Modeling loblolly pine (*Pinus taeda* L.) stand dynamics and its associated financial implications for forestland owners in the Southeastern U.S.

by

Héctor I. Restrepo

(Under the Direction of Bronson P. Bullock)

Abstract

Timberland investment performance depends on internal and external factors. Forestland owners can control the establishment location, the level of genetically improved seedlings, and management regimes that maximize the timber production profitability. Therefore, analyzing forest growth and yield factors, silvicultural responses (and potentially genetic gains of yield), and measures of the amount of merchantable volume is essential for forestland owner decision making. However, the effect of external economic factors such as the global market, interest rates, and macroeconomic stability may affect timberland investment returns as well. Thus, understanding the effect of the economy on timberland investments is also crucial. For instance, analyzing how the current economic outlook affects timber cutting contracts based on option pricing gives insight into the financial performance of timberland investments. Although the problem and justification addressed in this dissertation is globally applicable, we focused on loblolly pine in the southeastern United States because the combination of the species and region represents the most relevant timberland investment in the world. The general objective of this research was to determine the chief drivers of loblolly pine timber production, and to assess for the effect of the present external economic context on the timberland investments in the southeastern U.S. Regarding timber production drivers, the effect of the physiographic region, level of genetic improvement, level of management regime, stand density, and proportion of sawtimber were evaluated. Timber prices, timber price volatility, and interest rates were utilized to understand the effect of the current economic context on the estimated value of timber cutting contracts. In brief, this investigation gives insight to private and public forestland owners, forest products companies, timber investment management organizations (TIMOs) and real estate investment trusts (REITs) to make informed decisions on timber production and timberland investments in the southeastern U.S.

Index words: Forest growth and yield, silvicultural responses, timber product classes, long-term timber cutting contracts, option pricing.

Modeling loblolly pine (*Pinus taeda* L.) stand dynamics and its associated financial implications for forestland owners in the Southeastern U.S.

by

Héctor I. Restrepo

B.S.F.E., The National University of Colombia, Colombia, 2002

M.Fin., The Engineering School of Antioquia, Colombia, 2005

M.S.F.R., The National University of Colombia, Colombia, 2010

A Dissertation Submitted to the Graduate Faculty

of The University of Georgia in Partial Fulfillment

of the

Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2019

Modeling loblolly pine (*Pinus taeda* L.) stand dynamics and its associated financial implications for forestland owners in the Southeastern U.S.

by

Héctor I. Restrepo

Approved:

Major Professor: Bronson P. Bullock

Committee: Cristian R. Montes
Michael Kane
Richard Mei

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2019

Dedication

To my lovely wife and daughter, Adriana and Sofia, whom were supportive in the difficult moments and cheering me on during the good times of my doctoral studies. To my family, especially my mother, whom has given me the required financial and emotional support to succeed.

Acknowledgments

I want to express my sincere gratitude to my major professor, Dr. Bronson P. Bullock. His guidance and support during these four years were essential to accomplishing my doctoral studies. His mentoring style, approachability and kindness will continue to inspire me for many years to come.

I want to thank my committee members, Drs. Cristian Montes, Michael Kane, and Richard Mei. Thank you for making time in your busy schedules to attend meetings, review manuscripts, and provide suggestions to improve my research.

I am very grateful to the Harley Langdale Center for Forest Business, and its Director, Mr. Bob Izlar, for funding my doctorate; the Plantation Management Research Cooperative (PMRC), and its faculty and staff for taking the forest measurements and furnishing me some of the data to conduct part of my research; and the Warnell School of Forestry and Natural Resources family, professors for the wonderful classes I took, and staff, specially Kate deDufour, for being available at any time to answer my questions and solve administrative issues.

I am so thankful to Rafael de la Torre and his family for their invaluable support, and appreciative to my fellow graduate students for their selfless contributions to my research.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction and literature review

1.1 Introduction

Timber production and timberland investment profitability relies on internal and external factors. Internal factors are the technical and financial decisions that a landowner can make to maximize forest production, such as species and genetic improvement selection, location of the forest stands, and management regimes. On the other hand, economic external factors include the influence of the global market, domestic and foreign interest rates, and macroeconomic stability. Although the exogenous factors are important to evaluate in the potential success of forest investments, their associated risks cannot be effectively managed, unlike internal factors, which are under the control of the landowner.

Loblolly pine is the most commercially important forest species in the southeastern U.S. Loblolly pine planted area in this region consists of 14.4 million hectares, of which 80% are planted with half-siblings or open-pollinated families, 18% are planted with mass control pollinated or full-siblings families, and 1% are planted with clones. These genetically improved loblolly pine families are planted across 13 states and 4 physiographic regions with contrasting environmental conditions. A wide array of silvicultural practices are applied in conjunction with the genetic improvements, from low levels of inputs to very intensive forest management composed of bedding, fertilization, competing vegetation control, and thinning. Since most existing forest growth and yield models estimate total volume or biomass, proportions of the commercial timber product classes, i.e., sawtimber, chip-n-saw, and pulpwood, are required to accurately calculate timber values. The analysis of the forest growth and yield and measures to estimate volume merchantability as part of the internal forest production factors give insight into the potential financial performance of timberland investments.

Although external factors are beyond the control of the forest landowner, the effect of the economic outlook on timberland investments can be evaluated. Thus, the current forest business environment, characterized by relatively low timber prices, low timber price volatility, and low interest rates, contrasts with the past 30-years conditions. This research addressed the problem of the long-term timber cutting contracts by updating the primer valuation framework using call options. Results of this research would be valuable for private and public forestland owners, forest products companies, timber investment management organizations (TIMOs), and real estate investment trusts (REITs) that look for strategic information to maximize returns and minimize risks.

1.2 Literature review

Forest growth can be thought of as the outcome of two opposing factors, the unlimited trend of growth or biotic potential, and the growth constraints imposed by the environment (Zeide, 1993). As an expansion of this conceptualization, along with the stand age, there are three well-known factors related to forest growth: i) site quality or productive potential; ii) the stand's intrinsic characteristics, such as density and genetics; and iii) silvicultural treatments (Clutter et al., 1992).

Productive potential was the main topic in forestry research four decades ago. However, expected increases in per unit area production of 50-year pine improvement programs are gaining attention (Roth et al., 2007). The current volume gains of the second generation of genetically improved loblolly pine compared to unimproved planting stock range from 10-30% (Allen et al., 2005), which can be doubled if improvements in stem form and disease resistance are added (McKeand et al., 2006a). Likewise, there are contrasting findings about the interaction of genotype and environment (G x E) (Roth et al., 2007; McKeand et al., 2006b; Smith et al., 2014), which suggests uneven genetic responses across the southeastern U.S.

Intensive forest management has also contributed to increase timber productivity in the southeastern U.S. A wide range of silvicultural treatments such as mechanical site preparation, competing vegetation control, fertilization, and irrigation has been applied to maximize timber production (Allen et al., 2005). These silvicultural practices, in conjunction with stand age and density, result in a general expression of forest yield (Clutter et al., 1992). Volume gains due to intensive forest management can be derived from forest yield models by isolating the effect of remaining variables.

Forest yield estimation should be adjusted to take into account the portion of timber that can be sold. Frequencies of the number of trees per diameter classes are determined using probability distribution functions and taper equations to estimate proportions of merchantable volume per product class. These probability distribution functions, and models to estimate proportions for product classes, have been estimated as a function of age, height, density (Strub et al., 1986), and thinning (Burkhart, Bredenkamp, 1989). Moreover, proportions for product classes can include measurements of stem quality such as fork, broken top, sweep, and diseased trees (Choi et al., 2008; Buford, Burkhart, 1987).

Timberland financial performance depends greatly on external economic factors, which are beyond the control of the landowner (Reeves, Haight, 2000; Wan et al., 2013; Gao, Mei, 2013). Understanding the economic context is particularly important for timberland investments because returns are reaped many years after establishing a forest plantation (Ashton et al., 2001; Hildebrandt, Knoke, 2011). Thus, the effects of timber prices, timber price volatility, and interest rates on timberland investment performance represent obligated topics to be investigated.

1.3 Rationale and significance

The ultimate goal of biometrics, from the timber management perspective, is an accurate estimation of the merchantable volume. This is a vital input for forest financial return calculations. Moreover, profitability analyses are performed considering growth factors as fixed in spite of their variability. In this sense, thorough timber management analyses should consider the inherent uncertainty nature of volume yield, silvicultural responses and genetic gains, and timber product class distribution.

Mixed effects models contribute to understand the effect of forest growth drivers on yield at stand and landscape levels. As mentioned, there are three factors related to forest growth that should be included in the forest growth and yield model: site quality, intrinsic stand characteristics (seed source genetics and density), and management. Since forest growth of loblolly pine is different across physiographic regions of the southern U.S., a factor is required that takes into account such variability in forest yield model. There are also three broad genetically improved entries, i.e., open pollinated, mass control pollinated, and clones, and a continuum of management regimes, which should be considered in the estimation of loblolly pine growth and yield. In particular, silvicultural responses can be evaluated by taking the partial derivatives of the estimated mixed effects models with respect to the associated levels of management.

The determination of the proportion of timber products as a function of diameter, form, and stem quality assessment may differ across stands. The environment and management regimes may have an effect on the proportions of timber product classes. Hence, the tree diameter, location and the intensity level of management applied are the inputs to estimate volume merchantability.

Finally, the evaluation of the effect of the economic external factors on the timberland investment profitability may be informative for forestland owners, TIMOs, and REITs in the southeastern U.S. For instance, a research is needed to assess for the effect of the current risk-free rates of return, timber prices, and timber price volatility on the option pricing valuation of timber contracts.

1.4 Goals, objectives and hypotheses

1.4.1 General goal

To determine the chief drivers of loblolly pine timber production, and to assess for the effect of external economic context on the timberland investments in the southeastern U.S.

1.4.2 Specific objectives and hypotheses

1. To evaluate the difference in forest productivity among levels of genetically improved loblolly pine (half-sib, full-sib, clones), physiographic regions (Lower Coastal Plain, Upper Coastal Plain, Piedmont), intensity level of forest management (low, moderate and high), and stand density in the southeastern US.

   Hypothesis: No differences in loblolly pine performance will be present among the levels of factors evaluated.

   Methodological objective: A comprehensive evaluation of the forest growth drivers was conducted by searching for relevant research papers in the southeastern U.S. and statistically synthesized them using meta-regression.

2. To determine silvicultural responses in yield of loblolly pine in the southeastern U.S.

   Hypothesis: No differences between the levels of the management intensity will be present.

   Methodological objective: The contribution of silviculture to loblolly pine growth and yield was determined by partially differentiating the estimated meta-regression model obtained in objective one with respect to the moderate and high levels of management regimes.

3. To estimate the proportion of sawtimber at year 18 as a function of the location (as a proxy of the environmental factor), intensity of forest management, planting density, and size (DBH) in the southeastern U.S.

   Hypothesis: The proportion of trees with sawtimber potential at year 18 will not differ among locations,levels of management regimes, planting density, and DBH.

   Methodological objective: Theoretical Bayesian frameworks for binomial, hierarchical, and logit models were proposed and utilized to estimate the proportion of trees with sawtimber potential at year 18 as a function of locations, management intensity, planting density, and tree size.

4. To determine the potential effect of external economic factors on timberland investment returns in the southeastern U.S.

   Hypothesis: The external economic context does not affect timberland investments.

   Methodological objective: This objective was addressed by updating the primer long-term timber cutting contract framework by Shaffer Jr. (1984). The value of call options was estimated using the Black-Scholes (European) and binomial models (American and European). The required timber price volatility for this valuation was estimated from GARCH models.

Chapter 2

Growth and yield drivers of loblolly pine in the southeastern U.S.: A meta-analysis [1]

[1]Restrepo, H.I, B.P. Bullock, C.R. Montes. 2019. Growth and yield drivers of loblolly pine in the southeastern U.S.: A meta-analysis. Accepted by *Forest Ecology and Management*. Reprinted here with permission of publisher.

Abstract

An abundant amount of information has been accumulated over the past century on loblolly pine. However, few studies have been aimed at assembling this information. Three possible approaches can be used to synthesize available information on loblolly pine: a review paper in the form of a narrative discussion, systematic review compiling data in tables, and meta-analysis to statistically summarize data. The purpose of this research is to statistically synthesize suitable loblolly pine yield data in the southeastern United States using meta-analysis. There were 18 studies selected out of approximately 500 peer-reviewed papers, and three high-quality studies (one conference proceedings, one M.S. thesis, and one Ph.D. dissertation) evaluated, from which a database was compiled. Since forest growth has several drivers (i.e., age, site quality, genetics, density, and management) the use of meta-regression, a meta-analysis technique to account for variability associated with covariates, was used. Thus, meta-regression linear mixed effects yield models using the log-transformed Schumacher form, at the whole-stand level, were estimated as a function of the mentioned forest growth factors for diameter at breast height (DBH), height (Ht), basal area (BA), and volume (V). Overall, the estimated models suggest that these forest growth factors successfully explain yield variability. The Raudenbush's pseudo-R^2, which measures the amount of variation explained by the covariates, were 97, 94, 97, and 91%, for DBH, Ht, BA, and V models, respectively. However, the 95% confidence intervals (CI) of yield curves associated with some growth factor levels overlapped their corresponding reference level, suggesting no statistical differences at certain ages. In this sense, the CI's width is driven mainly by the number of studies, and their number of replicates, available for factor levels. Thus, the lack of information of factor levels, and their combinations, was identified and suggested to be investigated in future research in order to achieve narrower CIs. Meta-analysis and meta-regression are promising techniques to be applied in forestry research to give insight into the effect of growth factors on forest yield.

Key words: *Pinus taeda* L., Schumacher model, site quality, spatially explicit models, genetically improved trees, silvicultural regimes.

2.1 Introduction

Loblolly pine plantations occupy an area of 14.4 million of hectares across the southeastern United States (South, Harper, 2016), having a continuum of silvicultural regimes from low (operational) to high (intensive forest management) levels of inputs such as mechanical site preparation, chemical vegetation control, fertilization, and thinning (Allen et al., 2005). Current typical intensive forest management practices include establishment with genetically improved seedlings (McKeand et al., 2003, 2006a; McKeand, 2015). These forest management practices, in addition to stand age, density, and environmental conditions (site quality and productive potential) allow for an extensive expression of forest growth (Clutter et al., 1992).

Forest growth factors for loblolly pine, along with other relevant ecological and physiological species features, have been researched for a century (Wakeley, 1969). A search on the Web of Science database returned 2,700 papers with a title containing loblolly pine or *Pinus taeda* with a 5-year publication rate of 70 papers per year. Despite such a large amount of information related to loblolly pine, few studies have compiled and synthesized findings (Jokela et al., 2004; Thomas et al., 2017; Zhao et al., 2016), probably because raw data and summary statistics suitable for meta-analysis are rarely available in forestry research, if at all.

The abundant amount of information on loblolly pine growth and yield can be synthesized using meta-analysis, a statistical technique that allows for summarizing data to get parameter estimates related to a phenomenon, hypothesis or research question. The term meta-analysis was coined by Glass (1976) as the statistical analysis of a large collection of results from individual studies for integrating the findings as a rigorous alternative to the traditional narrative discussion. Meta-analyses are common in medicine, social sciences and education, but seldom published in natural resources, ecology and forestry. The earliest applications of combining results made by Fisher, Cochran, and Pearson in the 1930's came from the agricultural sciences (Petitti, 2000).

Meta-regression is a meta-analysis technique that accounts for variability in the response variable due to differences in the levels of covariates through the fixed effects (Hedges, Olkin, 1985; Hedges, Vevea, 1998). Usually, random effects are also needed to account for the heterogeneity between studies, which in conjunction with the fixed effects configures a mixed effects model (Borenstein et al., 2010; Hedges, Olkin, 1985; Hedges, Vevea, 1998).

The aim of this paper was to identify the relevant drivers and their contributions to loblolly pine growth and yield in the southeastern United States using meta-regression. Therefore, covariates into forest growth factors and associated to forest yield variability (i.e. age, site quality accounting for the environmental effect, management, and stand characteristics such as genetics and density) were considered. This analysis is based on the use of spatially explicit mixed effects forest growth and yield models, at the whole-stand level, for diameter at breast height, total height, basal area, and total volume. A priori assumptions were not made about taper, height-diameter relationships, allometry, biomass partitioning, or performance of loblolly pine across physiographic regions or due to soil characteristics, management level, or level of genetic improvement. For that reason, we believe that site index is not uniformly scaled to be included in the models as an explanatory variable. Some missing covariates and yield data were imputed and estimated to increase the number of studies to be included in the meta-regression.

2.2 Methods

2.2.1 Data

Approximately 500 research articles directly related to loblolly pine growth and yield in the southeastern United States were obtained from public databases (i.e. Web of Science and Google Scholar). The inclusion/exclusion criterion was defined as a research paper or study that presents the yield mean, number of replicates (plots), and standard deviation or standard error of the treatments, at the whole-stand level, for diameter at breast height (DBH, cm), total height (Ht, m), basal area (BA, m^2ha^{-1}), and total volume (V, m^3ha^{-1}).

We did not rely on existing models to estimate total height, basal area or volume because those may confound the statistical relationship between forest yield and its drivers. However, some key and high-quality studies may present the required information but not in the form to be directly used in the meta-analysis. Thus, some expected DBH means (4 out of 105) and expected BA values (35 out of 176) were estimated. Expected mean DBH were estimated from histograms of DBH calculating weighted averages and multiplying the diameter classes by their relative frequencies. Likewise, expected BA values were estimated using the mean of DBH, the standard deviation of DBH, and the stand density as inputs, and taking advantage of the mathematical statistical relationship between DBH and BA via the quadratic mean diameter (Curtis, Marshall, 2000). A brief proof is provided as follows.

Let d_1, d_2, ..., d_n be independent and identically distributed random variables associated with individual tree diameters (cm) in a stand with a known number of trees per hectare (TPH), n. An assumption of the probability distribution is not required to get the expected BA, but let μ_d and σ_d^2 be its mean and variance, respectively. The basal area of a tree (m^2) is:

$$ba = \pi \left(\frac{d}{200} \right)^2 \tag{2.1}$$

with expected value:

$$\mathsf{E}[ba] = \left(\tfrac{\pi}{40000} \right) \mathsf{E}[d^2] \tag{2.2}$$

where $\mathsf{E}[d^2]$ is the expected value of the quadratic mean diameter, which corresponds to the second moment of the distribution of DBH (Curtis, Marshall, 2000; Wackerly et al., 2008):

$$\mathsf{E}[d^2] = \mathsf{Var}[d] + \mathsf{E}[d]^2 = \hat{\sigma}_d^2 + \hat{\mu}_d^2 \tag{2.3}$$

where $\hat{\sigma}_d^2$, and $\hat{\mu}_d^2$ are the reported mean and standard deviation in the original study, parameter estimates of σ_d^2, and μ_d^2. Whereby,

$$\mathsf{E}[ba] = \left(\tfrac{\pi}{40000} \right) \left(\hat{\sigma}_d^2 + \hat{\mu}_d^2 \right) \tag{2.4}$$

Since:

$$BA = \sum_{i=1}^{n} \mathsf{E}[ba_i] \tag{2.5}$$

where ba_i is the basal area of the ith tree. Then:

$$\widehat{BA} = n\mathsf{E}[ba] \tag{2.6}$$

where $\widehat{BA}$ is an unbiased estimator of the BA ($\mathrm{m^2ha^{-1}}$). Similar mathematical statistics procedures were derived to estimate asymptotic standard deviations of the 4 and 35 missing DBH and BA values, respectively. However, a method to impute sample standard deviation when missing from the research study to use the cumulative maximum recorded standard deviation over age criterion was applied. This imputing criterion can be mathematically expressed as:

$$S_A = \max_{S} \left\{ \bigcup_{a=1}^{A} S_a \right\} \tag{2.7}$$

where S_A is the imputed sample standard deviation corresponding to the stand age A (years), $S_a = \{S \in T_a : T_a$ is the set of recorded stardard deviations of treatments at age $a\}$ with S the recorded sample standard deviation, and $\left\{ \bigcup_{a=1}^{A} S_a \right\}$ is the resulting set of the union of all recorded sample standard deviations from the age one up to stand age A (years). These procedures were needed since no raw data were available to calculate the mean and standard deviation of DBH and BA, and were justified to increase the number of relevant studies and treatments included in the meta-regression models.

Likewise, studies to be included in the meta-regression should clearly state the study site or location, area of measurement plots, management (e.g., types of silvicultural treatments, and their frequency and age of applications), stand age, genetics, and current stand density (or planting density and survival rate). If thinning was not mentioned as a treatment or feature of the stand, the plots were assumed to be non-thinned. Some stand densities (N_2) were also estimated using a survival equation as a function of the age and planting density (N_1) (Rose et al., 2002):

$$\hat{N}_2 = 2.5 + (N_1 - 2.5)(1 + 0.68A)^{1.46}(1 + A)^{-1.35} \exp[-5.9 \times 10^{-4} A^2] \tag{2.8}$$

This equation was used for taking into account the effect of mortality in yield curves. Along with the continuous covariates, age (A, years) and stand density (*Den*, TPH), forest growth factors were specifically categorized as:

- Levels of genetic improvement (*Gen*)
  - Unimproved or unknown (UU)
  - Half-sibling or open pollinated (HS)
  - Full-sibling or mass control pollinated (FS)
  - Clone (C)
- Levels of management regimes (*Mgmt*)
  - Low (L): basic level of inputs with one or two silvicultural practices at establishment
  - Moderate (M): *Mgmt*(L) + moderate amount of inputs at early ages or midrotation practices
  - High (H): either considerable amount of inputs in quantity and frequency throughout the rotation or *Mgmt*(L) + moderate amount of inputs at early ages and midrotation practices
- Physiographic regions (*Reg*)
  - Piedmont (P)
  - Upper coastal plain (UCP)
  - Upper coastal plain or piedmont (UCPP); combined term used when the research study did not discern between UCP and P
  - Lower coastal plain (LCP)

All this information was tabulated in a database, which also includes (if provided), climatic description of the location, site index (m, base age 25 years), references of previous or related studies, and relevant remarks from the study. Since studies typically have more than one treatment applied, the basic unit of analysis was treatment within study (e.g. where all replications of a specific combo from a research study would make up one record in the meta-regression database) representing a combination of the previously mentioned factors. Some repeated measures were included in the dataset and treated as independent observations since the model parameter estimates are assumed to be unbiased and unaffected by the misspecification of the error structure (Rencher, Schaalje, 2008).

2.2.2 Meta-regression model

The mixed-effects meta-regression model accounts for variation in the response variable as a function of covariates and studies through fixed and random effects, respectively. Let the mixed effects model be (Schwarzer et al., 2015; Viechtbauer, 2010):

$$\begin{aligned} y_i &= \theta + u_i + \varepsilon_i \\ & \quad u_i \sim \mathcal{N}(0, \tau^2); \varepsilon_i \sim \mathcal{N}(0, \sigma_i^2); \mathsf{cov}(\varepsilon_i, u_i) = 0 \end{aligned} \tag{2.9}$$

where y_i is the observed mean associated with treatment i, θ represents the fixed effect term; u_i represents the random effects term, assumed normally distributed with mean zero and variance τ^2; and ε_i represents the error term, assumed normally distributed with mean zero and variance σ_i^2 and independent of the random effects. Parameter θ is assumed to be a function of covariates:

$$\theta = \mathbf{X}\boldsymbol{\beta} \tag{2.10}$$

where $\mathbf{X}$ is the matrix of covariates and $\boldsymbol{\beta}$ is a vector of parameters associated with the covariates.

2.2.3 Forest yield model

The Schumacher model (Schumacher, 1939) was considered suitable to estimate forest yield models using a logarithmic transformation on the dependent variable and the inverse of the age. In the Schumacher model, the intercept works as an asymptote and the slope represents the growth rate of accumulation process. The raw data were not available to get the parameter estimates from the logarithm of individual values. Thus, the mean and standard deviation of the dependent variable were transformed as (Thomopoulos, 2013):

$$y_i^* = \ln(y_i^2) - \tfrac{1}{2}\ln(S_i^2 + y_i^2) \tag{2.11}$$

$$S_i^* = \sqrt{\ln\left(1 + \frac{s_i^2}{y_i^2}\right)} \tag{2.12}$$

The linear mixed effects Schumacher model is then:

$$y_i^* = \theta^* + u_i^* + \varepsilon_i^* \tag{2.13}$$

$$u_i^* \sim \mathcal{N}(0, \tau^{*2}); \varepsilon_i^* \sim \mathcal{N}(0, \sigma_i^{*2}); \mathsf{cov}(\varepsilon_i^*, u_i^*) = 0$$

where the star (*) denotes the log-transformation on the dependent variable.

2.2.4 Model estimation

Meta-regression linear mixed effects models were estimated using the *metafor* package (R Development Core Team, 2018; Viechtbauer, 2010). The parameter τ^{*2} was estimated using maximum likelihood, and the parameter θ^* was estimated using weighted least squares, with weights as (Viechtbauer, 2010):

$$w_i^* = \frac{1}{S_i^{*2} + \hat{\tau}^{*2}} \tag{2.14}$$

where w_i^* is the weight for the ith treatment, S_i^{*2} is the sample variance of the ith treatment, and $\hat{\tau}^{*2}$ is the maximum likelihood estimator of τ^{*2}, all terms in logarithmic units. Let the fixed effect of the reduced log-transformed Schumacher model, named as M0, be:

$$\hat{\theta}^* = \hat{\alpha}_0^* + \frac{\hat{\beta}_0^*}{A} \tag{2.15}$$

where $\hat{\theta}^*$ is an estimator, in logarithmic units, of the fixed effect θ^* for each of the primary yield variables, i.e. DBH (cm), Ht (m), BA (m^2ha^{-1}), and V (m^3ha^{-1}) as a function of the stand age (years); and $\hat{\alpha}_0^*$, $\hat{\beta}_0^*$ are parameter estimates related to the asymptote and growth rate, respectively. Then, nested models were estimated adding one factor, i.e. genetics, physiographic region and management, and a continuous variable (density) at a time using simple effects and their interaction with the inverse of age. The full model, termed as MI, was estimated as:

$$\begin{aligned} \hat{\theta}_{jkl}^* &= \hat{\alpha}_0^* + \hat{\alpha}_{1j}^* + \hat{\alpha}_{2k}^* + \hat{\alpha}_{3l}^* + \hat{\alpha}_4^*(Den) + \frac{1}{A}\left(\hat{\beta}_0^* + \hat{\beta}_{1j}^* + \hat{\beta}_{2k}^* + \hat{\beta}_{3l}^* + \hat{\beta}_4^*(Den)\right) \qquad (2.16) \\ & \quad j = 1, 2, 3; \;\; k = 1, 2; \;\; l = 1, 2, 3 \end{aligned}$$

where $\hat{\theta}_{jkl}^*$ is the estimate of the fixed effect, in logarithmic units, of the jth level of genetic improvement under the kth level of management regime in the lth physiographic region; $\hat{\alpha}_0^*$ represents the reference level of the asymptote estimate of unimproved genetics (Gen(UU)), under the low management regime ($Mgmt$(L)), planted in the upper coastal plain or piedmont (combined term) physiographic region (Reg(UCPP)), and with a density of 300 TPH (the minimum stand density reported for the oldest stand included in the meta-regression model that corresponds to a very common final harvest target in the southeastern U.S.); $\hat{\alpha}_{1j}^*$ is the parameter estimate of the jth level of genetic improvement (when Gen(HS), Gen(FS), or Gen(C), j equals 1, 2, or 3, respectively); $\hat{\alpha}_{2k}^*$ is the parameter estimate of the kth level of management regime (when $Mgmt$(M), or $Mgmt$(H), k equals 1, or 2, respectively); $\hat{\alpha}_{3l}^*$ is the parameter estimate of the lth physiographic region (when Reg(UCP), Reg(P), or Reg(LCP), l equals 1, 2, or 3); and $\hat{\alpha}_4^*$ is the parameter estimate associated with the stand density in excess of 300 TPH. The overall slope ($\hat{\beta}_0^*$) has the same interpretation as the overall asymptote term regarding the basic levels of the factors. Parameter estimates for β_{1j}^*, β_{2k}^*, β_{3l}^*, and β_4^* represent the marginal contribution to the growth rate in the form of the interaction between each factor and age.

The statistical contribution of covariates in the explanation of the variability was tested using the likelihood ratio test (LRT) (Rencher, Schaalje, 2008; Wackerly et al., 2008):

$$LRT = \frac{\max\limits_{\{(\boldsymbol{\beta}^*,\tau^{2*})\in\omega\}} \mathsf{L}(\boldsymbol{\beta}^*, \tau^{2*}|\mathbf{y}^*)}{\max\limits_{\{(\boldsymbol{\beta}^*,\tau^{2*})\in\Omega\}} \mathsf{L}(\boldsymbol{\beta}^*, \tau^{2*}|\mathbf{y}^*)} \tag{2.17}$$

where $\boldsymbol{\beta}^*$ represents the vector of parameters of the asymptote and growth rate in logarithmic units (α^*'s, and β^*'s), ω and Ω represent the parameter space of the reduced and full models, respectively, and $\mathsf{L}(\boldsymbol{\beta}^*, \tau^{2*}|\mathbf{y}^*)$ represents the likelihood function. For a large number of observations, $-2\ln(LRT)$ has approximately a χ^2 distribution with degrees of freedom equal to the difference of degrees of freedom of both models and rejection region defined as $\mathsf{RR} : \{-2\ln(LRT) > -2\ln(c) = c^*\}$. The goal of the LRT is to analyze the trade-off between the marginal increment in the explanation of variance in the model and the number of degrees of freedom spent to explain that additional amount of variance.

This variable selection process was conducted forward from the M0 model to the MI full model and backward from the MI model to the final parsimonious model, named MII. In the forward selection, the contribution of each factor (or continuous variable) was tested (p-value<0.05) until the MI model was achieved. Then, in the backward elimination process, non-significant variables or factor levels were removed in turn. This process began collapsing non-significant interaction factor levels to the reference level of the slope ($\hat{\beta}_0^*$), one at a time. Then, any non-significant variable or factor level was dropped from the asymptote term if its counterpart in the slope was deleted in the previous step. This procedure is in accordance with the principle of hierarchy, which states that variables should be present as a simple effect term to be considered for an additional term as an interaction (Kutner et al., 2005).

2.2.5 Fit evaluation and diagnostics

Parameter estimate τ^{*2} accounts for the variability among the treatments included in the meta-regression. The null hypothesis of heterogeneity of treatments $\tau^{*2} = 0$ can be tested using Cochran's Q-test, with the following statistic (Schwarzer et al., 2015):

$$Q^* = \sum_{i=1}^{T} w_i^* \left(\hat{\theta}_i^* - \frac{\sum_{i=1}^{T} w_i^* \hat{\theta}_i^*}{\sum_{i=1}^{T} w_i^*} \right) \overset{\text{Ho}}{\sim} \chi^2_{(T-1)} \tag{2.18}$$

where T is the number or treatments. To assess the amount of variance explained in the parsimoniuos (MII) model, the Raudenbush‘s pseudo R^2 was calculated, that represents the amount of variance explained by the mixed models compared to the random effects only (Schwarzer et al., 2015):

$$R^{*2}_{pseudo} = \frac{\tau^{*2}_R - \tau^{*2}_M}{\tau^{*2}_R} \tag{2.19}$$

where τ^{*2}_R is an estimator of the between-study variance of the random effects model and τ^{*2}_M is an estimator of the between-study variance of the mixed effects model.

Likewise, multicollinearity was assessed using the mean variance inflation factor, $\overline{VIF}$, and compared to the threshold of 10, which indicates moderate multicollinearity (Kutner et al., 2005):

$$\overline{VIF} = \frac{trace(r^{-1}_{\mathbf{X}})}{p} \tag{2.20}$$

where $r^{-1}_{\mathbf{X}}$ is the inverse of the correlation matrix of the design matrix $\mathbf{X}$, excluding the corresponding column of the intercept, and p is the number of parameter estimates excluding $\hat{\alpha}_0^*$. The bias induced by the logarithmic transformation was calculated using Snowdon's γ (Snowdon, 1991):

$$\gamma = \frac{\sum_{i=1}^{T} y_i}{\sum_{i=1}^{T} \hat{y}_i} \tag{2.21}$$

where y_i and $\hat{y}_i$ are the observed and estimated yield, in the original scale, for DBH, Ht, BA, and V in the ith treatment.

Moreover, studentized residual plots, Q-Q plots, funnel graphs (assessing the publication bias in meta-analysis), and Cook's distance plots were evaluated for additional fit diagnostics. Hence, Cook's distances larger than the 30th or 50th percentile of the $F_{(p,T-p)}$, were considered to have, respectively, moderate and high influential effects (Kutner et al., 2005).

Table 2.1: Number of treatments, number of measurement plots, and the summed total area of measurement plots over the selected research studies in the meta-regression of loblolly pine growth and yield in the southeastern United States.

	No. treatments (T)	No. plots	Area (ha)
DBH	105	1288	79
Ht	97	1344	81
BA	176	1476	70
V	111	1012	86

2.3 Results

2.3.1 Selected studies

Based on the selection criterion outlined above, 21 studies were included in the meta-analysis: 18 peer-reviewed scientific papers, one proceedings paper, one Ph.D. dissertation and one master of science thesis. Included studies constitute a representative sample size of loblolly pine yield (Table 2.1) over a wide range of environmental conditions of more than 41 counties located in 10 states across the southeastern United States (Figure 2.1). Additional information about included studies, their locations and characteristics are presented in Table 2.2. The number of studies containing clonal information for BA, and full-sibling and clonal information for V was considered insufficient to be included. The cut-off did not correspond to a certain preestablished threshold of the amount of data to be included in the meta-regression models. It was assessed in the first stages of the modeling process when the statistical procedures and their algorithms in R were unable to estimate some of parameter estimates for levels of genetic improvement. Therefore, these genetic improvement levels were omitted in the BA and V models.

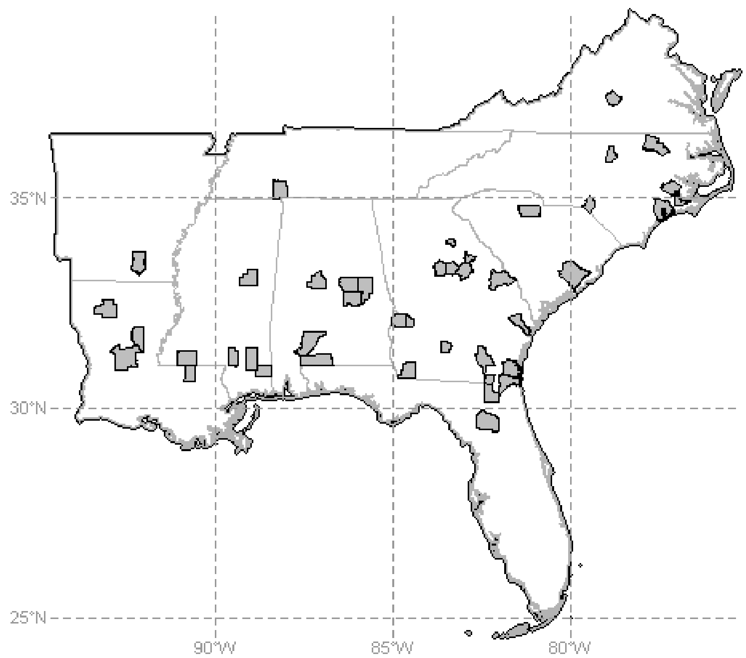

Figure 2.1: Southeastern United States counties in which studies have been conducted that were utilized in this research.

Table 2.2: Summary of studies selected to include in the meta-regression of loblolly pine growth and yield in the southeastern United States.

