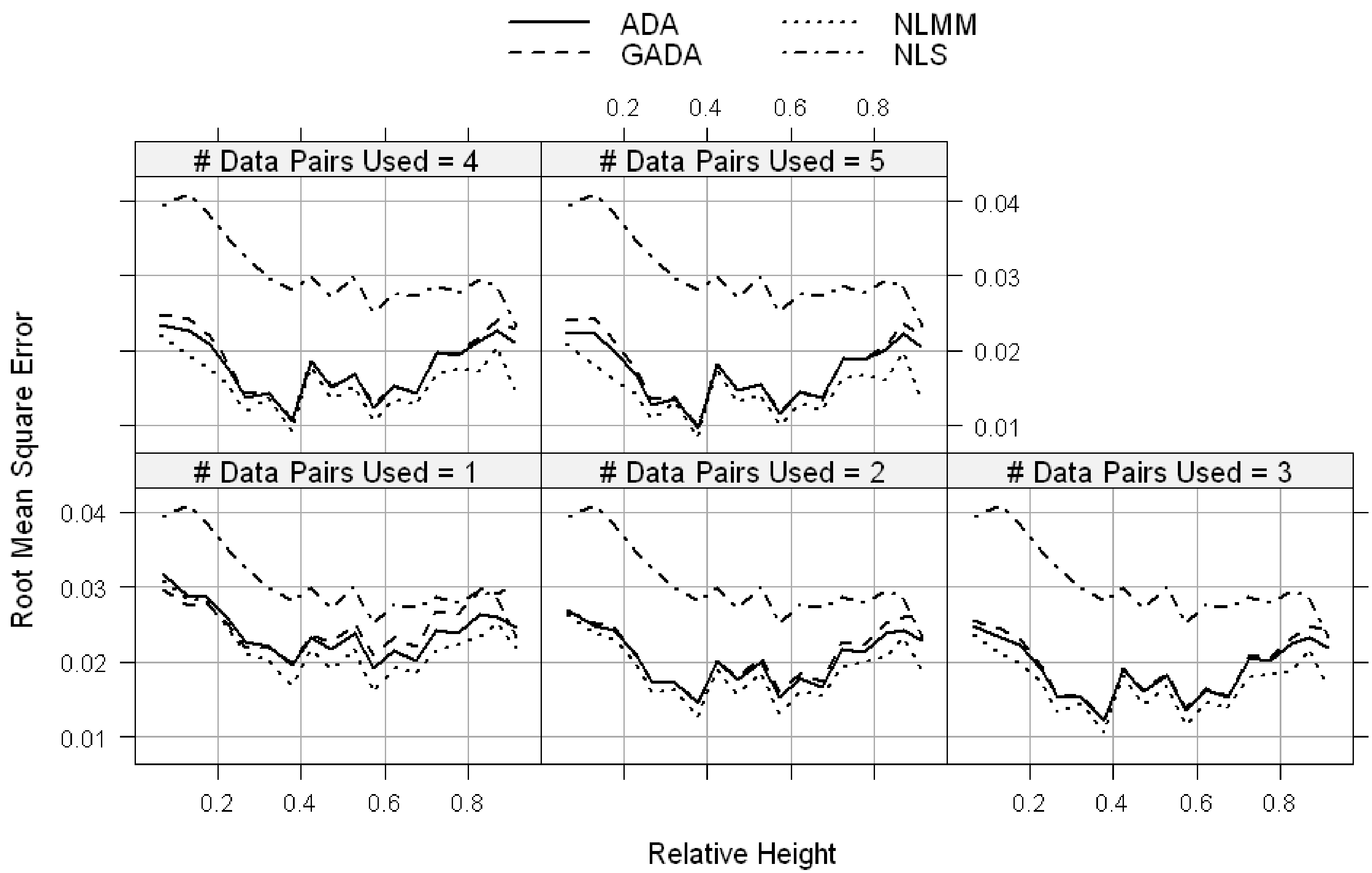

Figure 4.3: Root Mean Square Residual from M-fold cross validation with relative height by method used from model 1b

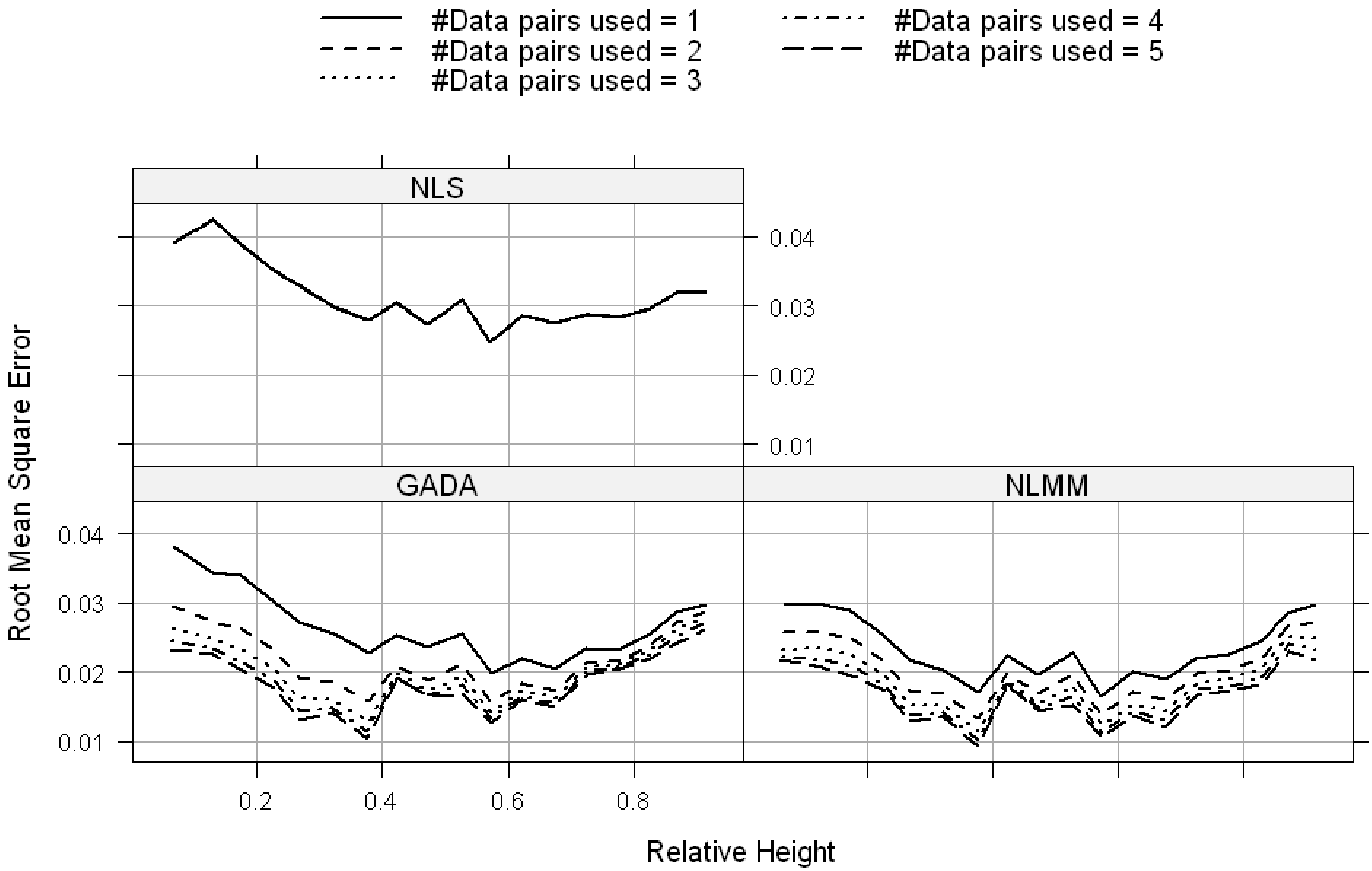

Figure 4.4: Root Mean Square Residual from M-fold cross validation with relative height by number of data pairs used from model 1a

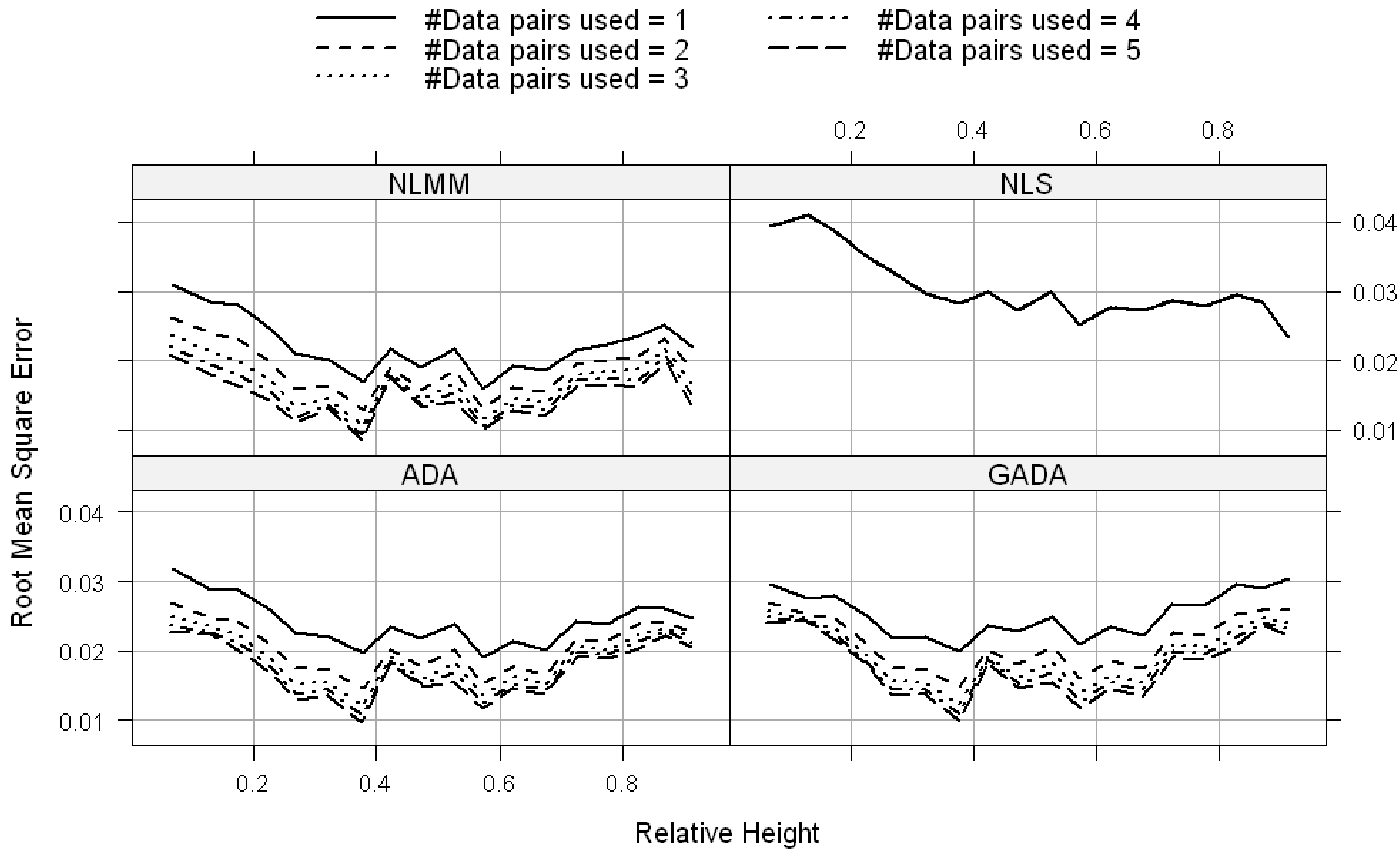

Figure 4.5: Root Mean Square Residual from M-fold cross validation with relative height by number of data pairs used from model 1b

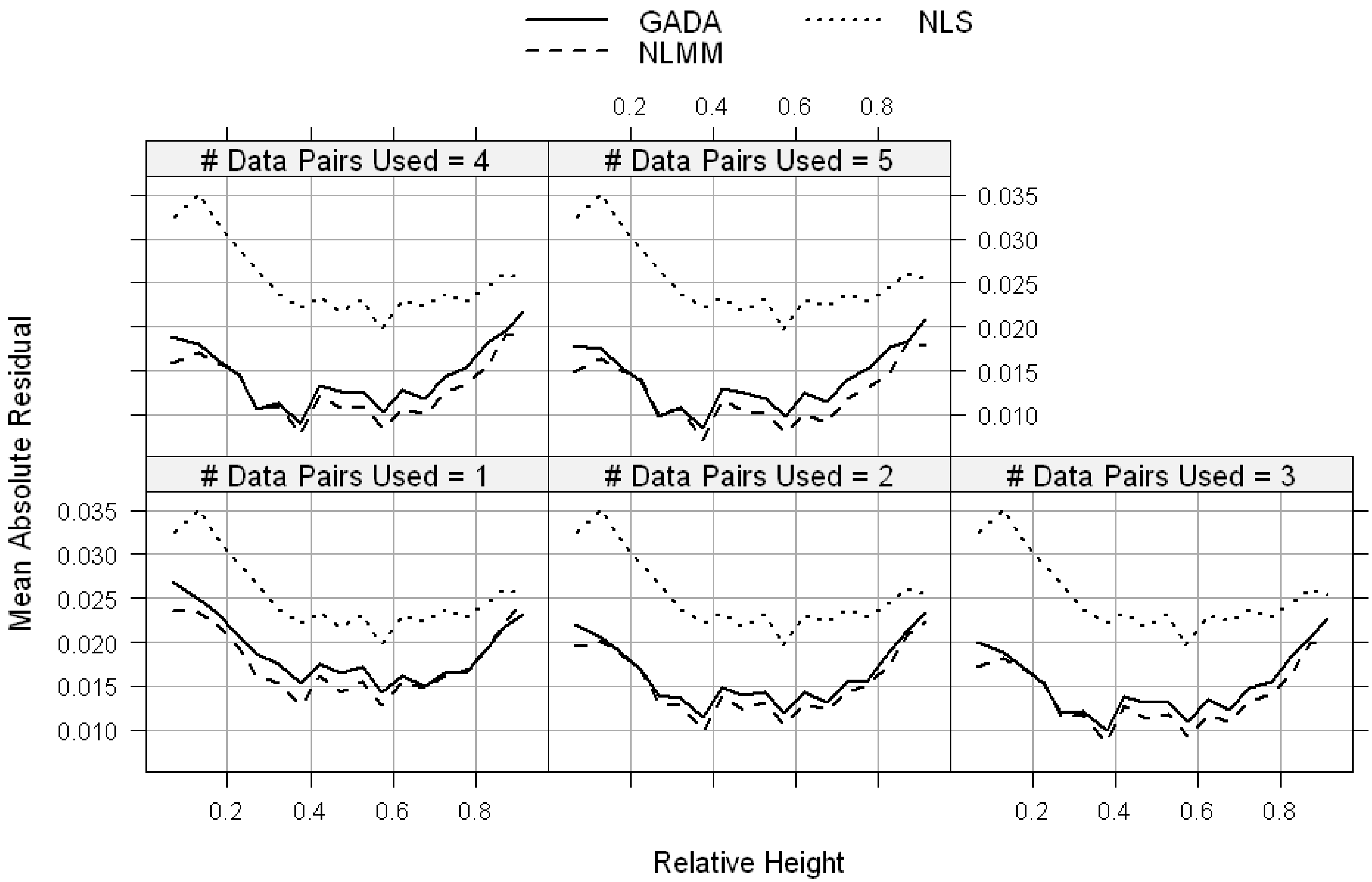

Figure 4.6: Mean Absolute Residual from M-fold cross validation with relative height by method used from model 1a

