A GROWTH AND YIELD SYSTEM FOR SLASH PINE INCLUDING RESPONSES TO SILVICULTURAL TREATMENTS

by

LAURA RAMIREZ QUINTERO

(Under the Direction of Cristian R. Montes)

ABSTRACT

Slash pine (*Pinus elliottii* Engelm) is the second most important commercial species in the southeastern United States, it is usually established in poorly drained flatwoods where it outperforms other common commercial pine species. Modeling slash pine growth and how it responds to silvicultural treatments is of interest to forest managers wanting to maximize growth and profits within the industry. In this dissertation, a system of differential equations is proposed to model slash pine growth including the effect of silvicultural treatments (i.e., bedding, and vegetation control). Data for this model came from a long-term study (30 years) established by the Plantation Management Research Cooperative (PMRC) across Georgia, Florida, and South Carolina.

To construct this growth and yield system, a dominant height model was first proposed in which the effect of bedding and vegetation control was closely evaluated. This model guided the construction of the mortality model, in which a modeling approach including height increments instead of time increments when using differential equations was evaluated. Building upon the mortality model, the whole system of differential equations was proposed after adding the basal area component. The model system describes the trajectory of three state variables: dominant height, survival/mortality, and basal area. Treatments effects were incorporated into the dominant height and basal area models by using parameter modifiers and dummy variables associated with each of the treatments. Survival was not affected by the studied treatments, but the presence of fusiform rust was found to be essential to determine the stand density trajectories for the evaluated stands. The parameters for the growth and yield system were simultaneously estimated using maximum likelihood and the variance-covariance was modeled within the system. The use of stochastic differential equations applied to these types of models in forestry was evaluated and summarized in the last part of this dissertation.

INDEX WORDS: basal area, bedding, differential equations, dominant height, growth and yield system, maximum likelihood, mortality, slash pine, stochastic, vegetation control.

A GROWTH AND YIELD SYSTEM FOR SLASH PINE INCLUDING RESPONSES TO SILVICULTURAL TREATMENTS

by

LAURA RAMIREZ QUINTERO

B.S., National University of Colombia, Colombia, 2016

M.S., National University of Colombia, Colombia, 2019

M.S., The University of Georgia, 2022

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

2023

© 2023

Laura Ramirez Quintero

All Rights Reserved

A GROWTH AND YIELD SYSTEM FOR SLASH PINE INCLUDING RESPONSES TO

SILVICULTURAL TREATMENTS

by

LAURA RAMIREZ QUINTERO

Major Professor: Cristian R. Montes
Committee: Bronson P. Bullock
Patrick C. Green
Ricardo M. Holdo

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
May 2023

DEDICATION

To my lovely husband, Leo, who accompanied me and supported me during all the ups and downs of this journey.

ACKNOWLEDGEMENTS

Many individuals and organizations contributed to the realization and finalization of this dissertation and doctoral program. I want to thank first my major professor, Cristian Montes, for first at all, convincing me to start this program at UGA when I thought there was no room from me in academia. I greatly appreciate his kindness and patience when explaining things to me, as well as his encouragement to always go further and use all my skills and strengths even when I did not know they were there. I also want to thank Dr. Bronson Bullock for always providing me with great perspective and helping me feel very welcomed as a Plantation Management Research Cooperative (PMRC) graduate student. I am also thankful for the insights and support provided by Dr. Corey Green and Dr. Ricardo Holdo who together with Dr. Montes and Dr. Bullock completed my doctoral committee.

I am immensely grateful to the PMRC. It was a great experience to be part of the PMRC as a graduate student and I really cannot think on how this could have been better. The opportunities and financial support provided by the PMRC really elevated my experience as a graduate student. I wish all other graduate students could land within such an incredible organization that cares about all the people involved and promotes the kind of personal and professional growth that I experienced here. Many thanks to all the organizations and individuals that support the PMRC and have helped all over these years to collect data and promote research in the region.

I also want to thank the Warnell School of Forestry and Natural Resources for providing me with financial support and also for promoting all sort of activities that made me engage with other students that today I consider great friends. Although my experience was a little different thanks to the global pandemic, I am also grateful to all my fellow graduate students with whom I crossed paths with. I will always remember the times when we shared our common problems and got together to share a meal or drink. Conference trips would not have been the same without them.

Finally, I want to thank all my family here and abroad, who provided support and encouragement and always were patient whenever I had to put school first. I want to specially thank my sister Karen with whom I have shared this experience from the very beginning and can relate to me on every single aspect of it. When things were overwhelming, at least we had each other to share a meal! And one more time, I want to thank my husband Leo for always listening to me. I know he knows more biometrics now than what he should know thanks to our discussions on why my analyses where not going as planned. Really, Thank you!

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1. INTRODUCTION

The forestry industry is one of the most important industries in the southeastern United States, it generates almost $25 billion dollars in revenue and employs more than 55,000 people for the state of Georgia alone, ranking second in employment after the food processing industry (GFC, 2021). This industry relies on forests (natural or planted) that are intensively managed to increase growth and improve product quality. Intensive forest management includes activities such as site preparation techniques to improve soil physical and chemical conditions, herbicides application to remove competing vegetation, fertilization to correct nutrient deficiencies, and midrotation treatments like thinings to control product size and quality (Fox et al., 2007; Jokela et al., 2010). These activities have been studied and tailored into silvicultural practices that are applied differentially depending on the tree species and site with the main goal of maximizing growth gains, which can justify the additional costs associated to these practices. It becomes then necessary to have growth and yield (G&Y) models/systems that can accurately predict gains from these silvicultural treatments and can therefore support managers when making decisions about forest management in the region.