Study	Location	A (years)	*Gen*	*Mgmt*	*Reg*	*Den* (TPH)	Variables
Akers et al. (2013)	Burke, GA Hancock, GA Jasper, GA	13	HS	M	UCPP	707-3410	DBH, Ht, BA, V
Albaugh et al. (2003)	Coosa, AL	12	HS	L	UCP	1581	Ht, BA
	Durham, NC	14	HS	L	UCP	1212	Ht, BA
	Tallapoosa, AL	10	HS	L	P	1584	Ht, BA
	Halifax, NC	12	HS	L	LCP	640	Ht, BA
	Webster, GA	12	HS	L	UCP	1985	Ht, BA
	Bibb, AL	18	HS	L	UCP	538	Ht, BA
	Chester, SC	22	HS	L	UCP	498	Ht, BA
	Stewart, GA	14	HS	L	UCP	1508	Ht, BA
Albaugh et al. (2004)	Scotland, NC	7-16	UU	L-M	UCP	1105-1200	BA
Amateis et al. (2004)	VA and NC	19	HS	L	UCPP	1660	DBH, Ht
Antony et al. (2011)	Bryan, GA	15	HS	L	LCP	987[a]	DBH, Ht
	Decatur, GA	15	HS	L	UCP	940[a]	DBH, Ht
	Amite, MS	15	HS	L	UCP	940[a]	DBH, Ht
	Escambia, AL	15	HS	L	UCP	940[a]	DBH, Ht
	Saint Helena, LA	15	HS	L	UCP	940[a]	DBH, Ht
	La Salle, LA	15	HS	L	UCP	940[a]	DBH, Ht
	Elmore, AL	15	HS	L	UCP	940[a]	DBH, Ht
	Tallapoosa, AL	15	HS	L	P	940[a]	DBH, Ht
	Jasper, GA	15	HS	L	P	940[a]	DBH, Ht
	Appomattox, VA	15	HS	L	P	940[a]	DBH, Ht
	Bienville, LA	15	HS	L	UCP	1086[a]	DBH, Ht
	Bradley, AR	15	HS	L	UCP	940[a]	DBH, Ht
	Hardin, TN	15	HS	L	UCP	940[a]	DBH, Ht
Aspinwall et al. (2011)	Onslow, NC	2,3	UU-FS	M	LCP	420-1023	BA, V
Borders and Bailey (2001)	Ware, GA	12	HS	L - H	LCP	1414-1661	BA, V
	Tift, GA	11	HS	L - H	UCP	741-1532	BA, V
	Putnam, GA	11	HS	L - H	P	1354-1582	BA, V
	Clarke, GA	10	HS	L - H	P	1117-1473	BA, V
Borders et al. (2004)	Ware, GA	15	HS	L-H	LCP	1250-1650	DBH, BA[c]
Clark (2013)	Taliaferro, GA	6	HS	L-M	P	1533	DBH, Ht, BA, V
Colbert et al. (1990)	Alachua, FL	4	HS	L-H	LCP	1440-1479	DBH, Ht, BA[c], V
Cumbie et al. (2011)	Monroe, AL	10	C	L	UCP	801[a]	Ht
	Nassau, FL	9	C	L	LCP	807[a]	Ht
Jones et al. (2010)	George, MS Lamar, MS Perry, MS	3-5	HS	L-H	LCP	1187-1207	DBH, Ht, BA[c]
Land et al. (2004)	Winston, MS	5-17	HS	L	UCP	796-4048	DBH, Ht, V
McCrady and Jokela (1996)	Berkeley, SC	3,4	HS	L	LCP	2687	DBH, Ht, V
Pile et al. (2016)	Berkeley, SC	9	C	L	LCP	910-1222	DBH
Roth et al. (2007)	Baker, FL Camden, GA	2-5	FS	M, H	LCP	1227-2990	DBH, Ht, BA[c]
Sayer et al. (2004)	Rapides, LA	17	UU	L, M	UCP	1059-2457	DBH[b]
Sharma et al. (2002)	VA and NC	5-16	HS	L	UCPP	1938-2173	DBH, Ht
Smith et al. (2014)	Scotland, NC	14	HS	L, H	UCP	2399-2867	DBH, Ht, BA, V
Smith (2010)	Scotland, NC	6	HS	L, H	UCP	2775-2883	DBH, Ht, BA, V
Williams and Farrish (2000)	North-central LA	25-30	UU	L	UCP	316-319	DBH, Ht, BA, V

Annotations for Table 2.2: A: Age in years, *Gen*: UU: unimproved, HS: half-sibling, FS: full-sibling, C: clone; *Reg*: physiographic region, UCP: Upper Coastal Plain, P: Piedmont UCPP: Upper Coastal Plain-Piedmont, LCP: Lower Coastal Plain; *Mgmt*: management, L, M, H: low, moderate and high, respectively; *Den*: density, TPH: trees per hectare; DBH: diameter at breast height (cm); Ht: total height (m); BA: basal area (m^2ha^{-1}); V: volume (m^3ha^{-1}).

a Imputed (Rose et al., 2002)

b estimated DBH from histogram of frequencies

c estimated BA from DBH

The imputation of the standard deviation is presented in Figure 2.2. This procedure allowed for the inclusion of studies that present only the mean of the yield variables and the number of plots measured, along with the associated covariates. In total, 25% of the standard deviations were imputed. We determined this was justified for the purpose of increasing the number of high-quality research studies to be included in the meta-regression.

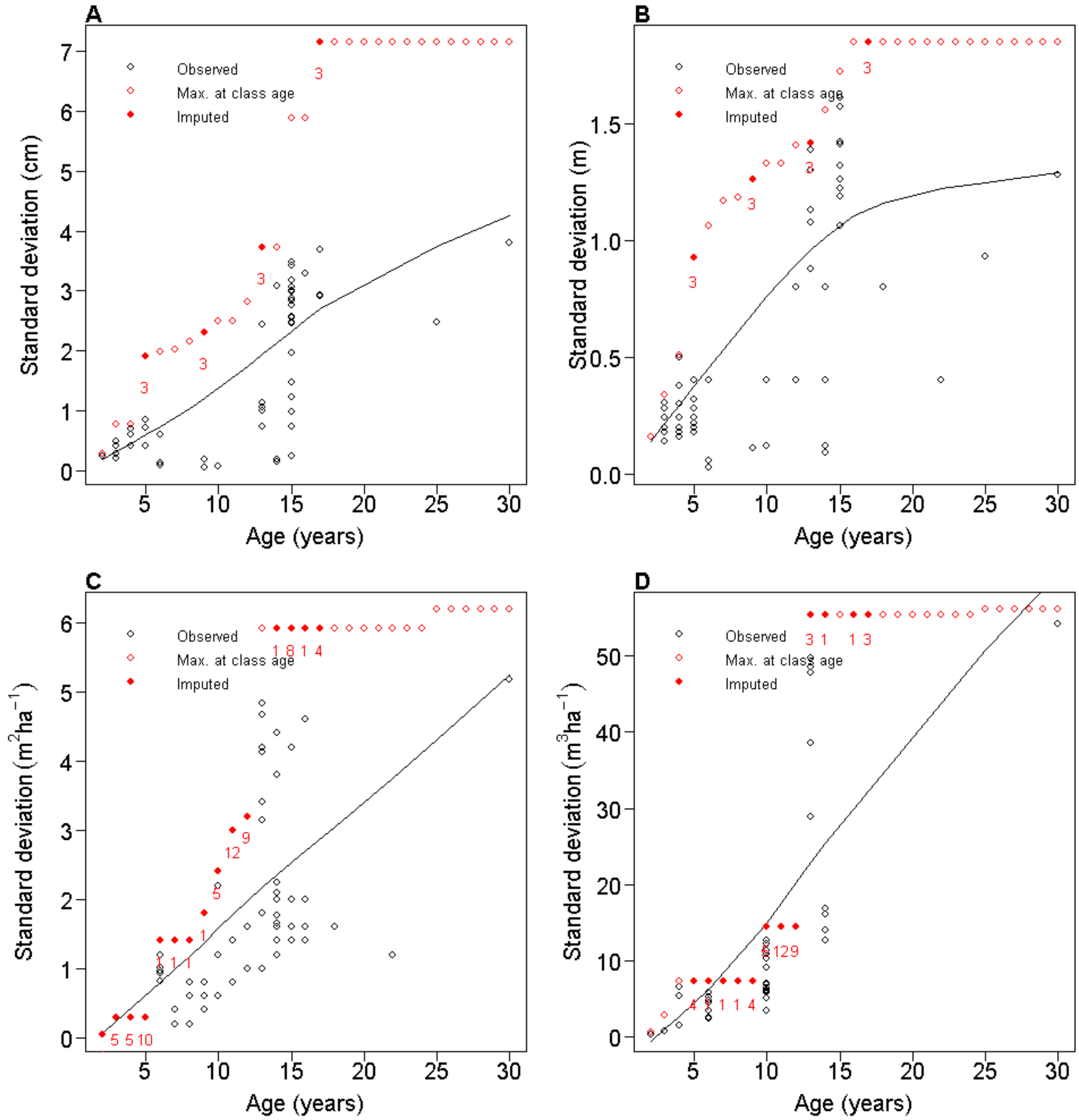

Figure 2.2: Cumulative maximum recorded standard deviation over age criterion for DBH (A), Ht (B), BA (C) and V (D). The number below the imputed standard deviation corresponds to the number of observations (combination of treatments) imputed. The solid line corresponds to smooth spline.

2.3.2 Variable selection and final parsimonious models

Full (MI) models for DBH, Ht, BA and V were fitted adding one factor in turn. Intermediate models from reduced (M0) to full (MI) models, including one growth factor at a time, had highly significant LRT (p-value<0.0001), except the V model including genetic improvement (only *Gen*(HS)) (p-value=0.2). Then, the parsimonious (MII) models were estimated excluding non-significant factor levels by collapsing those into the corresponding factor reference level. Most of the parameter estimates in the parsimonious (MII) models for DBH, Ht, BA, and V were highly significant (p-value<0.01 to <0.0001) (Table 3). However, two parameter estimates in the model for Ht, and two parameter estimates in the model for BA were statistically significant (p-value<0.05) (Table 3). The final estimated model for V is a function of the management and physiographic region, since genetics was dropped in the forward selection and density was dropped in the backward selection.

2.3.3 Fit evaluation and diagnostics

The variance associated with the random effects in the four parsimonious (MII) models were statistically significant ($\tau^{*2} \neq 0$, Q's p-value<0.0001) (Table 2.4). Therefore, the mixed effects model, with the random effects accounting for heterogeneity between treatments, seems reasonable for this meta-regression. The F statistic suggests that covariates in all models explained the variability of the yield (p-value<0.0001); in other words, by including the covariates, a significant statistical reduction in the sum squares of error was achieved (Table 2.4). Likewise, the Raudenbush's pseudo-R^2 ranged from 88 to 97% (Table 2.4).

The risk of multicollinearity can be considered relatively small for DBH, Ht, and BA ($\overline{VIF}$ <10), and moderate for V ($\overline{VIF} \approx 10$) (Table 2.4). The bias induced by the logarithmic transformation was negligible in all models ($\gamma \approx 1$) (Table 2.4). All estimated models approximately met the assumptions of constant variance and normality (Figure 2.3). Moreover, Cook's distance and funnel plots were evaluated. Few outliers were detected, and evidence of publication bias was not found (figures not shown).

Table 2.3: Parameter estimates of the parsimonious (MII) Schumacher models (logarithmic units) of loblolly pine for DBH, Ht, BA and V in the southeastern United States using meta-regression.

Variable	Estimate	Standard error	t value	p-value	CI LB	CI UB
DBH						
Intercept	3.58	0.052	68.912	<0.0001	3.477	3.683
Gen(FS)	0.217	0.054	4.05	<0.0001	0.111	0.323
Gen(C)	0.214	0.056	3.814	<0.0001	0.103	0.326
Mgmt(M)	0.149	0.035	4.305	<0.0001	0.08	0.218
Mgmt(H)	0.295	0.041	7.29	<0.0001	0.215	0.375
Reg(LCP)	-0.199	0.061	-3.291	0.001	-0.319	-0.079
Den	-2.9×10^{-4}	3.0×10^{-5}	-9.66	<0.0001	-3.5×10^{-4}	-2.3×10^{-4}
A^{-1}	-7.36	0.453	-16.238	<0.0001	-8.259	-6.46
Reg(LCP) $\times$ A^{-1}	1.702	0.443	3.84	<0.0001	0.822	2.582
Den $\times$ A^{-1}	6.4×10^{-4}	1.4×10^{-4}	4.464	<0.0001	3.5×10^{-4}	9.2×10^{-4}
Ht						
Intercept	3.41	0.061	56.259	<0.0001	3.289	3.53
Gen(FS)	0.192	0.072	2.675	0.009	0.049	0.334
Gen(C)	0.278	0.118	2.349	0.021	0.043	0.512
Mgmt(M)	0.163	0.048	3.377	0.001	0.067	0.259
Mgmt(H)	0.211	0.058	3.668	<0.0001	0.097	0.326
Reg(LCP)	-0.706	0.095	-7.466	<0.0001	-0.893	-0.518
Den	-1.6×10^{-4}	3.8×10^{-5}	-4.165	<0.0001	-2.3×10^{-4}	-8.2×10^{-5}
A^{-1}	-9.615	0.568	-16.917	<0.0001	-10.745	-8.485
Reg(LCP) $\times$ A^{-1}	4.828	0.581	8.31	<0.0001	3.673	5.983
Den $\times$ A^{-1}	4.8×10^{-4}	1.9×10^{-4}	2.586	0.011	1.1×10^{-4}	8.6×10^{-4}
BA						
Intercept	3.513	0.116	30.376	<0.0001	3.285	3.742
Gen(HS)	0.305	0.08	3.805	0.0002	0.147	0.463
Gen(FS)	0.439	0.113	3.873	0.0002	0.215	0.663
Mgmt(M)	0.386	0.062	6.19	<0.0001	0.263	0.51
Mgmt(H)	0.479	0.128	3.757	0.0002	0.227	0.731
Reg(UCP)	0.736	0.145	5.059	<0.0001	0.448	1.023
Den	1.8×10^{-4}	7.1×10^{-5}	2.493	0.0137	3.7×10^{-5}	3.2×10^{-4}
A^{-1}	-11.154	0.342	-32.597	<0.0001	-11.83	-10.478
Mgmt(H) $\times$ A^{-1}	1.328	0.599	2.218	0.0281	0.145	0.251
Reg(UCP) $\times$ A^{-1}	-11.56	1.28	-9.037	<0.0001	-14.087	-9.033
Den $\times$ A^{-1}	1.0×10^{-3}	2.7×10^{-4}	3.846	0.0002	5.0×10^{-4}	1.6×10^{-3}
V						
Intercept	6.634	0.185	35.926	<0.0001	6.266	7.001
Mgmt(M)	0.378	0.142	2.664	0.0093	0.096	0.661
Mgmt(H)	0.731	0.164	4.45	<0.0001	0.404	1.058
Reg(LCP)	-1.004	0.263	-3.825	0.0003	-1.527	-0.482
A^{-1}	-19.053	1.549	-12.303	<0.0001	-22.135	-15.972
Reg(LCP) $\times$ A^{-1}	9.351	1.706	5.481	<0.0001	5.956	12.745

Table 2.4: Fit evaluation of the MII Schumacher model (logarithmic units) of loblolly pine for DBH, Ht, BA and V in the southeastern United States using meta-regression.

Statistic	DBH	Ht	BA	V
Q^*	11270	6643	19404	12093
	(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)
$\hat{\tau}^{*2}$	0.01	0.021	0.075	0.202
$\hat{\tau}^{*}$	0.101	0.145	0.274	0.449
F	194	130	430	141
	(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)
R^{*2}_{pseudo}	96.61	94	96.93	90.76
$\overline{VIF}$	7.018	7.545	4.594	10.49
γ	1.011	1.005	1.009	1.08

Annotations for Table 2.3: A: Age (years); Gen(HS): half-sibling, Gen(FS): full-sibling, Gen(C): clone; Mgmt(M): moderate management; Mgmt(H): high management; Reg(UCP): upper coastal plain, Reg(LCP): lower coastal plain, CI: 95 % confidence interval; CI LB: 95% CI lower bound, CI UB: 95% CI upper bound.

Annotations for Table 2.4: Q^* is the test for residual heterogeneity; $\hat{\tau}^{*2}$ is the residual heterogeneity; $\hat{\tau}^{*}$ is the squared root of $\hat{\tau}^{*2}$; F is the test of statistical significance of covariates; R^{*2}_{pseudo} is the Raudenbush's pseudo-R^2; $\overline{VIF}$ is the mean variance inflation factor; and γ is the bias due to the logarithmic transformation. Values in parentheses represent the p-value of the test.

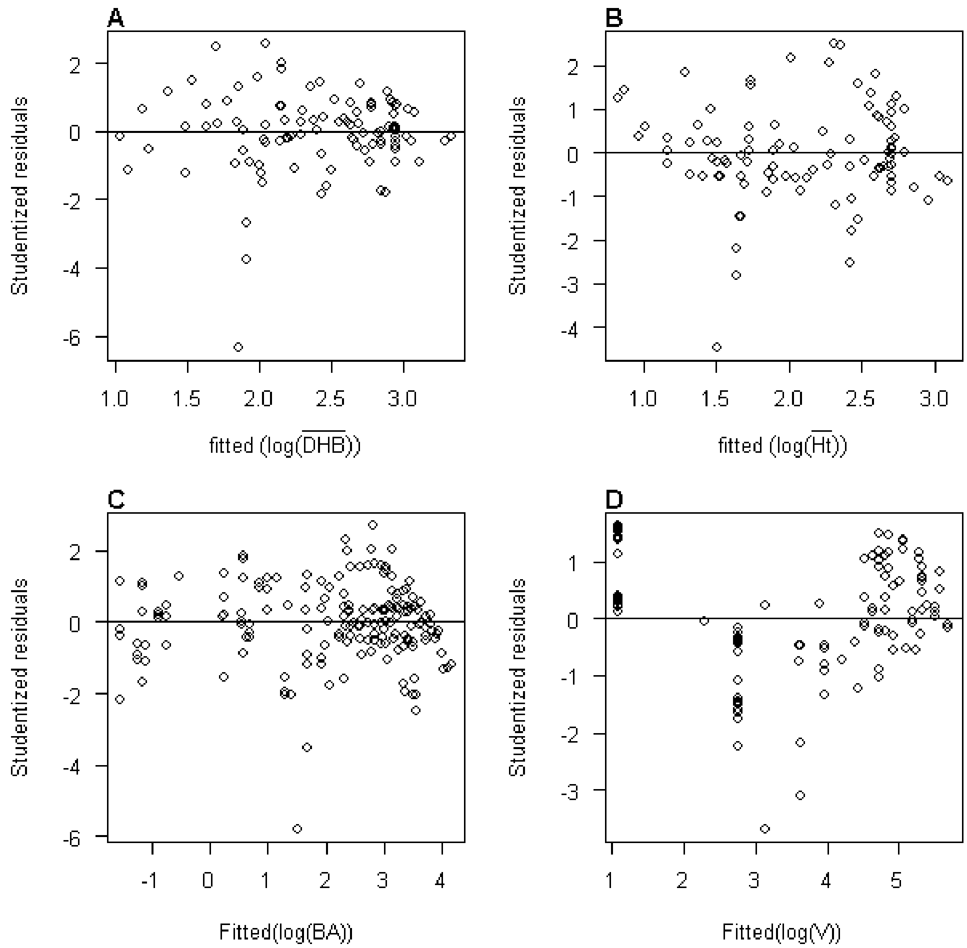

Figure 2.3: Standardized residuals of the meta-regression parsimonious (MII) Schumacher models for loblolly pine in the southeastern United States. (A) diameter at breast height (DBH), (B) total height (Ht), (C) basal area (BA), and (D) total volume (V)

2.3.4 Yield curves

Overall, more advanced genetics (i.e. *Gen*(FS) and *Gen*(C)), and higher management levels resulted in higher expected mean yields. Yield curves are presented by forest growth factors in Figures 2.4 to 2.7. The Upper Coastal Plain and Piedmont regions had higher expected mean yields than the Lower Coastal Plain. Moreover, the higher the density, the lower the expected mean of DBH, and Ht; and the higher the BA. Despite the statistical significance of the parameter estimates in the models, some of the 95% confidence intervals (CI) of the curves drawn for each of the growth factors suggest no statistical difference with respect to their reference levels (*Gen*(UU), *Mgmt*(L), *Reg*(UCPP), and Den was estimated utilizing Equation 2.8 as a function of planting density, assumed equal to 1500 TPH, and stand age) (Figures 2.4 to 2.7). In these figures, the only cross factor is stand density to account for the mortality effect.

For DBH (Figure 2.4), *Gen*(HS) was collapsed into the reference level, *Gen*(UU), due to its parameter estimate being non-significant (p-value>0.05), the 95% CI of advanced genetics (*Gen*(FS) and *Gen*(C)) did not overlap the 95% CI of the corresponding reference level (*Gen*(UU-HS)) but the 95% CI of *Gen*(FS) and *Gen*(C) are virtually identical overlapping each other; therefore, the performance of these two genetic entries is very similar over time. The 95% CI of *Mgmt*(M) just overlapped the 95% CI of the reference level (*Mgmt*(L)) after age 25 years, and the 95% CI of *Mgmt*(H) overlapped the 95% CI of *Mgmt*(M). The 95% CI of the significant level of physiographic region, *Reg*(LCP), overlapped the corresponding reference level (*Reg*(UCPP)). The yield curve for *Reg*(LCP) was truncated at 15 years to reflect the maximum age data available for the region.

For Ht (Figure 2.5), *Gen*(HS) was collapsed into the reference level *Gen*(UU) due to its parameter estimate being non-significant (p-value>0.05), the 95% CI of advanced genetics (*Gen*(FS) and *Gen*(C)) overlapped the reference level (*Gen*(UU-HS)), and the 95% CI of *Gen*(C) overlapped the 95% CI of *Gen*(FS). The moderate and high management levels (*Mgmt*(M) and *Mgmt*(H)) overlapped the reference level *Mgmt*(L) and the 95% CI of *Mgmt*(H) overlapped the 95% CI of *Mgmt*(M). The 95% CI of *Reg*(LCP) overlapped the reference level *Reg*(UCPP) between ages 5 and 8 years, and exhibited a different trend compared to the reference level *Reg*(UCPP) (Figure 2.5C). The yield curve for *Reg*(LCP) was truncated at 15 years to reflect the maximum age data available for the region.

For BA (Figure 2.6), the 95% CI of *Gen*(FS) and *Gen*(HS) overlapped the reference level after age 20 years, and the 95% CI of *Gen*(FS) overlapped the 95% CI of *Gen*(HS). The 95% CI of *Mgmt*(M) and *Mgmt*(H) did not overlap the reference level, but the 95% CI of *Mgmt*(H) overlapped the 95% CI of *Mgmt*(M). *Reg*(LCP) was collapsed into the reference level *Gen*(UCPP) due to its parameter estimate being non-significant (p-value>0.05), and the 95% CI of *Reg*(UCP) overlapped its reference level all the way between 10 and 30 years, but having different trends.

For V (Figure 2.7), the 95% CI of *Mgmt*(M) overlapped the reference level *Mgmt*(L) and the 95% CI of *Mgmt*(H) overlapped the 95% CI of *Mgmt*(M). The 95% CI of *Reg*(LCP) overlapped the reference level *Reg*(UCPP) between ages 6 and 24 years, but having different trajectories. The yield curve for *Reg*(LCP) was truncated at 12 years to reflect the maximum age data available for the region.

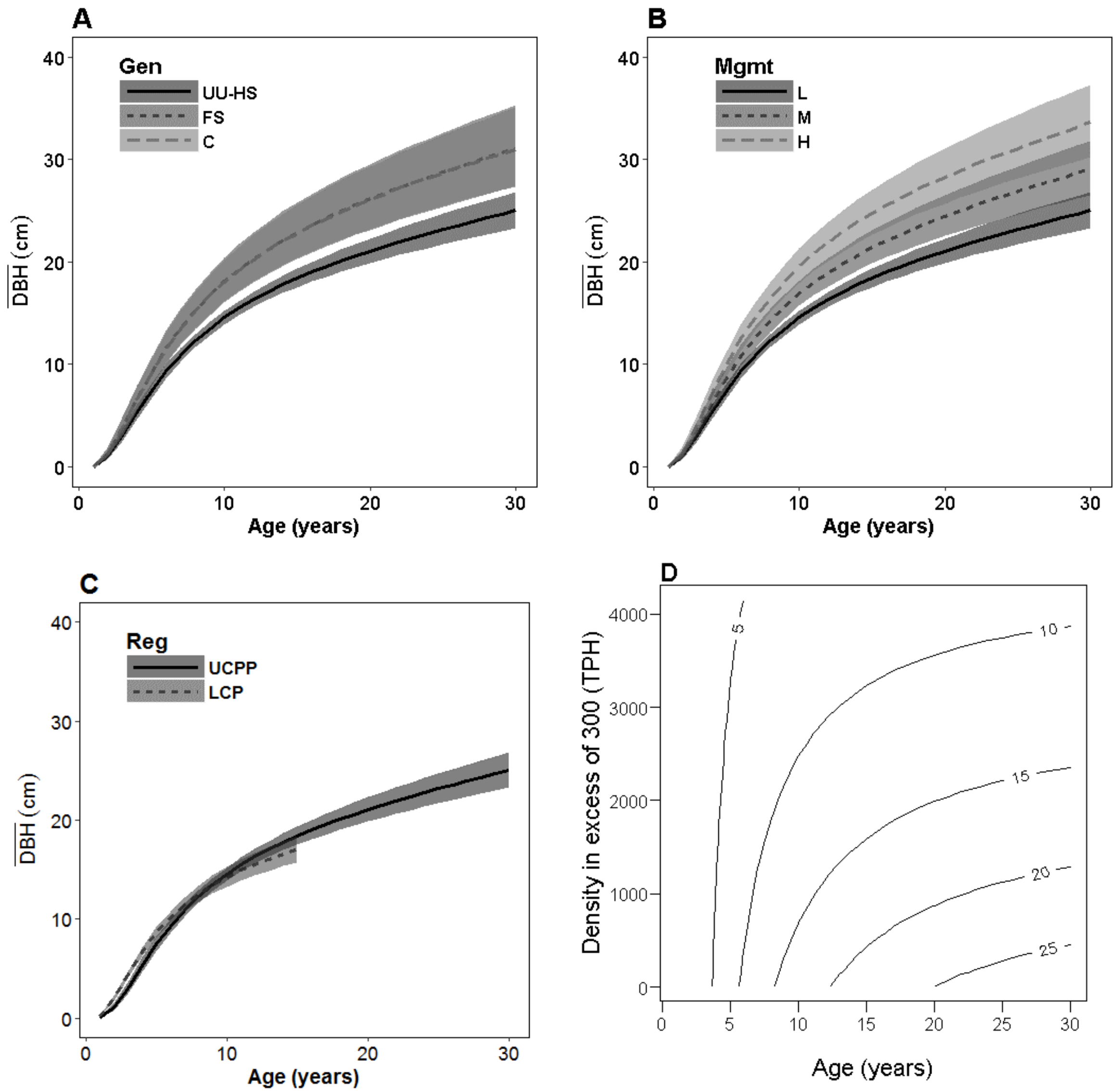

Figure 2.4: Mean and 95% confidence interval (CI, shaded region) for average diameter at breast height (DBH, cm) of loblolly pine with planted density equal to 1500 TPH (for A, B, and C) in the southeastern United States over age by (A) Genetics, UU: unimproved, HS: half-sibling, FS: full-sibling, C: clone; (B) Management, L, M, H: low, moderate and high, respectively; (C) Physiographic region, UCPP: Upper Coastal Plain-Piedmont, LCP: Lower Coastal Plain; (D) contour plot of DBH as function of age (years) and density (trees per hectare). Note that in (A) *Gen*(HS) was collapsed into the reference level, *Gen*(UU), due to non-significance (p-value<0.05) of its parameter estimate, the 95% CI *Gen*(FS) and *Gen*(C) completely overlapped; and in (C) the curve of *Reg*(LCP) was truncated at age 15. Since the model includes stand density, the mortality effect in (A), (B), and (C) was accounted for by Equation 2.8.

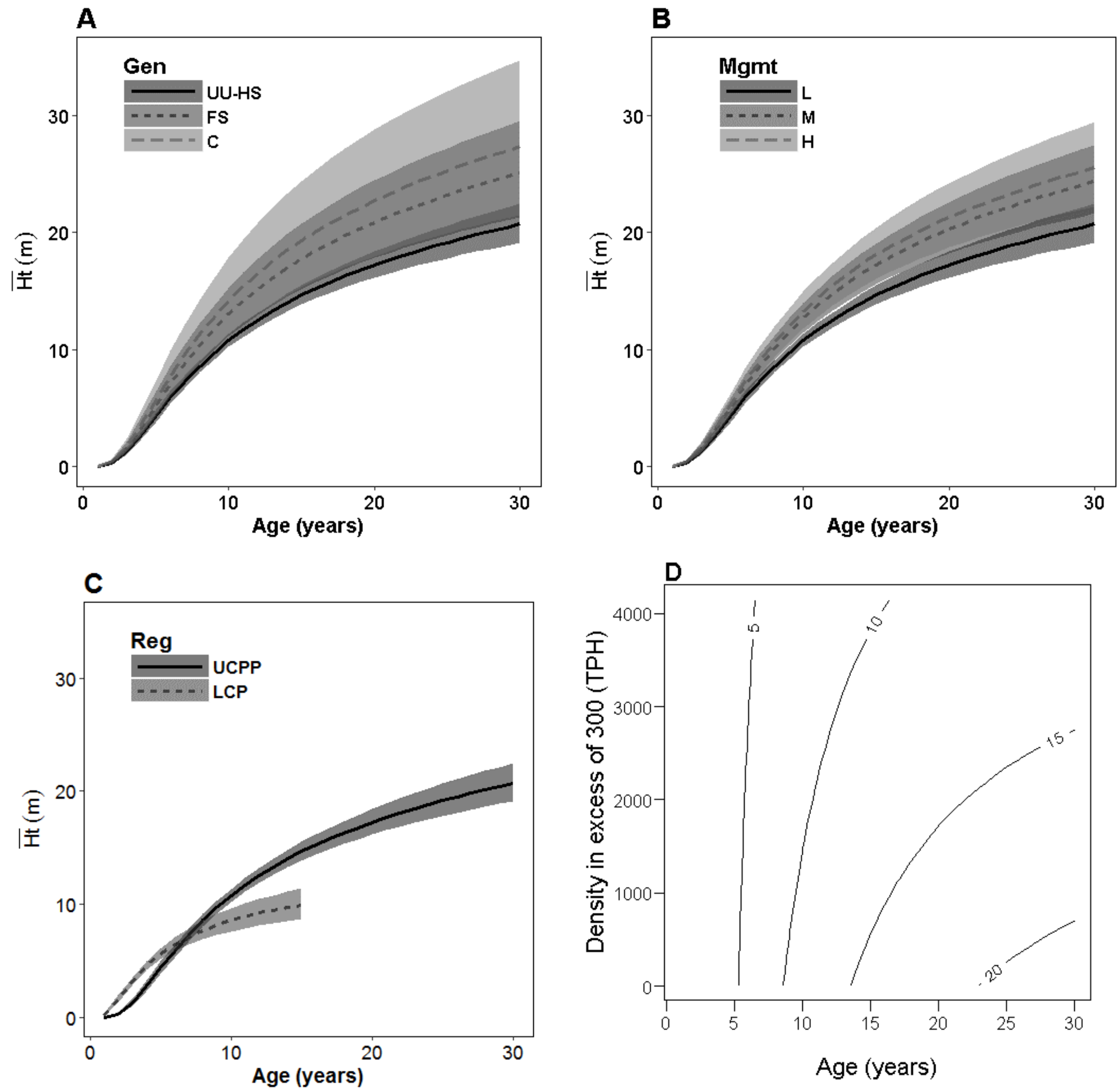

Figure 2.5: Mean and 95% confidence interval (shaded region) for average total height (Ht, m) of loblolly pine in the southeastern United States over age by (A) Genetics, UU: unimproved, HS: half-sibling, FS: full-sibling, C: clonal; (B) Management, L, M, H: low, moderate and high, respectively; (C) Physiographic region, UCPP: Upper Coastal Plain-Piedmont, LCP: Lower Coastal Plain; (D) contour plot of Ht as function of age (years) and density (trees per hectare). Note that in (A) *Gen*(HS) was collapsed into the reference level, *Gen*(UU), due to non-significance (p-value<0.05) of its parameter estimate; and in (C) the curve of *Reg*(LCP) was truncated at age 15. Since the model includes stand density, the mortality effect in (A), (B), and (C) was accounted for by Equation 2.8.

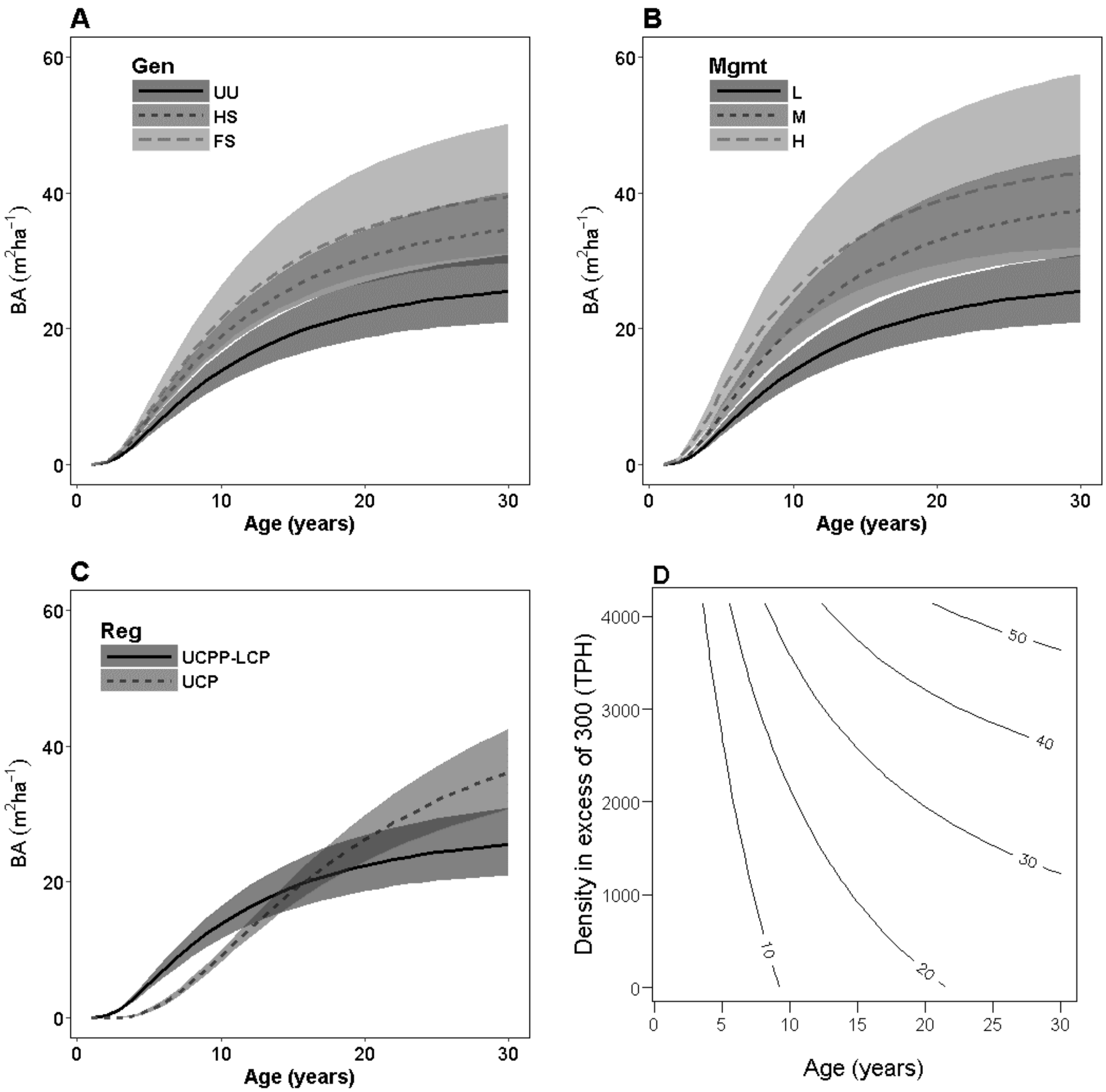

Figure 2.6: Mean and 95% confidence interval (shaded region) for basal area per hectare (BA, m^2ha^{-1}) of loblolly pine in the southeastern United States over age by (A) Genetics, UU: unimproved, HS: half-sibling, FS: full-sibling, clone was not considered; (B) Management, L, M, H: low, moderate and high, respectively; (C) Physiographic region, UCPP: Upper Coastal Plain-Piedmont, LCP: Lower Coastal Plain, UCP: Upper Coastal Plain; (D) contour plot of BA as function of age (years) and density (trees per hectare). Note that in (C) *Reg*(LCP) was collapsed into the reference level *Reg*(UCPP) due to non-significance (p-value<0.05) of its parameter estimate. Since the model includes stand density, the mortality effect in (A), (B), and (C) was accounted for by Equation 2.8.