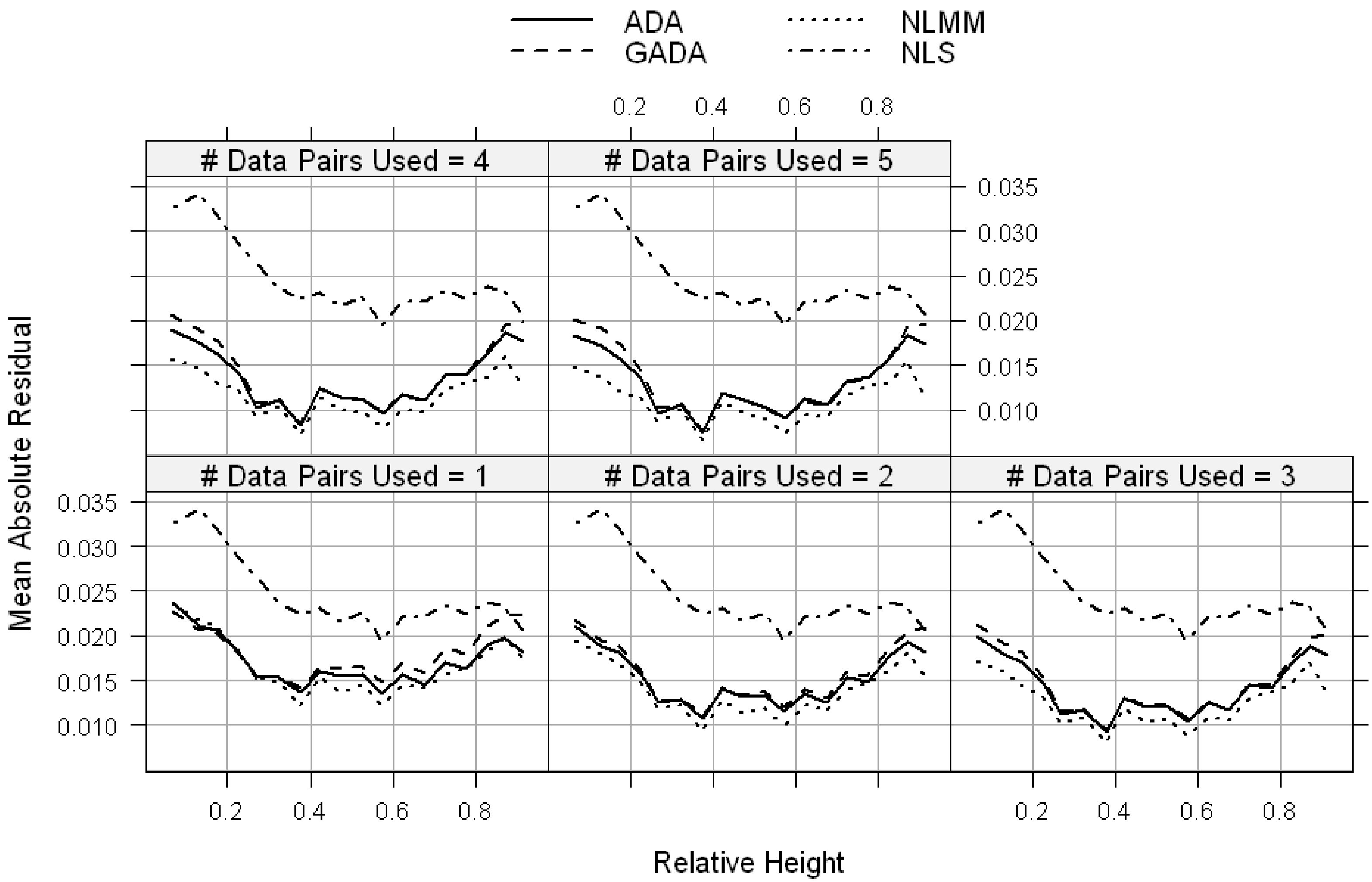

Figure 4.7: Mean Absolute Residual from M-fold cross validation with relative height by method used from model 1b

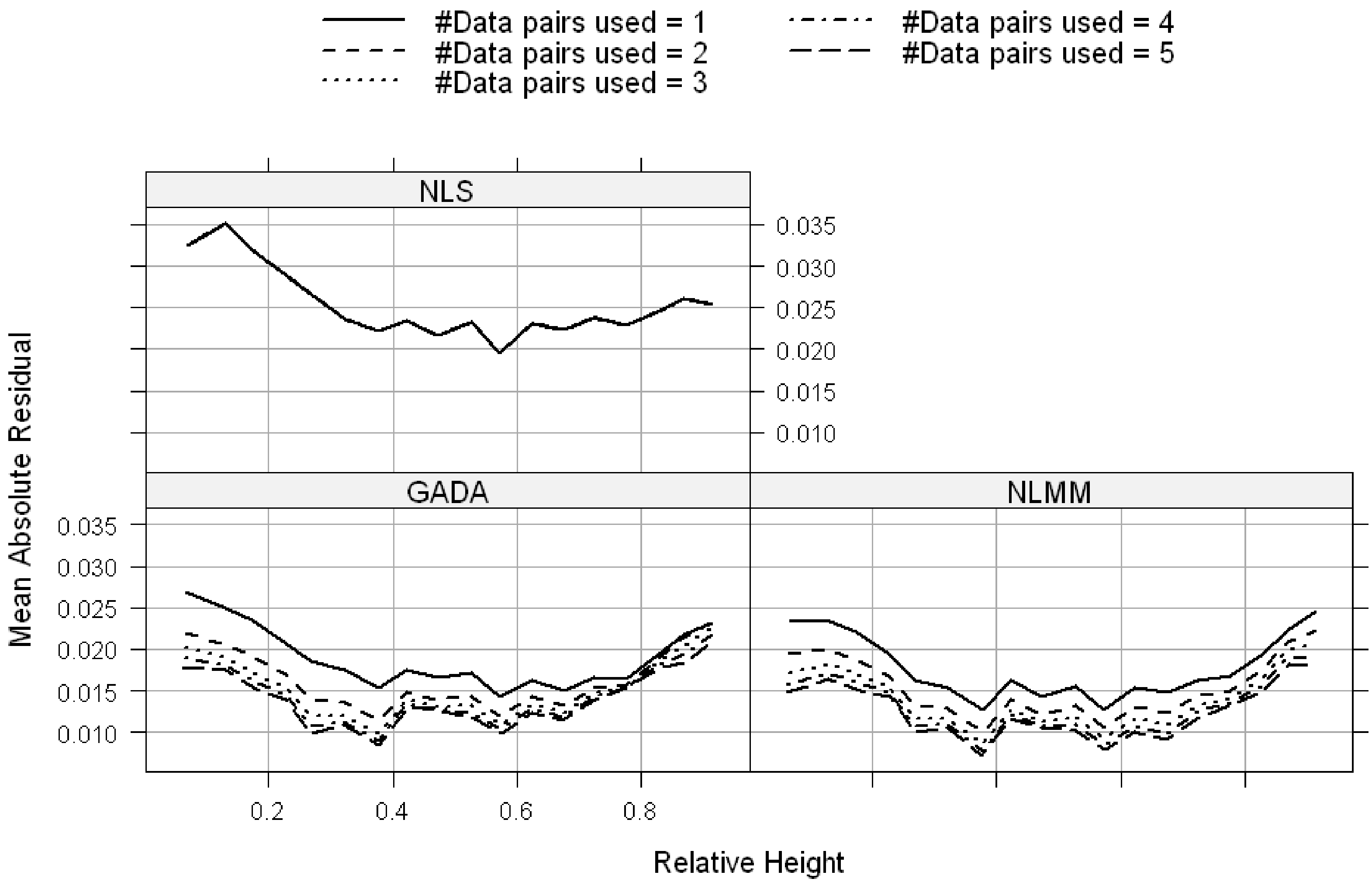

Figure 4.8: Mean Absolute Residual from M-fold cross validation with relative height by number of data pairs used from model 1a

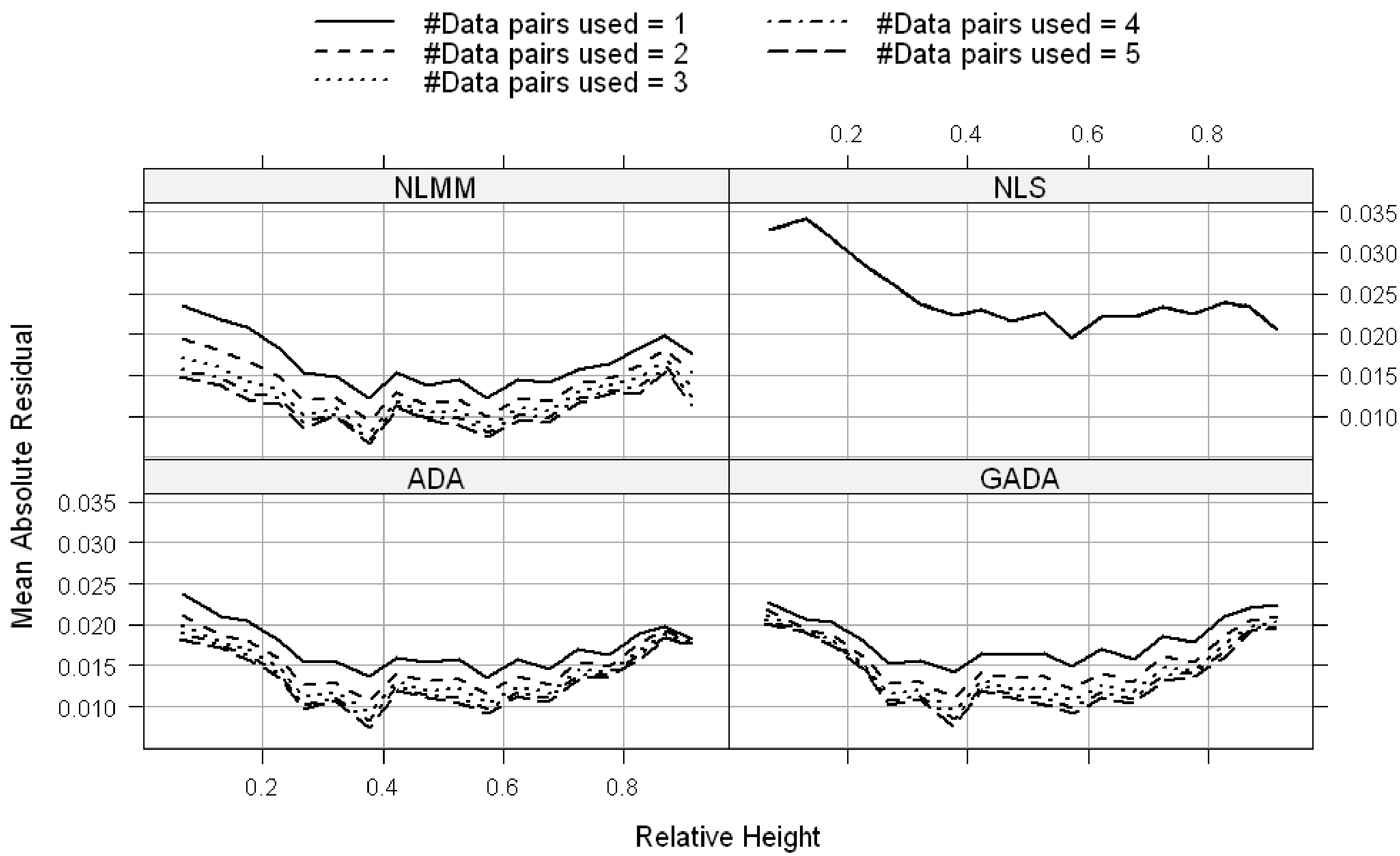

Figure 4.9: Mean Absolute Residual from M-fold cross validation with relative height by number of data pairs used from model 1b

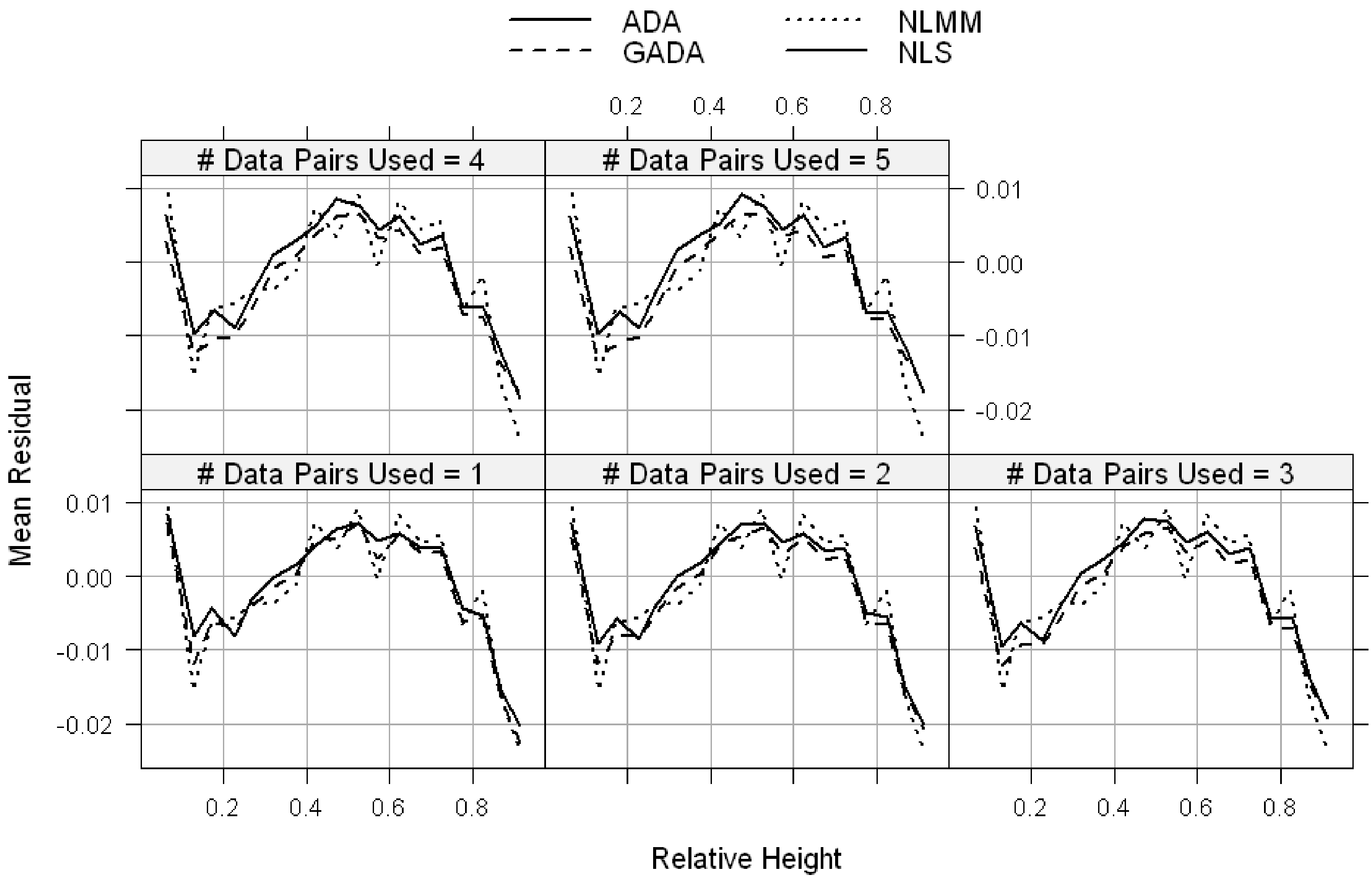

Figure 4.10: Mean Residual from M-fold cross validation with relative height by method used from model 1a