In this dissertation, silvicultural treatments were studied and incorporated into growth and yield models for the second most important commercial species in the Southeastern United States, slash pine (*Pinus elliottii* Engelm.) (Barnett and Sheffield, 2004). The focus was on two silvicultural treatments, bedding and complete vegetation control of competing vegetation. These two treatments are usually applied during site preparation with bedding aiming to improve rooting, soil moisture and nitrogen availability (Morris and Lowery, 1988) and vegetation control aiming to release competition for the crop tree and maximize the resources available on the site. To evaluate the effect of these treatments on slash pine growth, a long-term study established in 1979 by the Plantation Management Research Cooperative (PMRC) of the University of Georgia was used. This study was comprised of 72 different plots installed in the Lower Coastal Plain of Georgia, Florida and South Carolina, in which slash pine seedlings were planted and different silvicultural treatments were applied during site preparation. Measurements of diameter, height and other variables of interest were taken every three years starting at age five, with most plots having observations until year 26 and a few plots having a last measurement at year 31.

Growth was modeled for the three main variables composing a stand growth and yield system, dominant height, mortality/survival, and basal area. The long-term effects of the mentioned treatments were first evaluated for dominant height and are presented in Chapter 2. In Chapter 3 a new dominant height model is presented using differential equations with the aim of coupling this model with a mortality model in which height increments instead of time increments were tested. In addition in this chapter, mortality was modeled including the effect of fusiform rust (*Cronartium quercuum* Berk.) infection on the mortality rate. Fusiform rust is a common pathogen affecting slash pine plantations and it has been associated with higher mortality rates before (Devine and Clutter, 1985; Nance et al., 1981). The final G&Y system is presented in Chapter 4 with the three mentioned variables of interest. This chapter was built upon Chapter 3, using the same dominant height and mortality models explained in Chapter 3, and incorporating the additional component of basal area. The final G&Y system proposed was an intercorrelated system of differential equations in which dominant height and mortality are independent, but the basal area model uses these two variables as predictors. Both dominant height and basal area models were modified to include the effect of bedding and vegetation control. The mortality model did not include these modifications after finding non-significant effect of these treatments on the survival of the trees evaluated.

When using differential equations to model forest growth as in Chapter 3 and Chapter 4, a natural approach to incorporate uncertainty into the models has been to use stochastic differential equations (SDEs) (Garcia, 1979). In the last part of this dissertation (Chapter 5), a critical review of the use of SDEs in forestry is presented with the intention of explaining its advantages and limitations when incorporating process and observation error into forestry G&Y systems. Examples are presented to clarify some concepts and a different approach, also using differential equations but without the stochastic framework, is proposed to incorporate uncertainty into G&Y systems.

1.2. LITERATURE REVIEW

Slash pine (*Pinus elliottii* Engelm.) is the second most important commercial species in the southeastern United States (Barnett and Sheffield, 2004), it outperforms other commercial species when established on its natural habitat where poorly drained flatwoods are predominant (Barnett and Sheffield, 2004). As it is common in this region of the United States, slash pine plantations are accompanied by prescribed silvicultural treatments like bedding, vegetation control, and fertilization, which together aim to increase productivity over time (Fox et al., 2007; Jokela et al., 2010). Responses to silvicultural treatments have been studied for this species, reporting growth gains even at mature ages (Zhao et al., 2009). These gains can be incorporated into growth models in different ways. These strategies can usually be grouped in one of the following streams: 1. Look-up tables (e.g., Montes, 2001; Logan and Shiver, 2006). 2. The age-shift method (e.g., South *et al.*, 2006; Carlson *et al.*, 2008). 3. Separate equations for untreated and treated plots (e.g., Pienaar and Rheney, 1995). 4. A single equation that incorporates a multiplier function describing the treatment effect (e.g., Hynynen, 1995). 5. Modifications of the model parameters to account for treatment effect (e.g., Mason and Milne, 1999).

The most common approach to incorporate gains from silvicultural treatments into growth models has been to add a response function to a base model, as proposed by Pienaar and Rheney (1995). Nevertheless, when following this approach, the variability observed on the different plots is passed to the response function, ignoring the site-specific factors that interact with the treatments applied (Fang and Bailey, 2001). To avoid this issue, both the control and the treatment functions must be fitted simultaneously as Hynynen et al. (1998) proposed, or a relative response can be used, as proposed by Scolforo et al. (2020). Another approach, which avoids these issues and where few assumptions are required regarding the expected response form, is to modify the parameters of the base model according to the treatment applied (e.g., Mason and Milne, 1999; Salas, Stage and Robinson, 2008; Gyawali and Burkhart, 2015). This approach was tested and used in this dissertation to incorporate the effects of bedding and vegetation control on dominant hegith and basal area growth for slash pine.

Although some authors have found for slash pine that bedding is associated with a lower mortality rate when stands are located in poorly drained soils (Gent et al., 1986; Pritchett, 1979) and higher survival rates have been reported after vegetation control applications in some specific sites (Creighton et al., 1987), these silvicultural treatments are rarely mentioned when evaluating mortality in slash pine plantations. The effect of these treatments on mortality rates was evaluated as part of this dissertation. Nevertheless, the most important factor when assessing survival/mortality in slash pine plantations is the presence of fusiform rust (*Cronartium quercuum*

Berk.) infection. This pathogen has been associated with higher mortality rates in slash pine plantations and its absence/presence and magnitude of the infection are useful predictors when modeling mortality in slash pine (Bailey and Burgan, 1989; Devine and Clutter, 1985; Nance et al., 1981). Although infection by *C. quercumm* does not directly cause tree death, it produces cankers located in the stem or branches of the tree, which debilitate the tree and increase the probability of the tree dying compared to healthy trees (Jones, 1972; Sluder, 1977). The inclusion of average fusiform rust infection rates as a predictor of mortality was then tested for the slash pine stands evaluated in this research.