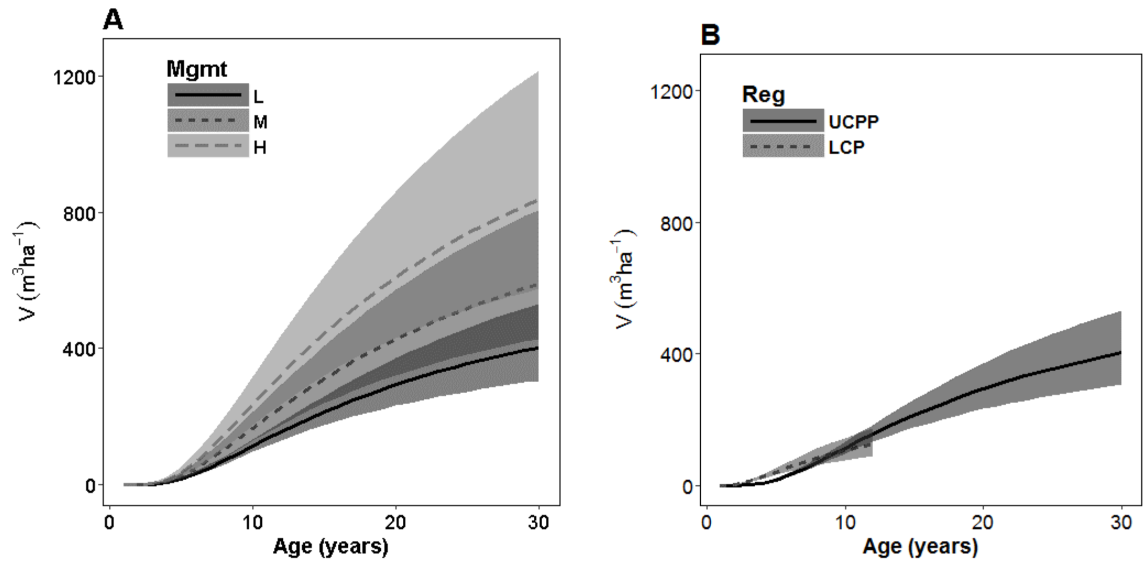

Figure 2.7: Mean and 95% confidence interval (shaded region) for volume per hectare (V, m^3ha^{-1}) of loblolly pine in the southeastern United States over age by (A) Management, L, M, H: low, moderate and high, respectively; (B) Physiographic region, UCPP: Upper Coastal Plain-Piedmont, LCP: Lower Coastal Plain. Note that in (B) the curve of *Reg*(LCP) was truncated at age 12.

2.3.5 Knowledge gaps

Width differences in the 95% confidence intervals between factor levels in the estimated curves suggests a lack of information for specific treatments or combinations of factor levels of available and selected research studies. Diagonal elements of the panels in Figure 2.8 (A to D) show frequency histograms of treatments by factor levels for the four response variables. Hence, most of the available data are for *Gen*(HS) (69%); whereas *Gen*(UU) (12%), *Gen*(FS) (13%), and *Gen*(C) (6%) were less well represented. Most of the advanced genetics (*Gen*(FS) and *Gen*(C)) information is related with DBH and Ht; whereas there is lack of studies presenting the effect of advanced genetics on the stand level basal area and volume. On the other hand, most of the treatments for DBH and Ht were for *Mgmt*(L) (62%), with low proportions of studies in *Mgmt*(M) (23%), and *Mgmt*(H) (15%); for BA and V, *Mgmt*(L) (45%) and *Mgmt*(M) (44%) were almost equally represented with a low proportion of *Mgmt*(H) (11%). In the same sense, most of the studies were conducted in

Reg(LCP) (45%) followed by *Reg*(UCP) (35%) with some studies in *Reg*(UCPP) (11%) and *Reg*(P) (9%).

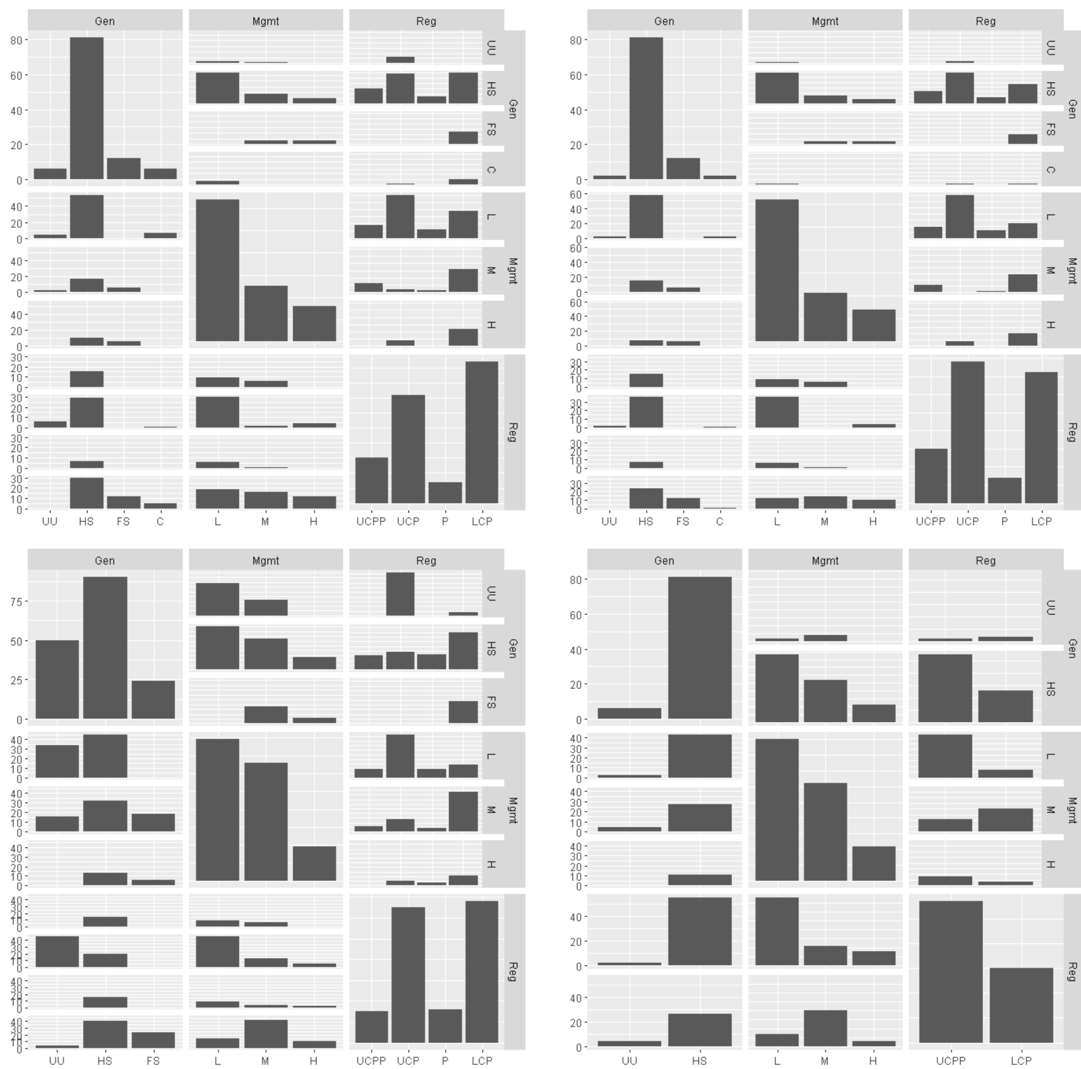

Figure 2.8: Frequency histogram of treatments by factor levels (diagonal panels) and combinations of factors levels (off-diagonal panels) for the response variables diameter at breast height (A), total height (B), basal area (C), and volume (D).

The off-diagonal elements of the panels in Figure 2.8 (A to D) present frequency histograms by combinations of factor levels, information that can be utilized to prioritize the allocation of resources to research a specific combination of treatments for which there is a lack of information. Thus, for example, no treatment considers the combination *Gen*(C), and *Mgmt*(M) or *Mgmt*(H). Since the only combination of *Gen*(FS) is with *Reg*(LCP), there is a lack of evidence on the performance of this level of genetic improvement in the other physiographic regions.

2.4 Discussion

The cornerstone of forestry research has been experimental designs to address specific forest growth factors and interactions. Thus, most of the loblolly pine growth and yield models may be locally restricted and density and genetic composition dependent based on the experimental designs used for the data to build the models. An ideal generalization using the traditional experimental-modeling approach would require such an investigational base, covering a broad range of ages, environmental conditions, management practices, and levels of genetic improvement. The amount of resources involved in such an experimental design would make the research technically challenging and economically unfeasible. The noted shortcoming of not having a large experimental base can be overcome with conclusions drawn out from a large collection of studies using meta-regression.

Mixed effects models are the most suitable approach to estimate a meta-regression of loblolly pine yield in the southeastern United States. The fixed component addresses the overall response driven by covariates, and the random term contributes to explaining the variability between the treatments (studies) (Borenstein et al., 2010). Moreover, selected studies constitute a random sample of the population of all existing studies and even those that will be conducted in the future (Viechtbauer, 2010), suggesting that the number of included treatments suffices to draw an overall conclusion about loblolly pine growth and yield in the southeastern United States. The low to medium values of the variance inflation factors suggests a relative independence between (among) the forest yield factors in the explanation of variability. Hence, the relative independence of factors may indicate that the yield mean can be estimated even if a particular combination of factor levels was not included in this meta-regression (e.g. *Gen*(FS) in *Reg*(UCP)).

Even though knowledge gaps may not affect point yield estimates, they impact yield confidence intervals. The width of a CI is mainly dependent on the number of available and selected studies, and treatments. In this sense, imputing missing standard deviation and stand density values, and estimating DBH means and BA values were justified to increase the number of treatments to be included in the meta-regression models. We propose novel methods to retrieve yield estimates by using mathematical statistical transformations and imputations. More precise yield estimates can be achieved by increasing the number of studies and treatments, especially in the factor levels under-represented in this current research. This represents a good opportunity for researchers, universities, forest research cooperatives, private companies and policy makers to initiate investigations in these areas of knowledge gaps.

To the best of our knowledge, this is the first time that the method of imputing standard deviation by utilizing the cumulative maximum recorded standard deviation over age criterion is used in a meta-analysis. We believe the method of imputing standard deviation over age is required given the heteroscedastic nature of forest yield data. Little statistical research has been done imputing standard deviations when researchers of the original studies to be used in a meta-analysis failed to report those values. The existing literature suggests that it is safe imputing standard deviation when missing in studies (Furukawa et al., 2006). Furthermore, a meta-analysis in education reported a rate of up to 80% of imputed standard deviations based on local and regional studies (Borman et al., 2003).

The chief use of standard deviations in a meta-analysis is to weight the mean of a treatment or study. Thus, by assigning the cumulative maximum recorded standard deviation over age to a study for which its standard deviation is missing, we conservatively gave the same credibility as the more variable (less reliable) studies. We think the trade-off of imputing missing standard deviations is positive. The potential negative impact of applying the imputation method on the meta-regression model performance is negligible, but including one additional treatment combination is worthy, especially for those combinations that have few data points.

Forest growth factors that were included as simple terms in the asymptote, or as interaction terms with age in the growth rate, explained the variability of loblolly pine yield in the southeastern United States. Improved genetic categories had higher asymptotes (timber production potential) than unimproved categories in the models for DBH, Ht and BA. Based on the 95% CI of yield curves, the performance of *Gen*(FS) and *Gen*(C) were statistically different compared with the performance of *Gen*(UU-HS) (collapsed factor level) for DBH; and between *Gen*(FS) and *Gen*(HS) compared with *Gen*(UU) for BA (at least until age 20 years). However, the 95% CI of advanced genetics (*Gen*(FS) and *Gen*(C)) and the 95% CI of the reference level were not statistical different for Ht. Genetic background (only *Gen*(HS)) was not a significant factor for V, which is contrary to the reported volume gains of genetically improved loblolly pine over unimproved planting stock (Allen et al., 2005; McKeand et al., 2006a).

Moderate and high levels of management were statistically significant in estimated models, having a positive effect on the asymptote and on the growth rate (only for BA). This is in accordance with previous findings indicating that silvicultural practices such as mechanical site preparation, vegetation control, fertilization and irrigation enhance timber production (Allen et al., 2005). Based on the 95% CI of the V curves, the high level of management (*Mgmt*(H)) showed a statistically difference over time compared to the low level of management (*Mgmt*(L)). However, the 95% CI of moderate level of management (*Mgmt*(M)) overlapped the reference level, suggesting that there is no statistical difference over time.

Clearly, the performance of loblolly pine is not the same across the southeastern United States. Estimated models showed statistical differences in the performance of loblolly pine associated with the environmental effect or site quality factors (climate, topography, and soil) accounted for in general by the physiographic region. Most of the existing forest yield models for loblolly pine in the southeastern United States do not include physiographic region as a covariate. Therefore, the statistical significance of this important driver on forest yield models cannot be directly tested. In that sense, all estimated models in this research include physiographic region as significant factor (p-value<0.05). Statistical differences in the performance of loblolly pine over time by physiographic region were found for Ht. The 95% CI of Ht over time in *Reg*(LCP) was different from the reference level (*Reg*(UCPP)), which suggests that trees in the Lower Coastal Plain region are shorter than those in the Upper Coastal Plain or Piedmont regions (combined term). A similar pattern regarding physiographic region is noted for V, with stands in the Lower Coastal Plain having a lower volume asymptote (production potential) than stands in the Upper Coastal Plain Piedmont region.

Stands in the Lower Coastal Plain exhibited the lowest forest yield (DBH, Ht, BA, and V) of the considered physiographic regions. Since the estimated models can isolate the effect of forest growth factors, it was possible to tease apart the effect of *Reg*(LCP) from the management factor. In the Lower Coastal Plain, bedding, herbicide application, and fertilization have traditionally been mandatory silviculture practices because of the high water table, the presence of competing vegetation and nutrient deficiencies, which corresponds to *Mgmt*(M) in our definition of management. Although it is well-known that poor sites under moderate or high management would have higher marginal yield responses than good sites, it is possible that the perception of the yield responses in the Lower Coastal Plain are the product of environment and management interaction. To clarify the trends based on our meta-regression models, we truncated the curves associated with D, H, and V for the *Reg*(LCP) to reflect the age range of data used in estimation. In contrast, we consider that there is no such a lack of information for BA because *Reg*(LCP) was collapsed into the *Reg*(UCPP) that has older stands represented in the dataset.

Stand density was the last factor considered in the estimated models, statistically significant in the DBH, Ht and BA models, but not significant for the V model. Density has a direct effect on DBH and BA, the higher the number of trees, the lower the mean DBH and the higher the BA. However, height has historically been assumed to be independent of density (Clutter et al., 1992), but evidence about the interaction between density and height in loblolly pine stands in the southern United States has been well documented (Antón-Fernández et al., 2011; Henskens et al., 2001; Land et al., 2004; MacFarlane et al., 2000). Although basal area increases as stand density increases, total height decreases as stand density increases since the *Den* parameter estimate in the Ht model has a negative sign. Volume is a function of random variables (form factor, height and basal area). In this simple volume expression, the inverse relationship between stand density and height may negate the positive relationship between basal area and stand density. However, it is difficult to foresee the result of the product of these three random variables without a thorough mathematical proof, which is outside of the scope of this research. Pienaar, Turnbull (1973) did not note any relationship between the asymptote of volume yield and the number of trees per unit area over a wide range of densities.

The purpose of this paper was to give insight into the factors impacting forest growth for loblolly pine while trying to keep the estimated models simple. However, there are some alternative procedures and covariate selection that may improve yield estimation and prediction ability. Although registered from studies when available, more precise data about the levels of genetic improvement (i.e. breeding generation or specific family), silvicultural practices (i.e. mechanical site preparation, fertilization, and vegetation control), and environmental conditions (temperature, precipitation, soil series, and CRIFF soil classification) were not used.

In that sense, physiographic region has traditionally been used to represent the environmental component in the yield models. However, other variables like water deficit and excess indices, soil physical or chemical properties, depth to the water table and restrictive layers, may explain additional variability. On the other hand, all silvicultural treatments were implicitly included in the management factor. Although it would be more accurate to consider each of the silvicultural treatments independently in addition to other factors, the resulting model will not be parsimonious as intended. In that regard, the level of application (e.g. amount of fertilizer), timing (age of application), and year of response (years after treatment), could be considered, but it would add more complexity to the model.

To the best of our knowledge, this is the first attempt to use meta-regression to estimate forest yield. Meta-regression seems to be a very promising technique to analyze forest growth and yield, a multifactorial phenomenon that requires advanced methods to strengthen findings and conclusions. Applications of this statistical method should contribute to obtaining more reliable estimates of forest yield and timber production; to elicit Bayesian priors for forest yield models; to implement accurate estimations for forestry planning; to measure, manage and reduce the uncertainties associated with yield; and to enhance financial and risk analyses of timberland investments. However, more steps will be taken in future works to improve the estimation and prediction ability of models: increasing the number of selected studies, trying different factors or covariates arrangements, assessing the effect of subfactors (e.g. water deficit and soil classes within environment; mechanical site preparation, fertilization, and vegetation control within management; and family generation within genetics), evaluating the effect of standard deviation imputation methods comparing the criterion used in this research and the expected value of the standard deviation, modeling yield and growth with nonlinear procedures, and including factor and variable interactions (e.g. genetics and environment, environment and management, and genetics and management).

2.5 Acknowledgments

The authors would like to thank the Harley Langdale Jr. Center for Forest Business and the Plantation Management Research Cooperative (PMRC) at the University of Georgia for supporting this research. We also would like to thank all of the researchers who installed, maintained, measured, and analyzed all the research trials we used in this research. The comments of the Associate Editor and anonymous reviewers greatly helped to improve the manuscript.

Chapter 3

Contribution of silviculture to loblolly pine growth and yield in the southeastern United States: A meta-analysis[1]

[1]Restrepo, H.I, B.P. Bullock, C.R. Montes. 2018. Contribution of silviculture to loblolly pine growth and yield in the southeastern United States: A meta-analysis. Accepted by Proceedings of the 19th biennial southern silvicultural research conference. e-Gen. Tech. Rep. SRS-234. Reprinted here with permission of publisher.

Abstract

There has been an increase in loblolly pine production driven by forest management practices like intensive silviculture and improved genetics. Some reported yield gains have been modeled using meta-regression mixed effects models accounting for the potential contribution of the four factors related to forest growth: age, site quality (environment), establishment culture and management, and stand intrinsic characteristics (genetics and initial planting density). The aim of this research was to describe a methodology that allows for the derivation of response equations from yield models for diameter at breast height, stand average height, basal area, and total volume in the Southeastern United States. When compared to low-level silviculture, moderate and intensive silviculture show volume gains at age 20 of 221 and 314 m^3/ha, respectively. Likewise, moderate and intensive management consistently performed better over time as compared to low management for all response variables. These management response curves and their associated mathematical expressions can be used to perform financial marginal analyses to improve forest land decision making.

Key words: *Pinus taeda* L., Schumacher model, silvicultural responses, volume gain.

3.1 Introduction

Loblolly pine (*Pinus taeda* L.) is the most commercially important forest species in Southeastern United States. The timber production in this region has been enhanced using genetically improved seedlings and a wide range of silvicultural treatments such as mechanical site preparation, vegetation control, fertilization, and irrigation (Allen et al., 2005). These intensive forest management practices, in conjunction with other key factors like initial planting density, stand age, and environmental conditions, result in the expression of forest yield (Clutter et al., 1992). Hence, volume gains resulting from intensive management practices can be analyzed by isolating the effect of age, environment, and density using growth and yield models. However, most of the existing models account only for one of the forest growth factors, either management, genetics, or environment. Thus, most of the research on the effect of intensive silviculture on loblolly pine growth and yield is locally restricted and/or density- and genetic composition-dependent. Therefore, those conclusions cannot be easily generalized.

An ideal response generalization would require a large experimental base, covering a wide range of ages, environmental conditions, management practices, and genotypes. The amount of resources involved in such an experimental base may make this kind of research technically challenging and economically unfeasible. In other research areas (like medicine), scientists have overcome the problem of not having a large experimental base with conclusions drawn out of a large collection of independent studies using meta-analysis. Meta-analysis is a statistical technique utilized to compile information for the purpose of integrating the findings as a rigorous alternative to the traditional narrative discussion (Schwarzer et al., 2015).

There is a meta-analysis on forest yield of loblolly pine in the Southeastern United States that accounts for all growth factors (Restrepo et al., 2019), i.e., effect of age, environment (through physiographic region), genetics, density, and management as explanatory variables for diameter at breast height (DBH), average height (Ht), basal area (BA), and total volume

(V). Therefore, these yield models can be used to derive silvicultural responses isolating the effect of the remaining factors, which is the purpose of this paper.

3.2 Methods

3.2.1 Four-factor forest yield models

The mean, standard deviation, and number of observations of the response for DBH, Ht, BA, and V were extracted from 21 studies selected out of 500 studies in the Southeastern United States in a meta-analysis framework (Restrepo et al., 2019). Included studies constitute a representative sample size of loblolly pine yield (Table 3.1) over a wide range of environmental conditions from 44 counties located in 10 States across the Southeastern United States (Figure 3.1). The model is termed a four-factor model because it considers covariates from the four factors of forest growth:

- Age: age of the stand in years
- Genetics: genetically improved categories of loblolly pine [unimproved or unknown (UU), half-sibling (HS), full-sibling (FS), and clone (C)]
- Mgmt: management intensity [low (L), moderate (M), and high (H) in quantity and frequency of inputs and applications]
- Region: physiographic regions [Upper Coastal Plain and Piedmont together (UCPP), Upper Coastal Plain (UCP), Piedmont (P), and Lower Coastal Plain (LCP)]
- Density: surviving density in stems/ha

Table 3.1: Number of treatments, number of measurement plots, and the summed total area of measurement plots over the selected research studies in the meta-regression of loblolly pine growth and yield in the southeastern United States.

Response variable	No. treatments	No. plots	Area (ha)
DBH	105	1288	79
Ht	97	1344	81
BA	176	1476	70
V	111	1012	86

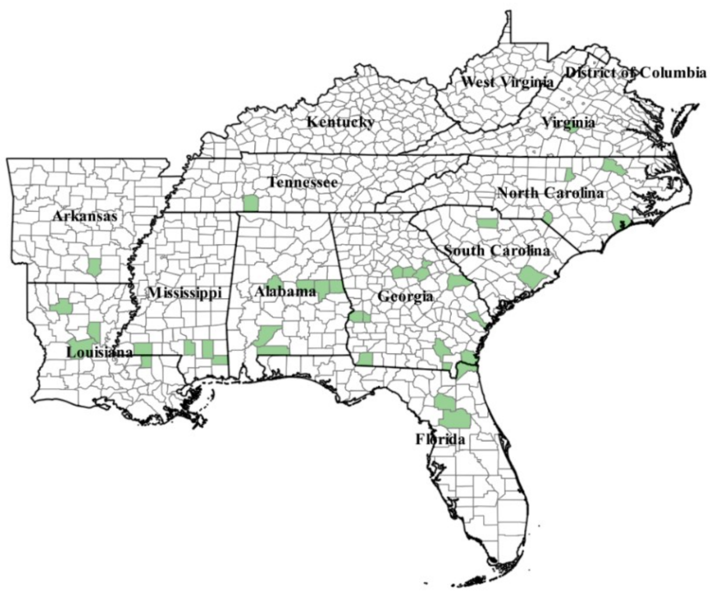

Figure 3.1: Southeastern United States counties in which studies have been conducted that were utilized in this research.

Forest yield models correspond to the log-transformed Schumacher model (Schumacher, 1939) estimated using linear mixed effects models (Restrepo et al., 2019) (Table 3.2):

$$y_i = \theta + u_i + \varepsilon_i \quad (3.1)$$

$$u_i \sim \mathcal{N}(0, \tau^2); \varepsilon_i \sim \mathcal{N}(0, \sigma_i^2); \mathsf{cov}(\varepsilon_i, u_i) = 0$$

where:

θ= the fixed effects term

Table 3.2: Summary of the preliminary loblolly pine yield models for the Southeastern United States using meta-regression (Restrepo et al., 2019). Significance codes: *** = p-value<0.0001; ** = p-value<0.001; * = p-value<0.01; . = p-value<0.05; blank = p-value<0.1.

Source	Estimates (log) and significance level							
	DBH		Ht		BA		V	
Intercept	3.67	***	3.46	***	3.48	***	5.67	***
Genetics(HS)					0.29	***		
Genetics(FS)	0.22	***	0.19	***	0.42	***		
Genetics(C)	0.21	**	0.28	*				
Mgmt(M)	0.15	***	0.16	**	0.39	***	0.67	***
Mgmt(H)	0.29	***	0.21	***	0.49	***	0.82	***
PhyRegion(UCP)					0.72	***		
PhyRegion(LCP)	-0.2	***	-0.71	***				
Density	-3×10-4	***	-1.6×10-4	***	-1.7×10-4	*		
1/Age	-7.36	***	-9.62	***	-11.53	***	-10.73	***
Genetics(HS)×1/Age								
Mgmt(H)×1/Age					1.31	***		
Region(UCP)×1/Age					-11.52	***		
Region(LCP)×1/Age	1.7	***	4.83	***				
Density×1/Age	6.4×10-4	***	4.8×10-4	*	1.1×10-3	***		

u_i = the random effects term assumed normally distributed with zero mean and variance τ^2

ε_i = the error term assumed normally distributed with zero mean and variance σ_i^2 and independent to random effects

An estimator for θ is:

$$\begin{aligned}\hat{\theta}^*_{jkl} = \quad & \hat{\alpha}_0 + \hat{\alpha}_{1j}(Genetics) + \hat{\alpha}_{2k}(Mgmt) + \hat{\alpha}_{3l}(Region) + \hat{\alpha}_4(Density) \\ & + \frac{1}{A}\left(\hat{\beta}_0 + \hat{\beta}_{1j}(Genetics) + \hat{\beta}_{2k}(Mgmt) + \hat{\beta}_{3l}(Region) + \hat{\beta}_4(Density)\right) \\ & j = 1,2,3;\ \ k = 1,2;\ \ l = 1,2,3\end{aligned} \tag{3.2}$$

where:

$\hat{\theta}^*_{jkl}$ = an estimator of the fixed effects of DBH, Ht, BA, or V in logarithmic units of the genetics j (HS=1, FS=2, C=3) with management regime k (M=1, H=2) in the physiographic region l (UCP=1, P=2, LCP=3)

α (asymptote) = parameter estimate

β (slope) = parameter estimate

Yield model for V did not consider the effect of Genetics(FS) and Genetics(C) due to the lack of observations of those levels of genetics.

3.2.2 Silvicultural responses

Responses for DBH, Ht, BA, and V associated with moderate and high levels of management were derived from the yield model with respect to the low level of management:

$$\begin{aligned}\frac{\partial\hat{\theta}_{jkl}}{\partial Mgmt(M)} &= \left(\hat{\alpha}_{21} + \frac{\hat{\beta}_{21}}{Age}\right)\hat{\theta}\\ \frac{\partial\hat{\theta}_{jkl}}{\partial Mgmt(H)} &= \left(\hat{\alpha}_{22} + \frac{\hat{\beta}_{22}}{Age}\right)\hat{\theta}\end{aligned} \tag{3.3}$$

These partial derivatives with respect to the low level of management were fixed to HS and UCPP levels of genetics and physiographic region, respectively, and the surviving density based on an arbitrary planting density of 1,500 trees/ha was estimated using the following equation (Rose et al., 2002):

$$\widehat{Den} = 2.5 + (1500 - 2.5)(1 + 0.68A)^{1.46}(1 + A)^{-1.35}\exp[-5.9\times 10^{-4}A^2] \tag{3.4}$$

3.3 Results and discussion

Responses for DBH, BA, Ht, and V were consistently ranked over time from Mgmt(M) to Mgmt(H) (Figure 3.2). Thus, at age 20 Mgmt(M) added 3.6 cm, 3.3 m, 16 m^2/ha, and 221 m^3/ha of DBH, Ht, BA, and V, respectively, with respect to Mgmt(L); whereas, at the same age, Mgmt(H) added 8.3 cm, 4.5 m, 27 m^2/ha, and 314 m^3/ha to the corresponding variables with respect to Mgmt(L). Basal area response curves are flat up to age 5 when response curves started exhibiting a linear-looking trend up to age 20. Partial derivatives of the yield models with respect to Mgmt(M) and Mgmt(H) are:

$$
\begin{aligned}
\frac{\partial DBH}{\partial Mgmt(M)} =& 0.15 \exp\left(3.67 + 0.15 - 3 \times 10^{-4} Density + \frac{1}{Age}(-7.36 + 6.4 \times 10^{-4} Density)\right) \\
\frac{\partial DBH}{\partial Mgmt(H)} =& 0.29 \exp\left(3.67 + 0.29 - 3 \times 10^{-4} Density + \frac{1}{Age}(-7.36 + 6.4 \times 10^{-4} Density)\right) \\
\frac{\partial Ht}{\partial Mgmt(M)} =& 0.16 \exp\left(3.67 + 0.16 - 1.6 \times 10^{-4} Density + \frac{1}{Age}(-9.62 + 4.8 \times 10^{-4} Density)\right) \\
\frac{\partial Ht}{\partial Mgmt(H)} =& 0.21 \exp\left(3.67 + 0.21 - 1.6 \times 10^{-4} Density + \frac{1}{Age}(-9.62 + 4.8 \times 10^{-4} Density)\right) \\
\frac{\partial BA}{\partial Mgmt(M)} =& 0.39 \exp(3.48 + 0.29 + 0.39 + 0.72 - 1.7 \times 10^{-4} Density) \\
& \times \exp\left(\frac{1}{Age}(-11.53 - 11.52 + 1.1 \times 10^{-3} Density)\right) \\
\frac{\partial BA}{\partial Mgmt(H)} =& \left(0.49 + \frac{1.31}{Age}\right) \exp(3.48 + 0.29 + 0.49 + 0.72 - 1.7 \times 10^{-4} Density) \\
& \times \exp\left(\frac{1}{Age}(-11.53 - 11.52 + 1.1 \times 10^{-3} Density)\right) \\
\frac{\partial V}{\partial Mgmt(M)} =& 0.67 \exp\left(5.67 + 0.67 + \frac{1}{Age}(-10.73)\right) \\
\frac{\partial V}{\partial Mgmt(H)} =& 0.82 \exp\left(5.67 + 0.82 + \frac{1}{Age}(-10.73)\right)
\end{aligned}
$$

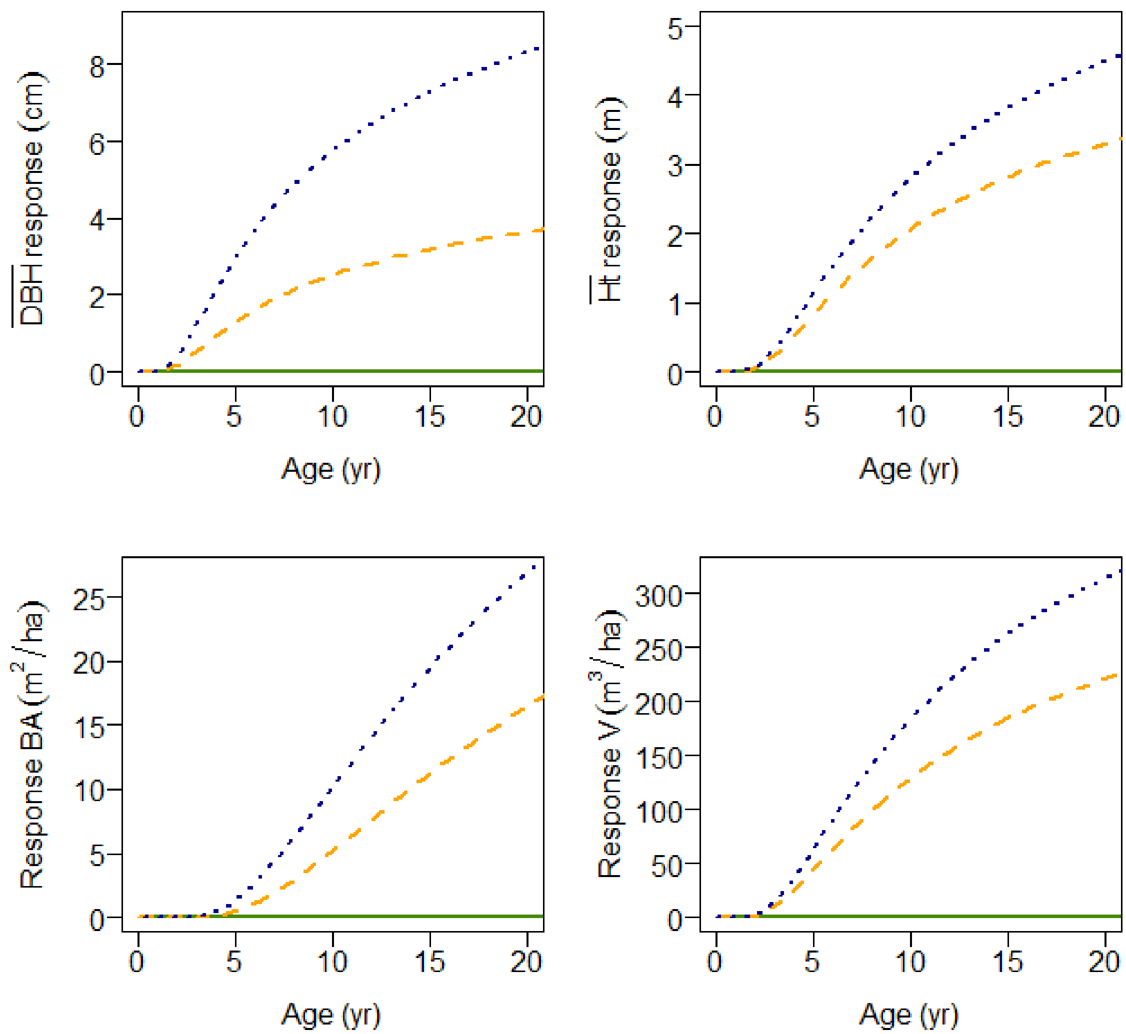

Figure 3.2: Loblolly pine silvicultural responses relative to low level of management [Mgmt(L)] in the Southeastern United States for diameter at breast height (DBH), total height (Ht), basal area (BA), and volume (V) keeping the genetics fixed as half-siblings planted and the physiographic region as Upper Coastal Plain Piedmont. Dashed line represents the response of moderate management [Mgmt(M)], and dotted line represents the response of high management [Mgmt(H)]. Solid, dashed and dotted lines represent Mgmt(L), Mgmt(M), and Mgmt(H).