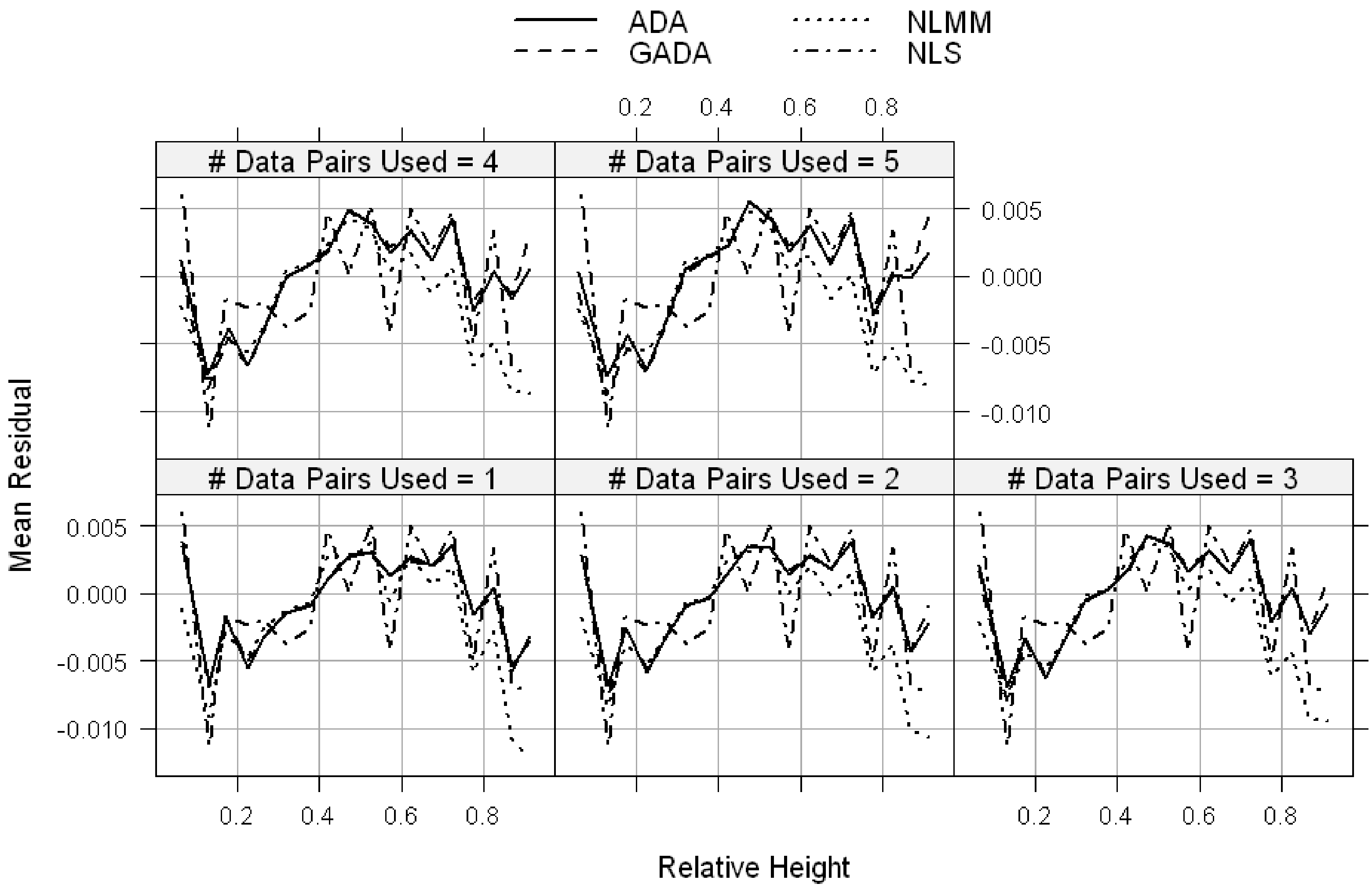

Figure 4.11: Mean Residual from M-fold cross validation with relative height by method used from model 1b

making predictions increased from one to five, the RMSE (Figure 4.4 and 4.5) and MAR (Figure 4.8 and 4.9) consistently improved for ADA, GADA and NLMM methodologies at all heights. The greatest differences in RMSE and MAR for the GADA and NLMM approaches were found when one prior observation was used for making predictions (Figure 4.4, 4.5, 4.8 and 4.9). The RMSE and MAR plots showed a specific pattern with relative height. Higher RMSE and MAR values were observed at the base and top of the tree. The MR plot of the GADA and NLMM forms of model 1a showed specific patterns with relative height; a positive bias at the very base of tree, becoming a negative bias at around a relative height of ~0.1, followed by a positive bias from a relative height of ~0.4 upwards until a relative height of 0.8 where it became negative bias again (Figure 4.10). The pattern in MR with relative height for model 1b was not as conspicuous as with model 1a (Figure 4.11).

4.5 Discussion

NLMM and GADA are two well known methods used to model longitudinal data collected in forestry. GADA methodology is basically a model parameterization technique used in the forestry literature to accommodate subject specific variation in curve shapes. NLMM is a statistical approach applied widely in different fields of science, including forestry. Foresters use these two methods to make predictions of the future forest conditions (eg: predicting height, basal area, or volume). We aimed to empirically compare the predictive performance of these two methodologies using disk SG data collected longitudinally (at different heights) in loblolly pine. This is the first attempt to understand the predictive performance of these two approaches not using conventional height-age data. It was found that subject specific methods, GADA (and ADA) and NLMM, perform better than traditional NLS where no subject specific information is used for making predictions. It was also observed that the error in prediction using GADA and

NLMM was reduced considerably as the number of prior data points used to make predictions increased from 1 to 5.

The GADA approach was proposed to explain variation in curve shapes with change in site/subjects productivity using an unobserved continuous variable χ, which measures the productive potential of a site/subject. In the present study χ measures the tree-to-tree change in curve shape. The unobserved variable χ is not measurable. Here, χ can be any variable which can explain the unexplained tree-to-tree variation in curve shapes. For example in height-age modeling, height measured at any pre-defined age of a stand or a tree (referred as site index in the literature) is used as a measure of the third dimension χ in the model. Similarly in the present study, we used disk SG measured at a pre-defined height level as a measure of χ in the model. Thus the flexibility of a GADA model depends on the assumptions made to define the functional relations used to entangle the subject-specific parameters with χ, if and only if χ is assumed to be a continuous variable. On the other hand, NLMM does not assume any functional relationship between subject-specific parameters with the subject productivity variable χ, but it estimates parameters specific to each subject (tree) as random effects (deviation from the fixed effect) which are usually assumed to follow a normal distribution with zero mean and a variance estimated while fitting. Here, the estimated random effect explains the deviation in curve shapes with respect to individual tree effect.

In the present study, we observed that NLMM performs better than GADA in terms of RMSE and MAR irrespective of the number of data points used for making predictions. The crux of GADA was to explain the change in curve shapes between response *y* and variable *t* with productive potential of subject χ using theoretically meaningful relationships between dependent parameters in the model with the unobserved theoretical variable χ (the subject productivity variable). However, the implementation of such relationships reported to invokes restrictions or constraints between the subject dependent parameters and in the estimable parameter space (Jordan et al. 2010; Stewart et al. 2010; Wang et al. 2008a). In the present study, defining relationships of subject specific parameters with unobserved subject productivity variable as; $\beta_{1i} = \chi_i$ and $\beta_{2i} = \gamma_1 + \gamma_2 \chi_i$ (Eq. 7a) ; $\beta_{0i} = \chi_i$ and $\beta_{2i} = \gamma \chi_i$ (Eq. 7b); invokes a restriction in the estimates of parameters β_{1i} and β_{2i} in Eq. 7a and β_{0i} and β_{2i} in Eq. 7b. For NLMM, such relations between subject specific parameters are absent and it is assumed that the random effects follow a multivariate normal (here a bivariate normal) distribution because two parameters are considered random in the two NLMM models fitted in this study. As indicated earlier, the flexibility of GADA models depends on *the functional relations used to entangle the subject-specific parameters with* χ . Thus in terms of prediction, the performance of the GADA model depends upon the functional relationship used in model development. However, the question is whether we can define such a flexible functional relationship which can describe a multidimensional parameter space (which is equivalent to the NLMM parameter space).

Most forestry related studies have aimed to make future predictions from a single pair of observations available in the present (measurements from the past are not used). Based on this study, it is evident that using single data pairs restricts information available for prediction irrespective of the method used (GADA or NLMM). The decrease in RMSE and MAR observed when an increasing number of data pairs were used for estimating the subject specific parameters in both GADA and NLMM supports this conclusion. Recent results reported by Stewart et al. (2010) and Jordan et al. (2010) using height growth data also supports these findings.

ADA models represent one parameter family of curves, i.e. the curves generated using the model are determined by one parameter which is assigned as a subject specific parameter from the base model. In relation to statistical literature, ADA is fixed effect regression where each subject has its own estimate of the subject specific parameter. ADA is a special case of GADA with one parameter in the base model allowed to vary with site/subject productivity χ. In this study, the ADA version of model 1b performed better than the GADA version in terms of RMSE and MAR. ADA might perform better than GADA in some situations, as in this study, where the assumed and/or attributed parameter relationship between subject specific parameters in GADA is based on false assumptions and lead to a model with biased predictions.

In summary, we observed an improvement in predictive performance of NLMM compared to GADA irrespective of the number of data points used for estimating the subject specific effect. We found a significant improvement in prediction for both GADA and NLMM as the number of data points used to estimate the subject specific effect increased from one to two. Based on this empirical study, it is advised to use at least two prior observations to estimate the subject specific effect and to make future predictions. Although we found an improvement in prediction for NLMM in this study, it needs to be investigated further both theoretically and empirically.

4.6 Acknowledgements

Thanks to the Wood Quality Consortium for providing the data utilized in this study.

4.7 References

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

Berhe, L. 2009. An overview of mixed effects models in forestry. Umeå University.

Borders, B.E., R.L. Bailey, and K.D. Ware. 1984. Slash pine site index from a polymorphic model by joining (splining) nonpolynomial segments with an algebraic difference method. Forest Science 30(2):411-423.

Calegario, N., R.F. Daniels, R. Maestri, and R. Neiva. 2005. Modeling dominant height growth based on nonlinear mixed-effects model: a clonal Eucalyptus plantation case study. Forest Ecology and Management 204(1):11-21.

Cieszewski, C.J. 2002. Comparing fixed and variable base age site equations having single versus multiple asymptote. Forest Science 48(1):7-23.

Cieszewski, C.J. 2003. Developing a well-behaved dynamic site equation using a modified Hossfeld IV function $Y^3 = (ax^m)/(c + x^{m-1})$, a simplified mixed-model and scant subalpine fir data. Forest Science 49(4):539-554.

Cieszewski, C.J., and R.L. Bailey. 2000. Generalized algebraic difference approach: Theory based derivation of dynamic site equations with polymorphism and variable asymptote. Forest Science 46(1):116-126.

Cieszewski, C.J., and M. Strub. 2008. Generalized algebraic difference approach derivation of dynamic site equations with polymorphism and variable asymptotes from exponential and logarithmic functions. Forest Science 54(3):303-315.

Davidian, M., and D.M. Giltinan. 1995. Nonlinear Models for Repeated Measurement Data. Chapman and Hall, London. 359 p.

Davidian, M., and D.M. Giltinan. 2003. Nonlinear Models for Repeated Measurement Data: An Overview and Update. Journal of Agricultural, Biological, and Environmental Statistics 8:387-419.

Fang, Z. 1999. A simultaneous system of linear and nonlinear mixed effects models for forest growth and yield prediction (Ph.D dissertation), University of Georgia, Athens. 177 p.

Fox, T.R., E.J. Jokela, and H.L. Allen. 2007. The development of pine plantation silviculture in the southern United States. Journal of Forestry 105(7):337-347.

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

Hall, D.B., and M. Clutter. 2004. Multivariate multilevel nonlinear mixed effects models for timber yield predictions. Biometrics 60(1):16-24.

He, R. 2004. Mixed effects modeling of wood properties of loblolly pine in the southeastern United States (Ph.D dissertation), University of Georgia, Athens. 213 p.

Jordan, L., K. Berenhaut, R. Souter, and R.F. Daniels. 2005. Parsimonious and completely compatible taper, total, and merchantable volume models. Forest Science 51(6):578-584.

Jordan, L., R. Souter, B. Parresol, and R.F. Daniels. 2006. Application of the algebraic difference approach for developing self-referencing specific gravity and biomass equations. Forest Science 52(1):81-92.

Jordan, L., M. Strub, T.G. Matney, R.A. Souter, E.B. Schultz, and B. Stewart. 2010. Comparing a taper model developed under the generalized algebraic difference approach and nonlinear mixed effect methodologies. Forest Science Submitted.

Lindstrom, M.J., and D.M. Bates. 1990. Nonlinear Mixed Effects Models for Repeated Measures Data. Biometrics 46(3):673-687.

Loague, K., and R.E. Green. 1991. Statistical and graphical methods for evaluating solute transport models: overview and application. Journal of Contaminant Hydrology 7:51-73.

Mayer, D.G., and D.G. Butler. 1993. Statistical validation. Ecological Modelling 68:21-32.

Phillips, K.M. 2002. Modeling within tree changes in wood specific gravity and moisture content for loblolly pine in Georgia (M.S. Thesis), University of Georgia, Athens. 95 p.

Stewart, B., L. Jordan, M. Strub, and R. Li. 2010. Comparing ADA, GADA, and site index models versus effects models: Restrictions in models formulation. Forest Science Accepted.

Wang, M.L., B.E. Borders, and D. Zhao. 2008a. An empirical comparison of two subject-specific approaches to dominant heights modeling: The dummy variable method and the mixed model method. Forest Ecology and Management 255(7):2659-2669.

Wang, M.L., K. Rennolls, and B.E. Borders. 2008b. Base-Age Invariant Site Index Models from a Generalized Algebraic Parameter Prediction Approach. Forest Science 54(6):625-632.

Wear, D.N., and J.G. Greis. 2002. The Southern Forest Resource Assessment: Summary Report. USDA Forest Service, Southern Research Station Report No. 53. 635p

Zhang, Y.J., and B.E. Borders. 2004. Using a system mixed-effects modeling method to estimate tree compartment biomass for intensively managed loblolly pines - an allometric approach. Forest Ecology and Management 194(1-3):145-157.

Zhang, Y.J., B.E. Borders, and R.L. Bailey. 2002. Derivation, fitting, and implication of a compatible stem taper-volume-weight system for intensively managed, fast growing loblolly pine. Forest Science 48(3):595-607.

Chapter 5

Effect of fertilization on growth and wood properties of thinned and unthinned midrotation loblolly pine (*Pinus taeda* L.) stands

5.1 Abstract

Growth and wood properties were measured on breast height cores collected from two stands, New Bern and Bertie, located in the lower Coastal Plain of North Carolina. The New Bern site was thinned before fertilizer application and the Bertie site was not. The study was laid out in a randomized complete block design with each treatment replicated in four blocks at New Bern and two blocks at Bertie. The treatments were different levels of nitrogen fertilization: Control – no nitrogen, 112 kg/ha, 224 kg/ha and 336 kg/ha. In addition 28 kg/ha of phosphorous was included with each treatment. The objective of this study was to examine the response in growth and wood properties to midrotation fertilization in a thinned versus an unthinned stand. A significant decrease in latewood specific gravity was observed following nitrogen fertilization in the thinned stand, but not in the unthinned stand. Whole ring width, latewood width and earlywood width, significantly increased following nitrogen fertilization at New Bern, but not at Bertie. Whole ring specific gravity, early wood specific gravity, latewood percent and earlywood:latewood ratio did not show any change due to fertilization in either stand. Responses in both growth and wood characteristics lasted for 2-3 years following fertilization and depended upon the amount of fertilizer applied. The response to nitrogen application was significant for the thinned stand only and not in the unthinned stand.

Key Words: Wood density, specific gravity, ring width, repeated measure, mixed-effects model

5.2 Introduction

Loblolly pine (*Pinus taeda* L.) is the most important plantation species in the southern United States (US) with a planted area of more than 13 million ha (Schultz 1997). Intensive management is commonly used to improve the productivity of loblolly pine plantations throughout its growing range (Jokela et al. 2004). Midrotation thinning and fertilization is a widely utilized silvicultural practice and by year 2004 a total of 6.5 million ha of southern pine plantations had received fertilization with a peak annual midrotation fertilization of 0.52 million ha in 2002 (Albaugh et al. 2007).

Water availability and low soil nutrient availability are considered to be the two principal resources limiting pine productivity in the southern US (Albaugh et al. 2004; Allen et al. 2005). Of these, low availability of soil nutrients such as nitrogen (N), phosphorous (P), potassium (K) and boron (B) are more important in limiting growth than water stress (Albaugh et al. 2004). Hence supplementing nutrient supply is an important silvicultural tool for increasing pine productivity (Allen et al. 2005; Fox et al. 2007). In studies established by the Forest Nutrition Cooperative throughout the southern US, over 80 % of intermediate-aged pine stands responded positively to the addition of N+P fertilizers (Allen et al. 2005), with reported growth gains of 1.6 tons/acre/year averaged over a period of eight years following the application of 224 kg/ha of N and 28 kg/ha of P (Fox et al. 2007).

Forest products industries are concerned with the potential effects of thinning and fertilization on wood properties including specific gravity (SG), microfibril angle, strength (Modulus of Rupture) and stiffness (Modulus of Elasticity). Early age fertilization (at ages 1 and 4) of loblolly pine did not change whole core SG (Mora 2003), while Jokela et al. (2004) found that fertilization combined with weed control decreased SG (compared to the control) for loblolly pine on some of the sites they examined. Albaugh et al. (2004) reported a decrease in ring specific gravity and an increase in the earlywood:latewood ratio of rings produced immediately following fertilization of 8-year old loblolly pine stand. A similar decrease in specific gravity of rings produced immediately following fertilization has been reported by others (Love-Myers et al. 2009; Zobel et al. 1961). In comparison, operational thinning does not appear to produce a significant change in the specific gravity of loblolly pine (Burton and Shoulders 1974; Tasissa and Burkhart 1998).