Growth and yield (G&Y) systems for forest stands usually include three state variables, dominant height, survival/mortality, and basal area. Differential equations are commonly used to define these systems, usually as an interdependent system of equations where some variables appear in both the left and the right-hand side of the equations. These systems require estimation techniques in which parameters are estimated simultaneously to get unbiased estimates of the parameters of the system (Borders and Bailey, 1986; Goelz and Burk, 1996). For slash pine, some authors have developed G&Y systems including silvicultural treatments effects, although using independent models for each component (Bailey and Burgan, 1989; McTague, 2009; Pienaar and Rheney, 1995), and when using simultaneous estimation, the focus has been on the baseline treatments, which do not consider treatment response (Borders and Bailey, 1986; Gallagher et al., 2019; Murphy and Sternitzke, 1979; Pienaar and Harrison, 1989; Sullivan and Clutter, 1972). Fewer authors have included both response to silvicultural treatments and simultaneous estimation (Fang et al., 2001; Martin et al., 1999). In this dissertation, a G&Y system for slash pine plantations using a system of differential equations where silvicultural treatments were included, and simultaneous estimation was used is presented.

Ideally, G&Y systems should be able to incorporate fluctuations generated by either random environmental factors or by errors during the measurements and sampling procedure (Sandland and McGilchrist, 1979). Different approaches have been proposed in the forestry literature to incorporate this variation. The most common approach, adapted to forestry systems from the econometry literature, is to model the variance by defining different variance/covariance matrices according to the expected error and correlations in the data. These matrices are then used in techniques known as two or three least squares (2SLS-3SLS) (Borders and Bailey, 1986; Pienaar and Harrison, 1989), or mixed effect models (Fang et al., 2001; Gallagher et al., 2019). LeMay (1990), for example, proposed an approach to fit simultaneous, contemporaneously correlated systems of equations with both serial correlated and heteroscedastic error terms called multistage least squares (MSLS).

Another approach to incorporate uncertainty into G&Y systems is the one proposed by Garcia (1979). Garcia proposed the use of stochastic differential equations (SDEs) to model growth in forestry. In this approach random variations are added as Wiener or Brownian motion process. Garcia proposed this application in which the process error (generated by stochasticity in the growth process itself) is separated from the observation error, attributed to measurements and sampling errors only. Although Garcia has modified and extensively worked on simplifying his proposed approach, and other authors have followed this approach (Donnet et al., 2010; Orrego et al., 2021; Rupšys, 2019; Zhang and Borders, 2001), SDEs are not commonly used in forestry as they are perceived as complex and hard to implement (Burkhart and Gregoire, 1994). The last portion of this dissertation was designed to provide a general context on the SDEs and explain the advantages and limitations of this approach also when compared with a different and simpler approach proposed here.

1.3. Objectives

1.3.1. Main objective

The main objective of this dissertation is to provide a growth and yield system for slash pine plantations in which silvicultural treatments are incorporated. The specific objectives of each one of the chapters are presented below.

1.3.2. Specific objectives

Chapter 2:

a) Characterize the long-term effects of bedding and vegetation control on the dominant height of slash pine plantations.

b) Construct a dominant height and site index model that could account for treatment effects and their interaction.

Chapter 3:

a) Develop a survival/mortality model for slash pine plantations including the effect of silvicultural treatments and the effects of fusiform rust infection on the mortality rate.

b) Determine if mortality models developed from differential equations in which height increments instead of time increments are used, are more accurate to describe mortality in slash pine.

Chapter 4:

a) Incorporate the effects of bedding and vegetation control into a growth and yield system for slash pine composed of dominant height, mortality, and basal area.

Chapter 5:

a) Explain SDEs when incorporating uncertainty into a growth and yield system in forestry.

b) Compare the use of SDEs to the alternative option of modeling variance directly when using maximum likelihood.

1.4. References

Bailey, R.L., Burgan, T.M., 1989. Fertilized midrotation-aged slash pine plantations-stand structure and yield prediction models. South. J. Appl. For. 13, 76–80.

Barnett, J.P., Sheffield, R.M., 2004. Slash pine: characteristics, history, status and trends. In: Dickens, E.D.; Barnett, J.P.; Hubbard, W.G.; Jokela, E.L., eds. Slash pine: still growing and growing! Proc. slash pine Symp. Gen. Tech. Rep. SRS-76. Asheville, NC U.S. Dep. Agric. For. Serv. South. Res. Station. 1–6.

Borders, B.E., Bailey, R.L., 1986. A compatible system of growth and yield equations for slash pine fitted with restricted three-stage least squares. For. Sci. 32, 185–201. https://doi.org/10.1093/forestscience/32.1.185

Burkhart, H.E., Gregoire, T.G., 1994. 11 Forest biometrics. Handb. Stat. 12, 377–407. https://doi.org/10.1016/S0169-7161(05)80013-6

Carlson, C.A., Fox, T.R., Allen, L.H., Albaugh, T.J., 2008. Modeling mid-rotation fertilizer responses using the age-shift approach.pdf. For. Ecol. Manage. 256, 256–262.

Creighton, J.L., Zutter, B.R., Glover, G.R., Gjerstad, D.H., 1987. Planted pine growth and survival responses to herbaceous vegetation control, treatment duration, and herbicide application technique. South. J. Appl. For. 11, 223–227. https://doi.org/10.1093/sjaf/11.4.223

Devine, O.J., Clutter, J.L., 1985. Prediction of survival in slash pine plantations infected with fusiform rust. For. Sci. 31, 88–94.