Overall, these responses are consistent with the expected management outcomes. In general, the higher the inputs and the frequency of the applications, the higher the resulting stand growth and yield (Albaugh et al., 2004; Aspinwall et al., 2011; Borders et al., 2004; Roth et al., 2007). Moreover, Mgmt(M) and Mgmt(H) are additive terms to a basic yield curve [Mgmt(L)], in a similar way that Pienaar, Rheney (1995) modeled silvicultural treatments.

High levels of inputs in quantity and frequency may adjust to asymptotic response curves, whereas low levels of management exhibit parabolic-looking curves (Snowdon, 2002). Thus, there is a possibility that high-order terms of Mgmt or interactions such as Mgmt x Genetics and/or Mgmt x Region were missing in the yield models. The use of first-order terms in the model, as a way to simplify the number of inputs, may cause the management response curves for moderate management to not exhibit a parabolic form and rather attain a peak and then decrease. Despite this mathematical limitation, these management responses give insight into the size of the effect of the three simple levels of management considered here. Hence, economic tradeoffs of operational and intensive forest management can be analyzed. Likewise, since yield models account for the effect of genetics, environment, and density, the model and their derived responses can be also utilized to analyze the effect of a combination of factors.

3.4 Conslusions

Forest growth factors have successfully explained loblolly pine yield (Restrepo et al., 2019). In those models, moderate and high levels of management were statistically different (superior) to the low level of management. Using this information, partial derivatives were taken to analyze silvicultural response equations. Volume at age 20 for moderate and high levels of management can be as much as 221 and 314 m^3/ha higher, respectively, as compared to the low level of management. Hence, yield models that consider the four factors of growth can be used to derive silvicultural responses isolating the effect of genetics, environment, and density. The same framework can be applied to determine a potential volume increase associated with genetically improved seedlings and differences in yield associated to the environment (physiographic region). This model could be used to perform a financial marginal analysis, characterizing the cost associated with the levels of management regimes and determining the profitability associated with each level.

Chapter 4

Loblolly pine sawtimber potential in the Lower Coastal Plain of the southeastern U.S.: A Bayesian approach [1]

[1]Restrepo, H.I, B.P. Bullock, N. Lazar. To be submitted to *Forest Science*.

Abstract

The stand timber value is a function of the merchantable volume and timber price. Finding the proportions of timber in each of the commercial pine product classes (i.e., sawtimber, chip-n-saw, and pulpwood) is a key component in calculating the stand timber value. These proportions can weigh either the total volume or timber prices to obtain the merchantable volume or blended price, respectively. The product class distribution is often reduced to a binary response variable focusing the attention on sawtimber, the timber class with the current highest price. Three theoretical Bayesian frameworks were proposed to estimate the proportion of trees with sawtimber potential: binomial, hierarchical and logit models. Data from a designed research trial evaluating the effect of stand density and management on stand dynamics of loblolly pine in the southeastern U.S. were used to estimate the proportion of trees with sawtimber potential. The dataset includes stem quality assessments, which contribute to a more realistic estimation of the sawtimber proportion. Although this analysis corresponds to a snapshot of the sawtimber proportion at year 18 (most recent remeasurement), it gives insight into the sawtimber potential at the rotation age. Furthermore, timber product class proportions can be used to optimize financial returns by performing marginal analysis to evaluate the effect of silvicultural practices.

Key words: Timber product class proportions, Bayesian models, Sawtimber potential.

4.1 Introduction

Two main forest growth and yield modeling approaches are usually utilized in the southeastern U.S.: size-class, and whole-stand models (Burkhart et al., 2018). Size-class models recognize the stand structure and growth dynamics based on diameter at breast height (DBH) classes (Poudel, Cao, 2013). Such a disaggregation of stand characteristics and structure by diameter classes in the size-class cohort model of the stand-table projection method allows for quantifying the volume in each of the most common commercial timber product classes, i.e., sawtimber, chip-n-saw, and pulpwood (Burkhart, 1979). Conversely, whole-stand models, which corresponds to most of the existing forest growth and yield models for loblolly pine, allow for estimating total volume or biomass at given age as a function of site index and stand characteristics (Burkhart et al., 2018). Whole-stand models make the differentiation of volume into the commercial timber classes somewhat cumbersome. This problem has been traditionally addressed by finding the proportions to weigh the total volume, and hence calculating the merchantable volume in each of the timber product classes. Likewise, these proportions can be used to calculate the blended timber price (Klemperer, 2003), resulting in a simplification of financial calculations.

One of the first attempts calculated the individual-tree probability of merchantability in old-field loblolly pine plantations as a function of DBH (Strub et al., 1986). Burkhart, Bredenkamp (1989) estimated the proportion of trees in pulpwood, sawtimber, and peelers by DBH classes using an extension of Strub et al. (1986) modeling approach. Teeter, Zhou (1998) estimated multinomial models to predict timber product proportions to distribute per-acre total volume within four categories, i.e., softwood pulpwood, softwood sawtimber, hardwood pulpwood, and hardwood sawtimber as a function of DBH and volume. Likewise, stem quality assessments such as fork, broken top, sweep, and disease incidence may improve estimation of the timber product class proportions (Choi et al., 2008; Buford, Burkhart, 1987).

Most monetary value of timber in intensive managed pine plantations in the southeastern United States corresponds to solid wood, the aggregation of sawtimber and chip-n-saw classes (Amateis, Burkhart, 2005). Forestland owners put considerable effort and invest a relative large amount of money to increase the proportion of solid wood in the stand volume as a strategy to maximize financial returns. Moreover, since collecting data for all mentioned commercial timber product classes is expensive, stem quality assessments have primarily been focused on sawtimber, the most valuable commercial timber product class in the southeastern U.S. Such an assessment of the proportion of sawtimber, and hence its complement, non-sawtimber, configures a binary response variable.

The binomial response model typifies the statistical approach to estimate the proportion of sawtimber. In an extension of the binomial model, theoretical probabilities are expressed as a function of covariates, mathematical setting that corresponds to hierarchical and logit models (McCulloch, Searle, 2001; Demidenko, 2013). In all these model types, the Bayesian approach has proven to be more intuitive, flexible, and powerful over the Frequentist counterpart (Gelman et al., 2013).

The objective of this paper was to propose a Bayesian theoretical framework for the binomial, hierarchical and logit models for estimating the proportion of trees with sawtimber potential. As a motivation to the problem of estimating proportions, the first approach presents a simple analysis without considering the effect of covariates. The Bayesian hierarchical model formulation considers the effect of factors on the proportion of trees with sawtimber potential: i) environment conditions accounted by the location effect, ii) the intensity of management practices (i.e. operational, and intensive), iii) planting density (trees per acre), iv) thinning, and v) the size of the tree (diameter at breast height, DBH, in inches). Finally, a Bayesian, and a brief comparison to the Frequentist approach, logit model scheme is presented to provide an estimation/prediction tool of the proportion of trees with sawtimber potential as a function of management and tree size. These mathematical formulations were applied to provide insight into the proportion of loblolly pine (*Pinus taeda* L.) trees with sawtimber potential in the southeastern U.S. We focused on just one genetically improved seedling type, the open pollinated family 7-56, a commonly planted well-tested half-sibling family in the southeastern U.S. We considered the last-available-measurement data subset for both unthinned and thinned plots because thinning is typically an obligated silvicultural practice for pine plantations with a sawtimber objective, and year 18 data, represents a good proxy of the pine sawtimber potential at the rotation age.

4.2 Estimation framework

4.2.1 Simple approach

Let y be the number of trees with sawtimber potential out of total number of trees, n. Consider the prior, likelihood and posterior distribution taking advantage of the conjugacy:

Jeffreys prior

$$\begin{aligned} \theta &\sim Beta(\tfrac{1}{2}, \tfrac{1}{2}) \\ p(\theta) &\propto \theta^{1/2}(1-\theta)^{1/2} \end{aligned} \tag{4.1}$$

Likelihood

$$\begin{aligned} y|\theta &\sim Binomial(n, \theta) \\ p(y|\theta) &\propto \theta^{y}(1-\theta)^{n-y} \end{aligned} \tag{4.2}$$

Posterior

$$\begin{aligned} p(\theta|y) &\propto p(\theta)p(y|\theta) = \theta^{y+1/2}(1-\theta)^{n-y+1/2} \\ \theta|y &\sim Beta(y+3/2\ ,\ n-y+3/2) \end{aligned} \tag{4.3}$$

4.2.2 Hierarchical model

Let y_i be the number of trees with sawtimber potential in the ith location out of the number trees in the ith location, n_i. Consider the prior, hyperprior, and posterior distribution taking advantage of the conjugacy:

$$
\begin{aligned}
y_i|\theta_i &\sim Binomial(n_i, \theta_i) \\
p(y_i|\theta_i) &\propto \theta_i^{y_i}(1-\theta_i)^{n_i-y_i}
\end{aligned}
\tag{4.4}
$$

with prior

$$\theta_i \sim Beta(\alpha, \beta) \tag{4.5}$$

and hyperparameters

$$
\begin{aligned}
\omega &= \frac{\alpha - 1}{\alpha + \beta - 2} \\
\kappa &= \alpha + \beta
\end{aligned}
\tag{4.6}
$$

therefore,

$$
\begin{aligned}
\alpha &= \omega(\kappa - 2) + 1 \\
\beta &= (1-\omega)(\kappa - 2) + 1
\end{aligned}
\tag{4.7}
$$

Hence, the prior can be rewritten as:

$$\theta_i \sim Beta\left(\omega(\kappa-2)+1\ ,\ (1-\omega)(\kappa-2)+1\right) \tag{4.8}$$

An hypterprior can be:

$$
\begin{aligned}
\omega &\sim Beta(\alpha, \beta) \\
p(\omega) &= \frac{\Gamma(\alpha_\omega + \beta_\omega)}{\Gamma(\alpha_\omega)\Gamma(\beta_\omega)}\omega^{\alpha_\omega - 1}(1-\omega)^{\beta_\omega - 1} \\
\kappa &\sim Gamma(\alpha_\kappa, \beta_\kappa) \\
p(\kappa) &= \frac{\beta_\kappa^{\alpha_\kappa}}{\Gamma(\alpha_\kappa)}\kappa^{\alpha_\kappa - 1}\exp(-\beta_\kappa \kappa)
\end{aligned}
\tag{4.9}
$$

With ω, and κ assumed independent. The joint posterior distribution is:

$$p(\boldsymbol{\theta}, \omega, \kappa|\mathbf{y}) \propto p(\omega, \kappa)p(\boldsymbol{\theta}|, \omega, \kappa)p(\mathbf{y}|\boldsymbol{\theta}, \omega, \kappa) \tag{4.10}$$

with

$$
\begin{aligned}
p(\omega, \kappa) &= \left(\frac{\Gamma(\alpha_\omega + \beta_\omega)}{\Gamma(\alpha_\omega)\Gamma(\beta_\omega)}\omega^{\alpha_\omega - 1}(1-\omega)^{\beta_\omega - 1}\right)\left(\frac{\beta_\kappa^{\alpha_\kappa}}{\Gamma(\alpha_\kappa)}\kappa^{\alpha_\kappa - 1}\exp(-\beta_\kappa \kappa)\right) \\
p(\boldsymbol{\theta}|\omega, \kappa) &= \prod_{i=1}^{M} \frac{\Gamma(\kappa)}{\Gamma(\omega(\kappa-2)+1)\Gamma((1-\omega)(\kappa-2)+1)}\theta_i^{\omega(\kappa-2)}(1-\theta_i)^{(1-\omega)(\kappa-2)} \\
p(\mathbf{y}|\boldsymbol{\theta}, \omega, \kappa) &= \prod_{i=1}^{M} \binom{n_i}{y_i}\theta^{y_i}(1-\theta_i)^{n_i - y_i}
\end{aligned}
\tag{4.11}
$$

Thus, the posterior is:

$$p(\boldsymbol{\theta}, \omega, \kappa|\mathbf{y}) \propto \left(\omega^{\alpha_\omega - 1}(1-\omega)^{\beta_\omega - 1}\kappa^{\alpha_\kappa - 1}\exp(-\beta_\kappa \kappa)\right) \times \left(\prod_{i=1}^{M} \frac{\Gamma(\kappa)}{\Gamma(\omega(\kappa-2)+1)\Gamma((1-\omega)(\kappa-2)+1)} \theta_i^{\omega(\kappa-2)+y_i}(1-\theta_i)^{(1-\omega)(\kappa-2)+n_i-y_i}\right) \tag{4.12}$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_M)$, the vector of θ's, and $\mathbf{y} = (y_1, y_2, ...y_M)$, the vector of the number of trees with sawtimber potential in each of the M groups.

4.2.3 Bayesian logit model

Consider a sample of n trees, and let y_i be the binary response with $y_i \overset{\text{iid}}{\sim} \mathit{Bernoulli}(\theta)$, where iid denotes *independent and identically distributed*, and Bernoulli(θ) is the *Bernoulli distribution* with mean θ (Wackerly et al., 2008). Thus, y_i takes the value of one if the ith tree has been assessed to have sawtimber potential or the value zero otherwise:

$$\begin{aligned} P(y_i = 1) &= \theta \\ P(y_i = 0) &= 1 - \theta \end{aligned}$$

Therefore, $\mathsf{E}(y_i) = 1(\theta) + 0(1-\theta) = \theta$, and $\mathsf{Var}(y_i) = \theta(1-\theta)$ (Wackerly et al., 2008). Incidentally, moments of all orders are equal to θ (McCullagh, Nelder, 1989). Because it is reasonable to think that there is a considerable variation among trees, an expansion of this model, called saturated model, allows for $y_i \overset{\text{id}}{\sim} \mathit{Bernoulli}(\theta_i)$. Such a variation can be accounted by covariates. For the sake of simplicity just one arbitrary tree variable, x, accounts by for the variation in the model:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad y_i = 0, 1$$

where β's represent unknown parameters to be estimated and ε_i is the error term. The expected response $\mathsf{E}(y_i)$ has a special meaning in this case. Since $\mathsf{E}(\varepsilon_i) = 0$ (Kutner et al., 2005):

$$E(y_i) = \beta_0 + \beta_1 x_i = \theta_i$$

The principal objective of estimating this model is to investigate the relationship between the response probability and the covariate (McCullagh, Nelder, 1989). Binary response variables, however, poses serious problems in comparison to the linear model (Kutner et al., 2005):

- Nonnormal error terms: because y_i can be only one or zero, each error term $\varepsilon_i = y_i - (\beta_0 + \beta_1 x_i)$ can be either $\varepsilon_i = 1 - (\beta_0 + \beta_1 x_i)$ or $\varepsilon_i = -(\beta_0 + \beta_1 x_i)$, which certainly does not follows the normal distribution.

- Nonconstant error variance: since $\mathsf{Var}(\varepsilon_i) = \mathsf{Var}(y_i - \theta_i) = \mathsf{Var}(y_i)$, because θ_i is a fixed unknown parameter, therefore:

$$\begin{aligned}\mathsf{Var}(\varepsilon_i) &= \theta_i(1-\theta_i) \\ &= \mathsf{E}(y_i)(1-\mathsf{E}(y_i)) \\ &= (\beta_0 + \beta_1 x_i)(1 - \beta_0 + \beta_1 x_i)\end{aligned}$$

  which depends on the level of x. Therefore, the variance is not constant.

- Constraints on response function: because the response function represents the probability of a event, the domain of the mean responses is constrained to $0 \leq \mathsf{E}(y_i) = \theta_i \leq 1$.

Unless restrictions are imposed on β's in the general linear model, we have $-\infty \leq y_i \leq \infty$. Thus, expressing θ as a linear combination of the covariates would be inconsistent with the laws of probability. An effective way of solve this problem is the use of transformation that maps the unit interval onto the whole real line $(-\infty, \infty)$ (McCullagh, Nelder, 1989). Thus, four transformations to estimate generalized linear models (GLM) have been proposed for modeling binary responses (McCullagh, Nelder, 1989):

- The logit or logistic mean response function: $f(\theta) = \ln(\theta/(1+\theta))$

- The probit mean response or inverse normal function: $f(\theta) = \Phi^{-1}(\theta)$

- The complementary log-log response function: $f(\theta) = \ln(-\ln(1-\theta))$
- The log-log function: $f(\theta) = -\ln(-\ln(\theta))$

Although the behaviors of the logit and probit are very similar (almost linearly related in the range $(0.1, 0.9)$, we focus mostly with the logit because its simpler theoretical properties and interpretation as the natural logarithm of the odds ratio ($\ln[\theta_i/(1-\theta_i)]$) (McCullagh, Nelder, 1989). The logit function assumes that errors follow the logistic distribution with probability distribution function (PDF) and cumulative distribution function (CDF) (Kutner et al., 2005):

$$\begin{aligned} f(y) &= \frac{\exp(y)}{[1+\exp(y)]^2} \\ F(y) &= \frac{\exp(y)}{1+\exp(y)} \end{aligned}$$

Therefore, the logistic mean response function is then:

$$\begin{aligned} P(y_i = 1) = \theta_i &= F(\beta_0 + \beta_1 x_1) \\ &= \frac{\exp(\beta_0 + \beta_1 x_i)}{1+\exp(\beta_0 + \beta_1 x_i)} \\ &= [1+\exp(-\beta_0 - \beta_1 x_i)]^{-1} \end{aligned}$$

Applying the inverse of the cumulative distribution function we obtain the *logit* transformation:

$$\begin{aligned} F^{-1}(\theta_i) = \ln\left(\frac{\theta_i}{1-\theta_i}\right) &= logit(\theta_i) && (4.13) \\ &= \beta_0 + \beta_1 x_i && (4.14) \end{aligned}$$

Note that the right hand side of the equation has a linear form, allowing for a linear estimation of the parameters. Several characteristics can be noticed about the logistic mean response function (Kutner et al., 2005): i) is bounded between zero and one; ii) the higher the parameter β_1, the higher the slope of the curve; iii) the function is monotonic increasing (decreasing) when the sign of β_1 is positive (negative); iv) the value of β_0 shifts the curve horizontally; v) it is symmetric, meaning that it would be the same to use the response variable y_i or an arbitrary variable $w_i = 1 - y_i$, in which case the signs of coefficients will be reversed.

Consider now the model with an additional covariate and the three different forms to present the model:

$$logit(\theta_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} \tag{4.15}$$

$$\frac{\theta_i}{1+\theta_i} = \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}) \tag{4.16}$$

$$\theta_i = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})}{1+\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})} \tag{4.17}$$

The interpretation of the magnitude of the effects in Equations (4.15 - 4.17) differs. Assume that all β's are positive. In Equations (4.15) and (4.16), the logit of θ and odds increase (decrease) β_1 and $\exp(\beta_1)$ units, respectively, with one unit of increase (decrease) of x_1. An interesting feature of Equations (4.15) and (4.16) is that the effect of the variables can be assessed independently if the other variable, x_2, is held fixed. The interpretation of parameters in Equation (4.17), however, is more complicated because the effect of a unit change in x_1 depends on the values of both x_1 and x_2, $\frac{\delta\theta_i}{\delta x_1} = \theta_i(1-\theta_i)\beta_1$ (McCullagh, Nelder, 1989).

An expansion of the model with k covariates, $\mathbf{x_i} = (1, x_{i1}, x_{i2}, ..., x_{ik})$ and $k+1$ parameters, $\boldsymbol{\beta}^\top = (\beta_0, \beta_1, \beta_2, ..., \beta_k)$, can be mathematically expressed as follows (Rencher, Schaalje, 2008; Zhang et al., 2011; McCullagh, Nelder, 1989):

$$y_i|\mathbf{x_i} \overset{\text{id}}{\sim} Bernoulli(\theta_i), \quad \theta_i = \mathsf{E}(y_i|\mathbf{x_i}), \quad logit(\theta_i) = \mathbf{x_i}\boldsymbol{\beta}, \quad i = 1, 2, ..., n \tag{4.18}$$

4.3 Case study

4.3.1 Experimental design description and data used

The Plantation Management Research Cooperative (PMRC) at the University of Georgia, Athens, GA, established 17 study sites in the Lower Coastal Plain (LCP) of Georgia, Florida, and South Carolina during the 1995/96 dormant season to test the effect of planting density and management intensity on loblolly pine growth and yield (Harrison, Kane, 2008). Planting density in all study sites were 300, 600, 900, 1200, 1500, and 1800 trees per acre (TPA); whereas the two management intensity treatments were operational and intensive. A complete description of the cultural treatments applied is presented in Appendix A. All installations were planted with loblolly pine first generation, open-pollinated family 7-56, an especially fast grower. At the time of year 18 measurement of the Coastal Plain Culture / Density study, 13 installations remained (Zhao et al., 2014) (Figure 4.1). At year 12, 600, 900, and 1200 TPA plots in four installations (7, 8, 12, and 13) were thinned to the current TPA on their 300-TPA counterparts. A third row thinning with low thinning on leave rows was implemented.

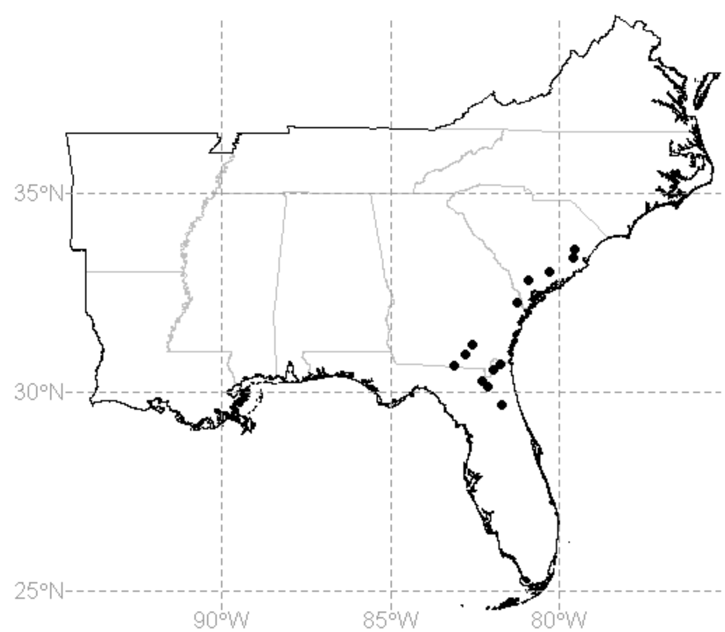

Figure 4.1: Locations of the remained Coastal Plain Culture / Density study sites at year 18.

The experimental design corresponds to a split-split plot design. The first level is the soil CRIFF classes (e.g. A, B1, B2, C, D, E, F, G) (Jokela, Long, 2012). The second level is the installation, in which plots were split for management intensity, and then the six planting densities were randomly established within each of the management plots (Harrison, Shiver, 1999). Forest measurements were taken at years 2, 4, 6, 8, 10, 12, 15, and 18. At each measurement, all 4.5-ft tall or taller trees were measured for diameter at breast height (DBH), and after the fourth growing season, total heights (Ht) were measured or estimated with site-specific height-diameter allometric models. In addition to DBH and Ht measurements, assessments of the tree crown class, forest health (rust infection and moth), defects, and stem quality (in terms of sawtimber potential) were made. Specifically, the field sawtimber potential assessments were made using the following codes:

- No defects, good sawtimber potential (sawtimber)
- Sawtimber reject for stem fork in first log (non-sawtimber code 1)
- Reject for crook or sweep (non-sawtimber code 2)
- Reject for *Cronartium* in first log (non-sawtimber code 3)
- Ugly tree (non-sawtimber code 4)

4.3.2 Simple approach

As a first step, the overall proportion of the trees with sawtimber potential was estimated. Based on the number of trees with sawtimber potential (y = 4,829) and the total number of trees in the dataset (n = 10,219), the posterior distribution of θ is Beta(4830.5,5391.5), with mean (μ_θ), mode ($\tilde{\mu}_\theta$), and variance (σ^2_θ):

$$\begin{aligned}
\mu_\theta &= \frac{y+3/2}{n+3} = \frac{4830.5}{10222} = 0.4726 \\
\tilde{\mu}_\theta &= \frac{y+1/2}{n+1} = \frac{4829.5}{10222} = 0.4726 \\
\sigma^2_\theta &= \frac{(y+3/2)(n-y+3/2)}{(n+3)^2(n+4)} = \frac{(4830.5)(5391.5)}{(10222^2)(10223)} = 2.438 \times 10^{-5}
\end{aligned} \tag{4.19}$$

Overall, 47% of the trees had sawtimber potential. The 95% highest posterior density credible interval is (0.463,0.482) (Figure 4.2).

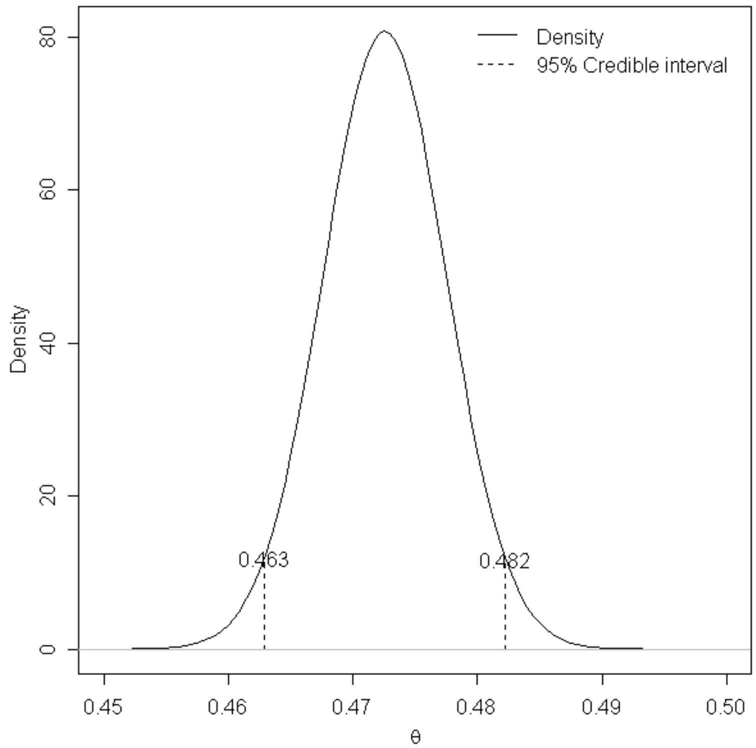

Figure 4.2: Posterior distribution of the proportion of loblolly pine trees with sawtimber potential in the southeastern United States. Dashed lines represent the lower and upper bound of the 95% highest posterior density region of its credible interval.

4.3.3 Hierarchical models

Four hierarchical models were evaluated to assess for the effect of factors on the sawtimber potential:

- Location or installation, accounting for the environmental effect, M=13
- Management, i.e., operational (O), and intensive (I), M=2
- Planting density, i.e., 300, 600, 900, 1200, 1500, and 1800 trees per acre, M=6
- Thinning, i.e., unthinned and thinned stands, M=2
- Discretized DBH from 1 to 17 inches, M=17

Table 4.1: Posterior distribution of the proportion of loblolly pine trees with sawtimber potential in 13 locations across the Lower Coastal Plain U.S. LB and UB are the lower and upper bounds, respectively, of the 95% credible interval.

Location	Observed proportion	Posterior $\mathsf{E}(\theta_i \vert y_i)$	LB (2.5%)	UB (97.5%)
1	0.472	0.472	0.436	0.508
4	0.559	0.559	0.528	0.59
6	0.463	0.463	0.43	0.497
7	0.396	0.396	0.354	0.439
8	0.35	0.35	0.316	0.385
9	0.326	0.326	0.291	0.363
11	0.516	0.516	0.485	0.547
12	0.486	0.485	0.448	0.523
13	0.479	0.479	0.439	0.52
14	0.525	0.525	0.489	0.562
15	0.374	0.374	0.343	0.406
16	0.586	0.586	0.555	0.617
17	0.508	0.508	0.475	0.541

Location

Since the likelihood overwhelmed the hyperprior, there was no effect of the selection of hyperparameters on the posterior distribution (Table 4.1). In Table 4.1, the observed proportion of trees with sawtimber potential for each of the locations is exactly the same that the corresponding posterior expected value of θ. Locations 4, and 16 had the highest proportions of sawtimber potential, whereas locations 8, and 9 had the lowest proportions of sawtimber potential (Figure 4.3, Table 4.1). Table 4.1 can be used to test for statistical differences between pairs of locations. If their 95% credible intervals did not overlap, we can conclude that the sawtimber proportions of them were statistically different. For instance, sawtimber potential of locations 4 and 16 were not statistically different; whereas the sawtimber potential of locations 4 and 8 were statistically different.

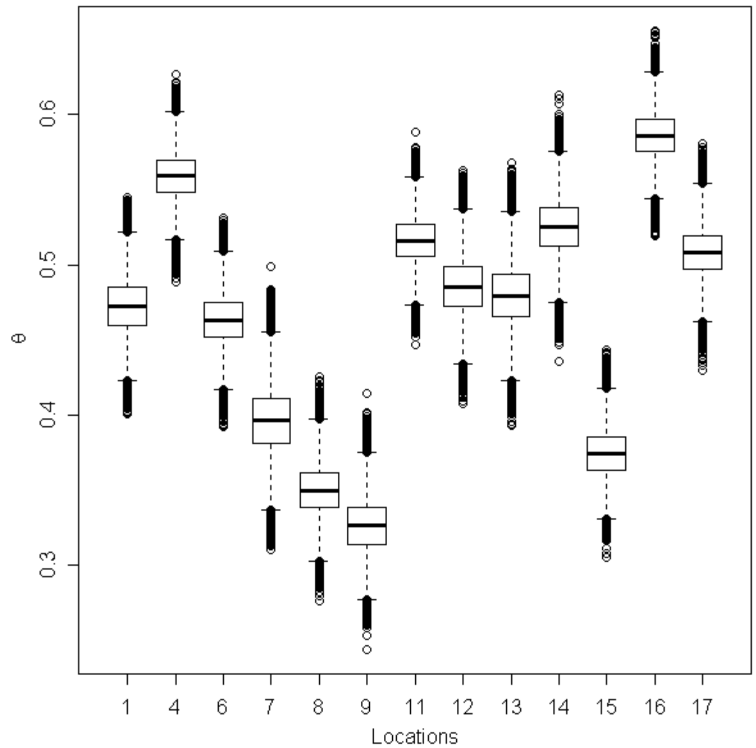

Figure 4.3: Posterior distribution of the proportion of loblolly pine trees with sawtimber potential across 13 locations in the southeastern United States.

Management

Management regimes have an effect on the sawtimber potential of loblolly pine, being the operational level superior over the intensive management regime. Since the 95% credible interval of the two management levels did not overlap, we can conclude that there was an statistical difference between them (Table 4.2, Figure 4.4).

Table 4.2: Posterior distribution of the proportion of loblolly pine trees with sawtimber potential under the management regimes operational and intensive in the Lower Coastal Plain U.S. LB and UB are the lower and upper bounds, respectively, of the 95% credible interval.

Management	Observed proportion	Posterior $\mathsf{E}(\theta_i \vert y_i)$	LB (2.5%)	UB (97.5%)
Intensive (I)	0.403	0.403	0.388	0.418
Operational (O)	0.524	0.524	0.511	0.537

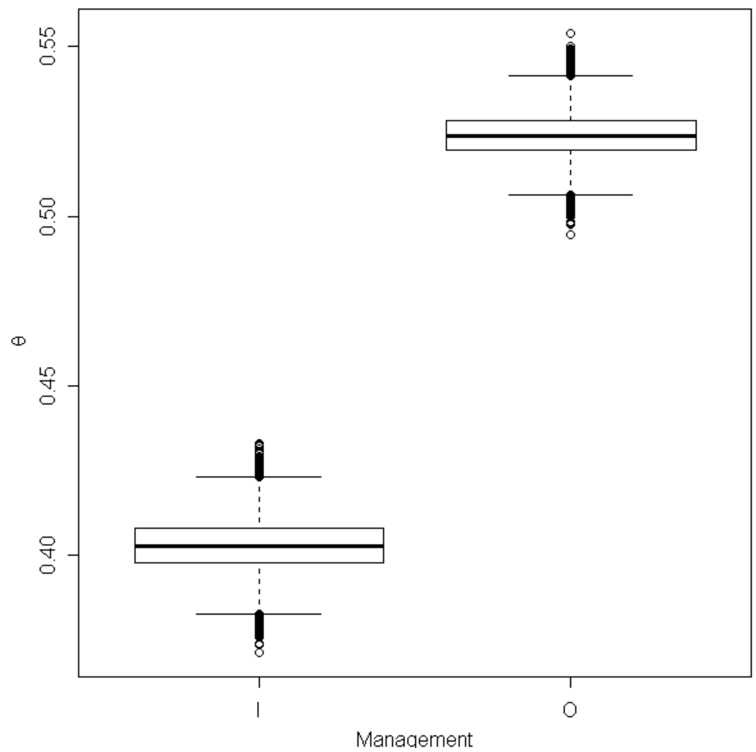

Figure 4.4: Posterior distribution of the proportion of loblolly pine trees with sawtimber potential by the management regimes operational (O), and Intensive (I) in the Lower Coastal Plain U.S.

Planting density

The lowest and highest sawtimber potential were for planting densities 600 and 900 TPA, respectively. Planting density, however, did not have an effect on the proportion of trees with sawtimber potential since the 95% credible intervals of the all six planting densities overlapped (Table 4.3, Figure 4.5).

Table 4.3: Posterior distribution of the proportion of loblolly pine trees with sawtimber potential by six planting densities in the Lower Coastal Plain U.S. LB and UB are the lower and upper bounds, respectively, of the 95% credible interval.

(trees per acre)	Observed proportion	Posterior $\mathsf{E}(\theta_i \mid y_i)$	LB (2.5%)	UB (97.5%)
300	0.477	0.477	0.453	0.501
600	0.46	0.46	0.432	0.487
900	0.484	0.484	0.458	0.51
1200	0.478	0.478	0.453	0.503
1500	0.472	0.472	0.451	0.493
1800	0.466	0.466	0.446	0.487

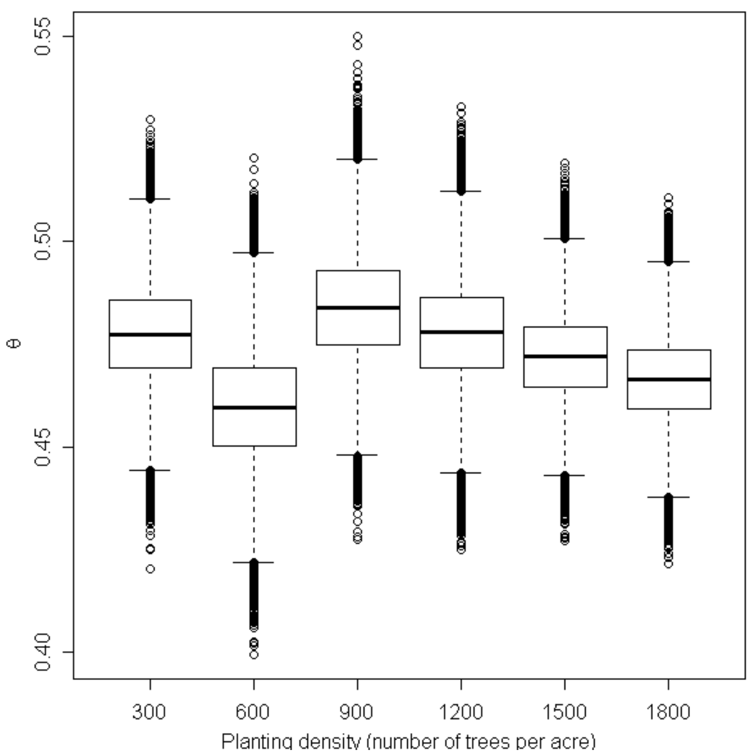

Figure 4.5: Posterior distribution of the proportion of loblolly pine trees with sawtimber potential by six planting densities in the Lower Coastal Plain U.S.