Successful application of intensive forest management practices, including midrotation thinning and fertilization, requires an understanding of three basic things: resource limitation which affect productivity, effects of different silvicultural treatments on the availability of limited resources and the final consequence of resource availability on productivity and wood quality (Allen et al. 2005). An understanding of the effects of different management practices on productivity and quality of wood is of great importance both for wood growers and wood buyers. Therefore, our objectives were to examine the effects of midrotation fertilization on growth and wood properties, including average whole ring SG (WRSG), latewood SG (LWSG), earlywood SG (EWSG), percentage latewood (RLWP), whole ring basal area (WRBA), earlywood basal area (EWBA), latewood basal area (LWBA) and earlywood:latewood ratio (ELWR), in a thinned and an unthinned stand and to understand how these properties changed with time.

5.3 Materials and methods

5.3.1 Origin of samples

The study was conducted on wood samples collected from two loblolly pine plantations, New Bern and Bertie, located in lower Coastal Plain of North Carolina. The plantations were part of the regionwide-13 study, an extensive field trial established throughout the southeast US by the

Forest Nutrition Cooperative (North Carolina State University) to identify the outcome of N and P fertilization in midrotation loblolly pine stands. The New Bern and Bertie stands were planted in 1970 and 1977 respectively with approximately 1482 trees per ha. The study was laid out in a randomized complete block design with four blocks at New Bern and in two blocks at Bertie. The treatments used in this study were Control – no Nitrogen (000N), 112 (112N), 224 (224N), and 336 (336N) kg/ha of N with all the treatments receiving 28 kg/ha of P. The New Bern stand was thinned to 605 trees per ha in 1983 and treated with different rates of N fertilizer in March, 1984 at the age of 14. No thinning was conducted at Bertie but the stand was treated with similar rates of N in April 1987 at the age of 10. All plots at the New Bern stand were thinned to 346 trees per ha in 1995 and a second fertilization treatment with 224 kg/ha of N and 28 kg/ha of P was applied in 1996. The two stands were harvested in 2003.

Increment cores (12 mm diameter) were collected from a subsample of 9 trees in each treatment plot from the two sites using a hydraulically driven borer. A total of 144 cores (36 cores per treatment) were collected from New Bern and 72 cores (18 cores per treatment) were collected from Bertie. Only 66 cores were analyzed from the 72 cores collected from Bertie (1 core from 112N and 5 cores from 224N were not used).

5.3.2 Wood property analysis

Radial strips (1.6 mm thick) were sawn from the breast height cores and conditioned to 8 % moisture content. All the radial strips were then read on a scanning X-ray densitometer (Quintek Measurement Systems™) at a resolution of 0.006 mm to determine earlywood, latewood and whole ring width and SG. The densitometry data were also used to determine radial growth and percent latewood in each annual ring. A SG of 0.48 was used to distinguish between earlywood and latewood. Specific gravity values were based on a green volume and oven dry weight basis.

5.3.3 Statistical analysis

The data collected from each tree can be considered as repeated measurements taken over time from an individual subject. A repeated measures analysis of variance with the main effects of treatment and time and their interaction was used to analyze this dataset. The data collected from the 5- year post fertilization period were used for the analysis, i.e. 1984-1988 for New Bern and 1987-1991 for Bertie. One year pre-fertilization measurement was used as covariate to adjust for any pre-treatment differences (data from year 1983 in New Bern and 1986 in Bertie). The data collected from each site were analyzed separately since the effect of thinning was not considered as a treatment and there is a possibility of confounding the effects of thinning and location (the thinned and unthinned sites were in different locations). Separate analysis of variance was conducted for each growth and wood property. Tukey's Honestly Significant Difference (HSD) test was used to conduct pair-wise means comparisons of treatments where a significant result was obtained from ANOVA.

The full linear mixed model used for the analysis of the data can be written as:

$$y_{ijkl} = \mu + \beta y_{ijk0} + F_i + T_l + (FT)_{il} + b_j + (Fb)_{ij} + (FbT)_{ijl} + e_{ijkl} \qquad \textbf{(1)}$$

$$i = 1,\ldots,4,\ \ j = 1,\ldots,4,\ k = 1,\ldots.,9,\ l = 1,\ldots,5$$

where y_{ijkl} = the property measured from l^{th} ring, of the k^{th} tree, of the j^{th} block, receiving the i^{th} fertilization treatment; μ = the population mean; β is the coefficient associated with linear covariate y_{ijk0}, a pre-fertilization measurement; F_i = the i^{th} fertilization effect; T_l = the l^{th} ring/time effect; $(FT)_{il}$ = the interaction of the i^{th} fertilization and l^{th} ring/time effect; b_j = the random effect of the j^{th} block with $b_j \sim NID(0,\sigma_b^2)$; $(Fb)_{ij}$ = the random interaction effect of the i^{th} fertilization and j^{th} block effects with $(Fb)_{ij} \sim NID(0,\sigma_{Fb}^2)$, the true error term for testing the treatment effect; $(FbT)_{ijl}$ = the random interaction of the i^{th} fertilization j^{th} block and l^{th} ring/time effect with $(FbT)_{ijl} \sim NID(0, \sigma^2_{FbT})$, the true error term for testing the main effect of time and treatment by time interaction; and e_{ijkl} = subsampling error, with $e_{ijkl} \sim NID(0, \sigma^2)$.

Since measurements were taken from rings produced in adjacent years, we expect that correlation exists among the measurements taken from each sample tree. Observations closer together will tend to be more alike than observations farther apart. Correlation structures are used for modeling the dependence among observations. In the context of mixed-effects models, they are used to model the correlation among the within-subject errors. If correlation among the repeated measurements exists, its autocorrelation pattern can be modeled with an appropriate spatial or temporal correlation model depending on the nature of the correlation. Since we have equally spaced measurements taken over time, a temporal correlation structure may be used to account for the autocorrelation among measurements within a tree. The correlation structures used include: unstructured, compound symmetry, heterogeneous compound symmetry, Toeplitz, Heterogeneous Toeplitz, first-order autoregressive and heterogeneous first-order autoregressive. Equation (1) was fit to the data from each site with different temporal correlation structures. Final model selection was based on the improvement made on the model using the Akaike's information criteria (AIC) and Bayesian information criteria (BIC).

All the tests were conducted using the MIXED procedure with a restricted maximum likelihood estimation (REML) method available in SAS version 9.1.3 (SAS 2004). The level of significance used in all tests was 0.05, unless otherwise stated.

5.4 Results

Results of the ANOVA for both sites are presented in Tables 5.1 and 5.2 respectively. A time related trend was present for all growth and wood properties (significant time effect), except for

WRSG at New Bern, indicating large variation in growth and wood properties with time (Tables 5.1 and 5.2). Since our main interest was to identify the influence of different midrotation fertilization regimes on growth and wood properties in a thinned and unthinned site, we will restrict our results and discussion to the main effect of treatment and treatment by time interaction terms.

Table 5.1. Analysis of variance table for New Bern – Thinned fertilized site.

Property	**source**	**Numerator d.f.**	**Denominator d.f.**	**F-value**	**p-value**
WRSG	F	3	11.1	2.07	0.1627
	T	4	44.1	1.45	0.2346
	F*T	12	44.3	1.59	0.1288
LWSG	F	3	9.05	5.58	0.0192
	T	4	46.2	31.51	<0.0001
	F*T	12	46.1	2.81	0.0058
EWSG	F	3	8.41	1.1	0.4001
	T	4	30.2	11.7	<0.0001
	F*T	12	30.4	1.4	0.2183
RLWP	F	3	107	0.59	0.6247
	T	4	47.7	7.94	<0.0001
	F*T	12	48.2	1.08	0.3997
WRBA	F	3	11.9	10.61	0.0011
	T	4	47.3	5.43	0.0011
	F*T	12	49	3.81	0.0004
EWBA	F	3	14.4	5.89	0.0078
	T	4	41.8	6.69	0.0003
	F*T	12	43.2	3.23	0.0023
LWBA	F	3	10.4	12.59	0.0009
	T	4	42	4.78	0.0029
	F*T	12	42.8	2.22	0.0278
ELWR	F	3	112	0.58	0.6265
	T	4	49	5.79	0.0007
	F*T	12	49.8	0.83	0.6237

*WRSG – Whole ring specific gravity; LWSG – Latewood specific gravity; EWSG – Earlywood specific gravity; RLWP – Ring latewood percent; WRBA – Whole ring basal area; EWBA – Earlywood basal area; LWBA – Latewood basal area; ELWR – Earlywood:Latewood ratio; F – Fertilization effect; T – Time effect.

Table 5.2. Analysis of variance table for Bertie - Unthinned fertilized site.

Property	source	Numerator d.f.	Denominator d.f.	F-value	p-value
WRSG	F	3	3.04	0.64	0.6359
	T	4	16.1	19.71	<0.0001
	F*T	12	14.6	0.58	0.8225
LWSG	F	3	4.23	1.33	0.3779
	T	4	123	29.19	<0.0001
	F*T	12	166	0.47	0.9285
EWSG	F	3	2.68	0.23	0.8737
	T	4	17.3	3.44	0.0306
	F*T	12	15.7	0.64	0.7796
RLWP	F	3	3.03	0.5	0.7084
	T	4	15.5	14.63	<0.0001
	F*T	12	14.3	0.87	0.5887
WRBA	F	3	5.73	1.01	0.4537
	T	4	91.8	8.63	<0.0001
	F*T	12	138	1.38	0.1822
EWBA	F	3	4.38	0.5	0.6982
	T	4	89.8	7.55	<0.0001
	F*T	12	135	1.25	0.2582
LWBA	F	3	62	1.11	0.3519
	T	4	18	3.13	0.0405
	F*T	12	16.6	1	0.4905
ELWR	F	3	3.78	0.19	0.896
	T	4	13.9	9.45	0.0007
	F*T	12	12.5	0.64	0.7764

*WRSG – Whole ring specific gravity; LWSG – Latewood specific gravity; EWSG – Earlywood specific gravity; RLWP – Ring latewood percent; WRBA – Whole ring basal area; EWBA – Earlywood basal area; LWBA – Latewood basal area; ELWR – Earlywood:Latewood ratio; F – Fertilization effect; T – Time effect.

The treatment by time interaction was not significant for WRSG at either site (Tables 5.1 and 5.2). At New Bern, a decrease in WRSG of 0.027 and 0.026 was observed for 336N and 224N respectively compared to the control in the first year following fertilization. By the third year differences were only 0.006 (224N) and 0.009 (336N) (Figure 5.1). At Bertie changes in WRSG were negligible (Figure 5.2).

The treatment by time interaction was significant for LWSG at New Bern. In the first year following fertilization significant differences were observed for the 224N (p-value=0.0095) and 336N (p-value<0.0001) treatments from the control and for the 336N treatment from the 112N (p-value=0.0465) treatment. Compared to the control LWSG was estimated to decrease by 0.062 and 0.043 for 336N and 224N respectively (Figure 5.1), while the difference in LWSG for the 112N and 336N treatments was 0.037. For the second year following fertilization differences were observed among the treatments and the control, eg. a decrease of 0.028 in LWSG for trees receiving 336N compared to the control, however the differences were not statistically significant. With time, (2-3 years) LWSG of the fertilized and control trees became similar (Figure 5.1). At New Bern, the main effect of fertilizer treatment was significant with the 336N (p-value=0.0313) and 224N (p-value=0.0622) treatments different from the control. At Bertie, treatment by time interactions were not significant for LWSG. Differences of 0.031, 0.042 and 0.041 compared to the control were observed for the 112N, 224N and 336N treatments respectively in the first year following fertilization. A difference in LWSG of 0.03 from the control was observed for the 336N treatment in the second to fourth years after fertilization. Significant treatment by time interaction and a treatment main effect was absent for EWSG and RLWP at both sites.

A treatment by time interaction was present at New Bern for WRBA, but was absent at Bertie. At New Bern, a significant increase was observed in WRBA for all treatments compared to the control in two years following fertilization. Increases in WRBA of 7.1, 7.4, 9.0 (first year; with p-values 0.0189, 0.0142, 0.0007) and 8.4, 9.1, 15.2 cm^2/ha (second year; with p-values 0.0279, 0.0114, <0.0001) were observed for the 112N, 224N and 336N treatments respectively compared to the control. By the third year only the 336N treatment was significantly different

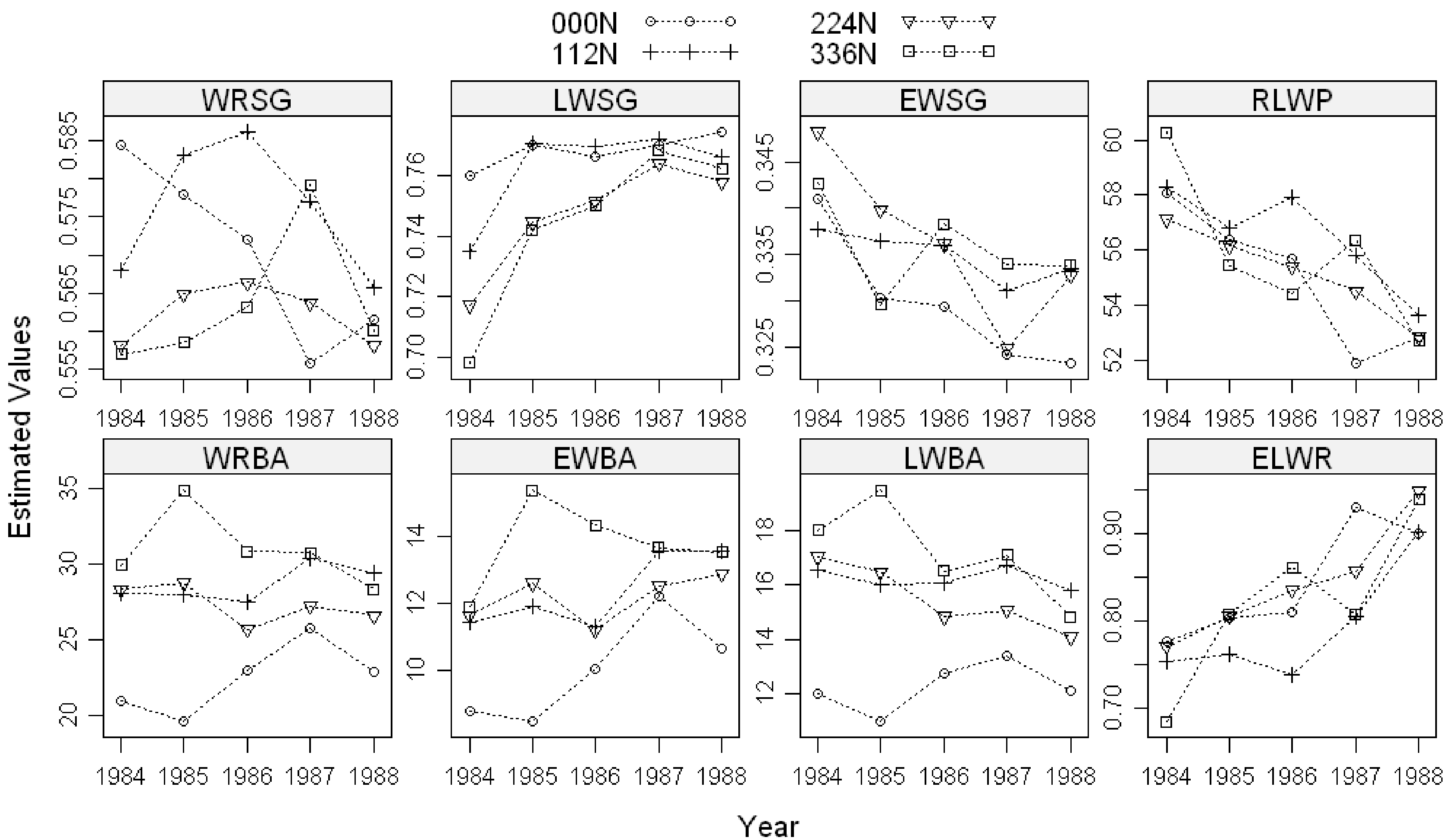

Figure 5.1. Plots of estimated growth and wood properties for different levels of nitrogen application by year at New Bern.