Donnet, S., Foulley, J.L., Samson, A., 2010. Bayesian analysis of growth curves using mixed models defined by stochastic differential equations. Biometrics 66, 733–741. https://doi.org/10.1111/j.1541-0420.2009.01342.x

Fang, Z., Bailey, R.L., 2001. Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. For. Sci. 47, 287–300. https://doi.org/10.1093/forestscience/47.3.287

Fang, Z., Bailey, R.L., Shiver, B.D., 2001. A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects. For. Sci. 47, 550–562. https://doi.org/10.1093/forestscience/47.4.550

Fox, T.R., Jokela, E.J., Allen, H.L., 2007. The development of pine plantation silviculture in the Southern United States. J. For. 105, 337–347. https://doi.org/10.1093/jof/105.7.337

Gallagher, D.A., Bullock, B.P., Montes, C.R., Kane, M.B., 2019. Whole stand volume and green weight equations for loblolly pine in the Western Gulf Region of the United States through Age 15. For. Sci. 65, 420–428. https://doi.org/10.1093/forsci/fxy068

Garcia, O., 1979. Modelling stand development with stochastic differential equations. Mensurat. Manag. Plan. Exot. For. Plant. New Zeal. For. Serv. FRI Symp. 315–333.

Gent, J.A., Allen, H.L., Campbell, R.G., Wells, C.G., 1986. Magnitude, duration, and economic analysis of loblolly pine growth response following bedding and phosphorus fertilization. South. J. Appl. For. 10, 124–128. https://doi.org/10.1093/sjaf/10.3.124

GFC, 2021. Economic Benefits of the Forest Industry in Georgia: 2021, Georgia Forestry Commission.

Goelz, J.C.., Burk, T.., 1996. Measurement error causes bias in site index equations. Can. J. For. Res. 26, 1585–1593. https://doi.org/10.16309/j.cnki.issn.1007-1776.2003.03.004

Gyawali, N., Burkhart, H.E., 2015. General response functions to silvicultural treatments in loblolly pine plantations. Can. J. For. Res. 45, 252–265. https://doi.org/10.1139/cjfr-2014-0172

Hynynen, J., 1995. Predicting the growth response to thinning for Scots pine stands using individual-tree growth models. Silva Fenn. https://doi.org/10.14214/sf.a9210

Hynynen, J., Burkhart, H.E., Allen, H.L., 1998. Modeling tree growth in fertilized midrotation loblolly pine plantations. For. Ecol. Manage. 107, 213–229. https://doi.org/10.1016/S0378-1127(97)00334-4

Jokela, E.J., Martin, T.A., Vogel, J.G., 2010. Twenty-five years of intensive forest management with southern pines: Important lessons learned. J. For. 108, 338–347. https://doi.org/10.1093/jof/108.7.338

Jones, J.E.P., 1972. Fusiform rust affects planted slash pine. J. For. 70, 350–352.

LeMay, V.M., 1990. MSLS: a linear least squares technique for fitting a simultaneous system of equations with a generalized error structure. Can. J. For. Res. 20, 1830–1839. https://doi.org/10.1139/x90-246

Logan, S.R., Shiver, B.D., 2006. Adjusting Slash Pine Growth and Yield for Silvicultural Treatments. Gen. Tech. Rep. SRS-92. Asheville, NC U.S. Dep. Agric. For. Serv. South. Res. Stn. 328–332.

Martin, S.W., Bailey, R.L., Jokela, E.J., 1999. Growth and yield predictions for lower coastal plain slash pine plantations fertilized at mid-rotation. South. J. Appl. For. 23, 39–45. https://doi.org/10.1093/sjaf/23.1.39

Mason, E.G., Milne, P.G., 1999. Effects of weed control, fertilization, and soil cultivation on the growth of Pinus radiata at midrotation in Canterbury, New Zealand. Can. J. For. Res. 29, 985–992. https://doi.org/10.1139/x99-087

McTague, J.P., 2009. Modelling the Response of Loblolly Pine to Juvenile Fertilization. Open For. Sci. J. 1, 80–88. https://doi.org/10.2174/1874398600801010080

Montes, C.R., 2001. A Silvicultural Decision Support System for Loblolly Pine Plantations. North Carolina State Unviersity.

Morris, L.A., Lowery, R.F., 1988. Influence of site preparation on soil conditions affecting stand establishment and tree growth. South. J. Appl. For. 12, 170–178.

Murphy, P.A., Sternitzke, H.S., 1979. Growth and Yield Estimation for Loblolly Pine in the West Gulf. Res. Pap. SO-154. New Orleans, LA U.S. Dep. Agric. For. Serv. South. For. Exp. Station. 12 p. 154.

Nance, W.L., Froelich, R.C., Shoulders, E., 1981. Effects of fusiform rust on survival and structure of Mississippi and Louisiana slash pine plantations. US Dep. Agric. For. Serv. South. For. Exp. Stn. Vol. 232.

Orrego, S., Montes, C., Restrepo, H.I., Bullock, B.P., Zapata, M., 2021. Modeling height growth for teak plantations in Colombia using the reducible stochastic differential equation approach. J. For. Res. 32, 1035–1045. https://doi.org/10.1007/s11676-020-01174-y

Pienaar, L., Rheney, J., 1995. Modeling stand level growth and yield response to silvicultural treatments. For. Sci. 41, 629–638. https://doi.org/10.1093/forestscience/41.3.629

Pienaar, L. V., Harrison, W.M., 1989. Simultaneous growth and yield prediction equations for pinus elliottii plantations in zululand. South African For. J. 149, 48–53. https://doi.org/10.1080/00382167.1989.9628992

Pritchett, W.L., 1979. Site preparation and fertilization of slash pine on a wet savanna soil. South. J. Appl. For. 3, 86–90.

Rupšys, P., 2019. Modeling dynamics of structural components of forest stands based on trivariate stochastic differential equation. Forests 10. https://doi.org/10.3390/f10060506

Salas, C., Stage, A.R., Robinson, A.P., 2008. Modeling Effects of Overstory Density and Competing Vegetation on Tree Height GRowth. For. Sci. 54, 107–122.