Thinning

There was no statistical difference in the posterior distribution of the proportion of loblolly pine trees with sawtimber potential at year 18 as a response of thinning applied at year 12 (Table 4.4). However, the uncertainty of the proportion was higher for thinned stands (Figure 4.6).

Table 4.4: Posterior distribution of the proportion of loblolly pine trees with sawtimber potential at year 18 as a response of thinning applied at year 12 in the Lower Coastal Plain U.S. LB and UB are the lower and upper bounds, respectively, of the 95% credible interval.

Management	Observed proportion	Posterior $\mathsf{E}(\theta_i \vert y_i)$	LB (2.5%)	UB (97.5%)
Unthinned	0.472	0.472	0.462	0.482
Thinned	0.484	0.484	0.447	0.522

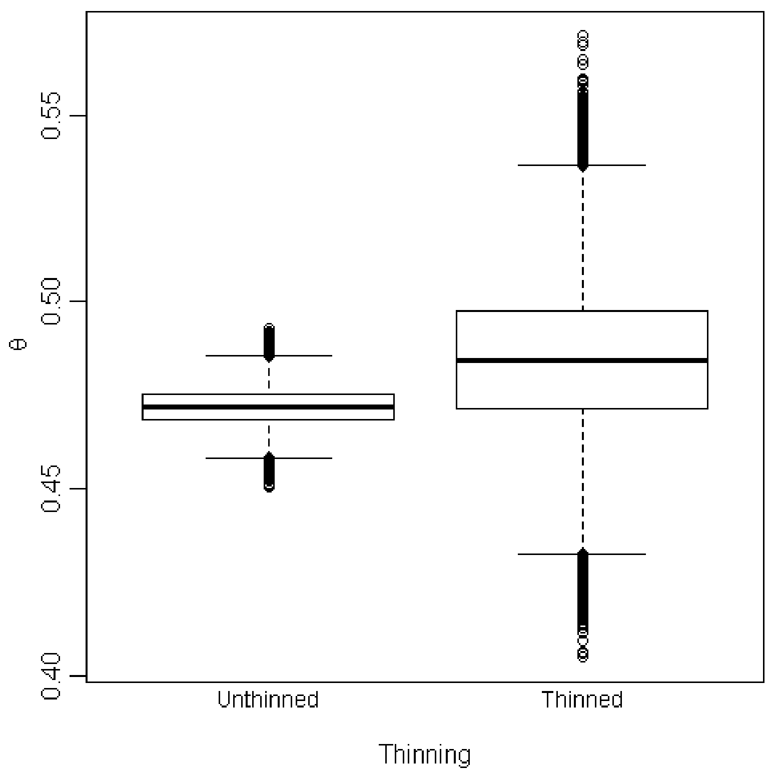

Figure 4.6: Posterior distribution of the proportion of loblolly pine trees with sawtimber potential at year 18 as a response of thinning applied at year 12 in the Lower Coastal Plain U.S.

DBH

The minimum and maximum sawtimber potential were for DBH classes 2 and 14 inches, respectively (Figure 4.7). Overall, the size of the trees had an effect on the proportion of sawtimber potential. There were statistical differences among some DBH classes. For instance, DBH class four inches was statistically different from DBH class seven inches, because their 95% credible intervals did not overlap (Table 4.5). Extreme DBH classes (lowest and highest) had wider 95% credible intervals and whiskers than intermediate DBH classes (Table 4.5, Figure 4.7), because the number of trees in the extreme DBH classes was relatively small (Figure 4.8). In other words, the lower the number of trees in a DBH class, the higher is the uncertainty of the proportion of sawtimber potential.

Table 4.5: Posterior distribution of the proportion of loloblly pine trees with sawtimber potential by 17 one-inch DBH classes in the Lower Coastal Plain U.S. LB and UB are the lower and upper bounds, respectively, of the 95% credible interval.

DBH (inches)	Observed proportion	Posterior $\mathsf{E}(\theta_i\|y_i)$	LB (2.5%)	UB (97.5%)
1	0.4	0.415	0.069	0.812
2	0.292	0.295	0.191	0.411
3	0.394	0.394	0.342	0.447
4	0.414	0.414	0.381	0.448
5	0.417	0.418	0.392	0.443
6	0.458	0.458	0.434	0.482
7	0.476	0.476	0.452	0.5
8	0.499	0.499	0.473	0.525
9	0.508	0.508	0.478	0.538
10	0.535	0.535	0.5	0.569
11	0.52	0.52	0.476	0.563
12	0.515	0.515	0.456	0.574
13	0.573	0.572	0.482	0.659
14	0.667	0.663	0.521	0.792
15	0.6	0.594	0.348	0.818
16	0.6	0.586	0.186	0.93
17	0.5	0.499	0.008	0.992

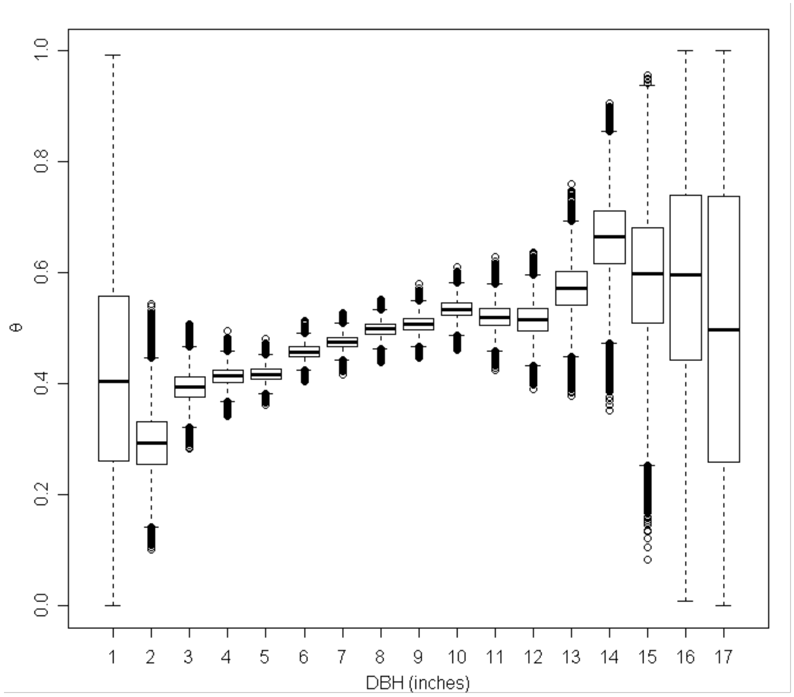

Figure 4.7: Posterior distribution of the proportion of loblolly pine trees with sawtimber potential by 17 one-inch classes of diameter at breast height (DBH) in the Lower Coastal Plain U.S.

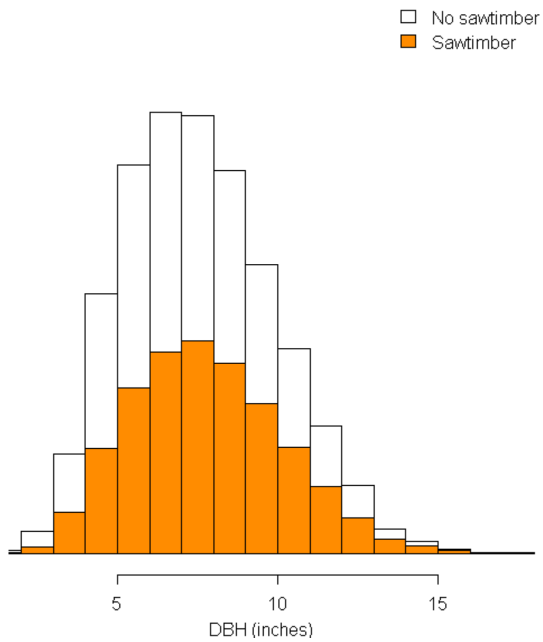

Figure 4.8: Histogram of frequencies of loblolly pine trees with sawtimber potential, and non-sawtimber potential, for the 17 one-inch classes of diameter at breast height (DBH) in the Lower Coastal Plain U.S.

4.3.4 Bayesian logit model

A Bayesian logit model for the proportion of loblolly pine trees with sawtimber potential as a function of management and DBH was estimated (Table 4.6). As mentioned, intensive forest management reduces the proportion of sawtimber potential whereas the higher the DBH, the higher the proportion of sawtimber potential. The trace plots and posterior distribution for each of the parameter estimated are shown in Figure 4.9. A logit Frequentist model was also estimate to compare it with the Bayesian logit model. The parameter estimates and their standard errors were virtually the same (Tables 4.6 and 4.7).

Table 4.6: Bayesian logit model to estimate the proportion of loblolly pine trees with sawtimber potential as a function of the management regime intensive (beta1) and diameter at breast height (DBH, beta2) in the Lower Coastal Plain U.S. SD is standard deviation, SE is standard error, and TS is time series.

Par.	Empirical mean and standard deviation				Quantiles				
	Mean	SD	Naive SE	TS SE	2.50%	25%	50%	75%	97.50%
β_0	-0.662	0.069	2.18×10^{-4}	0.0013	-0.797	-0.708	-0.662	-0.616	-0.526
β_1	-0.604	0.042	1.33×10^{-4}	0.0003	-0.687	-0.632	-0.604	-0.575	-0.521
β_2	0.106	0.009	2.81×10^{-5}	0.0002	0.088	0.100	0.106	0.112	0.123

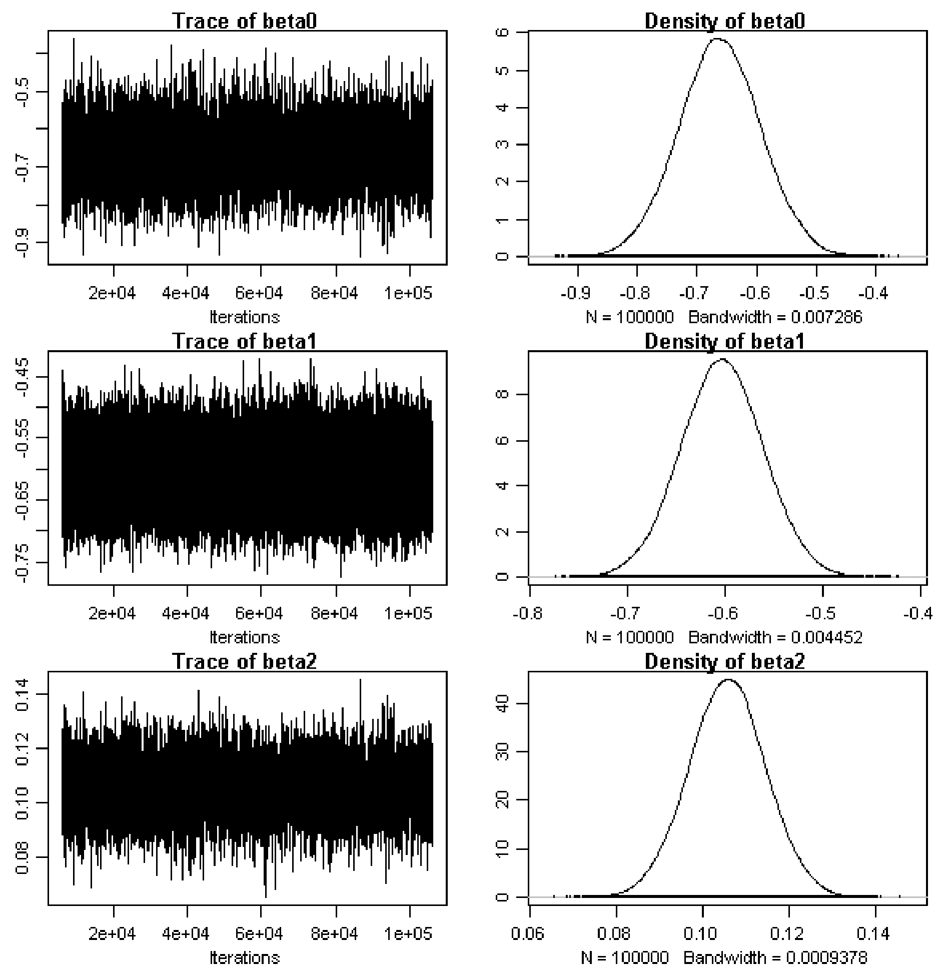

Figure 4.9: Trace plots and density plots of the Bayesian logit parameter estimates of the proportion of loblolly pine trees with sawtimber potential in the Lower Coastal Plain U.S.

Table 4.7: Frequentist logit model to estimate the proportion of loblolly pine trees with sawtimber potential as a function of the management regime intensive and diameter at breast height (DBH) in the Lower Coastal Plain U.S.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.661	0.068	-9.653	<0.0001
Intensive management	-0.603	0.042	-14.355	<0.0001
DBH	0.105	0.009	11.929	<0.0001

4.4 Discussion

A general Bayesian framework for theoretical probabilities was proposed and applied to estimate the proportion of loblolly pine trees with sawtimber potential in the Lower Coastal Plain U.S. Three mathematical approaches were presented: binomial, hierarchical and logit models. The binomial model results suggest that the overall sawtimber potential was almost a half of the stand timber volume, which positively contributes with the financial success of the timber production. This positive financial performance is due to the relative high price of sawtimber compared to chip-n-saw and pulpwood prices. However, a much better estimation of the proportion of sawtimber should consider factors like the environmental effect, management, planting density, thinning and tree size (DBH).

The effect of location on the proportion of trees with sawtimber potential may be due to contrasting environmental conditions. Site index, as the typical metric to assess for the environmental effect on forest productivity, is not a good descriptor of the timber product class distribution and timber merchantability potential (Burkhart, Bredenkamp, 1989). Conversely, stem form has been linked with soil properties. Deficiencies and excesses of macro, secondary, and micro nutrients may induce an expression of bad tree form; therefore affecting the tree sawtimber potential and stand timber value (Espinoza et al., 2012; Lehto et al., 2010). For instance, pines on slightly boron-deficient soils may have a thick stem base, and a low branch and needle mass to stem ratio; whereas a dramatic deficiency in boron results in the loss of the apical dominance (Lehto et al., 2010), which has serious consequences on stem quality and form.

The intensive management regime practiced on the locations in this research corresponds to a very high level of inputs (Appendix A). Thus, a high nitrogen to calcium ratio in soils, may result in a stem sinuosity of loblolly pine trees (Espinoza et al., 2012), affecting the sawtimber potential and stand timber value. This result has direct implications for the forestland owner. The effect of intensive management on tree form (sawtimber potential) (Green et al., 2018), may negate volume gains from silvicultural practices (Restrepo et al., 2018). Hence, forestland owners should evaluate the overall effect of intensive management on the stand timber value (Green et al., 2018).

Planting density did not affect the proportion of trees with sawtimber potential, which is consistent with results previously reported (Green et al., 2018; Burkhart, Bredenkamp, 1989). Although trees in high density stands have small branches (Borders, Volfovicz, 2010), which may increase the tree sawtimber potential, the effect of high stand density on tree size (Restrepo et al., 2019) may diminish the stand timber value. Unthinned and thinned stands presented similar sawtimber potential at year 18. This finding is consistent with results previously reported (Burkhart, Bredenkamp, 1989). The relative high uncertainty in the sawtimber proportion for thinned stands was due to the small sample size of thinned stands in comparison with unthinned stands.

Regarding size (DBH), one can expect that the higher the diameter, the higher the tree sawtimber potential. This result is consistent with the literature evaluating the effect of DBH on tree merchantability, and sawtimber proportion (Strub et al., 1986; Burkhart, Bredenkamp, 1989; Teeter, Zhou, 1998). However, other measures of tree size, like total height, may not explain volume merchantability (Burkhart, Bredenkamp, 1989). Realistically, only the trees with DBH greater to 11.6 inches would be in the sawtimber category at the rotation age. For that reason, although the stem quality and form may be good, small trees would be in the non-sawtimber category. Therefore, the total realized proportion of sawtimber trees may be lower than estimated in the analysis over the diameter distribution (Figure 4.7).

Estimated logit models can be used to estimate/predict the proportion of loblolly pine trees with sawtimber potential at year 18 as a function of the management intensity and DBH. The intensive management term corresponds to a bump down effect in the intercept, whereas the DBH term works as a slope in the model. The results from the logit models were consistent with the corresponding findings from the hierarchical models. Non informative priors were used in all Bayesian models. That is the reason why no differences were found between logit Bayesian and Frequentist models. Future work can elicit and utilize informative priors in the three Bayesian models explored in this research. Moreover, the model evaluated can be extended to a multinomial model considering the three pine commercial timber product classes.

Chapter 5

Long-term timber cutting contracts in the southeastern U.S.: Updating the primer valuation framework [1]

[1] Restrepo, H.I, R.B. Mei, B.P. Bullock. To be submitted to *Forest Science*.

Abstract

U.S. timberland ownership has drastically changed in the last decades, mainly driven by the divestitures of vertically integrated forest product companies. With the divestiture of land, forest product companies have exposed themselves to raw material risk. This risk is usually hedged by contractual options like long-term cutting contracts (LTTCs), which have been believed to represent a valuable asset for timber industry firms. However, since the mid 1980s, the methods, value, and implications of the LTTCs in the southeastern U.S. have not been updated. The forest business market as compared to the 1980s has changed in terms of timber prices, risk-free interest rates, and corporate risk-adjusted discount rates. The overall objective of this paper was to update the option pricing valuation framework for LTTCs proposed by Robert Shaffer Jr. in 1984. The estimation of volatility and use of information from the current financial/economic conditions was crucial to accomplishing the goal. In particular, the conditional volatility estimated from GARCH models was input into the Black-Scholes and binomial models (European and American) to estimate the value the call option of one LTTC. Contrary to Shaffer's result, our analysis suggests that LTTCs may not be profitable for forest product firms. This is primarily because timber price volatilities and the risk-free interest rate are relatively small. Thus, well-functioning wood markets not only preclude owning land by forest firms but also may diminish the value of LTTCs. Likewise, this result implies that forest companies will probably rely more on the open market and less on this type of legal agreement.

Key words: Timber market, Call options, Black-Scholes, Binomial model, GARCH, Conditional volatility.

5.1 Introduction

Timberland ownership in the United States has changed substantially in the past several decades (Yao et al., 2014). More than 50% of the forestland in the United States (70% in the South) shifted ownership in the period of 1980-2005. This was mainly driven by divestitures of vertically integrated forest product companies (VIFPCs), and acquisitions by timber investment management organizations (TIMOs) and real estate investment trusts (REITs) (Waggle, Johnson, 2009; Mei et al., 2013). Literature lists a wide array of motives for this land title transition from the TIMO, REIT, and VIFPC perspective. TIMOs and REITs' land acquisitions were motivated by favorable tax treatment and portfolio diversification goals (Lönnstedt, Sedjo, 2012). This diversification strategy includes timberlands, an asset characterized by favorable returns, low risk, and inflation hedge (Sun, Zhang, 2001; Switzer, 2006). On the other hand, VIFPCs' divestments were driven by a weak financial performance, accounting and tax disadvantages for owning land, land price appreciation, intensive forest management[2], the economics of specialization, and high carrying costs of land and forest management (Lönnstedt, Sedjo, 2012; Zhang et al., 2012).

Although all VIFPCs' reasons for divestiture deserve scrutiny, this research focuses on the two latter causes. In these scenarios, sufficient raw material in the open market and high property taxes motivated VIFPCs' divestitures (Lönnstedt, Sedjo, 2012). As VIFPCs rely less on their raw material and more on procured timber, they exposed themselves to a variety of risks. These risks have traditionally been hedged by contractual options, a rather common practice in forestry (Phelps, McCurdy, 1997; Zinkhan, 1991). These include long-term cutting contracts, buyout options in landowner assistance programs, options to buy timberland properties, and development rights sold by timberland owners to developers (Zinkhan, Cubbage, 2003). Of particular interest are the long-term timber contracts (LTTCs) that assure the required raw material and manage price volatility due to inter-annual and economic-cycle price fluctuations (Mei et al., 2013).

[2]higher per unit area timber production, which makes some land idle for VIFPCs.

The LTTCs have represented an essential asset for timber industry firms to cover their raw material risk. Shaffer Jr. (1984) presented a thorough valuation framework for LTTCs using the Black-Scholes model (Black, Scholes, 1973). Since the Shaffer's seminal work, the methods, value, and implications of engaging in LTTCs in the southeastern U.S. have not been updated. There is a new context for forest business compared to what it was in the mid 1980s in terms of timber prices (market), risk-free interest rates (economic context), and corporate risk-adjusted discount rates (cost of capital as a consequence of the current business environment) (Switzer, 2006; Yao, Mei, 2015; Mei et al., 2013). Likewise, statistical techniques to analyze time series, the sophistication of financial analyses, the post-recession economic outlook, and computational tools to accomplish solutions for complex problems have considerably evolved in the last 30 years (Yao, Mei, 2015). Therefore, we are convinced that updating Shaffer's framework would be highly beneficial and informative.

The stochastic properties of timber prices have been a great concern for timberland investors (Mei et al., 2010). Average timberland returns dropped from 14.3% (1982-1997) to 6.9% (1995-2010) as a result of declining timber prices (Mei et al., 2013). Timber prices in the southeastern U.S., categorized by the traditional commercial classes, i.e., sawtimber, chip-n-saw, pulpwood, have systematically been recorded since the late 1970s (Norris Fundation, 2018). Timber prices, and especially their corresponding logarithmic returns, behave like a financial time series (Zinkhan, Cubbage, 2003), characterized by time-varying volatility (Andersen et al., 2009).

There is no consensus in the forest finance literature whether the timber prices follow a geometric Brownian motion (GBM) or the mean reversion (MR) model (Chaudhari et al., 2016). Haight, Holmes (1991) suggested that quarterly timber prices follow a GBM, whereas monthly prices are MR. Likewise, the relative short sample period of the available time series from Timber-Mart South (TMS) does not provide enough evidence to conclude if timber prices are random or mean-reverting (Mei et al., 2010). The assumption of the nature of timber prices implies different risks and valuations of timberland business (Mei et al., 2013).

Conversely, the time-varying volatility feature refers to the fact that small (large) values are followed by small (large) values (Andersen et al., 2009). Moreover, volatility evolves continuously over time exhibiting clusters, within certain fix range, but unevenly leveraging the effect of positive and negative news on the prices (Tsay, 2010). Mentioned volatility features of the financial/economic time series have been extensively studied. There are several proxy volatilities such as absolute returns, squared returns, and stochastic volatility, but GARCH-type volatility is by far the most popular (Hwang, Valls, 2006). Engle (1982) proposed a theoretical framework to model the volatility named the autoregressive conditional heteroskedastic (ARCH) model. Bollerslev (1986) expanded the method to be more flexible and posses better mathematical properties, the so-called generalized ARCH (GARCH). This model has been applied to analyze volatility of returns of timberland investments and timber prices (Sun, 2013; Mei et al., 2010; Sun et al., 2013; Clements et al., 2017).

The overall objective of this paper was to update Shaffer's option pricing framework for LTTCs in the southeastern U.S. Volatility and information from the current financial/economic conditions were input into the Black-Scholes and binomial models to estimate the value call options for one LTTC. To assess for the effect of volatility on the valuation, three volatility measures were used to replicate Shaffer's LTTC: i) implied volatility or whole-sample-long standard deviation, ii) the conditional volatility estimated from an ARMA(p,q) - GARCH(1,1) model combination, and iii) the quasi-conditional volatility estimated from the moving window standard deviation.

5.2 Methods

5.2.1 Data

TMS is a non-profit organization that compiles and publishes quarterly prices of the three major commercial timber products (sawtimber, chip-n-saw, and pulpwood) for the southeastern United States (Norris Fundation, 2018). The timber price database is grouped into 22 regions, coded by the two-letter U.S. Postal Service state abbreviation and number assigned by TMS. For this research, a dataset of timber prices for sawtimber and pulpwood from 1977Q1 to 2018Q4 in the Georgia region two (GA2), Lower Coastal Plain, was obtained. In finance, it is common to conduct the analysis of the log-returns instead of using the raw prices. Thus, the log-returns were calculated as (Tsay, 2010):

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln(P_t) - \ln(P_{t-1}) \tag{5.1}$$

where r_t is the natural logarithm of the return at time t, P_t is the timber price at time t, and P_{t-1} is the price at time $t-1$. Hereafter, whenever returns are mentioned, they are considered log-returns, unless otherwise indicated.

5.2.2 Volatility measures and their estimation

Three methods were used to estimate the volatility of returns: the implied volatility estimated from the standard deviation (SD) of the time series, the conditional volatility (CSD) estimated from ARCH/GARCH models, and the quasi-conditional volatility estimated from moving window SD (QSD). SD is estimated by the sample standard deviation:

$$SD = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(r_t - \bar{r})^2} \tag{5.2}$$

where SD is the sample standard deviation, T is the total number of data, r_t is the return at time t, and $\bar{r}$ is the mean of the returns.

To mathematically formalize the ARCH/GARCH models, consider the conditional expectation and variance of the returns (Tsay, 2010):

$$\begin{aligned}
\mu_t(r_t) &= \mathsf{E}(r_t|\mathcal{F}_{t-1}) && (5.3)\\
\sigma_t^2(r_t) &= \mathsf{Var}(r_t|\mathcal{F}_{t-1}) && (5.4)\\
&= \mathsf{E}[(r_t - \mu_t(r_t))^2|\mathcal{F}_{t-1}] && (5.5)
\end{aligned}$$

where $\mu_t(r_t)$, or just μ_t, is the conditional expectation of r_t given $\mathcal{F}_{t-1}$; $\sigma_t^2(r_t)$, or just σ_t^2, is the conditional variance of r_t given $\mathcal{F}_{t-1}$; and $\mathcal{F}_{t-1}$ denotes the information set available at time $t-1$.

The time series of the returns is usually modeled using an ARMA(p,q) model, denoted the mean equation, estimated as (Tsay, 2010):

$$
\begin{aligned}
r_t &= \mu_t + a_t & (5.6)\\
\mu_t &= \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{i=1}^{q} \theta_i a_{t-i} & (5.7)\\
y_t &= r_t - \phi_0 & (5.8)
\end{aligned}
$$

where a_t is the *shock* or *innovation*, usually assumed $\mathcal{N}(0, \sigma_a^2)$.

The volatility, on the other hand, can be modeled using a function to describe the conditional variance as follows:

$$
\begin{aligned}
\sigma_t^2(r_t) &= \mathsf{Var}(r_t|\mathcal{F}_{t-1}) & (5.9)\\
&= \mathsf{Var}(a_t|\mathcal{F}_{t-1}) & (5.10)
\end{aligned}
$$

An estimation framework for the conditional volatility was first proposed by Engle in 1982, who called it as autoregressive conditional heteroscedastic (ARCH) model (Engle, 1982). In the ARCH model, the shock, a_t, is serially uncorrelated but serially dependent. If the values are truly independent, then nonlinear instantaneous transformations (such as taking logarithms, absolute values, or squaring) preserve independence. However, the same is not valid for correlation, as correlation is a measure of linear dependence. If the returns are independent and identically distributed (i.i.d.), then so are the absolute returns (Cryer, Chan, 2008). The dependence is described by a simple quadratic function of its lagged values (Tsay, 2010):

$$
\begin{aligned}
a_t &= \sigma_t \varepsilon_t & (5.11)\\
\sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2 & (5.12)
\end{aligned}
$$

where ε_t is a sequence of i.i.d. random variables with mean zero and variance one, usually assumed standard normal or standardized t-Student distribution. The coefficients must satisfy some regularity conditions to ensure that the unconditional variance of r_t is finite, i.e., $\alpha_0 > 0$, and $\alpha_i \geq 0$ for $i > 0$.

An extension of the ARCH model, so-called generalized autoregressive conditional heteroskedasticity (GARCH) was proposed by Bollerslev in 1986 (Bollerslev, 1986):

$$a_t = \sigma_t \varepsilon_t \tag{5.13}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2 \tag{5.14}$$

All the ARCH conditions apply to this model, in addition to $\beta_j \geq 0$; $\alpha_i = 0$ for $i > m$, and $\beta_j = 0$ for $j > s$; and $\sum_{i=1}^{\max(m,s)}(\alpha_i + \beta_i) < 1$ (Tsay, 2010). With these conditions, the unconditional variance of a_t is finite and the conditional variance σ^2 evolves over time. The GARCH(m,s) process is weakly stationary if and only if the persistence, $P = \sum_{i=1}^{m} \alpha_i + \sum_{j=1}^{s} \beta_j$, is less than unity (Teräsvirta, 2009). The stationary GARCH model has been slightly modified using the variance targeting method (Engle, Mezrich, 1996; Vaynman, Beare, 2014). With this approach, the condition that $\alpha_0 > 0$ is relaxed and this parameter is derived from the persistence rather than simultaneously estimated, as (Ghalanos, 2017):

$$\hat{\alpha}_0 = \tilde{\sigma}^2 (1 - \hat{P}) \tag{5.15}$$

where $\hat{\alpha}_0$ is an estimator for α_0, $\tilde{\sigma}^2$ is the unconditional variance of the squared residuals of the model ($\mathsf{E}(\varepsilon_t^2)$), and $\hat{P}$ is an estimator of the persistence.

In an attempt to provide a simple approach for the estimation of the volatility over time, the quasi-conditional volatility (QSD) or moving window standard deviation, is proposed:

$$QSD_{t(w)} = \sqrt{\frac{1}{w-1} \sum_{i=t-w+1}^{w} (r_i - \bar{r}_{t(w)})^2} \tag{5.16}$$

where $QSD_{t(w)}$, or just QSD, is the standard deviation at time t with a moving window w, and $\bar{r}_{t(w)}$ is the mean of returns at time t with moving window w. This measure of volatility copes with the weaknesses of the implied volatility, resembling the GARCH effect, but without the mathematical complexity of the GARCH models. A time-frame evaluation is required to obtain a meaningful QSD. Therefore, an array of windows from 4 to 40 quarters (one to ten years) was evaluated. The window, w, that maximizes the correlation between the QSD and CSD was chosen for the estimation of QSD.

5.2.3 Testing for autocorrelation and ARCH/GARCH effects

Consider the null hypothesis, H_0: $\rho_1 = \cdots = \rho_i = \cdots = \rho_k = 0$, against the alternative hypothesis, H_a: $\rho_i \neq 0$ for some $i \in \{1, ..., i, ..., k\}$, with ρ_i indicating the lag-i autocorrelation (Tsay, 2010). Thus, the significance of the autocorrelation may be tested using the Ljung-Box or modified Q-statistic (Tsay, 2010):

$$Q(k) = T(T+2) \sum_{i=1}^{k} \frac{\hat{\rho}_i^2}{T-i} \overset{\text{Ho}}{\sim} \chi^2_{(k)} \tag{5.17}$$

where T is the length of the time series, and $\hat{\rho}_i$ is the lag-i autocorrelation, and k is the order of the autocorrelation evaluated.

The test can be applied to the log-returns to identify if the time series is MR. Thus, insufficient evidence to reject the null hypothesis that the autocorrelations of log-returns are not different from zero would imply that the returns are MR. Likewise, the Ljung-Box statistic will have a chi-square distribution with k degrees of freedom under the assumption of no ARCH effect (Cryer, Chan, 2008). Thus, rejecting the null hypothesis that the autocorrelation of squared and absolute log-returns are statistically different from zero suggests the existence of the ARCH/GARCH effect (Zivot, 2009).

5.2.4 GARCH model building and estimation

There are several packages and functions to estimate GARCH models in R (R Development Core Team, 2018). However, one of the most widely used, stable and versatile packages is *rugarch* (Charles, Darné, 2019). The *rugarch* package (Ghalanos, 2017) jointly estimates an

ARMA(p,q) for the mean equation and a GARCH(m,s) for the conditional volatility. We selected the omnibus GARCH model, GARCH(1,1), which fits well to financial time series (Zivot, 2009). Thus, the only question is the order of the ARMA(p,q) that results in the best model. Therefore, all the combinations of ARMA models, orders $p = q = 0, 1, 2, 3, 4, 5$, e.g., (0,0), (1,0),..., (5,5), were tested in conjunction with the GARCH(1,1). A function to automate this estimation was written in R that prints the Bayesian Information Criterion (BIC) to identify the best combination of ARMA(p,q) and GARCH(1,1). The BIC was chosen as the primary criterion over the Akaike Information Criterion (AIC) and the log-likelihood (logLik) function because it is more suitable for small sample size datasets and takes into account the number of parameters, respectively. The models with the lowest BIC were selected to check for AIC, parameter estimate significance, and algorithm convergence issues. Likewise, the behavior of residuals such as QQ-plot (normality or t-Student) and empirical density of standardized residuals, and ACF of standardized residuals and squared standardized residuals were evaluated[3].

5.2.5 Option pricing models

An option is defined as the right, but not the obligation, to buy (sell) a given quantity of an asset on (before) a given date, at an agreed upon price today. Of particular interest to forest firms is the call option, which gives the right, but not the obligation, to buy a certain amount of timber from the forestland owner at a given date (or before in the case of American options) at the timber price agreed in the contract (Zinkhan, 1991). Both the forest firm and the landowner are liable for specific terms in the legal agreement.

Two models are usually used to estimate the value of options: the Black-Scholes and binomial. Black, Scholes (1973) were the first to propose a general equilibrium solution for the pricing of an option, as a risk-free portfolio hedge. The equilibrium value of a call option (V_0), the price at which it would trade in a perfectly efficient market, is mathematically

[3]for more detail about these diagnostics see Ghalanos (2017).

expressed as (Shaffer Jr., 1984):

$$V_0 = V_A\Phi(d_1) - E\Phi(d_2)e^{-rt} \tag{5.18}$$

where V_A is the current market price of the asset, E is the exercise price of the option, r is the annual continuously compounded risk-free rate of return, t is the length of the time (years) to the expiration of the option, and Φ is the Standard Normal cumulative distribution for d_1 and d_2, which are calculated as:

$$d_1 = \frac{\ln(V_A/E)+rt}{\sigma\sqrt{t}} + \frac{1}{2}\sigma\sqrt{t} \tag{5.19}$$

$$d_2 = d_1 - \sigma\sqrt{t} \tag{5.20}$$

where σ is the annualized volatility of the asset price.

The decision rule is that the annual value of the option (V_0) must exceed the contract's administrative yearly (or per period) cost (C) to be profitable for the forest firm. Therefore, a simple comparison between V_0 and C is often enough. When $(V_0 - C) > 0$ there is a profit for the forest firm, loss otherwise. However, when the contract is split into different periods or split by other criteria, the corresponding present value of V_0, PV_0, can be compared with the corresponding present value of C, PV_c, of all contract sections, all discounted at the firm's risk-adjusted rate (r_a) over n periods. In this case, PV_0, PV_c, and the net present value (NPV) of the contract are:

$$PV_0 = V_0\left[\frac{1-(1+r_a)^{-n}}{r_a}\right] \tag{5.21}$$

$$PV_c = C\left[\frac{1-(1+r_a)^{-n}}{r_a}\right] \tag{5.22}$$

$$NPV = PV_0 - PV_c \tag{5.23}$$

If the NPV is positive, the call option is profitable for the forest firm. A similar rationale is behind the binomial OPM (Rendleman Jr., Bartter, 1979). The binomial is a simple two-step OPM that allows for option pricing in a straightforward way. The payoff of a one-period call option can be expressed as:

$$C_t = (S_t - K) = \max\{0, S_t - K\} \tag{5.24}$$

where C_t is the payoff of a call option at time t, S_t is the stock price or asset price at time t, and K is the strike price of the call option. Let's consider S_t with the following piecewise distribution:

$$S_t = \begin{cases} S_{t_1} & w.p.\ \pi \\ S_{t_2} & w.p.\ (1-\pi) \end{cases} \tag{5.25}$$

The best approximation for S_{t_1} and S_{t_2} is that they are based on the stock price (asset price) at time zero as $S_{t_i} = S_0 e^{rt}$. Hence, the expected value of S_t is:

$$\begin{aligned} \mathsf{E}[S_t] = & \quad \mathsf{E}[\pi S_{T_1} + (1-\pi) S_{T_2}] && (5.26) \\ = & \quad \mathsf{E}[\pi(S_0 e^{rT}) + (1-\pi)(S_0 e^{rT})] && (5.27) \\ = & \quad S_0 e^{rT} && (5.28) \end{aligned}$$

Therefore, under the assumption of no arbitrage opportunity:

$$\begin{aligned} \mathsf{E}[S_T] - K = & \quad S_0 e^{rT} - K && (5.29) \\ \Longrightarrow & \quad K = S_0 e^{rT} && (5.30) \end{aligned}$$

The rationale behind this is that an individual has the opportunity to invest the money quantity S_0 at the risk-free interest rate r and receive after time t the same amount of money plus its corresponding interests. Both OPMs were estimated using DerivaGem - Version 2.01 (Hull, 2017) with four nodes for the binomial model.