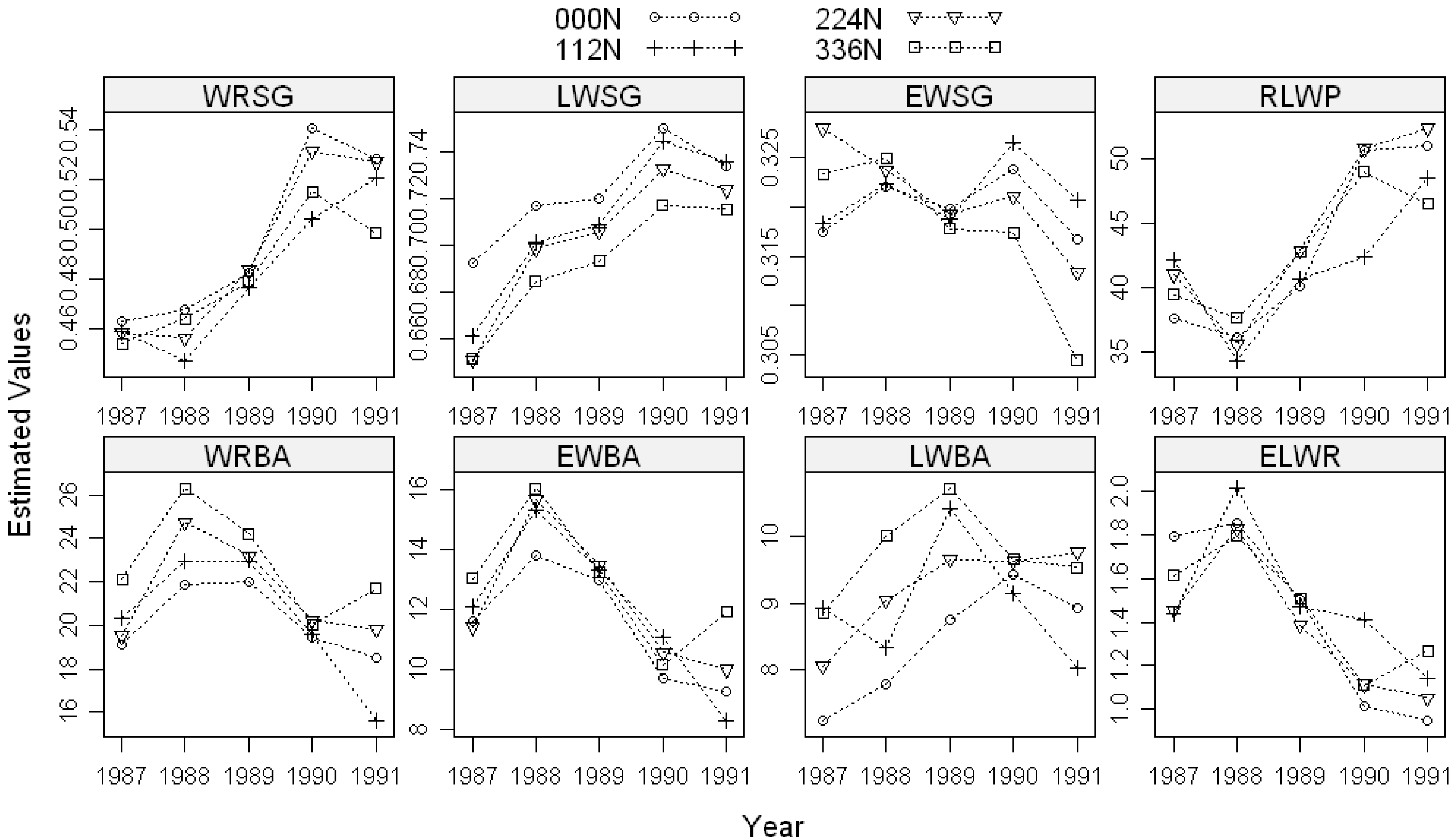

Figure 5.2. Plots of estimated growth and wood properties for different levels of nitrogen application by year at Bertie.

from the control (an increase of 7.9 cm^2/ha with p-value=0.0191). At Bertie, the WRBA of trees which received the 336N treatment increased by 3cm^2/ha compared to the control. Here, an increase of 2.8 and 4.4 cm^2/ha in WRBA was observed for treatments 224N and 336N from control in the second year post fertilization. By the third year, the response in WRBA was limited to the 336N treatment (an increase of 2 cm^2/ha from the control). None of these observed differences were statistically significant at Bertie.

A treatment by time interaction for EWBA was present at New Bern, but was absent at Bertie. At New Bern, significant differences in EWBA was absent in the first and third years following fertilization, but present in the second year post treatment. The EWBA of the 224N and 336N treatments increased significantly (4.1 and 6.9 cm^2/ha respectively) compared to the control (p-values 0.0569 and <0.0001) in the second year following fertilization. The main effect of treatment was significant at New Bern for the 336N treatment compared to the control (p-value=0.0045). At Bertie, an increase of 2.2 cm^2/ha for EWBA was observed for the 336N treatment from the control in the second year after fertilization.

At New Bern, a significant treatment by time interaction was present for LWBA. In the first year following treatment significant increases in LWBA (5 and 6 cm^2/ha) were observed for the 224N (p-value=0.0435) and 336N (p-value=0.0050) treatments compared to the control. In the second year all treatments provided significant increases in LWBA (5, 5.5 and 8.5 cm^2/ha for treatments 112N, 224N and 336N with p-values 0.0368, 0.0144 and <0.0001 respectively). At New Bern, no differences in LWBA were observed among treatments from the third year onwards. The interaction between treatment and time for LWBA was absent at Bertie with an estimated increase of only 2 cm^2/ha observed for the 336N treatment compared to the control in the second and third years post fertilization. The interaction between treatment and time was not found significant for ELWR at both sites.

5.5 Discussion

Midrotation fertilization in combination with thinning is recognized as a beneficial management practice in loblolly pine stands with several studies reporting positive responses in growth following these silvicultural practices (Amateis et al. 1996; Carlson et al. 2008; Haywood 2005; Haywood and Tiarks 2002). However, the comparison of growth and wood property responses to fertilization in thinned and unthinned stands has rarely been made. In this study we observed that the thinned and fertilized stand had a significant increase in basal area growth (both EWBA and LWBA) and a temporary reduction in LWSG and WRSG for two to three years immediately following the application of nitrogen. In the unthinned stand responses to fertilization were observed, but they were not as apparent as in the thinned stand

A general assumption is that treatments which positively affect growth rate will decrease ring SG (Jokela et al. 2004). In this study, both sites responded in accordance with this assumption. However, the magnitude of decrease in LWSG and increase in WRBA, EWBA and LWBA was larger for the thinned and fertilized site (New Bern) compared to Bertie which was only fertilized. The difference in response presumably is due to the additive effects of thinning and fertilization at New Bern. However, it is difficult to make definitive conclusions about this result because of differences in stand location, climatic conditions, and age. The lack of fertilization response on the unthinned site at Bertie suggests that that stand may not have been in a condition to respond to fertilization. High stand density levels in the unthinned stand may have restricted crown expansion and subsequent wood production (Amateis et al. 1996). Midrotation fertilization if often prescribed with thinning for this reason (Fox et al. 2007).

The response following silvicultural treatments can be classified as type A: a long term response following the application of a limiting resource and type B: a short term response following resource application (Nilsson and Allen 2003). Based on the results presented here, it is evident that the response in SG and basal area following midrotation fertilization is transient in nature and typical of a type B response. Responses for WRSG (not statically significant) and LWSG lasted for a maximum of 2-3 years following fertilization, especially in the thinned stand (Figure 5.1). A decrease in WRSG and LWSG was present in the unthinned stand, but was smaller in magnitude compared to the thinned site (Figure 5.2). A similar response was also observed in growth characteristics (Figure 5.1 and 5.2), which agrees with findings for loblolly pine (Antony et al. 2009; Love-Myers et al. 2009) and radiata pine (*Pinus radiata* D. Don) (Nyakuengama et al. 2002; Nyakuengama et al. 2003) where a decrease in WRSG and LWSG was observed in rings immediately following fertilization.

Our observations support the conclusion that the magnitude and duration of the response depends on the amount of N applied (Amateis et al. 2000) (Figure 5.1 and 5.2). The decline in LWSG for the fertilized trees compared to the control in the first year following fertilization was in the following order: 336N>224N>112N, with estimated values of 0.062, 0.043, and 0.025 respectively in the thinned stand and 0.041, 0.041, and 0.031 respectively in the unthinned stand. The response lasted for approximately three years following fertilization in trees which received 336N and slowly converged to the LWSG profile of control trees (Figure 5.1) in the thinned stand.

In a loblolly pine tree, earlywood (springwood) and latewood (summerwood) constitute an annual ring. The transition from earlywood to latewood largely depends upon changes in the concentration of the growth hormone auxin (Larson et al. 2001). Latewood formation commences when height growth (shoot elongation) ceases and new needles become mature allowing a large amount of photosynthetic material to become available for secondary wall thickening (Megraw 1985). Fertilization in a midrotation stand following thinning generally results in increased foliar growth with increased auxin production and a subsequent reduction in the availability of photosynthate for secondary cell wall thickening. An increase in the number of cells produced (increased auxin production) and a decrease in wall thickness (reduced photosynthate availability) might explain the temporary reduction in LWSG and subsequent reduction in WRSG with corresponding growth increases (WRBA, EWBA and LWBA) following midrotation fertilization.

In summary, we evaluated the effect of midrotation fertilization on growth and wood properties in thinned and unthinned loblolly pine stands. A temporary reduction in LWSG and WRSG was observed for two to three years immediately following the application of nitrogen in the thinned stand, but was not as apparent in the unthinned stand. WRBA, EWBA and LWBA also showed similar behavior with a distinct response (increase) in the thinned stand only. ELWR in both thinned and unthinned stands did not change significantly following midrotation fertilization.

5.6 Acknowledgements

The Wood Quality Consortium, University of Georgia, Athens, GA, USA funded this project. The USDA Forest Service, Athens, GA, USA provided valuable support in conducting this study. The wood samples were collected under a grant from the US Department of Energy. The Forest Nutrition Cooperative and Weyerhaeuser Company provided background information and access to the fertilization studies.

5.7 References

Albaugh, T.J., H.L. Allen, P.M. Dougherty, and K.H. Johnsen. 2004. Long term growth responses of loblolly pine to optimal nutrient and water resource availability. Forest Ecology and Management 192:3-19.

Albaugh, T.J., H.L. Allen, and T.R. Fox. 2007. Historical patterns of forest fertilization in the southeastern United States from 1969 to 2004. Southern Journal of Applied Forestry 31(3):129-137.

Allen, H.L., T.R. Fox, and R.G. Campbell. 2005. What is ahead for intensive pine plantation silviculture in the South? Southern Journal of Applied Forestry 29(2):62-69.

Amateis, R.L., J. Liu, M.J. Ducey, and H.L. Allen. 2000. Modeling response to midrotation nitrogen and phosphorous fertilization in loblolly pine plantation. Southern Journal of Applied Forestry 24(4):207-212.

Amateis, R.L., P.J. Radtke, and H.E. Burkhart. 1996. Growth and yield of thinned and unthinned plantations. Journal of Forestry 94(12):19-23.

Antony, F., L. Jordan, R.F. Daniels, L.R. Schimleck, A. Clark, and D.B. Hall. 2009. Effect of midrotation fertilization on growth and specific gravity of loblolly pine. Canadian Journal of Forest Research 39(5):928-935.

Burton, J.D., and E. Shoulders. 1974. Fast-grown, dense loblolly pine saw logs: A reality. Southern Journal of Applied Forestry 72(10):637-641.

Carlson, C.A., T.R. Fox, H.L. Allen, and T.J. Albaugh. 2008. Modeling mid-rotation fertilizer responses using the age-shift approach. Forest Ecology and Management 256(3):256-262.

Fox, T.R., H.L. Allen, T.J. Albaugh, R. Rubilar, and C.A. Carlson. 2007. Tree nutrition and forest fertilization of pine plantations in the southern United States. Southern Journal of Applied Forestry 31(1):5-11.

Haywood, J.D. 2005. Influence of precommercial thinning and fertilization on total stem volume and lower stem form of loblolly pine. Southern Journal of Applied Forestry 29(4):215-220.

Haywood, J.D., and A.E. Tiarks. 2002. Response of second-rotation southern pines to fertilizer and planting on old beds fifteenth-year results. P. 497-502 *in* Eleventh Biennial Southern Silvicultural Research Conference, Outcalt, K. (ed.). U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC.

Jokela, E.J., P.M. Dougherty, and T.A. Martin. 2004. Production dynamics of intensively managed loblolly pine stands in the southern United States: a synthesis of seven long-term experiments. For. Ecol. Manage. 192:117-130.

Larson, P.R., D.E. Kretschmann, A. Clark, and J.G. Isebrands. 2001. Formation and Properties of Juvenile Wood in Southern Pines A Synopsis, USDA Forest Service, General Technical Report FPL-GTR-129 42 p.

Love-Myers, K.R., A. Clark Iii, L.R. Schimleck, E.J. Jokela, and R.F. Daniels. 2009. Specific gravity responses of slash and loblolly pine following mid-rotation fertilization. Forest Ecology and Management 257(12):2342-2349.

Megraw, R.A. 1985. Wood quality factors in loblolly pine. TAPPI Press, Atlanta, GA.

Mora, C.R. 2003. Effects of early intensive silviculture on wood properties of loblolly pine (MS Thesis), North Carolina State University, Raleigh,North Carolina. 92 p.

Nilsson, U., and H.L. Allen. 2003. Short- and long-term effects of site preparation, fertilization and vegetation control on growth and stand development of planted loblolly pine. Forest Ecology and Management 175(1-3):367-377.

Nyakuengama, J.G., G.M. Downes, and J. Ng. 2002. Growth and wood density responses to later-age fertilizer application in *Pinus radiata*. IAWA Journal 23(4):431-448.

Nyakuengama, J.G., G.M. Downes, and J. Ng. 2003. Changes caused by midrotation fertilizer application to the fiber anatomy of *Pinus radiata* IAWA Journal 24(4):397-409.

SAS, I.I. 2004. SAS OnlineDoc Version 9.1.3 Cary, NC: SAS Institute Inc.

Schultz, R.P. 1997. Loblolly Pine : The Ecology and Culture of Loblolly Pine (*Pinus taeda* L.). Agriculture Handbook No. 713. USDA Forest Service, Washington DC 493 p.

Tasissa, G., and H.E. Burkhart. 1998. Modeling thinning effects on ring specific gravity of loblolly pine (Pinus taeda L.). Forest Science 44(2):212-223.

Zobel, B.J., G.F. Groggans, T.E. Maki, and F. Henson. 1961. Some effects of fertilizers on wood properties of loblolly pine. TAPPI 44(3):187-191.

Chapter 6

Modeling the effect of midrotation fertilization on specific gravity of loblolly pine (*Pinus taeda* L.)

6.1 Abstract

Ring specific gravity, earlywood and latewood specific gravity and percent latewood were measured on cores collected at breast height from a thinned and fertilized midrotation loblolly pine (*Pinus taeda* L.) plantation in the lower Coastal Plain of North Carolina. The study was laid out in a randomized complete block design receiving four levels of nitrogen fertilizer in 1984: Control- 000, 112, 224, and 336 kg/ha plus 28 kg/ha of phosphorous with each treatment. A consistent pattern of response was observed in latewood specific gravity (LWSG) following the application of different levels of fertilizer and used as a variable for modeling. The LWSG profiles of unfertilized trees followed a nonlinear relation with ring number from pith. A three parameter asymptotic function was used to explain the LWSG profile of unfertilized trees with ring number as a covariate. Application of N reduced LWSG and was modeled using a two parameter response function with year since fertilization as a covariate and separate parameter estimates for each fertilization level. Based on the model, the magnitude of maximum response was -0.025, -0.049, and -0.074 attained at 3.7, 1.9, and 0.8 years after fertilization for the 112, 224, and 336 kg/ha treatments respectively.

Key Words: Wood density, wood properties, repeated measure, nonlinear model, mixed effect model

6.2 Introduction

Midrotation fertilization is a widely used management practice in pine plantations of the southeastern United States of America. Based on a recent Forest Nutrition Cooperative (FNC) report approximately 0.6 million ha of southern pine plantations have received midrotation fertilization (FNC 2006). The widespread adoption of this practice can be attributed to increased wood production in both biologically and financially attractive ways.

Fertilization at midrotation, especially following precommercial thinning, in loblolly pine (*Pinus taeda* L.) stands has been found to have a strong positive influence on volume production. Data from various field trials established by the FNC have found that over 85 % of fertilized stands responded to a combination of nitrogen (N) and phosphorous (P) fertilization (one-time application of 224 kg/ha N and 28 kg/ha P) with an average growth gain of 30 % over a 6-year period (FNC 2006). There are also various reports that fertilization combined with other practices (e.g. thinning, vegetation control etc.) can considerably increase the volume production of loblolly pine (Borders and Bailey 2001; Haywood 2005). Due to the importance of fertilization, response models were proposed to describe the effect of midrotation fertilization on the productivity of loblolly pine plantations and incorporated them into available growth and yield prediction systems (Martin et al. 1999; Amateis et al. 2000).