Sandland, R.L., McGilchrist, C.A., 1979. Stochastic Growth Curve Analysis Author. Biometrics 35, 255–271.

Scolforo, H.F., Montes, C., Cook, R.L., Lee Allen, H., Albaugh, T.J., Rubilar, R., Campoe, O., 2020. A New Approach for Modeling Volume Response from Mid-Rotation Fertilization of Pinus taeda L. Plantations. Forests 11, 646. https://doi.org/10.3390/f11060646

Sluder, E.R., 1977. Fusiform rust in loblolly and slash pine plantations on high-hazard sites in Georgia. For. Serv. US Dep. Agric. Southeast. For. Exp. Stn. Vol. 160.

South, D.B., Miller, J.H., Kimberley, M.O., Vanderschaaf, C.L., 2006. Determining productivity gains from herbaceous vegetation management with “age-shift” calculations. Forestry 79, 43–56. https://doi.org/10.1093/forestry/cpi058

Sullivan, A.D., Clutter, J.L., 1972. A simultaneous growth and yield model for loblolly pine. For. Sci. 18, 76–86.

Zhang, Y., Borders, B.E., 2001. An iterative state-space growth and yield modeling approach for unthinned loblolly pine plantations. For. Ecol. Manage. 146, 89–98. https://doi.org/10.1016/S0378-1127(00)00448-5

Zhao, D., Kane, M., Borders, B., Harrison, M., 2009. Long-term effects of site preparation treatments, complete competition control, and repeated fertilization on growth of slash pine plantations in the flatwoods of the southeastern United States. For. Sci. 55, 403–410. https://doi.org/10.1093/forestscience/55.5.403

CHAPTER 2

LONG-TERM TERM EFFECT OF BEDDING AND VEGETATION CONTROL ON DOMINANT HEIGHT OF SLASH PINE PLANTATIONS IN THE SOUTHEASTERN UNITED STATES[1]

[1] Ramirez, L., Montes, C. R., & Bullock, B. P. 2022. *Forest Ecology and Management*. 522:120479. Reprinted here with permission of publisher.

Abstract

The long-term effect of bedding and vegetation control on dominant height in slash pine (*Pinus elliottii* Engelm) was evaluated using data from a site preparation study established in 1979 by the Plantation Management Research Cooperative (PMRC) at the University of Georgia in the southeastern United States. The experimental design corresponded to a 2 x 2 factorial with replications over 16 different locations, distributed over the natural range of slash pine. Our results show sustained gains in dominant height, reaching a peak increment around age 11, with values of 1.0, 2.2, and 2.9 m of average gain for the bedding, vegetation control, and combined (Bed + Veg.) treatments, respectively. At age 31, an approximate rotation age, these gains were no longer present for the bedding treatment, whose dominant height trajectory converged to the values of the untreated control and decreased to 1.9 m for both treatments involving competing vegetation control. These results are similar to previously reported results in the literature for these two treatments in slash pine. We proposed a modified Chapman-Richards type model to describe these trends. In this modeling approach, the base equation was modified using a set of dummy variables in the form of power functions to reflect the treatment effect. Both treated and untreated plots were simultaneously fitted in this model, and contrarily to the most common approach of adding an independent factor to a base model to account for the treatment response, our model does not assume the control plot to be error free. The flexibility of the proposed model allows practitioners to include observed gains in dominant height from these treatments. A slash pine site index model using the algebraic difference approach (ADA) was also derived.

2.1. Introduction

Slash pine (*Pinus elliottii* Engelm.) is the second most important commercial species in the southern United States (Barnett and Sheffield, 2004). On its natural habitat, characterized by poorly drained flatwoods and seasonally flooded areas, it outperforms other common commercial pine species, producing high quality timber that encompasses a large portion of the regional timber market (Barnett and Sheffield, 2004). Traditionally in this region of the United States, intensive silvicultural management has been prescribed for commercial forest plantations with the aim of increasing resources available to the crop trees and reducing competition as well as to increase end product value (Fang and Bailey, 2001; Martin et al., 1999). Common silvicultural treatments include bedding, herbaceous and/or woody vegetation control at establishment, fertilization with nitrogen and phosphorus, and thinning (Fox et al., 2007; Jokela et al., 2010). These treatments are long justified by different studies showing significant responses in height, basal area, and volume when they are applied at a juvenile stage (Colbert et al., 1990; Jokela et al., 2000; Zhao et al., 2008), with gains visible at mature ages (Fang and Bailey, 2001; Jokela et al., 2010; Zhao et al., 2009). With the increasing interest to maximize carbon capture, investigating whether these gains are maintained or reduced, close to or beyond traditional harvest ages for the species (25 – 30 years), becomes a question of interest, and one that only few studies can answer.

Snowdon and Waring (1984, 1981) studied the nature of silvicultural responses at early stages in other pine species (radiata pine). Their work provided the foundations to characterize growth rates of forest plantations after silvicultural treatments were applied, classifying them into two broad categories that link response to short- or long-term resource availability. Type I responses are the product of those treatments which temporarily increase the growth rate of the stand but do not have a sustained effect on site properties (e.g., nitrogen fertilization, weed control), while type II responses are the result of treatments such as phosphorus fertilization, or continuous nitrogen fertilization, which can generate a sustained change in site productivity (Snowdon, 2002). Other classifications in the forestry literature refer to the type I response as type B, and to the response type II as type A, and a third classification called type C is usually used to described treatments such as bedding, which generate an early gain that dissipates with time (Hughes et al., 1979; Morris, 1988).