5.2.6 Contract one summary (Shaffer Jr., 1984)

The LTTC under analysis corresponds to contract one presented by Shaffer Jr. (1984). All the details can be found in the original paper. However, some features are worth mentioning. The timber volume presented in Shaffer's contract one is in MBF and cords, but current timber prices are in dollars per ton ($/ton). Therefore, we used conversion factors from MBF to tons and cords to tons as 7.7 tons/MBF and 2.6 tons/cord, respectively (Dicke, Parker, 2016). The time horizon of the contract was 25 years, split into two periods, 1-14 and 15-25 years, depending on the potential harvest volume. As was structured in the original contract, the unit time of analysis is a semester. Thus, there were 28 periods for the first part of the contract and 22 periods for the second part of the contract. The annual risk-free rate of return and the firm's risk-adjusted discount rate by that time were 11.8 and 7 %, respectively. The total annual cost of administering the provisions of the contract, i.e., salaries, employee benefits, transportation, supplies, and office expenses, was \$64,800. Hence, the six-month-period cost was $C = \$32,400$. This value was chosen as the threshold to compare the value obtained from Black-Scholes and bionomial models to determine the profitability of the LTTC.

5.3 Results

5.3.1 Exploratory analysis, autocorrelation, and ARCH/GARCH effects

Sawtimber prices experienced a sinuous increasing trend from 1977Q1 to 1998Q2, where they reached their maximum at \$54.87/ton. They then stayed around the mean of \$47/ton until 2007Q1, when they began to steadily decline to a value of \$24.55/ton in 2012Q2[4]. Since then, the price has stayed relatively constant centered around a mean of \$28/ton (Figure 5.1 A).

Although the trend for pulpwood prices was similar to that of sawtimber during the period 1977Q1-1998Q2, reaching the maximum at \$21.13/ton, the overall variability was much higher. There was depreciation in pulpwood prices in the period 1998Q2-2003Q1, attaining a value at \$5.83/ton[5]. Since this time, pulpwood prices have increased and reached the current short-term mean around \$14/ton (Figure 5.1B). No further analysis was performed on timber prices given their non-stationary nature. Instead, the log-returns were studied.

[4]The ratio of this value to the all time lowest value, which was the first data collected in 1977Q1, was $\frac{\$24.55/ton}{\$14.67/ton} = 1.673$.

[5]The ratio with the all time lowest value, which was the price in 1977Q1, was $\frac{\$5.83/ton}{\$5.72/ton} = 1.019$.

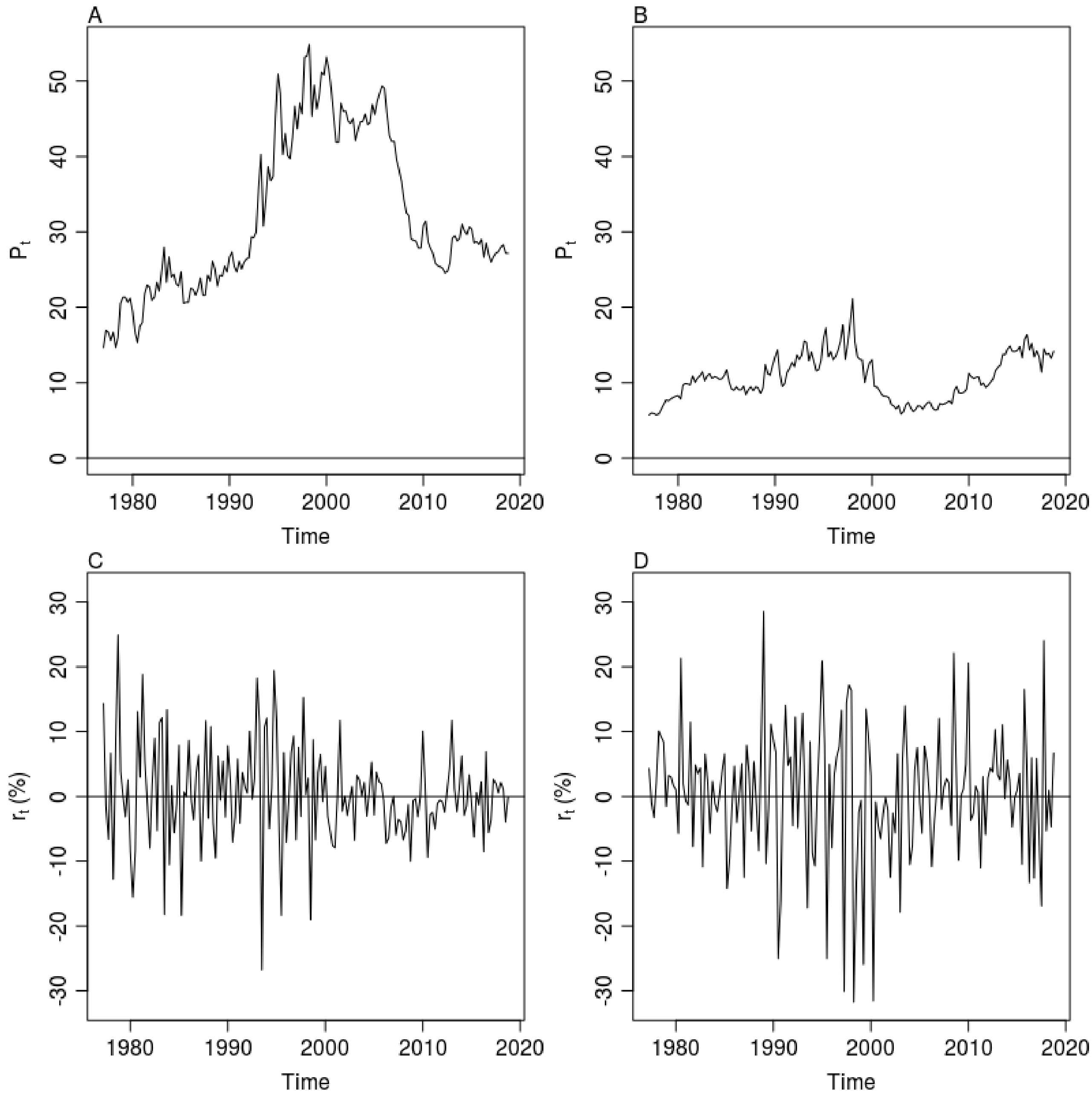

Figure 5.1: Quarterly timber prices (1977Q1 - 2018Q4, A & B), and log-returns (1977Q2 - 2018Q4, C & D) for sawtimber (left) and pulpwood (right) in the TMS Georgia region two (GA2), U.S.

Three periods can be identified for sawtimber returns (Figure 5.1C): i) a high variability at the beginning of the time series (1977Q2) with a gradual reduction up to 1992Q1; ii) a sudden increase in the variability led to the lowest value of -26.8% in 1993Q3 and again, a reduction of the variability until 2001Q1; and iii) a relative stabilization of returns from 2001Q1 to 2018Q4. For pulpwood, the pattern of variability can be described in three periods as well (Figure 5.1D): i) a period of mild variability with two noteworthy high positive returns in 1980Q3 (21.3%), and 1989Q1 (28.5%), but without much change on the negative side of the returns; ii) from 1990Q3, a return equal to -25.0% marked the beginning of a highly variable period in which the all time lowest value at -31.8% was recorded in 1998Q2; and iii) the third and last period started in 2001Q3, characterized by a relative constant variation. Incidentally, the absolute value of the lowest return in both sawtimber and pulpwood, i.e., 26.8 and 31.8%, exceeded their maximum positive values, 24.9 and 28.5%, respectively. This finding suggests an asymmetry in the distribution of the returns. Described periods can be identified in squared and absolute log-returns as well (Figure 5.2).

Sawtimber and pulpwood mean returns were not statistically different from zero (p-value>0.05) (Table 5.1). The distributional properties of the returns support the mentioned asymmetry and suggest deviations from the normality, being negatively skewed, heavy-tailed (positive excess of kurtosis), and with a large Jarque-Bera statistic (p-value<0.05) (Table 5.1). The hypothesis of equal variances of the log-returns between sawtimber and pulpwood was rejected (p-value<0.05), suggesting that their variances were statistically different.

The sample autocorrelation (ACF) and partial autocorrelation (PACF) for sawtimber returns present one significant spike at lag eight (Figure 5.2). A clear ARMA model cannot be inferred from the ACF/PACF pattern, but because the significant lag is far away, the overall autocorrelation may be negligible (p-value>0.05) (Table 5.2). For the pulpwood returns, the ACF spikes at lags 2, 4, 8, 12, 18, 19, and 20; whereas the PACF spikes at lags 2, 8, and 18. The ACF pattern and the spike of the second PACF lag suggests a possible best fit with MA(2) or ARMA(p,2) models.

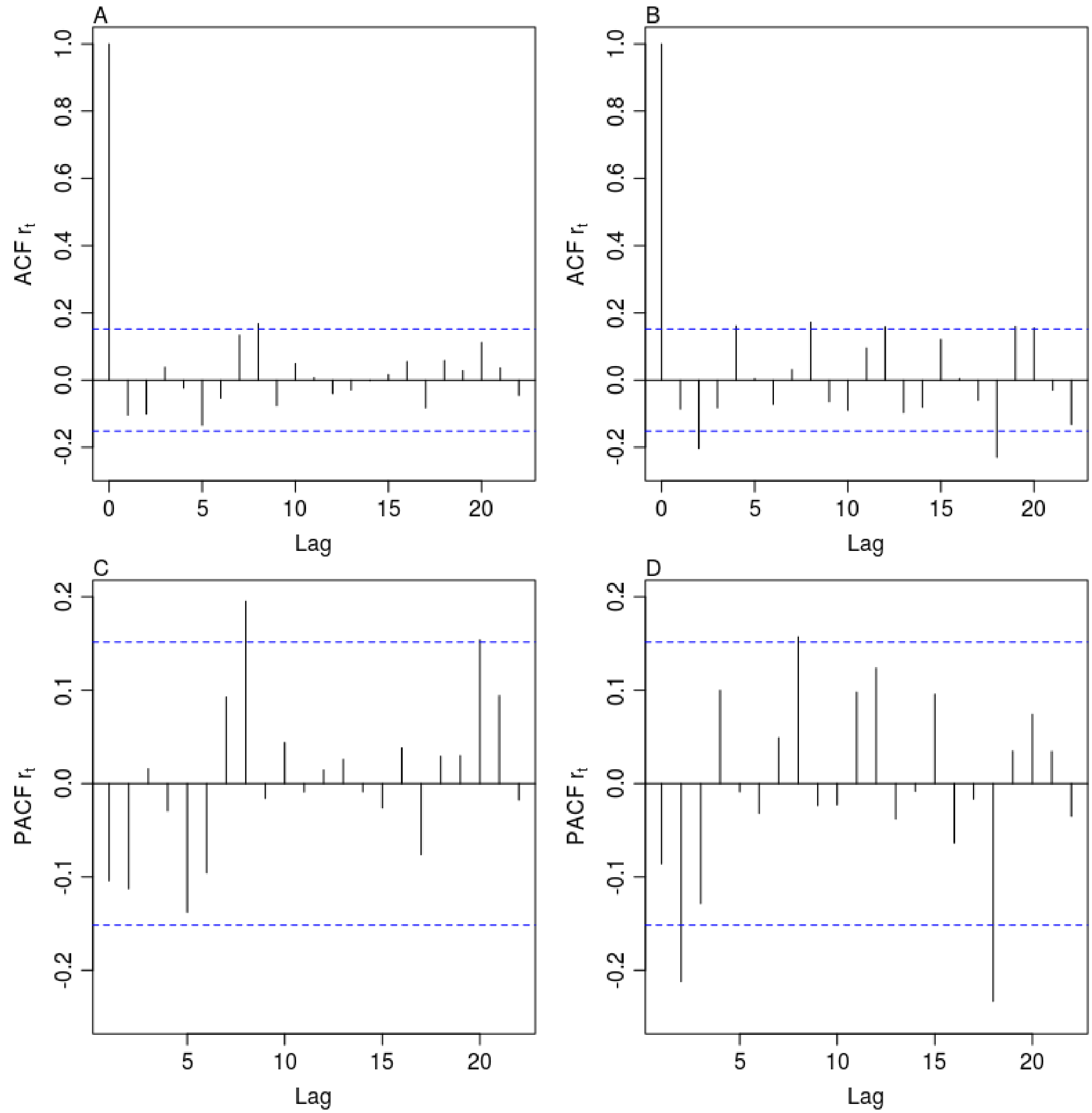

Figure 5.2: Autocorrelation function (ACF, A & B), and partial autocorrelation function (PACF, C & D) of the log-returns for sawtimber (left) and pulpwood (right) in the TMS Georgia region two (GA2), U.S.

Special attention, however, must be paid to the squared and absolute returns to achieve a meaningful model that considers volatility. Squared and absolute returns indicate some persistence, suggesting a possible ARCH effect (Figure 5.3, Table 5.2). In particular, squared

Table 5.1: Summary statistics of the log-returns for sawtimber and pulpwood in the TMS Georgia region two (GA2), U.S.

	Sawtimber		Pulpwood	
N	167		167	
Mean (%)	0.37		0.54	
t-Statistic	0.627	(0.532)	0.698	(0.486)
SD (quarterly,%)	7.6		10.0	
Skewness	-0.029		-0.532	
Excess kurtosis	1.263		1.390	
Jarque-Bera test	11.117	(0.004)	21.313	(<0.001)

and absolute returns show evidence of volatility clustering. There is a direct relationship between returns and volatility: the higher the return, the higher the volatility. The ACF of the squared returns presents some conspicuous spikes at lags 2, 8, 10 and 19 for sawtimber (Figure 5.4A), and lags 4 and 8 for pulpwood (Figure 5.4B). The ACF of the absolute returns spikes at lags 2, 7, 10, 19, and 21 for sawtimber (Figure 5.4C); whereas no evident spikes were present for pulpwood (Figure 5.4D). Similarly, the PACF of squared returns shows high significant values at lags 2 and 19 for sawtimber (Figure 5.5A), and 4 and 8 for pulpwood (Figure 5.5B). The PACF for absolute returns spikes at lags 2, 7 and 19 for sawtimber (Figure 5.5C) and 16 for pulpwood (Figure 5.5D).

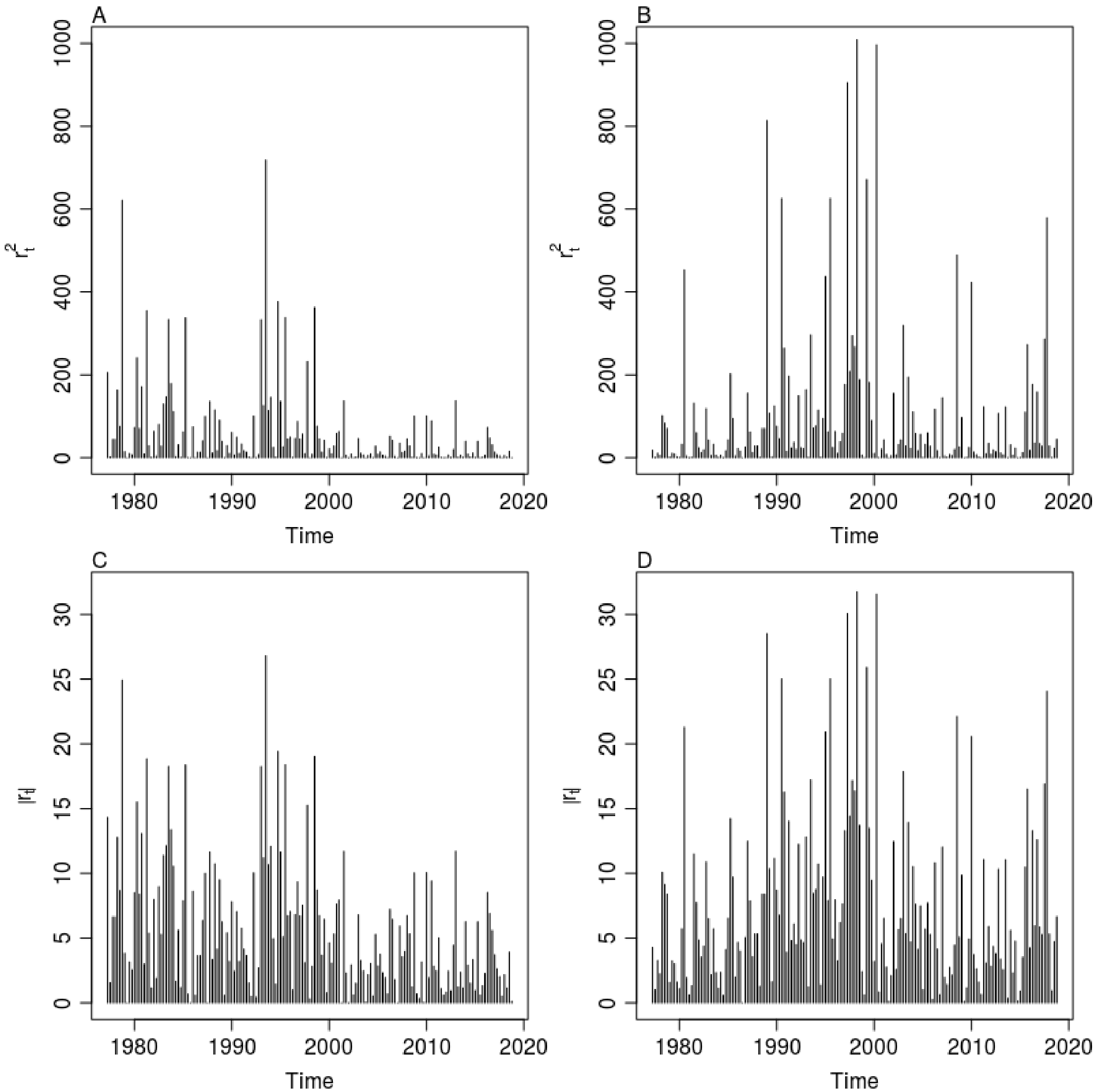

Figure 5.3: Quarterly (1977Q2 - 2018Q4) squared returns (A & B) and absolute value of the returns (C & D) of sawtimber (left) and pulpwood (right) in the TMS Georgia region two, U.S.

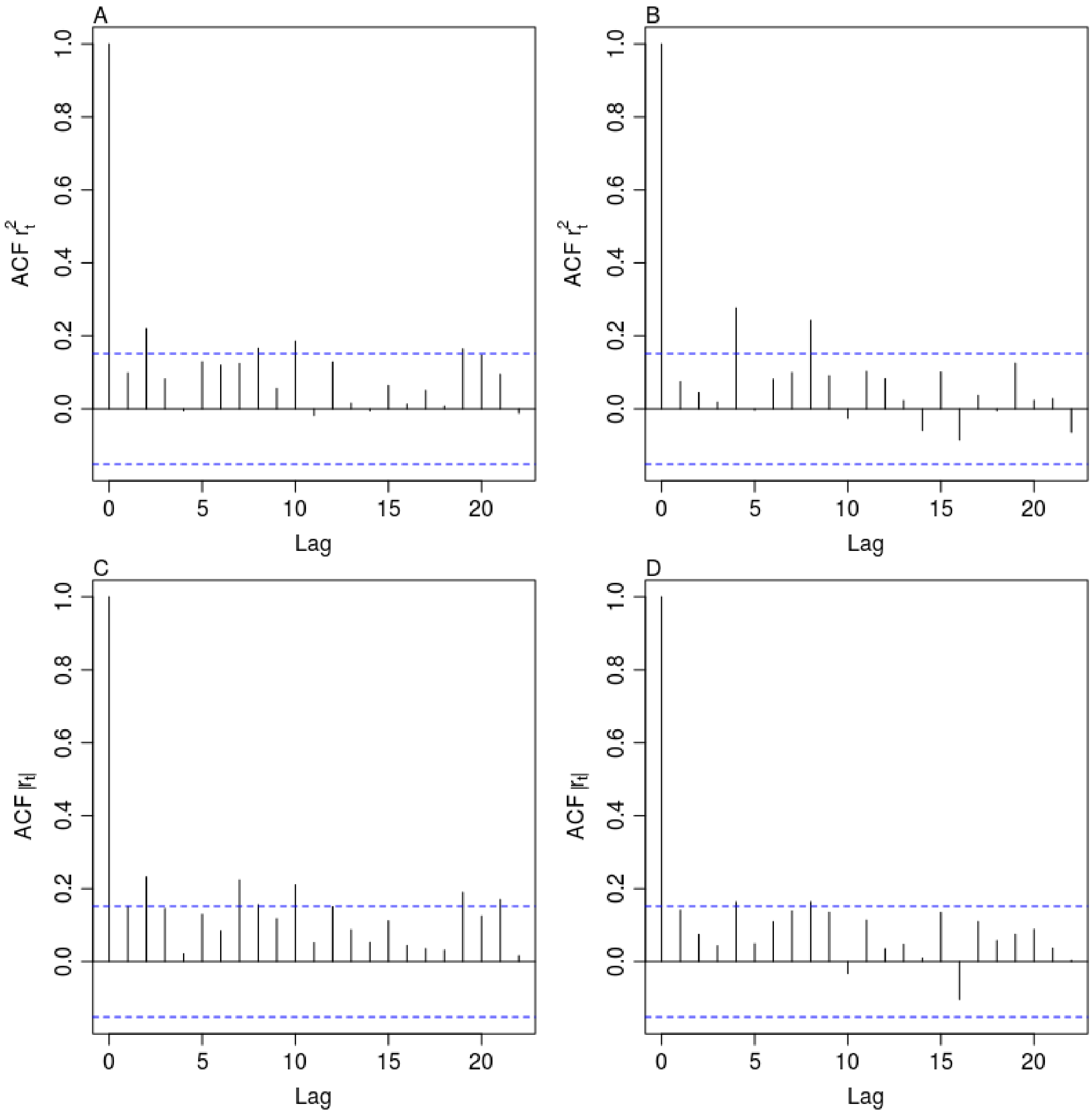

Figure 5.4: Sample autocorrelation function (ACF) of the squared (A & B) and absolute (C & D) log-returns of sawtimber (left) and pulpwood (right) in the TMS Georgia region two, U.S.

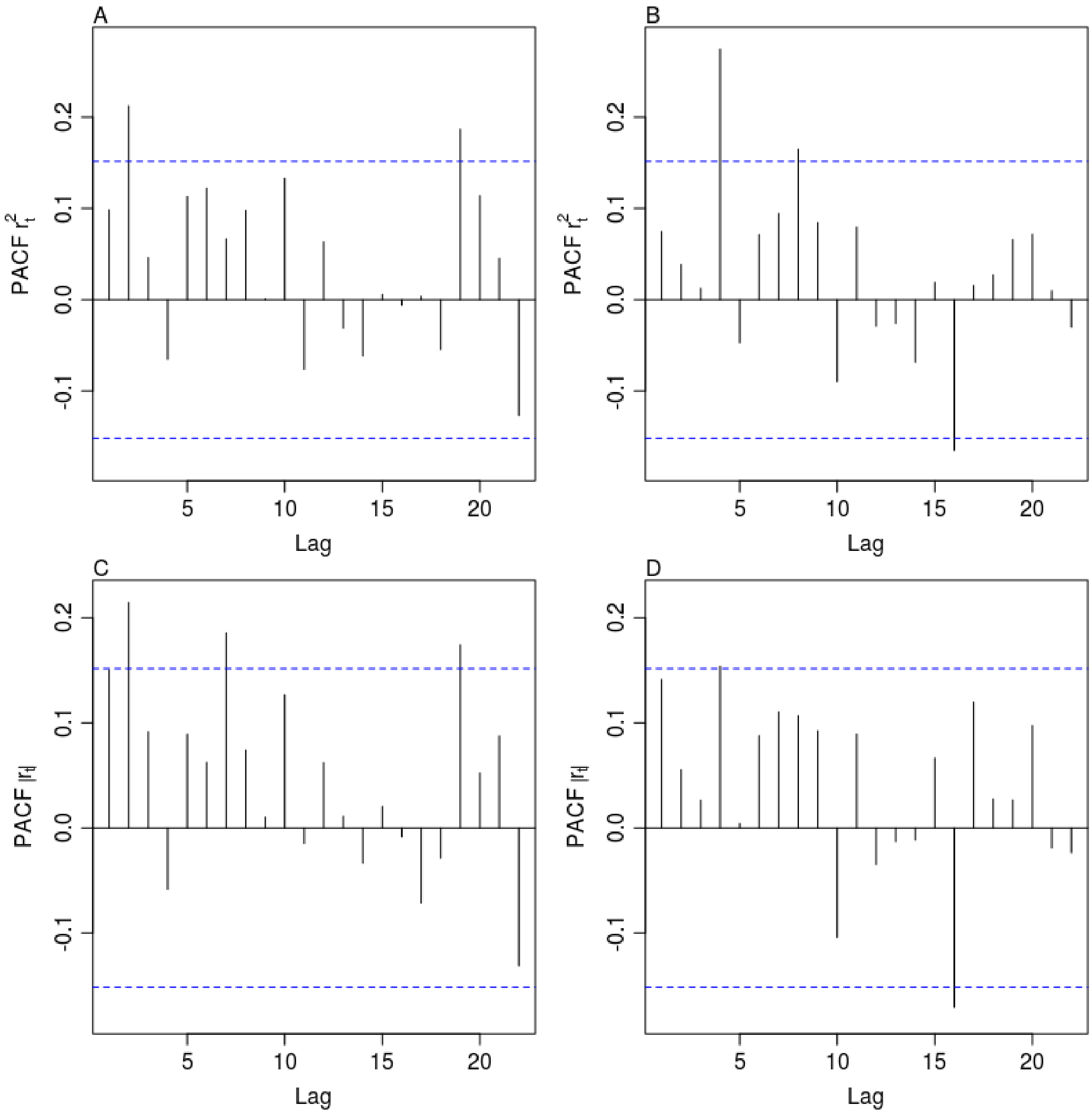

Figure 5.5: Sample partial autocorrelation function (PACF) of the squared (A & B) and absolute (C & D) log-returns of sawtimber (left) and pulpwood (right) in the TMS Georgia region two, U.S.

Table 5.2: Ljung-Box test (p-value) for the log-returns, squared log-returns and absolute log-returns of sawtimber and pulpwood in the TMS Georgia region two, U.S.

Lag	Sawtimber			Pulpwood		
	r_t	r_t^2	$\|r_t\|$	r_t	r_t^2	$\|r_t\|$
1	0.179	0.203	0.052	0.266	0.334	0.068
2	0.174	0.008	0.002	0.017	0.533	0.119
3	0.290	0.013	0.001	0.026	0.726	0.206
4	0.428	0.029	0.003	0.009	0.007	0.06
5	0.231	0.019	0.002	0.018	0.015	0.093
6	0.289	0.014	0.002	0.025	0.019	0.076
7	0.170	0.010	< 0.001	0.041	0.019	0.041
8	0.059	0.003	< 0.001	0.012	0.001	0.014
9	0.067	0.005	< 0.001	0.016	0.001	0.008
10	0.089	0.001	< 0.001	0.017	0.002	0.013
11	0.127	0.002	< 0.001	0.017	0.002	0.011
12	0.163	0.001	< 0.001	0.007	0.002	0.016
13	0.208	0.002	< 0.001	0.007	0.003	0.022
14	0.266	0.004	< 0.001	0.008	0.004	0.033
15	0.327	0.005	< 0.001	0.006	0.004	0.021
16	0.362	0.008	< 0.001	0.009	0.004	0.018
17	0.358	0.010	< 0.001	0.011	0.007	0.015
18	0.387	0.015	< 0.001	0.001	0.010	0.019
19	0.443	0.006	< 0.001	< 0.001	0.007	0.021
20	0.379	0.003	< 0.001	< 0.001	0.010	0.022

5.3.2 GARCH models

Because the ARCH/GARCH effects were detected for both sawtimber and pulpwood, the GARCH model was considered the appropriate statistical technique to estimate their volatility. In total 576 models were estimated to select the best combination of ARMA(p,q), GARCH(1,1), probability distribution, and variance targeting approach (Table B.1).

The best model without statistical issues for sawtimber (BIC and AIC equal to -2.443 and -2.611, respectively) was the combination of ARMA(4,3) without intercept, GARCH(1,1) assuming normal errors, and the utilization of the variance targeting approach (Appendix

B). All parameter estimates were very highly statistically significant (p-value<0.001) (Table 5.4). The estimated persistence was 0.972. The distributional properties of the standardized residuals were satisfactory, and the ACF of both the standardized and squared standardized residuals shows no remaining autocorrelation (Figure 5.6). The Ljung-Box test confirms the ACF result for the standardized and squared standardized residuals (p-value>0.05) (Appendix C), suggesting that the residuals are white noise without dependence. Therefore, the ARCH/GARCH effect was taken into account in the model.

Table 5.3: Best combination of ARMA(p,q) with and without intercept, and GARCH(1,1) assuming normal and t-Student errors with and without variance targeting (VT) for sawtimber in the TMS Georgia region two, U.S. μ indicates if the intercept was included in the ARMA model.

ARMA(p,q)			Distribution	VT	BIC	AIC	Statistical issues					
							Par. Significance		ACF		Ljung-Box	
p	q	μ					ARMA	GARCH	r_t	r_t^2	r_t	r_t^2
4	3	yes	Normal	yes	-2.461	-2.648			x	x		
2	3	no	Normal	yes	-2.448	-2.578			x		x	
4	3	no	Normal	yes	-2.443	-2.611						
3	2	no	Normal	no	-2.443	-2.592		x				
3	2	no	Normal	yes	-2.436	-2.566				x		
2	3	no	t-Student	yes	-2.428	-2.577			x		x	
2	3	no	Normal	no	-2.420	-2.569		x	x		x	
2	3	yes	Normal	yes	-2.418	-2.567			x		x	
0	0	no	Normal	yes	-2.415	-2.453		x		x		
5	5	no	Normal	yes	-2.413	-2.638						

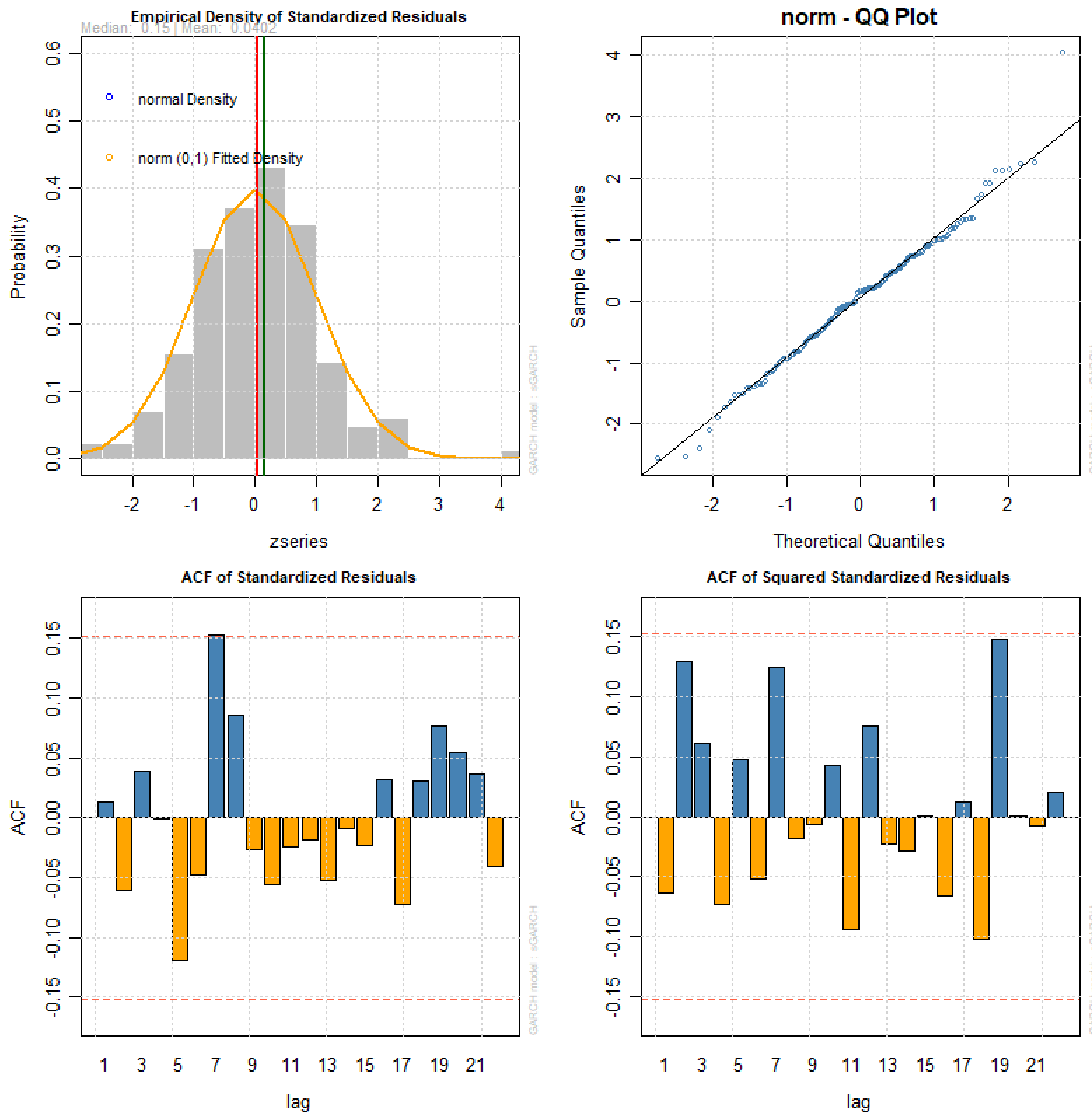

Figure 5.6: Diagnostics for the model ARMA(4,3) without intercept, and GARCH(1,1) assuming normal errors with variance targeting for sawtimber log-returns in the TMS Georgia region two, U.S.

The best model without statistical issues for pulpwood (BIC and AIC equal to -1.730 and -1.861, respectively) was the combination of ARMA(2,2) without intercept and GARCH(1,1) assuming t-Student errors with variance targeting (Table 5.5). All parameter estimates

Table 5.4: Parameter estimates for the combination of ARMA(4,3) without intercept, and GARCH(1,1) assuming normal errors with variance targeting for sawtimber log-returns in the TMS Georgia region two, U.S.