Wood properties, such as specific gravity (SG), latewood and earlywood SG, percent latewood, microfibril angle, modulus of elasticity, and modulus of rupture can all be potentially affected by fertilization. Of these, SG is considered as a surrogate of wood quality in various studies because of its high correlation with strength and stiffness of wood products and pulp yield (Panshin and de Zeeuw 1980). Several studies have found that the SG of growth rings produced immediately following midrotation fertilization decreased (Williams and Hamilton

1961; Zobel et al. 1961; Mallonee 1975; Morling 2002; Antony et al. 2009a, 2009b). However, the effect is transient and the reduction in SG is observed only for a short period of time after fertilization (reports vary from 2-5 years) before reverting back to values similar to unfertilized trees (Morling 2002; Nyakuengama et al. 2002; Antony et al. 2009a, 2009b). The change in ring SG is largely related to changes in latewood SG as observed by Clark et al. (2004) who observed a decrease in latewood SG following annual fertilization in a12-year-old loblolly pine plantation. Researchers have also reported a reduction in percent latewood for a few years following fertilization (Williams and Hamilton 1961, Clark et al. 2004). No specific pattern of change due to fertilization was present for earlywood SG.

Attempts to model wood property responses to midrotation fertilization are lacking from the literature. Considering the wide-spread adoption of midrotation fertilization in loblolly pine plantations, an understanding of wood property responses and the modeling of this response is of primary importance. The objectives of the present paper were two fold: 1) to model the response to mid-rotation fertilization on ring SG (RSG), latewood SG (LWSG), earlywood SG (EWSG) and percentage latewood (RLP) which will be operationally useful; 2) to test the effect of fertilization on these wood properties based on the modeled response profile.

6.3 Materials and methods

6.3.1 Sample origin

The study was conducted on wood samples collected from an even-aged loblolly pine plantation planted in 1970 (1482 tree per ha) at New Bern, North Carolina in the lower Coastal Plain. This was one of the 19 installations of the FNC Region-wide 13 study established across the southeast US in site-prepared loblolly pine stands between 1984-1987 (FNC Report No: 39 1997). The experimental design was a randomized complete block with four treatments replicated on four blocks (total of 16 plots). The treatments used in this study were control – no Nitrogen (000N), 112 (112N), 224 (224N), and 336 (336N) kg/ha with all treatments receiving 28 kg/ha of phosphorous. The stand was thinned to 605 trees per ha in 1983 and treated with the different rates of N fertilizer in March, 1984. All plots were thinned to 346 trees per ha in 1995 and received a second fertilizer application of 224 kg/ha N in 1996. The stand was harvested in 2003.

From each plot, nine trees were sampled in proportion to the diameter distribution of the trees in each plot. Increment cores (12 mm in diameter) were collected from the sampled trees at breast height (1.37 m above ground) from each treatment plot using a hydraulically driven increment core borer. Defective, suppressed, or infected trees were excluded from sampling. A total of 144 trees were sampled.

6.3.2 Sample preparation and data collection

Each core was dried at 50^{o}C for approximately 24 hr and glued to custom made core holders. Radial strips of 1.6 mm thick were sawn from these breast height cores and conditioned to 8 % moisture content for approximately 48 hr before scanning. All the radial strips were then read on a scanning X-ray densitometer at a resolution of 0.006 mm to determine earlywood, latewood, and whole ring SG. The densitometry data was also used to determine radial growth and percent of latewood in each annual ring. A SG of 0.48 was used to distinguish between earlywood and latewood (Jordan et al. 2008). SG measurements were based on a green volume and oven dry weight basis. Since all the trees received a second thinning and fertilizer application (224 kg/ha N) in 1995-1996, the data up to 1995 only were used in this modeling work.

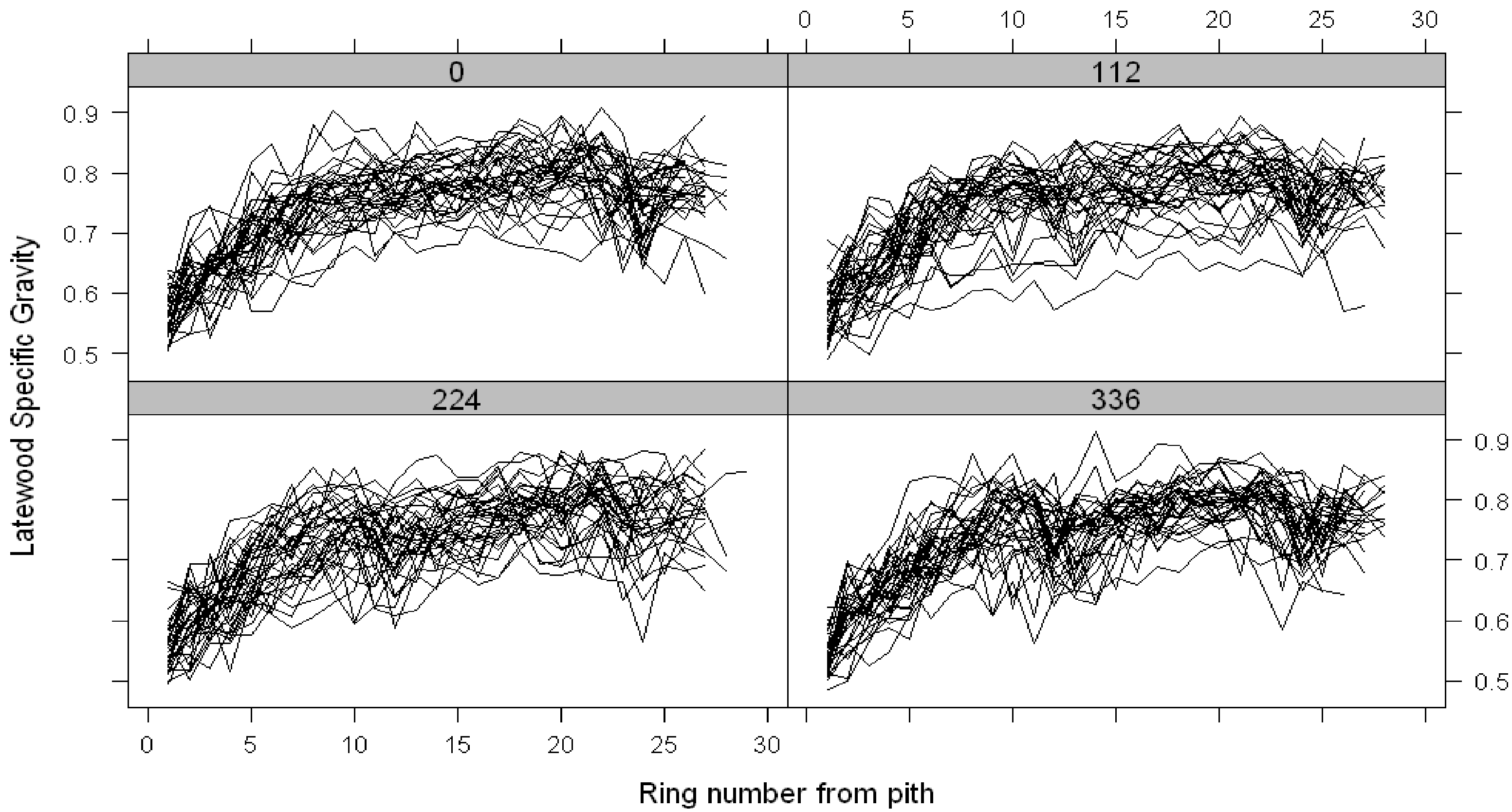

Figure 6.1: Subject specific latewood specific gravity profiles plotted against ring number by treatment group.

6.3.3 Model development

Earlier studies based on this data set reported a significant response in 4-year average RSG and LWSG following fertilization. For more information on these studies the reader is referred to Antony et al. (2009a, 2009b). Even though these studies found significant responses for average RSG and LWSG (4-year post fertilization average), operationally it will be of interest to model the ring-by-ring responses following fertilization in a predictable form. Subsequent analysis of the data demonstrated that the impact of a single dose of fertilizer on RSG at midrotation cannot be modeled successfully. However, we found specific patterns of responses in LWSG following fertilization that can be described by a response model. The effect of single dose midrotation fertilization on other properties such as EWSG and LWP was absent in this study.

LWSG shows a specific pattern with ring number from pith in loblolly pine. It rapidly increases in the first few rings and approaches an upper asymptote. The LWSG profile of individual trees for each treatment are presented in Figure 6.1 and demonstrates that large tree-to-tree variation exists in the LWSG profiles within each treatment group. To take account of the variation at different levels of the design in the modeling process, the hierarchical structure of the data (tree in plot in block) was maintained throughout the model building process via the inclusion of nested random effects.

The model building process involved: 1) modeling LWSG profiles of unfertilized trees (the base model) and 2) adding a response function to the base model to represent the response following application of different levels of nitrogen (the response model). A variety of base models were fitted to the grouped data and compared based on different fit statistics. A three-parameter asymptotic function was selected as the base model to represent the LWSG profiles of unfertilized trees which had the form

$$f(\beta, x) = \beta_1 + (\beta_2 - \beta_1)e^{[-e^{\beta_3} x]} \quad [1]$$

where x was the ring number from pith, β_1 represents the asymptote as x approaches ∞ and β_2 is LWSG when x is zero and β_3 is the logarithm of rate constant (Ratkowsky 1990).

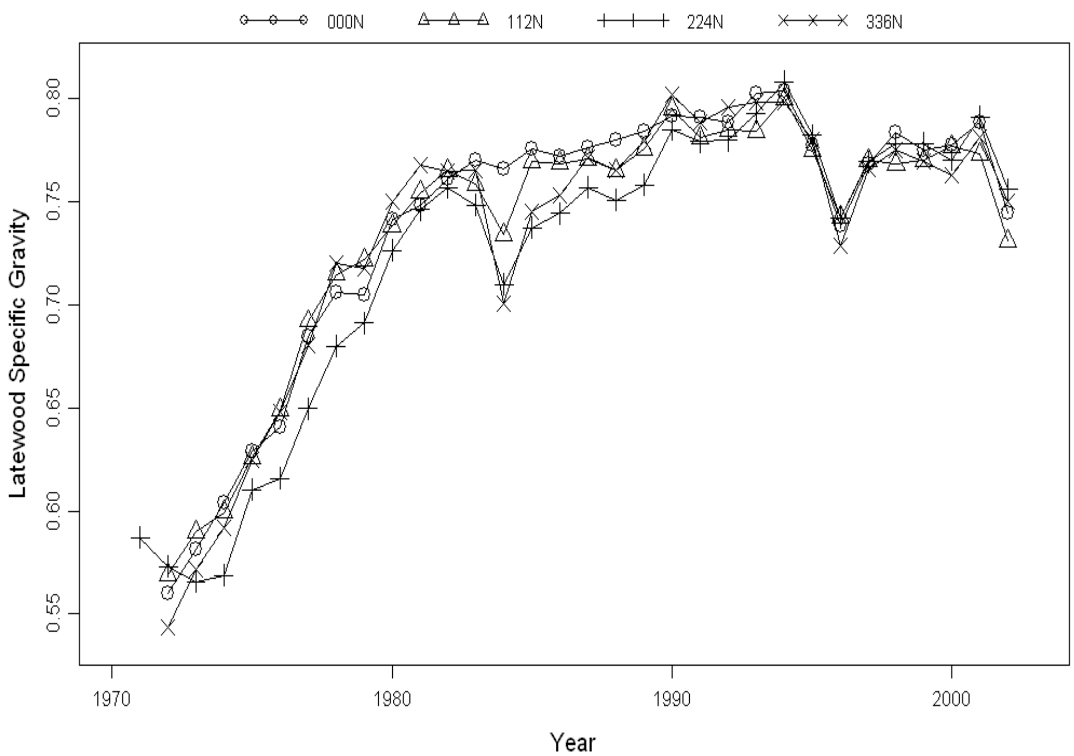

Figure 6.2: Plot of mean latewood specific gravity plotted with year by treatment. The drop in latewood specific gravity at two points of fertilizer application was evident at in 1984 and 1996.

It is evident from the mean plot (Figure 6.2) that immediately after fertilization in year 1984, LWSG dropped considerably, especially in trees which received a higher rate of fertilizer (224N and 336N) and took a few years to revert back to the profile of unfertilized trees. It is also clear from the mean plot that the intensity of response on LWSG depends on the amount of fertilizer applied. So the response model could be represented as a function which relates the drop in LWSG to the amount of fertilizer applied and time since fertilization. Since the response occurred at midrotation age immediately after the LWSG profiles of trees had started to plateau, it will be more meaningful and efficient to model the treatment response as an explicit function of interpretable parameters. To this extent we used a height growth response model proposed by Pienaar and Rheney (1995)

[2] $$f(\beta,t) = (\beta_4.t).e^{[-\beta_5.t]}$$

where '*t*' was time since fertilization, parameters β_4 and β_5 represents the magnitude and pattern of response respectively. These parameters were easily interpretable with $\left(\beta_4 / \beta_5\right)e^{-1}$ represents the magnitude of maximum response attained in $\left(1/\beta_5\right)$ years after treatment.

Parameter β_5 determines the longevity of the response.

Let y_{ijkl} represent the LWSG of *l*th ring from *k*th tree in the *j*th plot of the *i*th block; the nonlinear mixed model can be represented as

[3] $$y_{ijkl} = \beta_{1ijk} + (\beta_{2ijk} - \beta_{1ijk})e^{[-e^{\beta_{3ijk}} x_{ijkl}]} + (\beta_{4ijk}.t_{ijkl}).e^{\left[-\beta_{5ijk}.t_{ijkl}\right]} + \varepsilon_{ijkl}$$

Since the blocks, plots, and trees were considered as random samples taken from large population of blocks, plots, and trees, there is a potential need to account for block-to-block, plot-to-plot, and tree-to-tree heterogeneity. The mixed model framework (Pinheiro and Bates 2000) can be potentially used to accommodate this multilevel heterogeneity by expressing the parameters in the model as mixed effects as represented below

[4] $$\beta_{ijk} = A_{ijk}\beta + B_{ijk,1}b_i + B_{ijk,2}b_{ij} + B_{ijk,3}b_{ijk}$$

Here,

$$A_{ijk} = B_{ijk,1} = B_{ijk,2} = B_{ijk,3} = I_5;$$
$$\beta = \begin{pmatrix}\beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5\end{pmatrix}^T;$$
$$b_i = \begin{pmatrix}b_i^{(1)} & b_i^{(2)} & b_i^{(3)} & b_i^{(4)} & b_i^{(5)}\end{pmatrix}^T; b_{ij} = \begin{pmatrix}b_{ij}^{(1)} & b_{ij}^{(2)} & b_{ij}^{(3)} & b_{ij}^{(4)} & b_{ij}^{(5)}\end{pmatrix}^T \text{ and}$$
$$b_{ijk} = \begin{pmatrix}b_{ijk}^{(1)} & b_{ijk}^{(2)} & b_{ijk}^{(3)} & b_{ijk}^{(4)} & b_{ijk}^{(5)}\end{pmatrix}^T$$

where b_i, b_{ij} and b_{ijk} are the block, plot, and tree level random effects, $B_{ijk,1}, B_{ijk,2}$ and $B_{ijk,3}$ are the corresponding design matrices; β is the fixed-effect parameter vector; and A_{ijk} is the corresponding design matrix, I_5 is a 5 x 5 identity matrix with all the diagonal elements equal to 1.

The random effects and the within-plot error term were assumed to be distributed normally as $b_i \sim N(0, \mathbf{\Psi}_1)$; $b_{ij} \sim N(0, \mathbf{\Psi}_2)$; $b_{ijk} \sim N(0, \mathbf{\Psi}_3)$; and $\varepsilon_{ijkl} \sim N(0, \sigma^2 \Lambda_{ijk})$. Here $\mathbf{\Psi}_1$, $\mathbf{\Psi}_2$ and $\mathbf{\Psi}_3$ were variance-covariance matrices representing different levels of block, plot and tree random effects. A full model with random effects associated with all the parameters in the model considered first by assuming a diagonal variance-covariance matrix structure for random effects and an independent structure to within tree error. These assumptions were relaxed in the later stages of fitting by assuming different variance-covariance structures for the random effects. Several reduced models were also fitted by dropping the random effect terms associated with the parameters. The best model from these fittings was selected by comparing the fitted models using Akaike's Information Criteria (AIC).

The next step in the model building process was to incorporate any covariates, here the fertilization levels, into appropriate parameters in the model. Since β_4 and β_5 were the parameters associated with the fertilization response, these parameters were expressed in such a way that they can accommodate the influence of different levels of fertilization (112N, 224N and 336N). The second part, the response model, in Eq. 3 will be zero for trees in the control group and for trees before fertilization because the variable time since fertilization will carry a value of zero. It is also reasonable to assume that the parameters, β_4 and β_5, will be equal to zero for trees in the control group. Treating the fertilization levels as factors, the parameters β_4 and β_5 represented using dummy variable as β_{41} *I(112N)* + β_{42} *I(224N)* + β_{43} *I(336N)* and β_{51} *I(112N)* + β_{52} *I(224N)* + β_{53} *I(336N)*, where *I(112N)* = 1 if fertilized with 112kg/ha of nitrogen, 0 otherwise; *I(224N)*, and *I(336N)* are defined similarly.