Analysis of variance (ANOVA) and modeling, are two common approaches employed when analyzing and defining the observed response type, with modeling being a more comprehensive method that has become an essential tool to also determine the economic feasibility of silvicultural treatments (Snowdon, 2002). To include the effect of silvicultural treatments as part of growth and yield systems, practitioners have tried different strategies. These strategies can usually be grouped in one of the following streams: 1. Look-up tables (e.g., Montes, 2001; Logan and Shiver, 2006). 2. The age-shift method (e.g., South et al., 2006; Carlson et al., 2008). 3. Separate equations for untreated and treated plots (e.g., Pienaar and Rheney, 1995). 4. A single equation that incorporates a multiplier function describing the treatment effect (e.g., Hynynen, 1995). 5. Modifications of the model parameters to account for treatment effect (e.g., Mason and Milne, 1999).

Among these strategies, the one proposed by Pienaar and Rheney (1995) has been frequently used due to its ease of implementation (e.g., Mason and Milne, 1999; Quicke, Glover and Glover, 1999; Amateis et al., 2000). With this method, a response to a given treatment is first characterized as the cumulative difference between a control and a treated subject and later added to a baseline model that calculates gains at a stand or plot level. Nevertheless, this approach ignores the inherent variability in state variables (basal area, dominant height and stand density) between different plots of a given stand, assigning the same response based on the treatment applied, without any consideration of how this response would vary depending on specific site attributes (Fang and Bailey, 2001). Including treatments response in this way, implies that the variability in the control plots used to build the response function is passed to the response factor. The same issue can be present if a treatment modifier is added as a multiplicative factor to a base model and treated and untreated plots are fitted independently as in Gyawali and Burkhart (2015). Nevertheless, if both the base and treated plots are modeled simultaneously, as in Hynynen et al. (1998), or a relative response is used as in Scolforo et al. (2020), this problem can be overcome.

Although the other mentioned approaches have proven to be effective to model treatment response, they usually rely on assumptions about the response type and the interaction between treatments. This is the case of the age-shift approach, where it is assumed that the shape of the growth curve does not change with the inclusion of silvicultural treatments (South et al., 2006), being useful only when Type B responses are assumed. Look up tables as in Logan & Shiver (2006) have been used as a way to modify the response according to the base site index (i.e., dominant height at base age for the control treatment), although a linear relationship between site index (SI) and the expected gain was assumed by these authors. On the other hand, few assumptions regarding the treatment effect are necessary when treatment responses are modeled by including variation factors directly on the parameters of the original model (e.g., Mason and Milne, 1999; Salas, Stage and Robinson, 2008; Gyawali and Burkhart, 2015). When following this approach, both treated and untreated plots are modeled simultaneously, avoiding the assumption of an error free control plot. Therefore, a response type does not need to be assumed given that the estimated changes in the parameters account for the response trend.

One important aspect that must be considered when deriving site index equations from dominant height models that include responses to site preparation silvicultural treatments, is that the expected response (or gain) is influenced by the base SI. This relationship has been confirmed across several commercial species and silvicultural treatments (Fang and Bailey, 2001; Logan and Shiver, 2006; Zhao et al., 2016). Therefore, to accurately generate SI curves, the gain in dominant height as a function of base SI should be included so that accurate predictions are made for a given site.

Logan & Shiver (2006) acknowledged this relationship and proposed a response function to adjusts dominant height curves by using values of expected gain for different site indexes and silvicultural treatments. Nevertheless, these authors assumed an arbitrary linear relationship between base SI and gain for all the treatments. For the SI equations developed in the present research, the relationship between base SI and gain in dominant height was hypothesized to be inverse (decrease in gain with increasing SI), but non-linear.

We hypothesize that by modifying the parameters of a dominant height model according to the treatment applied, a flexible model will be generated such that the long-term response to treatments is accurately captured. Thus, the objectives of this research were to (i) characterize the long-term effects of bedding and vegetation control on the dominant height of slash pine plantations, and (ii) construct a dominant height and site index model that could account for treatment effects and their interaction.

2.2. METHODS

2.2.1. DATA

For this research we used a long-term slash pine study established in 1979 by the Plantation Management Research Cooperative (PMRC) at the University of Georgia in the southeastern United States. The study's main objective was to evaluate differences in growth response to site preparation silvicultural treatments. Treatments for this study included burning, chopping, bedding, competing vegetation control, and fertilization. Mid-rotation treatments, including thinnings, were not carried out on the study plots. The study layout comprised 20 installations across Georgia, Florida, and South Carolina, stratified equally over Spodosols and non-Spodosols.

From those installations, only 14 of them with measurements up to year 31 remain active. In addition to these installations, 2 more installations that are no longer active but were measured up to year 26, were included for this research. Figure 2-1 shows the distribution of the 16 installations used in this research.

From the mentioned study, a subset of the original treatments was taken to form a 2 x 2 factorial design with bedding and complete vegetation control as main treatments for a complete randomized factorial design. The vegetation control treatment refers to herbicide application targeting all competing vegetation until crown closure (Zhao et al., 2007). Chopping and burning were considered to be the operational site preparation treatments at the time of the trial installation, and plots receiving them were taken as the control plots for this research. Plot size corresponded to 0.2 ha, with a measurement plot of 0.08 ha. At the time of planting, seedlings were double planted to ensure an approximately homogeneous initial planting density. Double planted slots had the smaller seedling removed after one year. Measurements of diameter at breast height (DBH) for all the trees, and total height for a portion of the trees, were taken every 3 years from age 5 to age 31, with the latest measurement taken in 2010. No major mortality events, rather than natural mortality, were observed during the measurement periods. For more details on the study design, see Zhao et al., (2009) and Zhao et al., (2007).