	Estimate	Std. Error	t value	Pr(>|t|)
ϕ_1	-0.736	0.000	-4131.314	< 0.001
ϕ_2	0.662	0.000	3843.680	< 0.001
ϕ_3	1.066	0.000	4051.868	< 0.001
ϕ_4	0.165	0.000	3073.425	< 0.001
θ_1	0.722	0.000	3977.350	< 0.001
θ_1	-0.826	0.000	-3786.474	< 0.001
θ_1	-1.152	0.000	-3830.266	< 0.001
α_1	0.144	0.042	3.439	0.001
β_1	0.828	0.052	15.890	< 0.001
α_0	1.3×10^{-4}	NA	NA	NA

were very highly statistically significant (p-value<0.001) or highly statistically significant (p-value<0.01), except for the ARCH parameter, α_1, (p-value=0.06) (Table 5.6). The estimated persistence was 0.938. The distributional properties of the standardized residuals were satisfactory, and the ACF of both the standardized and squared standardized residuals identify one correlation that do not seriously affect the overall autocorrelation (Figure 5.7). This is confirmed by the Ljung-Box test (p-value>0.05) (Appendix C). Residuals are white noise without dependence; therefore the ARCH effect was taken into account.

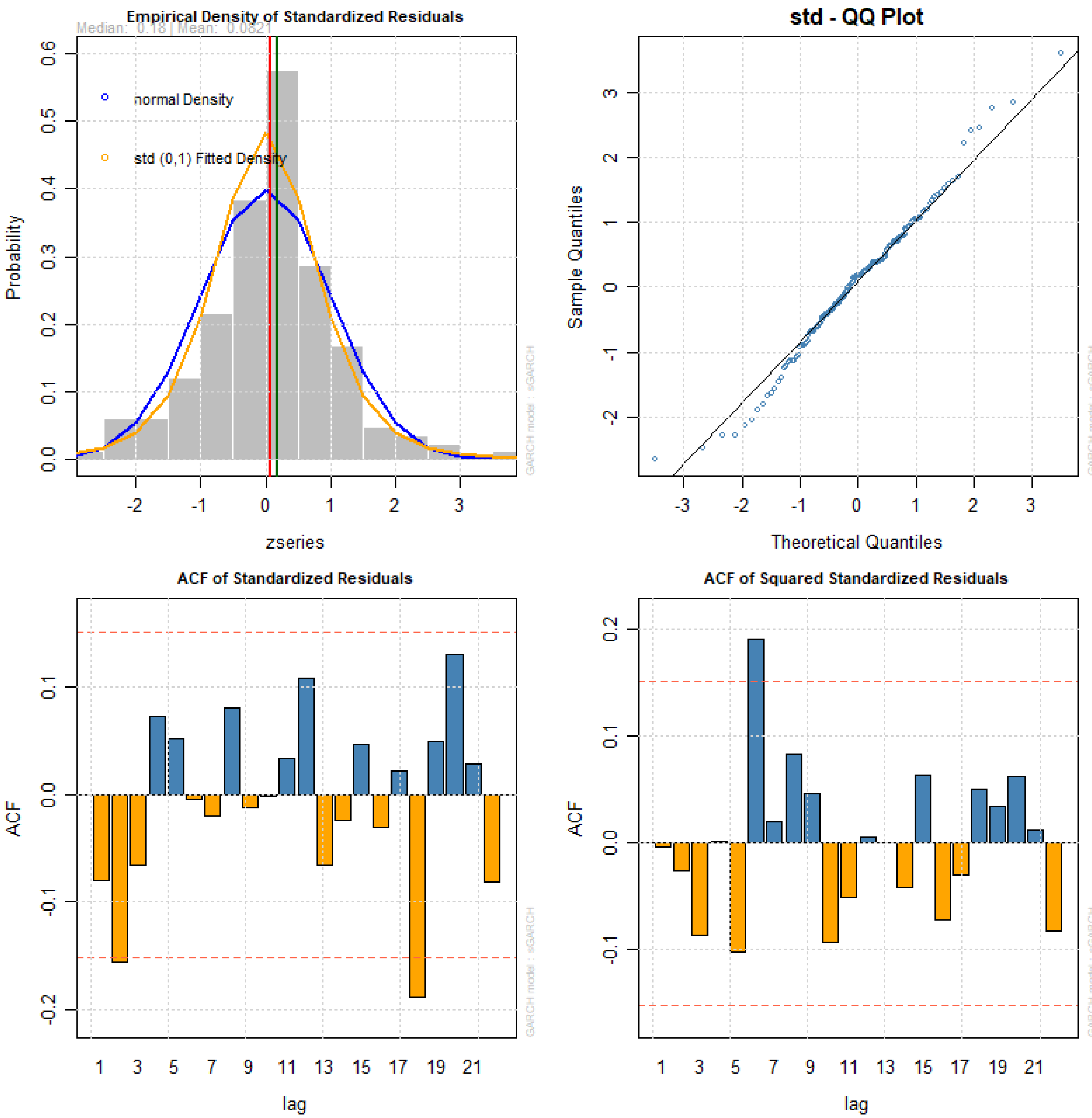

Figure 5.7: Diagnostics for the model ARMA(2,2) without intercept, and GARCH(1,1) assuming t-Student errors with variance targeting for pulpwood log-returns in the TMS Georgia region two, U.S.

Table 5.5: Best combination of ARMA(p,q) with and without intercept, and GARCH(1,1) assuming normal and t-Student errors with and without variance targeting (VT) for pulpwood in the TMS Georgia region two, U.S. μ indicates if the intercept was included in the ARMA model.

ARMA(p,q)			Distribution	VT	BIC	AIC	Statistical issues					
							Par. Significance		ACF		Ljung-Box	
p	q	μ					ARMA	GARCH	r_t	r_t^2	r_t	r_t^2
0	0	no	t-Student	yes	-1.765	-1.821		x	x		x	
0	0	no	Normal	yes	-1.755	-1.793		x	x		x	
0	1	no	t-Student	yes	-1.743	-1.817	x	x	x		x	
2	0	no	Normal	yes	-1.742	-1.817	x	x	x			
0	1	no	Normal	yes	-1.740	-1.796	x	x	x		x	
1	0	no	t-Student	yes	-1.740	-1.814	x	x	x		x	
0	2	no	Normal	yes	-1.737	-1.811	x	x				
0	0	no	t-Student	no	-1.736	-1.811	x	x	x		x	
2	0	no	t-Student	yes	-1.734	-1.828	x	x				
0	0	yes	Normal	yes	-1.732	-1.788	x		x		x	
0	2	no	t-Student	yes	-1.731	-1.824	x		x			
2	2	no	t-Student	yes	-1.730	-1.861						

5.3.3 Annualized volatility

The first measure of volatility considered was the historical or implied volatility (SD), estimated by Equation (5.2) and presented in Table 5.1 at the value of 7.6 and 10.0% for quarterly log-returns of sawtimber and pulpwood, respectively. The corresponding factor in calculating the annualized volatility is $\sqrt{4} = 2$. Therefore, the annualized implied volatility was 15.2, and 20.0% for sawtimber and pulpwood, respectively.

The implied volatility presupposes a constant variance over time, a hypothesis that was already rejected. Therefore, a more meaningful approach to estimate the volatility of log-returns for sawtimber and pulpwood is the GARCH model. The mean of the annualized predicted volatility of the four quarters ahead (2019Q1-2019Q4) was 7.7% and 18.6% for sawtimber and pulpwood, respectively.

Table 5.6: Parameter estimates for the combination of ARMA(2,2) without intercept, and GARCH(1,1) assuming t-Student errors with variance targeting for pulpwood log-returns in the TMS Georgia region two, U.S.

	Estimate	Std. Error	t value	Pr(>|t|)
ϕ_1	0.037	0.012	2.966	0.003
ϕ_2	-0.986	0.004	-248.807	< 0.001
θ_1	-0.068	0.009	-7.836	< 0.001
θ_2	1.011	0.005	187.057	< 0.001
α_1	0.075	0.041	1.830	0.067
β_1	0.863	0.076	11.397	< 0.001
shape	5.243	1.769	2.963	0.003
α_0	5.7×10^{-4}	NA	NA	NA

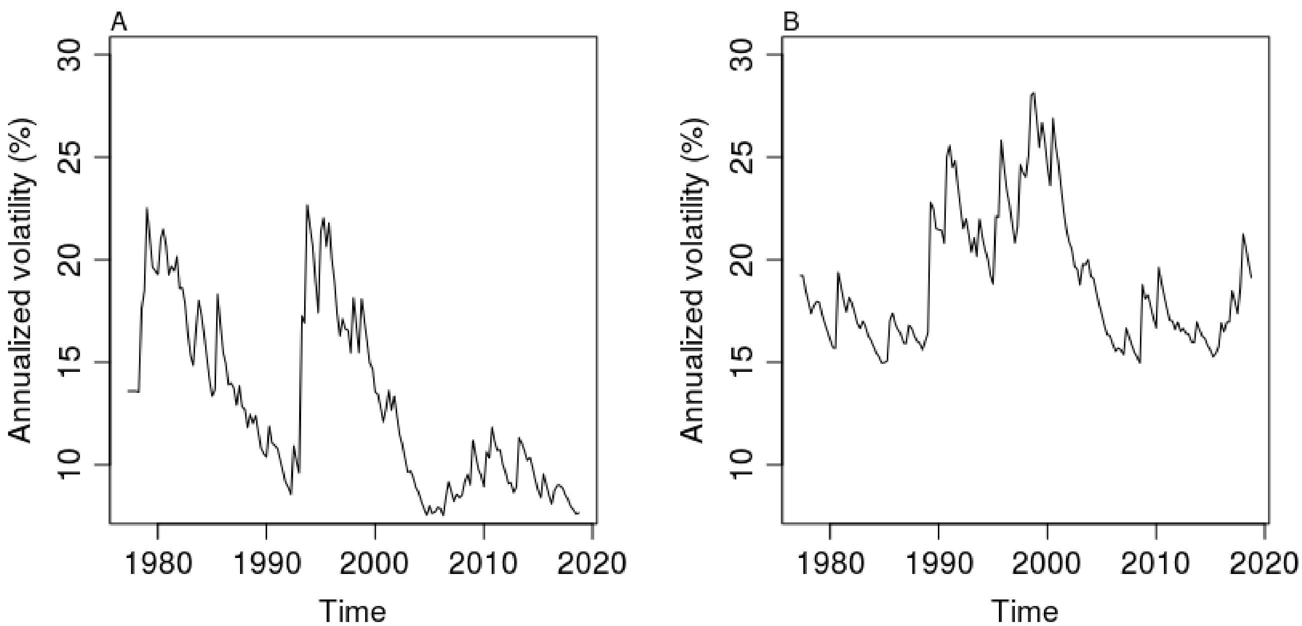

Figure 5.8: Conditional volatility (CSD) from estimated GARCH models of the log-returns of sawtimber (A), and pulpwood (B) in the TMS Georgia region two, U.S.

An intermediate alternative between the implied volatility approach and GARCH models was estimated (Equation 5.16), which captures the temporal changes of volatility within a specific time frame (Figure 5.9). The window at 12 quarters, equivalent to three years, maximized the correlation between the estimations of the QSD and the conditional volatility from the GARCH model at 93% for both sawtimber and pulpwood. The mean of the annualized volatility of the last 12 quarters (2016Q1-2018Q4) from QSD was 8.6% and 20.0% for sawtimber and pulpwood, respectively. The trends of the QSD and CSD can both be split into the same periods as the log-return's volatility, as described in §5.3.1. There are no clear differences between CSD and QSD for both sawtimber and pulpwood log-returns given that their boxes and whiskers overlapped (Figure 5.10). However, QSD has more variability than CSD.

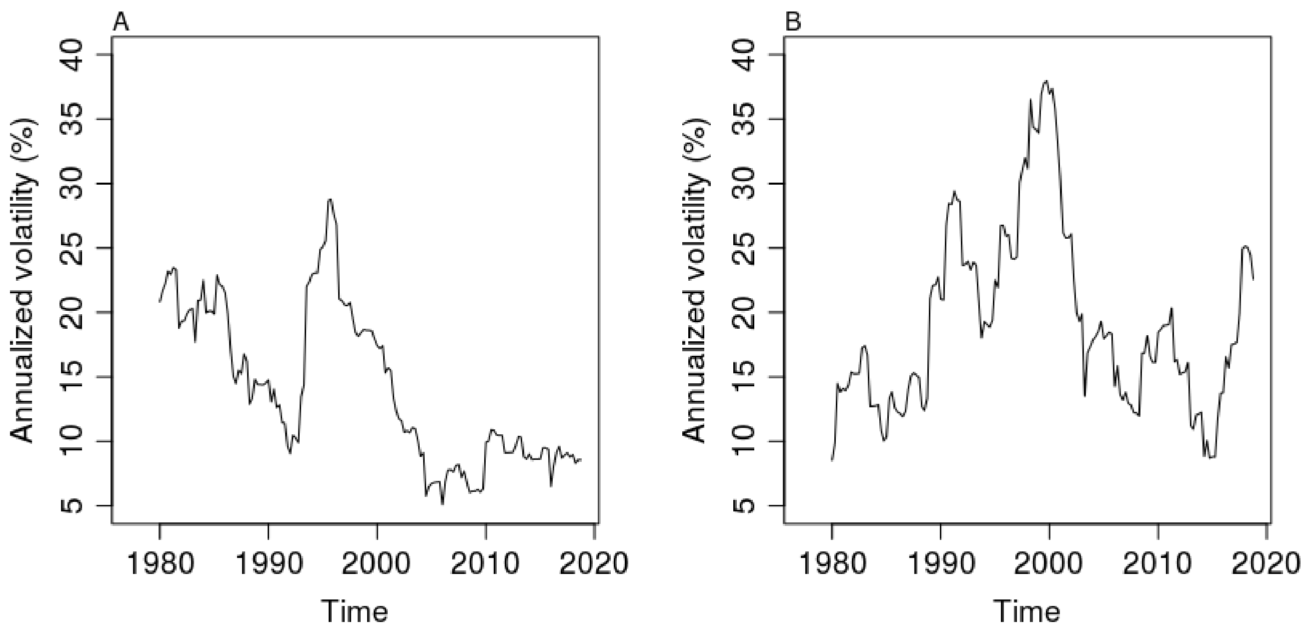

Figure 5.9: Quasi-conditional volatility or moving window standard deviation ($QSD_{t(12)}$, window = 12 quarters) of the log-returns for Sawtimber (A), and pulpwood (B) in the TMS Georgia region two, U.S.

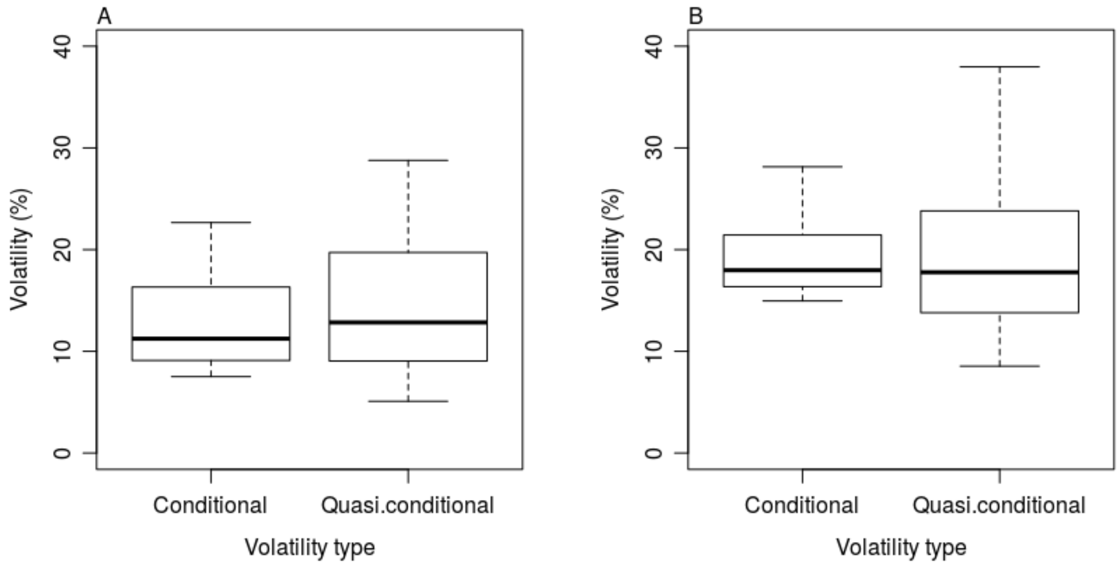

Figure 5.10: Comparison of the conditional volatility (CSD) and quasi-conditional volatility ($QSD_{t(12)}$, window = 12 quarters) of the log-returns for Sawtimber (A), and pulpwood (B) in the TMS Georgia region two, U.S.

5.3.4 Option princing of the LTTC

The inputs for the option pricing valuation are presented in Table 5.7. Timber volumes correspond to the conversion from MBF and cords to tons. Timber prices for sawtimber and pulpwood were from the last quarter of 2018 (latest available data). The timber value is the product of the timber volume and timber price and corresponds to strike and exercise price. The risk-free interest rate is the rate for U.S. Treasury Bonds with maturity at 10 years. Annualized volatility corresponds to each of the values for implied volatility (SD), conditional volatility (CSD) and quasi-conditional volatility (QSD). The risk-adjusted rate is the current appropriate firm's discount rate.

Results suggest that long-term timber contracts may not be profitable for forest product firms (Table 5.8). Most six-month equilibrium values of the call option (the sum of V_0 for sawtimer and pulpwood) per period were less than the threshold of the six-month cost of administering the provisions of the contract ($C = \$32,400$) by OPMs and volatility type.

Table 5.7: Summary of annualized value of the inputs for the option pricing model (OPM).

	Shaffer Jr. (1984)		Update or conversion	
	Sawtimber	Pulpwood	Sawtimber	Pulpwood
Timber volume	4,060 MBF	9,000 cords	31,262 tons	23,400 tons
	5,074 MBF	11,250 cords	39,070 tons	29,250 tons
Timber price	\$180.0/MBF	\$10.0/cord	\$27.16/ton	\$14.15/ton
Timber value	\$730,800	\$90,000	\$849,076	\$331,110
	\$913,320	\$112,500	\$1,061,136	\$413,888
Risk-free interest rate	11.8%		2.5%	
Annualized volatility	22.5%	22.6%	SD=15.2%	SD=20.0%
			CSD=7.7%	CSD=18.6%
			QSD=8.6%	QSD=20.0%
Risk-adjusted discount rate	7.0%		4.0%	
Costs (\$)	64,800		64,800	

No further analysis is required because the costs exceed the revenues. The six-month equilibrium value of the call option calculated with the implied volatility for the first period using both OPMs was higher than the threshold. In this case, the calculation of the NPV is required to evaluate the entire contract horizon (25 years). For both OPMs, the present value of the administering costs was PVc=\$1,018,125 (present value of the six-month payments of \$32,400 discounted at 2% per semester). For the Black-Scholes model with SD, PV_0=\$1,058,786, resulting in a positive NPV equal to \$40,661. Conversely, for the binomial model with SD, PV_0=\$1,002,718 resulting in a negative NPV equal to -\$15,407. Differences were found between the value of the call option estimated with the Black-Scholes and binomial models. Nevertheless, the binomial American and binomial European were equal. Therefore, there is no early exercise premium for the call option.

To identify the combination of a pair of input variables that make a LTTC beneficial, the value of the call option was plotted (V_0), compared against the six-month administrative contract cost ($C = \$32,400$), resulting in regions of loss ($C > V_0$) and profit ($V_0 > C$) (Figure 5.11). Volatility against the risk-free interest rate generates a parabolic threshold curve of the value of the call option for the LTTC (Figure 5.11 A). However, the timber value against the risk-free interest rate (Figure 5.11 B) and the timber value against volatility (Figure 5.11 C) yield a negative exponential threshold curve.

Table 5.8: Six-month equilibrium value of the call options (V_0) for sawtimber, pulpwood, and total (sawtimber + pulpwood). The option pricing models (OPMs) were Black-Scholes and binomial (American = European, no early exercise premium). The two periods of the contract were 1-14 and 15-25 years. The three volatility measures were implied volatility (SD), conditional volatility (CSD), and quasi-conditional volatility (QSD).

OPM	Period	Volatility type	Sawtimber ($)	Pulpwood ($)	Total ($)
Black-Sholes	1-14 years	SD	20,841	10,339	31,180
		CSD	12,040	9,694	21,734
		QSD	13,086	10,339	23,425
	15-25 years	SD	26,046	12,924	**38,970**
		CSD	15,047	12,117	27,164
		QSD	16,354	12,924	29,278
Binomial	1-14 years	SD	19,748	9,780	29,529
		CSD	11,476	9,174	20,650
		QSD	12,459	9,780	22,240
	15-25 years	SD	24,680	12,225	**36,906**
		CSD	14,343	11,467	25,810
		QSD	15,571	12,225	27,797

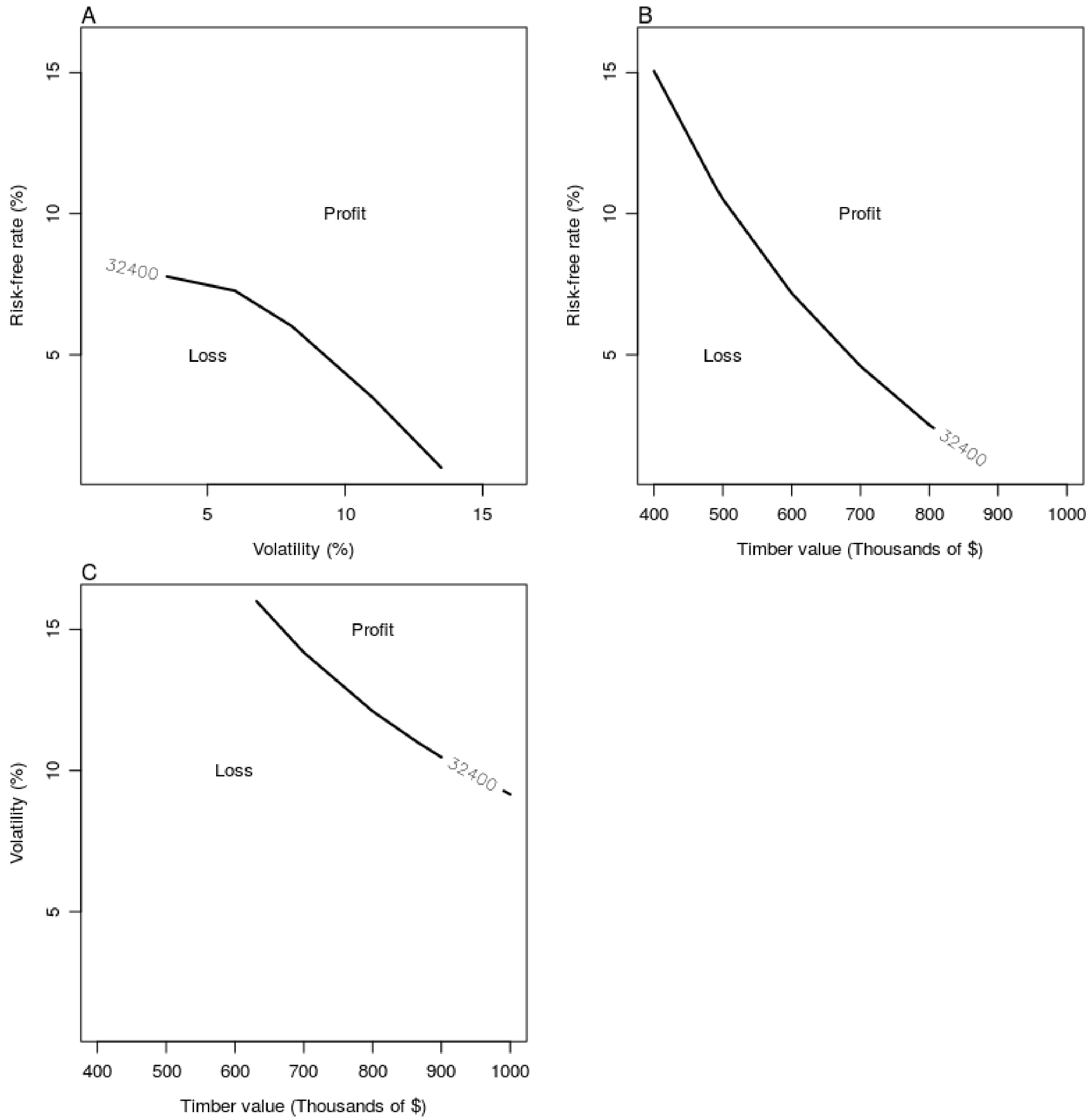

Figure 5.11: Regions of loss and profit as the result of the comparison between the six-month value of a call option (sum of V_0 for sawtimber and pulpwood) and the six-month administrative cost (C =\$32,400) by the pairs of the input variables in the Black-Scholes model. The threshold of C =\$32,400 defines the break even line. Volatility vs. the risk-free interest rate with fixed timber value at \$800k (A), the timber value vs. the risk-free interest rate with fixed volatility at 12% (B), and the timber value vs. volatility with fixed risk-free interest rate at 2.5% (C).

5.4 Discussion

The timberland investment environment has drastically changed in the past 30 years. Timber prices, volatility, and interest rates are all significantly different from when they were first examined as part of the LTTC framework by Shaffer Jr. (1984) (Mei et al., 2013). We analyzed the effect of the estimation of the current volatility on the value of the LTTC under present timber prices and interest rates. Based on the existing conditions, the analyzed LTTC is not profitable for forest product firms, mainly because volatility and the risk-free interest rate are relatively small. Indeed, the equilibrium value of the call option was less than the non-updated cost from 1984. Certainly, the cost of administrating a contract is greater than what it was 30 years ago (Callaghan et al., 2018).

The theory of the OPM dictates that the value of a call option increases (decreases) as the current market value (in this case the market value equals the strike price) increases (decreases), as volatility increases (decreases), and as the risk-free interest rate increases (decreases) (Zinkhan, 1991). However, the effect of the combination of a pair of variables on the value of the option is often unclear. We depicted such a relationship and showed the levels of volatility, risk-free interest rate, and timber value that would make the analyzed LTTC contract profitable.

A visual inspection of the timber prices and log-returns suggests marked differences in the sawtimber and pulpwood returns. Three different trends of prices and returns were identified for sawtimber and pulpwood. However, the time series changed at various times for sawtimber and pulpwood. Similarly, the post-recession behavior was very different for sawtimber and pulpwood. Sawtimber prices fell 50% while pulpwood prices appreciated 140%. Also, the variability of sawtimber at the beginning of the time series was relatively high and slowly decreased, while the variability of pulpwood was relatively constant with more positive returns than negative ones. However, at the end of the available time series, the variability for sawtimber was relatively lower than the variability of pulpwood.

The mean of the log-returns for both sawtimber and pulpwood was not statistically significantly different from zero. Mean reverting to zero returns are expected based on the efficient-market hypothesis (Cryer, Chan, 2008). This is the case of timberland investments and financial instruments with timber as the underlying asset, which mimic the financial market behavior (Zinkhan, Cubbage, 2003). As wood markets become more perfectly functioning, one of the principal reasons for holding timberlands disappears (Lönnstedt, Sedjo, 2012; Yao et al., 2014). Indeed, given the level of timber prices and timberland returns, new timberland investments are not optimal (Mei, Clutter, 2015). This helps to explain the trend of large scale divestitures in the past three decades. Under this premise, well-functioning wood markets will almost certainly provide raw material at a reasonable-competitive market price (Lönnstedt, Sedjo, 2012). Therefore, the contemporary wood market conditions and economic outlook not only preclude owning land but also may make LTTCs unprofitable for forest firms.

Often, data from time series of the log-returns are uncorrelated. In that case, the time series seems like white noise, from which one might incorrectly conclude that no further analysis is required (Andersen et al., 2009). However, the sample autocorrelation of both the squared and absolute value of log-returns does not vanish even for large lags, making evident that there is serial dependence beyond serial correlation (Andersen et al., 2009).

The implied volatility was larger than the conditional and quasi-conditional volatilities. Evidence supports this finding (Tsay, 2010). The higher value of the implied volatility may mislead to a conclusion that the contract is profitable when it is not. When the implied volatility was used into the Black-Scholes model, it resulted in a positive NPV of the LTTC. For that reason it is highly recommended to use the conditional volatility estimated using GARCH models or a similar estimation method that accounts for the conditional heteroscedasticity of the returns. Moreover, since the ARCH/GARCH effect was identified in both sawtimber and pulpwood, there is a high-order serially dependence or volatility clustering (Cryer, Chan, 2008). Volatility clustering also suggests that the returns may not be independent and identically distributed (Cryer, Chan, 2008). An intricate relationship exists between the mean of the log-returns and the volatility; high risk is often expected to lead to high returns (Zivot, 2009). Therefore, we consider that the conditional volatility (CSD) estimated using the GARCH model is the best estimation of volatility, since it takes into account its temporal relationship with the returns.

The quasi-conditional volatility approach was also proposed. This measure is useful for practitioners and forest companies that want a rule of thumb for timber return volatility. At first glance, the estimation of QSD seems like a very empirical measure of volatility. However, QSD has a fundamental explanation in the sense that it may reflect the effect of business cycles. Although timberland investments, as a whole, are thought uncorrelated with the business cycles (Waggle, Johnson, 2009), returns of timber products reflect the changes in the demand for goods and services in an economy (Yao, Mei, 2015). Thus, for the two analyzed time series, the current volatility is a result of the last three years' economic/financial performance of the wood market. Specifically, sawtimber CSD mirrored the business cycle trend[6].

A thorough exploration for the best combination of ARMA(p,q) and GARCH(1,1) setting was performed. The best model for sawtimber was an ARMA(4,3) without intercept and GARCH(1,1) assuming normal errors, and utilizing the variance targeting approach. The best model for pulpwood was an ARMA(2,2) without intercept and GARCH(1,1) assuming *t*-Student errors, and utilizing the variance targeting approach. The differences in the model orders and distributional properties confirm that returns of sawtimber and pulpwood behave differently. The time series of log-returns were modeled to simultaneously account for the mean equation and volatility. The mean of returns was modeled instead to assuming MR or GBM behavior. The time series of sawtimber returns was stationary in the mean suggesting that the log-returns are MR; whereas the corresponding for pulpwood was not true. This is important because the nature of timber prices implies different risks and valuations of

[6]Business cycles are depicted in Galvão (2002)

timberland investments (Mei et al., 2013). Evidence suggests that MR returns are less volatile than returns from random timber prices (Mei et al., 2013), in accordance with what we found. Moreover, the ACF of pulpwood returns showed a clear annual (four quarters) seasonal component in the time series; whereas the seasonality is not present in the sawtimber returns.

Given the small sample size of the time series of sawtimber and pulpwood, GARCH parameter estimates might be biased (Iglesias, Phillips, 2011; Hwang, Valls, 2006). However, the relative high persistence in the estimated models, especially for sawtimber, suggests that the possibility of bias is relatively low (Hwang, Valls, 2006). Evaluating the bias in the GARCH parameter estimates associated with small sample size needs to be undertaken in future work. Likewise, the seasonal component in the pulpwood returns suggests the existence of a $SARIMA_4$ process, an interesting time series feature to be analyzed. Moreover, although the existence of significant intervention has not been found for sawtimber log-returns (Mei et al., 2010), future work should address the effect of the 2008 recession on returns and volatilities of timber prices.

5.4.1 Conclusions

Well-functioning wood markets not only preclude owning land by forest firms, but also may diminish the value of LTTCs. The reduction in value of LTTCs is a result of the current low, risk-free interest rate and low volatility. The effect of three measures of volatility (sample standard deviation, GARCH volatility, and moving average standard deviation) were evaluated on the option pricing value of the LTTC. The best combination of the order for ARMA(p,q) with GARCH(1,1) was ARMA(4,3) and ARMA(2,2), identified for sawtimber and pulpwood returns, respectively. Quasi-conditional volatility depicted a similar trend to that which the conditional volatility estimated from GARCH models. Thus, given its simplicity, quasi-conditional volatility may be put into the option pricing models as an alternative to the GARCH-type volatility.

Chapter 6

Discussion and conclusions

Both internal and external factors influence financial timberland investment performance. The landowner has control on internal factors, such as species and genetic improvement selection, location of the forest plantation stands, and management regimes. Analyzing forest growth drivers, optimizing silvicultural practices to cost-effectively increase productivity, and determining the timber product class distribution are ways to increase profitability of timber production and manage associated risks of timberland investments. Conversely, external economic factors, e.g., global markets, interest rates, macroeconomic stability, may greatly influence the financial performance of timber production and timberland investments. Understanding the effect of these economic exogenous factors in the current forest business environment has also the most relevant importance.

Meta-analysis showed to be applicable to statistically combine, compile, and synthesize a large amount of information from research studies to analyze forest growth drivers. Such an intended investigation to determine the magnitude of effect of relevant forest growth factors on forest yield would require a massive experimental design. This type of research would be prohibitively expensive and technically unfeasible given the amount of resources involved.

Mixed effects meta-regression models for diameter at breast height (DBH), total height (Ht), basal area (BA), and volume (V) were estimated for loblolly pine (*Pinus taeda* L.) in the southeastern U.S. Both the fixed and random effects related with covariates and study identifiers (an arbitrary code), respectively, contribute to the explanation of forest yield variability. Overall, the level of genetic improvement, i.e., half-siblings, full-siblings, and clones; the US Southeastern physiographic region, i.e., Lower Coastal Plain, Upper Coastal

Plain, and Piedmont; the level of management regime, i.e., low, moderate, and high; and the stand density; successfully explain forest yield. The 95% confidence intervals over time were depicted for the levels within each of the forest growth factors to infer statistical differences among them. Likewise, given the relative orthogonality of the covariates in the estimated models, we can isolate the effect of each of the factors on forest growth and yield.

One of the main contributions of the meta-analysis conducted was an approach to impute standard deviations when missing in research studies. Thus, the cumulative maximum recorded standard deviation to account for the heteroscedastic nature of the forest yield variables was proposed. Moreover, knowledge gaps on forest growth factors were identified. These are associated with the lack of research studies to explain the effect of a single or a combination of drivers on loblolly pine growth and yield in the southeastern U.S.

Estimated models were utilized to determine the contribution of silviculture to forest yield of loblolly pine in the southeastern U.S. The effect of silvicultural practices on forest yield can be determined by taking partial derivatives of the meta-regression models for DBH, Ht, BA, and V with respect to the parameter estimates of moderate and high management regimes. The equations obtained represent the silvicultural response of applying moderate or high management regimes in comparison to the low management regime, which serves as the reference level. The responses can be utilized to conduct financial marginal analysis in determining the profitability of a level of management regime. The same approach of partial derivatives can be applied to determine the genetic gain or differences between pairs of physiographic regions (environmental effect) on forest growth and yield, and in conducting financial marginal analyses on these factors.

Finding the stand sawtimber (most valueble timber class) proportion is vital in calculating the stand timber value. The sawtimber proportion can weight either the total volume or timber prices to obtain the merchantable volume or the blended price, respectively. Three Bayesian approaches were proposed, described and applied in determining the proportion of loblolly pine trees with sawtimber potential in the southeastern U.S.: binomial, hierarchical and logit models. Overall, the sawtimber potential was almost a half of the stand timber volume. The tree size had a positive effect on the sawtimber potential, whereas the intensive management regime reduced the proportion of trees with sawtimber potential. Hence, although intensive management increases forest yield, forestland owners should consider the overall effect of silvicultural practices on the stand timber value. There were found some differences among locations (environmental effect), whereas the effect of planting density on sawtimber potential was negligible.

The economic context, or external factor, may considerably affect the timber production performance and timberland investment risks. In this sense, timber prices has been a great concern for timberland investors. The contemporary economic context considering the current forest business environment, timber markets, timber prices, timber price volatility, and rates of return was analyzed by updating the primer long-term timber cutting contract (LTTC) framework in the southeastern U.S. (Shaffer Jr., 1984). Modern Well-functioning wood markets not only preclude owning land by forest firms, but also may diminish the value of LTTCs. The reduction in value of LTTCs is a result of the current combination of low risk-free interest rate and low volatility.