After formulating the appropriate mean model and the random effect structure, the independent matrix structure associated with the within-tree error was relaxed. This will enable us to explain the heteroscedasticity in the data and serial correlation across measurements successfully. Different variance functions usually used in the growth modeling such as the power model $\left(Var(\varepsilon_{ijkl}) = \sigma^2 \left|\upsilon_{ijkl}\right|^{2\delta}\right)$, the exponential model $\left(Var(\varepsilon_{ijkl}) = \sigma^2 e^{2\delta\upsilon_{ijkl}}\right)$ and the constant power model $\left(Var(\varepsilon_{ijkl}) = \sigma^2\left(\delta_1 + \left|\upsilon_{ijkl}\right|^{\delta_2}\right)^2\right)$ were used to define any nonconstant variance within data with respect to the covariate υ_{ijkl}. The data collected represent SG measurements taken from increment cores collected at breast height of a tree. The data were collected from pith-to-bark and represent changes in SG over time and we can reasonably expect that correlation exists among the measurements. It is also reasonable to assume that observations closer together will tend to be more alike than observations farther apart. If correlation among the repeated measurements exists, its autocorrelation pattern can be modeled with an appropriate correlation model. The autoregressive models (AR(*p*)), moving-average models (MA(*q*)), and autoregressive with moving average models (ARMA(*p*,*q*)) were used with the data to account for dependence across repeated measurements within each tree. AIC criterion was used for checking significant changes in performance of the models. The nonlinear mixed models were implemented using the nlme package available in R (Pinheiro et al. 2009).

6.4 Results

Models with different combinations of random and fixed effects for parameters β_1, β_2, β_3, β_4, and β_5 were fitted on Eq. 3. After several trials, the model with plot-level random effects on parameters β_3 and β_4 and tree-level random effects on parameters β_1, β_2, β_3, and β_5 was selected as the best random effect model (AIC = -11566.12).

After identifying the random effect component, the parameters in the response function (ie β_4 and β_5) were expanded with indicator variables for each fertilization level (AIC = -11574.31). Based on a likelihood ratio test conducted between the full model (with treatment terms in it) and the reduced model (with no treatment terms), the treatment terms were found to be significant (LR = 16.19 with 4 df, p-value = 0.0028). All the plot and tree level random effects identified above were kept in the model at this stage. Before conducting any hypothesis tests on treatment level parameters, it is important to identify any violation of the constant variance assumption and autocorrelation across measurements.

Heterogeneity of residuals (here the residuals will be the difference between the observed and fitted specific gravity values at any point in time conditional on the best linear unbiased predictor of random effects) can be detected from plots of standardized residuals with fitted values or with covariates in the model (not presented here). We relaxed the homogeneous variance assumption of the model and sought significant improvements in the model fitting criteria. Based on the AIC criterion, the best model was one having an exponential variance function with ring number as a covariate (AIC = -11675.1).

The data collected here represent SG measurements taken over time from increment cores collected at breast height of a tree and we expect that correlation exists among the measurements. It is reasonable to assume that observations closer together will tend to be more alike than observations farther apart. We further relaxed the within residual structure of the current model, where the observation taken at two time points within an individual were considered independent of each other, by assuming different correlation structures. Based on the AIC criterion, an auto regressive moving average with (1, 1) order (ARMA (1, 1)) was selected as the best model to represent the within residual correlation (AIC = -11874.9).

The plot-level random effect associated with the parameter β_4 (AIC = -11876.94) and tree-level random effect associated with parameter β_2 (AIC = -11878.92) was found to be very small and dropped from further model building.

The objective of this study was to find a model that explained the changes in LWSG following midrotation fertilization. Based on the fitted model, all parameters in the base function were found to be significant (p-value <0.0001). A test of whether response model parameters were equal to zero (β_{41}=0; β_{42}=0; β_{43}=0; β_{51}=0; β_{52}=0; and β_{53}=0) i.e. different from control profiles, was conducted. Based on the final fitted model, all the parameters in the response model were significantly different from zero with critical value of 0.05. These tests indicate that the parameters were different from $\beta_{40} = 0$ and $\beta_{50} = 0$, the expected value of the parameter for control trees. A question of interest here is whether the response profiles of trees receiving treatments 112N, 224N, and 336N are different or not i.e. H_0: $\beta_{41} = \beta_{42} = \beta_{43}$ and $\beta_{51} = \beta_{52} = \beta_{53}$. It was found that the parameters in the response model were different from treatment to treatment (p-value < 0.0001 based on a 2 df F-test,). Based on the subsequent tests conducted, it was found that the response profile of 112N treatment was different from control, 224N, and

336N (p-value < 0.0001). It was also found that the 224N, and 336N treatments were significantly different from the control and that the response profile of these treatments were different from each other. Since response profiles of all treatments were found to be significantly different, the full mean structure with separate parameters for 112N, 224N, and 336N treatments was maintained.

Table 6.1: The estimated fixed effect parameters and the variance components for the nonlinear mixed effect model (denominator df = 2954 used for the t-test).

Parameter	Estimate	SE	t	p-value
β_1	0.8099	0.0039	206.25	<0.0001
β_2	0.5062	0.0052	97.31	<0.0001
β_3	-1.7180	0.0451	-38.09	<0.0001
β_{41}	-0.0179	0.0043	-4.21	<0.0001
β_{42}	-0.0721	0.0093	-7.73	<0.0001
β_{43}	-0.2380	0.0307	-7.75	<0.0001
β_{51}	0.2669	0.0534	4.99	<0.0001
β_{52}	0.5388	0.0564	9.55	<0.0001
β_{53}	1.1848	0.0999	11.86	<0.0001
$\sigma^2_{b^{(3)}_{ij}}$	0.0071			
$\sigma^2_{b^{(1)}_{ijk}}$	0.0011			
$\sigma^2_{b^{(3)}_{ijk}}$	0.0364			
$\sigma^2_{b^{(5)}_{ijk}}$	0.0301			
σ^2	0.0022			

After appropriately specifying the random effects, fixed effects, within residual covariance and correlation structure, the model with plot-level random effect $b^{(3)}_{ij}$ (on parameter β_3) and tree-level random effects $b^{(1)}_{ijk}, b^{(3)}_{ijk}$ and $b^{(5)}_{ijk}$ (on parameters β_1, β_3 and intercept of β_5) with an exponential function of ring number as the variance structure and an ARMA (1,

1) correlation structure was selected as the final model. The final model can be represented as follows:

[6] $$y_{ijkl} = f(x_{ijkl}, t_{ijkl}, \beta_{ijk}) + \varepsilon_{ijkl}$$

$$y_{ijkl} = \beta_{1ijk} + (\beta_{2ijk} - \beta_{1ijk})e^{\left[-e^{\beta_{3ijk}} x_{ijkl}\right]} + (\beta_{4ijk}.t_{ijkl}).e^{\left[-\beta_{5ijk}.t_{ijkl}\right]} + \varepsilon_{ijkl}$$

$$\beta_{1ijk} = \beta_1 + b_{ijk}^{(1)}$$

$$\beta_{2ijk} = \beta_2$$

$$\beta_{3ijk} = \beta_3 + b_{ij}^{(3)} + b_{ijk}^{(3)}$$

$$\beta_{4ijk} = \beta_{41} I(112N) + \beta_{42} I(224N) + \beta_{43} I(336N)$$

$$\beta_{5ijk} = \beta_{51} I(112N) + \beta_{52} I(224N) + \beta_{53} I(336N) + b_{ijk}^{(5)}$$

$$\varepsilon_{ijkl} \sim N(0, \sigma^2 \Lambda_{ijk})$$

$$\sigma^2 \Lambda_{ijk} = \sigma^2 G_{ijk}^{1/2}\left(x_{ijkl}, \delta\right) \Gamma_{ijk}(\phi, \theta) G_{ijk}^{1/2}\left(x_{ijkl}, \delta\right)$$

$$G_{ijk}\left(x_{ijkl}, \delta\right) = e^{2\delta . x_{ijkl}}$$

$$\Gamma_{ijk}(\phi, \theta) = ARMA(1,1)$$

$$b_i = 0$$

$$b_{ij} = b_{ij}^{(3)}$$

$$b_{ijk} = \left(b_{ijk}^{(1)}, b_{ijk}^{(3)}, b_{ijk}^{(5)}\right)^T$$

where $G_{ijk}\left(x_{ijkl}, \delta\right)$ is the variance function with a parameter δ (which has an estimated value of -0.02254) and $\Gamma_{ijk}(\phi, \theta)$, the serial correlation function (which has estimated parameters ϕ = 0.6827 and θ = -0.3747). All other parameters were defined previously.

The estimated parameters, corresponding standard errors and p-values for the fixed effects of the model are presented in Table 6.1. A plot of mean predicted latewood specific gravity under different fertilization regimes for a hypothesized tree with 30 rings was produced by assuming random effects estimates for $b_{ij}^{(3)}$, $b_{ijk}^{(1)}$, $b_{ijk}^{(3)}$ and $b_{ijk}^{(5)}$ as zero and using the estimates for fixed effects in Eq. 6 (Figure 6.3). Finally a plot of population level and subject specific prediction for randomly chosen subjects within each treatment is presented in Figure 6.4.

6.5 Discussion

The operational application of N and P at midrotation is an accepted means of increasing productivity and economic return from loblolly pine plantations. Depending on the level of fertilizer applied there is generally a significant growth response which can be influenced by a number of factors including the geographic location of the stand, stand age, whether the stand was thinned or not, soil chemical and physical properties, the availability of soil moisture and climatic conditions (Amateis et al. 2000). Typically wood properties are also affected by midrotation fertilization, but compared to research conducted on growth responses to midrotation fertilization have received little attention. Generally SG and stiffness are decreased and MFA increased by midrotation fertilization (Antony et al. 2009a), these changes are all detrimental to wood quality in general and can influence end product quality. Hence the impact of midrotation fertilization on wood properties is of concern to both wood growers and wood buyers. Quantitative models that adequately represent the response in wood properties are important for maximizing product categorization and utilization efficiency and also for understanding how wood properties may be influenced by differing levels of fertilizer. The purpose of this study was to model the response of ring LWSG (one of the most important wood properties in determining overall wood quality) following the midrotation application of N fertilizer based on samples collected from a stand located in the lower Coastal Plain of North Carolina. Our approach was to first model the overall trend in LWSG from pith-to-bark (the base model) and then model the response to fertilization (the response model).

A 3-parameter asymptotic model was used to explain the trend in LWSG with ring number from pith-to-bark (the base model). Nonlinear models are more appropriate for describing this process because of their parsimonious nature, flexibility of generated curve shapes, and interpretability of parameters. Apart from the nonlinearity of the modeled process, the data has multiple levels of heterogeneity among blocks, plots and trees and correlation of observations measured within a tree. The mixed effect modeling approach provides a flexible choice for fitting models with parameters having both fixed and random effects and with an array of within subject error structures (covariance and serial correlation). The advantages of such models is their ability to describe the mean structure of processes after taking account of the different levels of heterogeneity (block-to-block, plot-to-plot and tree-to-tree) in the data through random effects (Lindstrom and Bates 1990) and the serial correlation of measurements taken through time using appropriate variance-covariance structures. Using a model of the form represented in eq.6, we defined the changes in LWSG after taking account of the different sources of variation expected from the hierarchical structure of the experimental design (block /plot/tree) and within-tree error structure.

The 3-parameter asymptotic model used to represent change in LWSG of unfertilized trees is a continuous nonlinear model with the following parameter interpretations: β_1 (0.8099) representing the upper maximum for SG; β_2 (0.5062) representing SG of the first growth year, and β_3 (-1.7180) representing the log of rate of change in SG with time. The heterogeneity in the data was accounted for by tree-level random effects associated with parameters β_1 and β_5 and plot and tree-level random effects associated with parameter β_3. The estimated values are presented in Table 1 and found to be relatively high in magnitude. This indicates high tree-to-tree variability in LWSG profiles (through β_1 and β_3) and in the response (through β_5). This is evident from Figure 6.1.

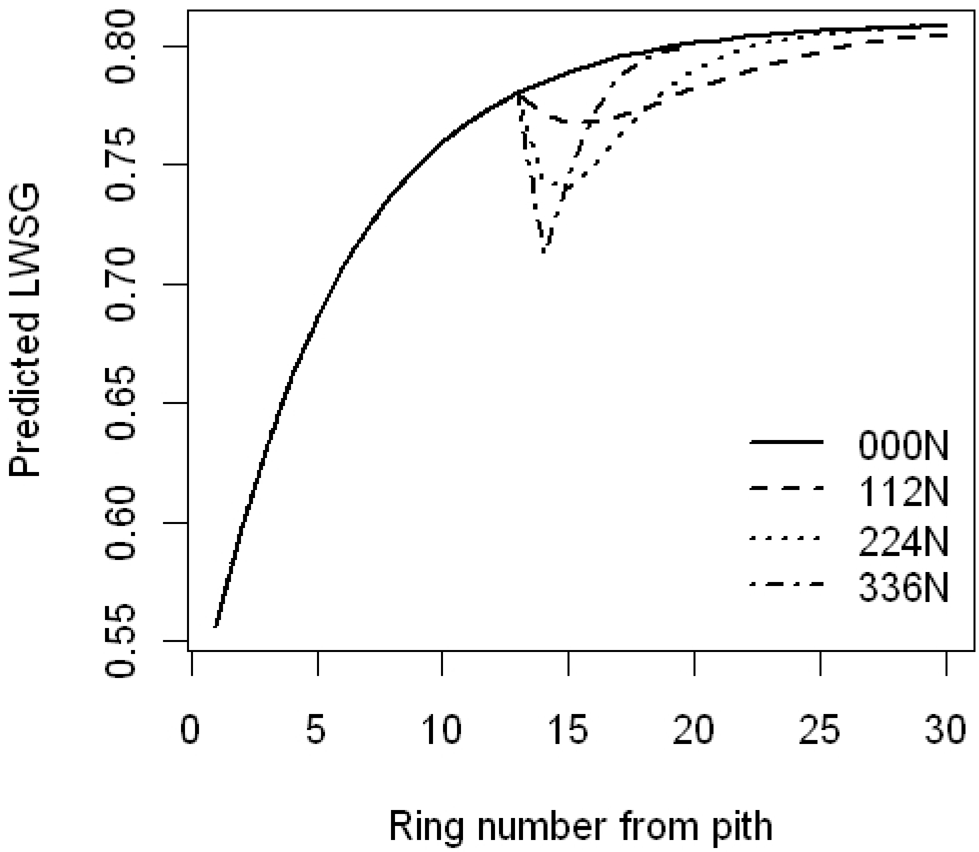

Figure 6.3. Plot of the mean predicted latewood specific gravity for each treatment group from the final model.

Martin et al. (1999) and Amateis et al. (2000) modeled the response of growth characteristics (dominant height, basal area and volume) following fertilization using varying forms of the response function proposed by Pienaar and Rheney (1995). In this study, we used a similar response function to that proposed by Pienaar and Rheney (1995) to describe the response in LWSG following fertilization. The proposed response model allowed us to successfully identify any significant difference among mean response profiles to different nitrogen levels with time. Addition of parameters to represent treatment levels showed that the response following the 112N treatment was smaller than the response observed for the 224N and 336N treatments (Figure 6.3). The mean response profile of trees following the application of 112N, 224N and 336N was found to be significantly different from that of the control. It is possible to find the magnitude of maximum mean decrease in LWSG and years taken after time of treatment to attain the maximum response from the estimated response function parameters. Based on the estimated parameters, the magnitude of maximum response was -0.025, -0.049, and -0.074 attained at 3.7, 1.9, and 0.8 years after fertilization for the 112N, 224N, and 336N treatments respectively. Based on this model, the magnitude of maximum decrease in LWSG for the 112N treatment is very small and is attained only 3.7 years after fertilization. However, the magnitude of maximum response increases and time to attain that response decreases with an increase in the rate of N applied (Figure 6.3).