Only a fraction of the tree heights was measured on every plot, therefore a DBH-height model was fitted for each plot at every measured age to estimate the remaining heights and then determine the average dominant height. The height model was:

$$H_{tot_{ijk}} - 1.4 = \beta_{0_{ij}}\, exp\left(\frac{\beta_{1_{ij}}}{DBH_{ijk}}\right)$$

Where $H_{tot_{ijk}}$ is the total height for the k^{th} tree in the i^{th} plot at age j. $\beta_{0_{ij}}$ and $\beta_{1_{ij}}$ are parameters to be estimated for each plot at every age, and DBH_{ijk} is the diameter at 1.4 m for the k^{th} tree in the plot i at age j. Once the total height was estimated for all the trees, dominant height was calculated as the per plot average height of the dominant and co-dominant trees (i.e., trees with DBH greater than the quadratic mean diameter). A summary of the dominant height values per treatment is presented in Table 2-1. Additional stand characteristics are presented in Table 2-2 for a subset of the measurement periods.

2.2.2. Dominant height model

To model the response to silvicultural treatments, a reparametrized Chapman-Richards (CR) type model was selected, including dummy variables that allow changes over the asymptote and the slope parameters depending on the treatment being applied. This model was compared with a model fitted following the Pienaar and Rheney (1995) modelling approach (PR). These authors developed a height growth and a basal area model including responses to silvicultural treatments by adding an independent treatment factor to base models for each variable. The two models compared are summarized below.

$DH_t = a_0 b_1^{z_1} b_2^{z_2} \times \left(1 - exp\left(-a_1 b_3^{z_1} b_4^{z_2}\ t\right)\right)^{a_2}$	(CR)	This paper
$DH_t = a_0(1 - exp(-a_1 t))^{a_2} + (b_0 Z_1 + b_1 Z_2) t e^{-b_2 YST}$	(PR)	(Pienaar and Rheney, 1995)

Where DH_t is dominant height (in meters), at age t (in years), YST is years since the treatments were applied, which is equivalent to t since all the treatments were applied at the establishment phase, $a_0, .., a_2$ and $b_0, .. b_4$, are parameters to be estimated, Z_1 is a dummy variable equal to 1 if bedding was applied and zero otherwise, and Z_2 is also a dummy variable that equals 1 if vegetation control was applied, or zero otherwise.

The models were fitted using non-linear least squares using the software R (R Core Team, 2018). The models' performance was evaluated using the adjusted coefficient of determination for non-linear models (R^2_{adj}), Root Mean Square Error (RMSE), and Akaike's information criterion (AIC), calculated as follows:

$$R^2_{adj} = 1 - \frac{(n-1)\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2}{(n-p)\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2}{n-p}}$$

$$AIC = n \ln \hat{\sigma}^2 + 2(p+1)$$

Where n is the total number of observations, p is the number of parameters in each model, Y_i is the observed dominant height, $\widehat{Y}_i$ is the estimated dominant height, and $\hat{\sigma}^2$ is the estimated mean square error of the model, calculated as follows:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2}{n}$$

A k-fold cross-validation with 5 folds was carried out to evaluate the predictive performance of the models. This procedure consisted of removing one fifth of the data points (chosen randomly) and fitting the proposed models with the remaining data. Afterwards, an estimation of the points previously excluded was made with the fitted model and the RMSE was calculated. The procedure was repeated five times using random plots selected without replacement and the RMSEs found for the five iterations were averaged.

2.2.3. SITE INDEX MODEL

Site index (SI) equations can be derived from dominant height models following the procedure highlighted by Bailey and Clutter (1974), commonly known as the algebraic difference equation approach (ADA). This method replaces one of the parameters in a yield equation by a local parameter, under the assumption that it will be related to the stand SI. Depending on which parameter is selected as local, either anamorphic or polymorphic SI curves can be generated. When more than one parameter is related to SI, a generalized difference equation approach (GADA) proposed by Cieszewski and Bailey (2000) is often followed. These methods have been widely used to derive SI equations for commercial plantations (Diéguez-Aranda et al., 2006, 2005).

SI models seldom include responses to silvicultural treatments. Some authors (e.g., Antón-Fernández et al., 2011; Sharma et al., 2002; Tymińska-Czabańska et al., 2022), have developed SI models that are sensitive to stand density, nevertheless, silvicultural treatments such as bedding or vegetation control are not frequently included in SI models. Both of the models presented could be used to derive a SI model which includes responses to silvicultural treatments by relating parameters a_0 or $a_1$to SI and deriving a dynamic equation. These formulations are presented in Table 2-3.

The dominant height model that presented the best fit (from section 2.2.2) was selected to derive the SI model by estimating locally the parameters $a_0$or a_1 using the dummy variable approach. The initial values required for the non-linear least squares' optimization were selected from the global values estimated for the dominant height model in section 2.2.2. The best SI model was selected using the R^2_{adj} and $RMSE$ metrics.

The SI model required to be modified to properly express gains in SI. We hypothesized that gains in SI are inversely related to the base SI. The modification consisted of adding a gain function that models this relationship. The gain function was constructed by fitting the dominant height model individually for each one of the 16 installations, calculating the SI for the control treatment (SI_{25}), and then calculating the gain in SI due to the treatments (G_{HDOM_i}). Different models were fitted to these values including a linear, a logarithm and an exponential function (Table 2-4). The model with higher coefficient of determination (R^2) was selected to be combined with the SI model.

The gain function was added to the SI model by replacing the value of DH_{t_0} in the dynamic equations in Table 2-3 by $\left(DH_{t_0} + G_{HDOM_i}\right)$. Adding this factor guarantees that dominant height predicted at the base age reflects treatments gains. This gain function is required regardless of the used approach (ADA anamorphic or polymorphic, or even GADA). Without this additional function, all the dominant height curves would yield the same dominant height at the point where $t = 25$ (base age used for slash pine). This would be incorrect since some treatments do show gains in dominant height at age 25 and the fact that the models predict the same dominant height at the base age is an artifice of the model generated by the way the dynamic equations are derived. This was evident in the work of Socha et al., (2021) who developed a dominant height model for Scot pine in Poland using the GADA approach. In their work, parameters were modified to account for regional differences, and the described issue is evident when plotting the dominant height trajectories for the different regions.