Bibliography

Albaugh Timothy J., Allen H. Lee, Dougherty Phillip M., Johnsen Kurt H. Long term growth responses of loblolly pine to optimal nutrient and water resource availability // Forest Ecology and Management. 2004. 192, 1. 3–19.

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

Amateis Ralph L, Burkhart Harold E. The influence of thinning on the proportion of peeler, sawtimber, and pulpwood trees in loblolly pine plantations // Southern Journal of Applied Forestry. 2005. 29, 3. 158–162.

Andersen Torben G., Davis Richard A., Kreiss Jens-Peter, Mikosch Thomas. Introduction // Handbook of financial time series. Berlin Heidelberg: Springer-Verlag, 2009. 1–13.

Antón-Fernández Clara, Burkhart Harold E., Strub Mike, Amateis Ralph L. Effects of initial spacing on height development of loblolly pine // Forest Science. 2011. 57, 3. 201–211.

Ashton M S, Mendelsohn R, Singhakumara B M, Gunatilleke C V, Gunatilleke I A, Evans A. A financial analysis of rain forest silviculture in southwestern Sri Lanka // Forest Ecology and Management. 2001. 154. 431–441.

Aspinwall Michael J., King John S., McKeand Steven E., Bullock Bronson P. Genetic effects on stand-level uniformity and above- and belowground dry mass production in juvenile loblolly pine // Forest Ecology and Management. 2011. 262, 4. 609–619.

Black Fischer, Scholes Myron. The Pricing of options and corporate liabilities // Journal of Political Economy. 1973. 81, 3. 637–654.

Bollerslev Tim. Generalized autoregressive conditional heteroskedasticity // Journal of Econometrics. 1986. 31. 307–327.

Borders B. E., Will R. E., Markewitz D., Clark A., Hendrick R., Teskey R. O., Zhang Y. Effect of complete competition control and annual fertilization on stem growth and canopy relations for a chronosequence of loblolly pine plantations in the lower coastal plain of Georgia // Forest Ecology and Management. 2004. 192, 1. 21–37.

Borders Bruce, Volfovicz Roberto. Impact of cultural treatments and planting density on branch characteristics of loblolly pine at seven and eight years of age. PMRC Technical Report 2010 - 2. At, 2010. 192.

Borenstein Michael, Hedges Larry V., Higgins Julian P.T., Rothstein Hannah R. A basic introduction to fixed-effect and random-effects models for meta-analysis // Research Synthesis Methods. 2010. 1, 2. 97–111.

Borman Geoffrey D, Hewes Gina M, Overman Laura T, Brown Shelly. Comprehensive school reform and achievement: a meta-analysis // Review of Educational Research. 2003. 73, 2. 125–230.

Buford Marilyn, Burkhart Harold E. Genetic improvement effects on growth and yield of loblolly pine plantations // Forest Science. 1987. 33, 3. 707–724.

Burkhart Harold E. Growth and yield of southern pines – State of the art // Proceedings of the Southern Forest Economists Workshop. Chapel Hill, NC, 1979. 11.

Burkhart Harold E., Avery Tomas Eugene, Bullock Bronson P. Forest measurements. Long Grove, IL, USA: Waverland Press, Inc, 2018. Sixth ed. 434.

Product-Class Proportions for Thinned Loblolly Pine Plantations. // . 1989. 13, 1987.

Callaghan Daniel W, Khanal Puskar N, Straka Thomas J. An Analysis of costs and cost trends for Southern forestry practices // Journal of Forestry. 2018. 117, 1. 21–29.

Charles Amélie, Darné Olivier. The accuracy of asymmetric GARCH model estimation // International Economics, https:/doi.org/10.1016/j.inteco.2018.11.001. 2019. In press, November.

Chaudhari U K, Kane M B, Wetzstein M E. The key literature of, and trends in, forestry investment decisions using real options analysis // International Forestry Review. 2016. 18, 2. 146–160.

Choi Jungkee, Burkhart Harold E., Amateis Ralph L. Modeling trends in stem quality characteristics of loblolly pine trees in unthinned plantations // Canadian Journal of Forest Research. 2008. 38, 6. 1446–1457.

Clements Sherwood, Tidwell Alan, Jin Changha. Futures markets and real estate public equity: Connectivity of lumber futures and Timber REITs // Journal of Forest Economics. 2017. 28. 70–79.

Clutter Jerome L, Fortson James C, Pienaar Leon V, Brister Graham H, Bailey Robert L. Timber management: a quantitative approach. New York, NY, USA: John Wiley & Sons, Inc., 1992. Reprint. 333.

Cryer Jonathan D., Chan Kung-Sik. Time series analysis with applications in R. New York, NY, USA: Springer, 2008. 2nd. 491.

Curtis R P, Marshall D D. Why quadratic mean diameter? // Western Journal of Applied Forestry. 2000. 15, 360. 137–139.

Demidenko Eugene. Mixed models: Theory and applications with R. Hoboken, New Jersey: John Wiley & Sons, Inc., 2013. Second. 717.

Dicke Stephen G., Parker Robert C. Pine timber volume-to-weight conversions. 2016. 1–4.

Engle Robert F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation // Econometrica. 1982. 50, 4. 987–1007.

Engle Robert F., Mezrich J. GARCH for groups // Risk1. 1996. 8. 36–40.

Espinoza J. A., Allen H. L., McKeand S. E., Dougherty P. M. Stem sinuosity in loblolly pine with nitrogen and calcium additions // Forest Ecology and Management. 2012. 265. 55–61.

Furukawa Toshi A., Barbui Corrado, Cipriani Andrea, Brambilla Paolo, Watanabe Norio. Imputing missing standard deviations in meta-analyses can provide accurate results // Journal of Clinical Epidemiology. 2006. 59, 1. 7–10.

Galvão Ana Beatriz. Can non-linear time series models generate US business cycle asymmetric shape? // Economics Letters. 2002. 77, 2. 187–194.

Gao Lei, Mei Bin. Investor attention and abnormal performance of timberland investments in the United States // Forest Policy and Economics. 2013. 28, February. 60–65.

Gelman Andrew, Carlin John B., Stern Hal S., Dunson David B., Vehtari Aki, Rubin Donald B. Bayesian data analysis. Boca Raton, FL: CRC Press, Taylor & Francis Group LLC, 2013. 3rd. 667.

Ghalanos Alexios. Introduction to the rugarch package. 2017. 1–50.

Glass G. Primary, secondary, and meta-analysis of research // Educational Researcher. 1976. 5, 10. 3–8.

Green P. Corey, Bullock Bronson P., Kane Michael B. Culture and density effects on tree quality in midrotation non-thinned loblolly pine plantations // Forests. 2018. 9, 2. 1–16.

Haight Robert G, Holmes Thomas P. Stochastic price models and optimal tree cutting: Results for loblolly pine // Natural Resource Modeling. 1991. 5, 4. 423–443.

Harrison W.M., Kane Michael B. PMRC coastal plain culture / density study: age 12 analysis. PMRC Technical Report 2008 - 1. Athens, GA, 2008. 74.

Harrison W.M., Shiver D.B. PMRC culture / density study: age 2 analysis. PMRC Technical Report 1999 - 2. Athens, GA, 1999. 24.

Hedges L.V., Olkin Ingram. Statistical methods for meta-analysis. Orlando, FL: Academic Press, Inc., 1985. 369.

Hedges L.V., Vevea J L. Fixed- and random-effects models in meta-analysis // Psychological Methods. 1998. 3, 4. 486–504.

Henskens F L, Battaglia M, Cherry M L, Beadle C L. Physiological basis of spacing effects on tree growth and form in Eucalyptus globulus // Trees-Structure and Function. 2001. 15, 6. 365–377.

Hildebrandt Patrick, Knoke Thomas. Investment decisions under uncertainty-A methodological review on forest science studies // Forest Policy and Economics. 2011. 13. 1–15.

Hull John C. DerivaGem: Fundamentals of futures and option markets. Boston, 2017. 598.

Hwang Soosung, Valls Pedro L. Small sample properties of GARCH estimates and persistence // The European Journal of Finance. 2006. 12, 6-7. 473–494.

Iglesias Emma M., Phillips Garry D.A. Small sample estimation bias in garch models with any number of exogenous variables in the mean equation // Econometric Reviews. 2011. 30, 3. 303–336.

Jokela Eric J., Dougherty Philip M., Martin Timothy A. Production dynamics of intensively managed loblolly pine stands in the southern United States: A synthesis of seven long-term experiments // Forest Ecology and Management. 2004. 192, 1. 117–130.

Using Soils to Guide Fertilizer Recommendations for Major Land Areas and CRIFF Soil. // . 2012. 1–12.

Klemperer W David. Forest resource economics and finance. 2003.

Kutner Michael H, Nachtsheim Christopher J, Neter John, Li William. Applied linear statistical models. 2005. 5th. 1396.

Land S B Jr., Roberts S D, Duzan H Jr. Genetic and spacing effects on loblolly pine plantation development through age 17 // 12th Biennial Southern Silvicultural Research Conference. Gen. Tech. Rep. SRS-71. 594p. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station, 2004. 413–419.

Lehto Tarja, Ruuhola Teija, Dell Bernard. Boron in forest trees and forest ecosystems // Forest Ecology and Management. 2010. 260, 12. 2053–2069.

Lönnstedt Lars, Sedjo Roger A. Forestland ownership changes in the United States and Sweden // Forest Policy and Economics. 2012. 14, 1. 19–27.

MacFarlane D W, Green E J, Burkhart H E. Population density influences assessment and application of site index // Canadian Journal of Forest Research. 2000. 30, 9. 1472–1475.

McCullagh P., Nelder J.A. Generalized linear models. London & New York: Chapman & Hall, 1989. Second. 511.

McCulloch C.E., Searle S.R. Generalized, linear, and mixed models. New York: John Wiley & Sons, Inc, 2001.

McKeand Steve, Mullin Tim, Byram Tom, White Tim. Deployment of genetically improved loblolly and slash pines in the south // Journal of Forestry. 2003. 101, 3. 32–37.

McKeand Steven E. The success of tree breeding in the southern US // BioResources. 2015. 10, 1. 1–2.

McKeand Steven E, Abt Robert C, Allen Lee, Li Bailian, Catts Glenn P. What are the best loblolly pine genotypes worth to landowners? // J. Forestry. 2006a. 104, 7. 352–358.

McKeand Steven E., Jokela Eric J., Huber Dudley A., Byram Thomas D., Allen H. Lee, Li Bailian, Mullin Timothy J. Performance of improved genotypes of loblolly pine across different soils, climates, and silvicultural inputs // Forest Ecology and Management. 2006b. 227, 1-2. 178–184.

Mei Bin, Clutter Michael, Harris Thomas. Modeling and forecasting pine sawtimber stumpage prices in the US South by various time series models // Canadian Journal Forest Research. 2010. 40. 1506–1516.

Mei Bin, Clutter Michael L. Evaluating timberland investment opportunities in the United States: A real options analysis // Forest Science. 2015. 61, 2. 328–335.

Mei Bin, Clutter Michael L, Harris Thomas G. Timberland return drivers and timberland returns and risks: A simulation approach // Southern Journal of Applied Forestry. 2013. 37, 1. 18–25.

Norris Fundation . Timber Mart-South. Athens, GA, 2018.

Petitti Diana B. Meta-analysis, decision analysis, and cost-effectiveness analysis: methods for quantitative synthesis in medicine. New York, USA: Oxford University Press, Inc., 2000. 2nd. 319.

Phelps John E., McCurdy Dwight. Twenty-year trends in practices of Illinois timber buyers // Forest Products Journal. 1997. 47, 9. 31–32.

Pienaar L. V., Turnbull K. J. The Chapman-Richards generalization of Von Bertalanffy's growth model for basal area growth and yield in even-aged stands // Forest Science. 1973. 19, 1. 2–22.

Pienaar Leon V, Rheney John W. Modeling stand level growth and yield response to silvicultural treatments // Forest Science. 1995. 41, 3. 629–638.

Poudel Krishna Prasad, Cao Quang V. Evaluation of methods to predict Weibull parameters for characterizing diameter distributions // Forest Science. 2013. 59, 2. 243–252.

R Development Core Team . R: A language and environmental for statistical computing. Vienna, Austria, 2018.

Reeves L H, Haight R G. Timber harvest scheduling with price uncertainty using Markowitz portfolio optimization // Annals of Operations Research. 2000. 95. 229–250.

Rencher Alvin C., Schaalje G. Bruce. Linear models in statistics. New Jersey, USA: Jhon Wiley & Sons, Inc., 2008. 2nd. 672.

Rendleman Jr. Richard J, Bartter Brit J. Two-state option pricing // The Journal of Finance. 1979. 34, 5. 1093–1110.

Restrepo Héctor I., Bullock Bronson P., Montes Cristian R. Growth and yield drivers of loblolly pine in the southeastern U.S.: A meta-analysis // Forest Ecology and Management. 2019. 435. 205–218.

Restrepo Héctor Iván, Bullock Bronson P, Montes Cristian R. Contribution of silviculture to loblolly pine growth and yield in the southeastern U.S.: A meta-analysis

// Proceedings of the 19th biennial southern silvicultural research conference. e-Gen. Tech. Rep. SRS-234. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station, 2018. 152–157.

Rose C, Hall D, Shiver B, Clutter M, Borders B. A generalized methodology for developing flexible whole stand survival models. Plantation Management Research Cooperative (PMRC) Technical Report 2002-2. University of Georgia, Athens, GA, 2002. 27.

Roth Brian E., Jokela Eric J., Martin Timothy A., Huber Dudley A., White Timothy L. Genotype x environment interactions in selected loblolly and slash pine plantations in the Southeastern United States // Forest Ecology and Management. 2007. 238, 1-3. 175–188.

Schumacher F X. A new growth curve and its application to timber-yield studies. // Journal of Forestry. 1939. 37. 819–820.

Schwarzer Guido, Carpenter James R, Rücker Gerta. Meta-analysis with R. 2015. 252.

Valuation of Certain Long-Term Timber Cutting Contracts. // . 1984. 30, 3. 774–787.

Smith B.C., Bullock B.P., Isik F., McKeand S.E. Modeling genetic effects on growth of diverse provenances and families of loblolly pine across optimum and deficient nutrient regimes // Canadian Journal of Forest Research. 2014. 44, 11. 1453–1461.

Snowdon P. A ratio estimator for bias correction in logarithmic regressions // Canadian Journal of Forest Research. may 1991. 21, 5. 720–724.

Snowdon Peter. Modeling Type 1 and Type 2 growth responses in plantations after application of fertilizer or other silvicultural treatments // Forest Ecology and Management. 2002. 163, 1-3. 229–244.

South David, Harper Richard. A decline in timberland continues for several Southern yellow pines // Journal of Forestry. 2016. 114, 2. 116–124.

Strub Mike R., Green Edwin J., Burkhart Harold E., Pirie Walter R. Merchantability of loblolly pine–An application of nonlinear regression with a discrete dependent variable // Forest Science. 1986. 32, 1. 254–261.

Sun C, Zhang D. Assessing the financial performance of forestry-related investment vehicles: capital asset pricing model vs. arbitrage pricing theory // American Journal of Agricultural Economics. 2001. 83, 3. 617–628.

Sun Changyou. Price variation and volume dynamics of securitized timberlands // Forest Policy and Economics. 2013. 27. 44–53.

Sun Changyou, Rahman Mohammad M., Munn Ian A. Adjustment of stock prices and volatility to changes in industrial timberland ownership // Forest Policy and Economics. 2013. 26. 91–101.

Switzer Toccoa. Money does grow on trees // National Real Estate Investor. 2006. 45, 5. 98–101.

Teeter Lawrence, Zhou Xiaoping. Projecting timber inventory at the product level // Forest Science. 1998. 45, 2. 226–231.

Teräsvirta Timo. An introduction to univariate GARCH models // Handbook of financial time series. Berlin Heidelberg: Springer-Verlag, 2009. 17–42.

Thomas R Quinn, Brooks Evan B, Jersild Annika L, Ward Eric J, Wynne Randolph H, Albaugh Timothy J, Dinon-aldridge Heather, Burkhart Harold E, Domec Jean-christophe, Fox Thomas R, Gonzalez-Benecke Carlos A, Martin Timothy A, Noormets Asko, Sampson David A, Teskey Robert O. Leveraging 35 years of Pinus taeda research in the southeastern US to constrain forest carbon cycle predictions: regional data assimilation using ecosystem experiments // Biogeosciences. 2017. 14, July. 3525–3547.

Thomopoulos Nick T. Essentials of Monte Carlo simulation: statistical methods for building simulation models. New York: Springer, 2013. 171.

Tsay Ruey S. Analysis of financial time series. Hoboken, New Jersey: John Wiley & Sons, Inc., 2010. 3rd. 677.

Vaynman Igor, Beare Brendan K. Stable limit theory for the variance targeting estimator // Essays in Honor of Peter C. B. Phillips (Advances in Econometrics, Volume 33). 2014. 639–672.

Viechtbauer Wolfgang. Conducting meta-analysis in R with the metafor package // Journal of Statitistical Software. 2010. 36, 3. 1–48.

Wackerly Dennis D., Mendenhall III William, Sheaffer Richard L. Mathematical statistics with applications. Belmont, CA, USA: Thomson Learning, Inc., 2008. 7th. 912.

Waggle Doug, Johnson Don T. An analysis of the impact of timberland, farmland and commercial real estate in the asset allocation decisions of institutional investors // Review of Financial Economics. 2009. 18. 90–96.

Results of southern pine planting experiments established in the middle twenties. // . 1969. April. 237–241.

Wan Yang, Clutter Michael L., Siry Jacek P., Mei Bin. Assessing the impact of macroeconomic news on the U.S. forest products industry portfolio across business cycles: 1963-2010 // Forest Policy and Economics. 2013. 28. 15–22.

Yao Wenjing, Mei Bin. Assessing forestry-related assets with the intertemporal capital asset pricing model // Forest Policy and Economics. 2015. 50. 192–199.

Yao Wenjing, Mei Bin, Clutter Michael L. Pricing timberland assets in the United States by the arbitrage pricing theory // Forest Science. 2014. 60, 5. 943–952.

Zeide Boris. Analysis of growth equations // Forest Science. 1993. 39, 3. 594–617.

Zhang Daowei, Butler Brett J., Nagubadi Rao V. Institutional timberland ownership in the US South: Magnitude, location, dynamics, and management // Journal of Forestry. 2012. 110, 7. 355–361.

Zhang Hui, Lu Naiji, Feng Changyong, Thurston Sally W., Xia Yinglin, Zhu Liang, Tu Xin M. On fitting generalized linear mixed-effects models for binary responses using different statistical packages // Statistics in Medicine. 2011. 30, 20. 2562–2572.

Zhao Dehai, Kane Michael, Teskey Robert, Fox Thomas R., Albaugh Timothy J., Allen H. Lee, Rubilar Rafael. Maximum response of loblolly pine plantations to silvicultural management in the southern United States // Forest Ecology and Management. 2016. 375. 105–111.

Zhao Dehai, Kane Michael, Wang Mingliang. Coastal plain culture / density study: age 18 analysis for loblolly pine. PMRC Technical Report 2014 - 3. Athens, GA, 2014. 43.

Zinkhan F. Christian. Option pricing and timberland ’ s land-use conversion option // Land Economics. 1991. 67, 3. 317–325.

Zinkhan F. Christian, Cubbage Frederick W. Financial analysis of timber investments // Forest in a market economy. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2003. 77–95.

Zivot Eric. Practical issues in the analysis of univariate GARCH models // Handbook of financial time series. Berlin Heidelberg: Springer-Verlag, 2009. 113–155.

Appendix A

Detail of installation plots

Site preparation and subsequent silvicultural treatments represent two levels of management intensity; operational and intensive culture. The operational treatment consisted of bedding in the spring followed by a fall herbicide treatment. The herbicide treatment consisted of 12 oz. Arsenal plus 1 qt. Garlon 4 per acre if competition was waxy-leafed species or 12 oz. Arsenal plus 1 qt. Accord per acre if the competition consisted mainly of grass. Herbicide was applied in a five-foot band over the rows. At planting, 500 lbs. of 10-10-10 fertilizer was applied. The intensive cultural treatment consisted of bedding in the spring followed by a fall herbicide application. The herbicide treatment was a broadcast application of 16 oz. Arsenal, 2 qts. Garlon 4 and 2 qts. Accord per acre. At planting, 500 lbs. of 10-10-10 fertilizer was applied. Fertilizer treatments, including the addition of micro nutrients, will be repeated at least every three years. Beginning in the spring of the first growing season (1996), the plots were sprayed with 4 oz. Oust per acre along with directed sprays to keep the plots free of competing vegetation. Insecticides designed to control tip moths were applied as often as necessary to maintain tip moth control through the first two growing seasons.

Appendix B

Model building for the GARCH models

Table B.1: Bayesian Information Criterion (BIC) of the combination of ARMA(p,q) and GARCH(1,1).

ARMA(p,q)			Sawtimber				Pulpwood			
			Normal		t-Student		Normal		t-Student	
p	q	Mean	VT	NVT	VT	NVT	VT	NVT	VT	NVT
0	0	0	-2.415	-2.385	-2.405	-2.375	-1.755	-1.725	-1.765	-1.736
1	0	0	-2.389	-2.359	-2.38	-2.35	-1.734	-1.703	-1.74	-1.711
2	0	0	-2.362	-2.331	-2.347	-2.317	-1.742	-1.712	-1.734	-1.705
3	0	0	-2.336	-2.306	-2.321	-2.291	-1.718	-1.687	-1.706	-1.676
4	0	0	-2.306	-2.276	-2.29	-2.26	-1.699	-1.669	-1.686	-1.656
5	0	0	-2.287	-2.257	-2.268	-2.239	-1.667	-1.636	-1.654	-1.624
0	1	0	-2.39	-2.36	-2.381	-2.35	-1.74	-1.71	-1.742	-1.713
1	1	0	-2.364	-2.334	-2.351	-2.32	-1.723	-1.693	-1.721	-1.691
2	1	0	-2.332	-2.302	-2.319	-2.289	-1.715	-1.684	-1.705	-1.675
3	1	0	-2.308	-2.278	-2.291	-2.261	-1.695	-1.665	-1.683	-1.654
4	1	0	-2.281	-2.252	-2.261	-2.23	-1.677	-1.646	-1.655	-1.626
5	1	0	-2.257	-2.227	-2.238	-2.209	-1.643	-1.612	-1.628	-1.598
0	2	0	-2.361	-2.331	-2.347	-2.317	-1.737	-1.706	-1.731	-1.701
1	2	0	-2.332	-2.301	-2.317	-2.287	-1.706	-1.675	-1.7	-1.67
2	2	0	-2.324	-2.295	-2.312	-2.283	-1.709	-1.678	-1.73	-1.702
3	2	0	-2.441	-2.394	-2.387	-2.277	-1.689	-1.658	-1.705	-1.676
4	2	0	-2.279	-2.247	-2.268	-2.238	-1.65	-1.619	-1.635	-1.606
5	2	0	-2.245	-2.218	-2.239	-2.207	-1.615	-1.585	-1.6	-1.571
0	3	0	-2.335	-2.305	-2.322	-2.291	-1.705	-1.674	-1.699	-1.67
1	3	0	-2.305	-2.275	-2.291	-2.261	-1.682	-1.651	-1.676	-1.646
2	3	0	-2.446	-2.42	-2.431	-2.281	-1.691	-1.661	-1.715	-1.643

Table B.1 continued from previous page

ARMA(p,q)			Sawtimber				Pulpwood			
			Normal		t-Student		Normal		t-Student	
p	q	Mean	VT	NVT	VT	NVT	VT	NVT	VT	NVT
3	3	0	-2.312	-2.325	-2.277	-2.233	-1.663	-1.633	-1.688	-1.659
4	3	0	-2.399	-2.389	-2.276	-2.363	-1.628	-1.602	-1.667	-1.582
5	3	0	-2.219	-2.194	-2.319	-2.177	-1.586	-1.574	-1.717	-1.712
0	4	0	-2.308	-2.278	-2.293	-2.263	-1.699	-1.669	-1.686	-1.656
1	4	0	-2.28	-2.25	-2.265	-2.235	-1.677	-1.647	-1.664	-1.634
2	4	0	-2.271	-2.243	-2.26	-2.23	-1.663	-1.632	-1.692	-1.662
3	4	0	-2.275	-2.323	-2.314	-2.199	-1.632	-1.602	-1.662	-1.633
4	4	0	-2.25	-2.199	-2.22	-2.187	-1.625	-1.564	-1.675	-1.598
5	4	0	-2.19	-2.165	-2.319	-2.284	-1.582	-1.554	-1.562	-1.533
0	5	0	-2.293	-2.264	-2.277	-2.247	-1.669	-1.639	-1.655	-1.625
1	5	0	-2.263	-2.233	-2.246	-2.216	-1.651	-1.621	-1.625	-1.596
2	5	0	-2.248	-2.22	-2.231	-2.204	-1.632	-1.602	-1.661	-1.632
3	5	0	-2.217	-2.193	-2.205	-2.175	-1.633	-1.575	-1.637	-1.608
4	5	0	-2.189	-2.159	-2.175	-2.146	-1.595	-1.542	-1.626	-1.53
5	5	0	-2.406	-2.174	-2.154	-2.125	-1.565	-1.577	-1.726	-1.573
0	0	1	-2.386	-2.356	-2.376	-2.346	-1.732	-1.701	-1.751	-1.724
1	0	1	-2.36	-2.33	-2.351	-2.321	-1.713	-1.682	-1.727	-1.7
2	0	1	-2.332	-2.302	-2.318	-2.288	-1.726	-1.695	-1.722	-1.693
3	0	1	-2.306	-2.276	-2.292	-2.262	-1.703	-1.672	-1.695	-1.666
4	0	1	-2.276	-2.247	-2.261	-2.231	-1.681	-1.65	-1.67	-1.641
5	0	1	-2.257	-2.228	-2.239	-2.21	-1.648	-1.617	-1.637	-1.608
0	1	1	-2.361	-2.331	-2.351	-2.321	-1.722	-1.691	-1.73	-1.703
1	1	1	-2.335	-2.304	-2.321	-2.291	-1.707	-1.677	-1.711	-1.683

Table B.1 continued from previous page

ARMA(p,q)			Sawtimber				Pulpwood			
			Normal		t-Student		Normal		t-Student	
p	q	Mean	VT	NVT	VT	NVT	VT	NVT	VT	NVT
2	1	1	-2.302	-2.272	-2.29	-2.259	-1.7	-1.669	-1.694	-1.665
3	1	1	-2.279	-2.249	-2.262	-2.232	-1.679	-1.649	-1.67	-1.642
4	1	1	-2.252	-2.222	-2.232	-2.201	-1.65	-1.62	-1.64	-1.611
5	1	1	-2.227	-2.198	-2.209	-2.18	-1.62	-1.59	-1.607	-1.578
0	2	1	-2.332	-2.301	-2.317	-2.287	-1.721	-1.691	-1.72	-1.692
1	2	1	-2.302	-2.271	-2.288	-2.258	-1.691	-1.66	-1.69	-1.661
2	2	1	-2.294	-2.265	-2.282	-2.253	-1.676	-1.646	-1.668	-1.639
3	2	1	-2.293	-2.386	-2.367	-2.247	-1.667	-1.637	-1.654	-1.624
4	2	1	-2.249	-2.248	-2.239	-2.209	-1.642	-1.612	-1.627	-1.742
5	2	1	-2.215	-2.189	-2.209	-2.179	-1.592	-1.562	-1.58	-1.551
0	3	1	-2.305	-2.275	-2.292	-2.262	-1.689	-1.659	-1.689	-1.66
1	3	1	-2.275	-2.245	-2.262	-2.232	-1.666	-1.636	-1.665	-1.637
2	3	1	-2.415	-2.389	-2.403	-2.38	-1.671	-1.64	-1.696	-1.626
3	3	1	-2.282	-2.291	-2.319	-2.29	-1.644	-1.613	-1.671	-1.643
4	3	1	-2.428	-2.242	-2.318	-2.217	-1.614	-1.68	-1.696	-1.7
5	3	1	-2.189	-2.164	-2.275	-2.147	-1.562	-1.646	-1.708	-1.682
0	4	1	-2.278	-2.248	-2.264	-2.234	-1.681	-1.651	-1.67	-1.641
1	4	1	-2.25	-2.22	-2.236	-2.206	-1.655	-1.624	-1.645	-1.616
2	4	1	-2.241	-2.214	-2.231	-2.201	-1.643	-1.613	-1.675	-1.647
3	4	1	-2.297	-2.278	-2.281	-2.194	-1.613	-1.582	-1.644	-1.616
4	4	1	-2.231	NA	-2.24	-2.21	-1.583	-1.552	-1.676	-1.643
5	4	1	-2.161	-2.135	-2.286	-2.25	-1.665	-1.526	-1.656	-1.63
0	5	1	-2.264	-2.234	-2.248	-2.218	-1.65	-1.62	-1.638	-1.61

Table B.1 continued from previous page

ARMA(p,q)			Sawtimber				Pulpwood			
			Normal		t-Student		Normal		t-Student	
p	q	Mean	VT	NVT	VT	NVT	VT	NVT	VT	NVT
1	5	1	-2.234	-2.204	-2.217	-2.187	-1.628	-1.589	-1.609	-1.581
2	5	1	-2.218	-2.19	-2.204	-2.174	-1.635	-1.582	-1.645	-1.617
3	5	1	-2.187	-2.164	-2.176	-2.146	-1.612	-1.582	-1.619	-1.591
4	5	1	-2.159	-2.129	-2.209	-2.128	-1.57	-1.538	-1.605	-1.576
5	5	1	-2.293	-2.105	-2.391	-2.096	-1.659	-1.625	-1.58	-1.552

Appendix C

Model outputs for the GARCH models

```
MODEL FOR SAWTIMBER
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model : sGARCH(1,1)
Mean Model : ARFIMA(4,0,3)
Distribution : norm

Optimal Parameters
------------------------------------
Estimate  Std. Error   t value Pr(>|t|)
ar1    -0.735872    0.000178 -4131.314 0.000000
ar2     0.661796    0.000172  3843.680 0.000000
ar3     1.066014    0.000263  4051.868 0.000000
ar4     0.164937    0.000054  3073.425 0.000000
ma1     0.722292    0.000182  3977.350 0.000000
ma2    -0.825891    0.000218 -3786.474 0.000000
ma3    -1.152035    0.000301 -3830.266 0.000000
alpha1  0.144261    0.041948     3.439 0.000584
beta1   0.827833    0.052097    15.890 0.000000
omega   0.000129          NA        NA       NA

Robust Standard Errors:
Estimate  Std. Error     t value Pr(>|t|)
ar1    -0.735872    0.000243 -3029.36754  0.00000
ar2     0.661796    0.001527   433.35651  0.00000
ar3     1.066014    0.001054  1010.92737  0.00000
ar4     0.164937    0.000873   188.91777  0.00000
ma1     0.722292    0.001222   590.90361  0.00000
ma2    -0.825891    0.002216  -372.73619  0.00000
ma3    -1.152035    0.002957  -389.64622  0.00000
alpha1  0.144261    1.083182     0.13318  0.89405
beta1   0.827833    1.492732     0.55458  0.57918
omega   0.000129          NA          NA       NA

LogLikelihood : 226.9985

Information Criteria
------------------------------------

Akaike       -2.6108
Bayes        -2.4427
```

```
Shibata      -2.6162
Hannan-Quinn -2.5426

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
statistic p-value
Lag[1]                           0.02724  0.8689
Lag[2*(p+q)+(p+q)-1][20]   7.98017  1.0000
Lag[4*(p+q)+(p+q)-1][34]  11.76821  0.9790
d.o.f=7
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
statistic p-value
Lag[1]                           0.6818  0.4090
Lag[2*(p+q)+(p+q)-1][5]    3.7764  0.2833
Lag[4*(p+q)+(p+q)-1][9]    5.6643  0.3386
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
Statistic Shape Scale P-Value
ARCH Lag[3]    0.6292 0.500 2.000  0.4277
ARCH Lag[5]    1.5707 1.440 1.667  0.5739
ARCH Lag[7]    3.0309 2.315 1.543  0.5077

Nyblom stability test
------------------------------------
Joint Statistic:  2.352
Individual Statistics:
ar1    0.01921
ar2    0.01891
ar3    0.01767
ar4    0.01851
ma1    0.01663
ma2    0.01988
ma3    0.02154
alpha1 0.25534
beta1  0.27472

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:       2.1 2.32 2.82
Individual Statistic:  0.35 0.47 0.75
```

```
Sign Bias Test
------------------------------------
t-value   prob sig
Sign Bias           1.3414 0.1817
Negative Sign Bias  0.3298 0.7419
Positive Sign Bias  0.6499 0.5166
Joint Effect        3.6722 0.2991


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
group statistic p-value(g-1)
1    20     33.60      0.02048
2    30     30.78      0.37566
3    40     44.02      0.26753
4    50     63.84      0.07556


Elapsed time : 2.399172
```

```
MODEL FOR PULPWOOD

*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model : sGARCH(1,1)
Mean Model : ARFIMA(2,0,2)
Distribution : std

Optimal Parameters
------------------------------------
Estimate  Std. Error   t value Pr(>|t|)
ar1     0.036548    0.012323    2.9658 0.003019
ar2    -0.986029    0.003963 -248.8072 0.000000
ma1    -0.068075    0.008687   -7.8361 0.000000
ma2     1.010579    0.005403  187.0574 0.000000
alpha1  0.075162    0.041068    1.8302 0.067220
beta1   0.863189    0.075737   11.3972 0.000000
shape   5.242813    1.769139    2.9635 0.003042
omega   0.000569          NA        NA       NA

Robust Standard Errors:
Estimate  Std. Error  t value Pr(>|t|)
ar1     0.036548    0.016603   2.2013 0.027713
ar2    -0.986029    0.017470 -56.4399 0.000000
ma1    -0.068075    0.009440  -7.2114 0.000000
ma2     1.010579    0.005570 181.4446 0.000000
alpha1  0.075162    0.033102   2.2706 0.023170
beta1   0.863189    0.049879  17.3058 0.000000
shape   5.242813    2.094603   2.5030 0.012314
omega   0.000569          NA       NA       NA

LogLikelihood : 162.3953

Information Criteria
------------------------------------

Akaike       -1.8610
Bayes        -1.7303
Shibata      -1.8644
Hannan-Quinn -1.8080
```

```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
statistic p-value
Lag[1]                         1.106 0.29292
Lag[2*(p+q)+(p+q)-1][11]      6.954 0.06193
Lag[4*(p+q)+(p+q)-1][19]      9.876 0.48901
d.o.f=4
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
statistic p-value
Lag[1]                    0.002571  0.9596
Lag[2*(p+q)+(p+q)-1][5]   1.237096  0.8041
Lag[4*(p+q)+(p+q)-1][9]   5.295955  0.3873
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
Statistic Shape Scale P-Value
ARCH Lag[3]     1.257 0.500 2.000  0.2621
ARCH Lag[5]     2.300 1.440 1.667  0.4087
ARCH Lag[7]     6.000 2.315 1.543  0.1412

Nyblom stability test
------------------------------------
Joint Statistic:  0.5877
Individual Statistics:
ar1    0.08028
ar2    0.05827
ma1    0.05166
ma2    0.05339
alpha1 0.06093
beta1  0.06364
shape  0.14542

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:       1.69 1.9 2.35
Individual Statistic:  0.35 0.47 0.75

Sign Bias Test
------------------------------------
t-value   prob sig
Sign Bias           0.3710 0.7111
Negative Sign Bias  0.1338 0.8937
```

```
Positive Sign Bias  0.9176 0.3602
Joint Effect        2.6732 0.4448


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
group statistic p-value(g-1)
1    20     19.95       0.3978
2    30     37.97       0.1230
3    40     33.00       0.7393
4    50     48.27       0.5027


Elapsed time : 0.31883
```