The effect of midrotation fertilization on SG is considered to be transient in nature and may last for a few years following treatment (Posey 1964; Ross et al. 1979). This is evident from our response model and Figure 6.3 where the LWSG profile of trees decreases immediately after fertilization and returns to normal (i.e. similar to the control) after a few years. The sudden decrease and the subsequent return to levels similar to the control depend upon the amount of

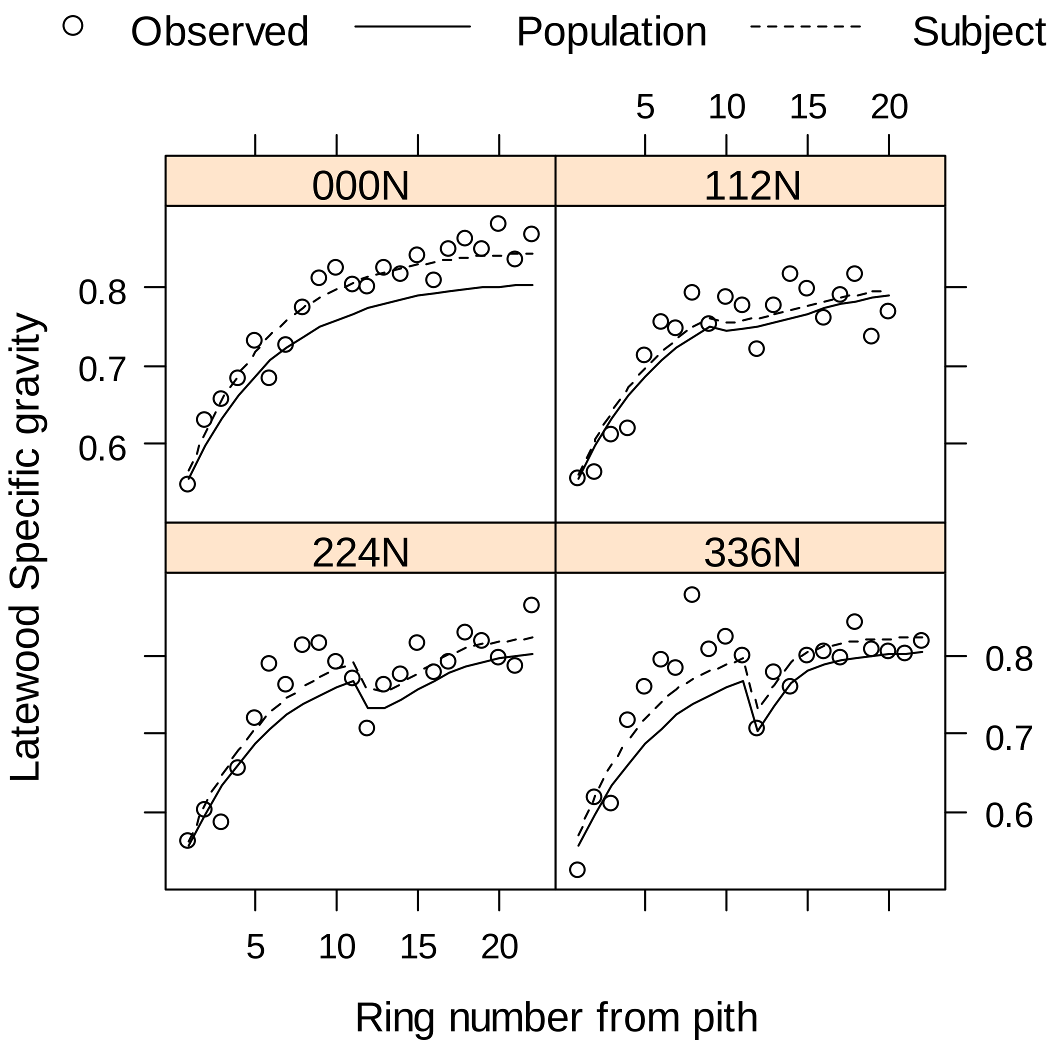

Figure 6.4. Plot of the predicted latewood specific gravity at population level (random effects set to zero) and individual subject level values along with the observed values for randomly selected subjects from each treatment.

fertilizer applied. This conclusion from the model agrees with the findings of Antony et al. (2009a, 2009b) where a significant drop in 4- year post treatment average LWSG was observed for the 224N and 336N treatments compared to the control, but not for the 112N treatment.

Significant change in LWSG following fertilization may impact ring SG considerably and hence the strength of wood produced and the pulp yield. For instance, according to Mitchell (1964) a SG difference of 0.02 units can lead to differences of 22.7 kg (50 lb) in pulp yield per ton of round wood, emphasizing the importance of changes in wood properties (here SG) on pulp productivity. The model presented here can be successfully used to predict the decrease in LWSG following midrotation fertilization. The change in LWSG might be attributed to the collective effects of changes in tracheid properties such as tracheid wall thickness, radial diameter, etc. This agrees with earlier findings where a decrease in density and a corresponding decrease in tracheid wall thickness were reported following midrotation fertilization in radiata pine (*Pinus radiata* D.Don) (Nyakuengama et al. 1991, 2002, 2003), Douglas fir (*Pseudotsuga menziesii* (Mirb.) Franco) (Erickson and Harrison 1974) and in loblolly pine (Clark et al. 2004). This short-term change in SG can be attributed to a large crown response (specifically needle formation) which produces a temporary change in wood formation i.e. a higher earlywood to latewood ratio (Larson et al. 2001).

The present study focused on modeling the mean LWSG trend from pith to bark and then the response of LWSG after midrotation fertilization. The mean LWSG profile of loblolly pine trees was successfully modeled using a 3-parameter asymptotic function. The mean response profile of the treatments (112N, 224N, and 336N) with time was represented using a 2-parameter response function with time since fertilization as a covariate. We found a significant difference in the mean response profiles of fertilized trees compared to that of the control trees. It is important to recognize the fact that trees vary in their response to fertilization. As noted earlier a number of factors influence the response and as a consequence each site can be expected to respond in a slightly different way. Hence the applicability of the models developed in this study is limited as they were based on a single site. However, the approach used in this study could be applied to the development of more general models. Ideally more general models would be available; however, this would require the inclusion of samples from multiple sites and regions.

6.6 Acknowledgements

The Wood Quality Consortium, University of Georgia, Athens, GA, USA funded this project. The USDA-Forest Service, Athens, GA, USA provided valuable support in conducting this study. The wood samples were collected under a grant from the US Department of Energy. The Forest Nutrition Cooperative and Weyerhaeuser Company provided background information and access to the fertilization study.

6.7 References

Amateis, R. L., J. Liu, M. J. Ducey, and H. L Allen.. 2000. Modeling response to midrotation nitrogen and phosphorous fertilization in loblolly pine plantation. Southern Journal of Applied Forestry 24(4):207-212.

Antony, F., L. Jordan, L. R. Schimleck., A. Clark, and R. F. Daniels. 2009a. The effect of mid-rotation fertilization on the wood properties of loblolly pine (*Pinus taeda* L.). IAWA J. 30 (1): 49-58.

Antony, F., L. Jordan, R. F. Daniels, L. R. Schimleck, A. Clark, and D.B. Hall. 2009b. Effect of midrotation fertilization on growth and specific gravity of loblolly pine (*Pinus taeda* L.). Canadian Journal of Forest Research 39(5): 928–935.

Borders, B. E. and R. L. Bailey. 2001. Loblolly pine- pushing the limits of growth. Southern Journal of Applied Forestry 25: 69–74.

Clark, A., B. E. Borders and R. F. Daniels. 2004. Impact of vegetation control and annual fertilization on properties of loblolly pine wood at age 12. Forest Products Journal 54: 90–96.

Erickson, H. D. and T. H. Harrison. 1974. Douglas-Fir wood quality studies part I: Effects of age and stimulated growth on wood density and anatomy. Wood Science and Technolology 8: 207–226.

Forest Nutrition Cooperative (FNC) Report No. 39. 1997. North Carolina Forest Nutrition Cooperative. Department of Forestry. North Carolina State University. Raleigh, North Carolina.

Forest Nutrition Coopertive. 2006. Forest Nutrition Cooperative Synthesis. Available online at http://www.forestnutrition.org/synthesis.htm; last accessed Apr. 11, 2007.

Haywood, J. D. 2005. Influence of precommercial thinning and fertilization on total stem volume and lower stem form of loblolly pine. Southern Journal of Applied Forestry 29: 215–220.

Jordan, L., A. Clark, L. R. Schimleck, D. B. Hall, and R. F. Daniels. 2008. Regional variation in wood specific gravity of planted loblolly pine in the United States. Canadian Journal of Forest Research 38:698-710.

Larson, P. R., D. E. Kretschmann, A. Clark., and J. G. Isebrands. 2001. Formation and Properties of Juvenile Wood in Southern Pines A Synopsis, USDA Forest Service, General Technical Report FPL-GTR-129 42p.

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

Mallonee, E. H. 1975. Effect of thinning and seasonal time of Nitrogen fertilization on the growth of pole sized loblolly pine (MS Thesis). Department of Forestry, North Carolina State University. Raleigh, North Carolina.

Martin, S.W., R. L. Bailey, and E. J. Jokela. 1999. Growth and yield prediction for lower coastal plain slash pine plantations fertilized at mid-rotation. Southern Journal of Applied Forestry 23: 39-45.

Mitchell. H. 1964. Patterns of variation in specific gravity of southern pines and other coniferous species. Tappi 47(5): 276-283.

Morling, T. 2002. Evaluation of annual ring width and ring density development following fertilization and thinning of scots pine. Annals of Forest Science 59: 29–40.

Nyakuengama, J. G. 1991. The physical and chemical properties of superior radiate pine (*Pinus radiata* D. Don) wood for thermo-mechanical pulping compared with standarad radiate pine. (B. Sc. Hons. Thesis). Department of Forestry, Australian National University. Canberra, Australia.

Nyakuengama, J. G., G. M. Downes and J. Ng. 2002. Growth and wood density responses to later-age fertilizer application in *Pinus radiata.* IAWA J. 23: 431–448.

Nyakuengama, J. G., G. M. Downes, and J. Ng. 2003. Changes caused by midrotation fertilizer application to the fiber anatomy of *Pinus radiata*. IAWA J. 24: 397–409.

Panshin, A.J., and C. deZeeuw. 1980. *Textbook of Wood Technology*. McGraw-Hill Book Co., New York. 705 p.

Pienaar, L. V. and J. W. Rheney. 1995. Modeling stand level growth and yield response to silvicultural treatments. Forest Science 41: 629–638.

Pinheiro, C. J. and M. D. Bates. 2000. Mixed-Effects Models in S and S-PLUS. Springer. 528 p.

Pinheiro, C. J., M. D. Bates, S. DebRoy , D. Sarkar, and the R Core team. 2009. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-92.

Posey, C. E. 1964. The effect of fertilization upon wood properties of loblolly pine. Technical Report No. 22. School of Forestry. N. C. State University. Raleigh, N.C. 63 p.

Ratkowsky, D. A. 1990. *Handbook of nonlinear regression models*. Mercel Dekker, New York. 237 p.

Ross, D., E. Buckner., H. Core and F. Woods. 1979. Nitrogen fertilization decreases yellow poplar wood density. Southern Journal of Applied Forestry 3: 119–122.

Williams, R. F. and J. R. Hamilton. 1961. The effect of fertilization on four wood properties of slash pine. Journal of Forestry 59: 662–665.

Zobel, B. J., G. F. Groggans, T. E. Maki, and F. Henson. 1961. Some effects of fertilizers on wood properties of loblolly pine. Tappi Journal 44(3):187–191.

Chapter 7

Summary and Conclusions

Wood quality is a multifaceted term with no concrete definition. Generally, wood quality is defined in terms of the quality of end product to which a piece of wood is put to use. Thus wood quality is a cumulative expression of anatomical, physical and mechanical properties of a piece of wood for a particular end product. Identifying and measuring these properties are important for proper utilization of timber. Loblolly pine (*Pinus taeda* L.) is a tree crop growing throughout the southeastern United States. Large variation in loblolly pine wood properties occurs across growing regions, between stands within a region, among trees within stands, within the tree and within annual rings. For forest products manufacturer, an understanding of variation in wood properties is important for efficient raw material segregation and optimization of manufacturing processes. The present dissertation was a comprehensive effort to model the variation in wood specific gravity (SG), imposed by both natural and/or silvicultural causes, in planted loblolly pine from the southern United States.

SG is considered as an important wood quality measure because of its high correlation with the strength of solid wood products and the yield and quality of pulp produced. Disk SG and moisture content (MC) were measured at different height levels from sampled trees growing in plantations across loblolly pine's natural range. The development and use of models for explaining the variation in disk SG and MC and for their prediction was covered in the first part of this thesis. A three segmented quadratic model and a semiparametric model were proposed to explain the vertical and regional variation in disk SG. Based on both models, the loblolly pine stem can be divided in to three segments with respect to the vertical variation in disk SG. The regional variation in disk SG was also identified using the proposed models. The mean trend in disk SG of trees from the southern Atlantic and Gulf Coastal Plain was observed to be higher than that for other physiographical regions (Upper Coastal, Hilly Coastal, northern Atlantic Coastal Plain and Piedmont). Regional variation in disk SG at specified height levels was explained using maps developed using stand level disk SG data and semiparametric approach.

SG and MC are highly negatively correlated ($\rho = -0.779$) properties: high SG is associated with low MC and vice-versa. A multivariate mixed model system was proposed to model the two properties simultaneously using stand average disk SG and MC. Regional variation in mean trend of the properties was incorporated in the model. Contemporaneous correlation between the two components in the model was accounted for by defining within stand error structure appropriately. The predictive performance of the multivariate model relative to univariate models for SG and MC was evaluated using root mean square prediction error (RMSPE) assuming that one variable was available to predict the other variable .Improved prediction was observed for multivariate model, 29 (SG) and 26 (MC) % improvement in RMSPE, after taking account of the contemporaneous correlation between the these two properties.

Two methods commonly used in forestry to deal with longitudinal data collected over time/ heights are: Generalized Algebraic Difference Approach (GADA) and nonlinear mixed models (NLMM). A study was conducted to evaluate the predictive performance of these two approaches using disk SG data collected from 81 loblolly pine trees (a subset of main data). GADA and NLMM forms of two base models were proposed and used to fit to the data. The predictive performance of the two approaches was assessed using M-fold cross validation, where

SG data from one tree was retained to validate predictions made using models fitted on the remaining trees. The predictive performance of models was assessed by using 1, 2, 3, 4 or 5 observed SG-height pairs as prior information available for the estimation of subject specific effects. Root mean square error (RMSE), mean absolute residual (MAR) and mean residual (MR) were used to evaluate the performance of the models. Based on this study, the NLMM found to be performed better than the GADA methodology in terms of both RMSE and MAR. A 3-19 % improvement in RMSE and MAR was observed for the NLMM approach compared to the GADA approach with improvement varying with the number of prior observations used to estimate the subject specific effect. RMSE and MAR for GADA and NLMM decreased considerably as the number of data points used for estimating the subject specific parameter increased from 1 to 5.

Wood is a secondary byproduct of growth. Silvicultural operation which makes any change in the growth processes imparts a change in the wood properties and thus the quality of wood produced. Various silvicultural practices have been identified as producing positive influence on growth and yield of loblolly pine. However, increased growth rates from silvicultural practices may result in an increase in juvenile wood formation and deterioration of product quality. A study was conducted to determine the effect of midrotation fertilization on growth and wood properties of loblolly pine. Growth and wood properties were measured on breast height cores collected from two stands, located in the lower coastal plain of North Carolina. The first site was fertilized following thinning and the second site was not thinned before fertilization. The study was laid out in a randomized complete block design with each treatment replicated in four blocks in both locations. The fertilization treatments were different levels of nitrogen: control - no nitrogen, 112 kg/ha, 224 kg/ha and 336 kg/ha, in addition 28 kg/ha of phosphorous was included with each treatment.

In the first study, we examined the response in growth and wood properties to midrotation fertilization in both thinned and unthinned stand. In the thinned stand, a significant decrease in latewood SG was observed immediately following nitrogen fertilization. Whole ring width, latewood width and earlywood width significantly increased following nitrogen fertilization in the thinned stand, but not in the unthinned stand. Whole ring SG, early wood SG, latewood percent and earlywood:latewood ratio did not show any change due to fertilization in either stands. The growth and wood property responses lasted for 2-3 years following fertilization. The magnitude of the response found to be depended upon the amount of fertilizer applied and differed between thinned and unthinned sites.

A consistent pattern of response was observed in latewood SG following the application of different levels of nitrogen fertilizer in the thinned stand. A three parameter asymptotic function was used to model the latewood SG profile of unfertilized trees with ring number as a covariate. Application of nitrogen at midrotation age reduced latewood SG and the response was modeled using a two parameter function with year since fertilization as a covariate and separate parameter estimates for each fertilization level. The response models was able to reproduce the latewood SG response from fertilization.