2.3. RESULTS

Bedding and vegetation control had a positive effect on dominant height through age 11. The average dominant height for all the treatments, as well as the average gain is shown in Figure 2-2. Vegetation control had a stronger effect compared to bedding alone, whereas the combined treatment generated the highest response of all the treatments. At age 31, the effect of bedding was no longer visible, converging to the control treatment. On the other hand, the vegetation control still had a positive effect on dominant height at this age. At age 31, stands where vegetation control was applied (with or without bedding) had a dominant height 1.9 m higher than those on the control treatment for the average condition.

From the analyzed data, the effect of bedding on dominant height can be categorized as a Type C response (initial increase, then decreasing over time), with an initial increase in dominant height compared to the control treatment, but a decrease to zero at age 31. The vegetation control treatment generated a type B response (initial increase, then sustained gain over time), attaining a maximum at age ~11 and remaining relatively stable until age 31. When combined with bedding, a type C response was observed, with values decreasing after reaching the maximum response. The combination of bedding and vegetation control showed no strong interaction. This was expected since the two treatments were targeting different resources on the site. While bedding improves soil physical conditions and increases runoff in poorly drained sites, providing an elevated environment out of saturated conditions for the seedlings (Morris, 1988), the vegetation control treatment reduces the loss of resources to competing vegetation and produces a major allocation of these resources to the crop-trees (Allen et al., 1990). In general, an additive response was observed. Major differences between the vegetation control and the combined treatment were observed at younger ages (<15 years), where the bedding effect was greater. At age 31, these two treatments generated a very similar response due to the almost null effect of bedding at this late age in the rotation.

2.3.1. Dominant height models

The estimated parameters for the CR and the PR model are presented in Table 2-5. The CR model was modified to include a variance stabilization parameter (β) to correct for heteroscedasticity. Weighted regression was used for estimating the parameters with weights equivalent to the inverse of the variance ($1/\sigma^{\beta}$). The statistics used to evaluate the models' performance are presented in Table 2-6. A very similar performance was found between the two models, with the same R^2_{adj}, and very similar values of RMSE. Both models also performed similarly when evaluating their performance using cross-validation. The AIC was the only criteria where there was a bigger difference between the models, favoring the PR model, most likely for having two less parameters compared to the CR model.

Even though both models had similar precision (Table 2-6) , when evaluating them by comparing the predicted and observed gain, the differences between the models become more apparent (Figure 2-3). The CR model more accurately predicts the average gain observed when bedding is applied, whereas the PR model underestimates the gain at early ages (<20 years) and overestimates the gain at later ages (>20 years). Nevertheless, when underpredicting, the difference between the average gain and the predicted gain for this model were not greater than 0.3 m, and when overpredicting, the maximum difference observed (at age 31) was less than 0.5 m.

The ability of a model to accurately predict gains from a given treatment depends on how accurately it predicts dominant height for both the control and the respective treatment. This can be better seen in Figure 2-4 for the CR model and Figure 2-5 for the PR model. In fact, both models predict with a low error (less than 0.2 m) the average dominant height (red line) for the control treatment up to age 26 but overestimate this value when approaching to age 31. Both models also accurately predict dominant height for the vegetation control treatment. The biggest differences observed in Figure 2-3 are the result of combining the errors for the control and the treatment predictions. These differences are magnified when the control is underestimated but the gain is overestimated. For example, if the dominant height is underestimated for the control, by 0.2 m, and the dominant height for the treatment is overestimated by the same 0.2 m, the gain would be overestimated by 0.4 m, which is what happened with the CR model when evaluating gains for the vegetation control treatment.

When comparing both models with the observed average dominant height (red lines in Figure 2-4 and Figure 2-5), the differences between them become less apparent, and both show good predictions especially over the range of 5-25 years. The variation in the observed dominant height for the different plots of the study (grey lines in Figure 2-4 and Figure 2-5) is hypothesized to be a consequence of the different local conditions. The inclusion of site index into the dominant height model was then tested in an attempt to include this variability into the dominant height model. The CR model form was chosen to test this hypothesis and further construct a dynamic dominant height model that allowed derivation of the site index model. The residual error distribution of this model is presented in Figure 2-6 and additional diagnostics plots for the CR model are presented in Figure 2-7. The predicted vs. observed plot shows how the original data is distributed equally along the 1 to 1 line. Although there is higher dispersion for the higher values of dominant height, there are not obvious patterns of underestimation or overestimation. The normal Q-Q plot shows slight deviation from the normal distribution, especially, in the tails. Nevertheless, since the main purpose of this model was not to do inference, these deviations were not considered a significant pitfall of the model.

2.3.2. Site Index model

To generate the SI model, the two dynamic equations presented in Table 2-3 for the CR model were fitted using local parameters (either a_0 or a_1) per installation. When a_0 was related to site index and fitted locally for each installation, the RMSE was 0.955 m, while when the growth rate (a_1) was related to site index, the RMSE was 1.123 m. Thus, the anamorphic dynamic equation using a_0 as the parameter related to site was used to generate the SI model by assuming $t_0 = 25$ and $DH_{t_0} = SI_{25}$, as follows:

$$DH_t = SI_{25}\left[\frac{1 - exp\left(-a_1 b_3^{z_1} b_4^{z_2}\, t\right)}{1 - exp\left(-a_1 b_3^{z_1} b_4^{z_2}\, (25)\right)}\right]^{a_2}$$

The estimated parameters for this model are presented in Table 2-7. (Only the global parameters are presented).

After exploring different models (Table 2-4), it was found that for the treatments involving chemical vegetation control, a linear model can be justified to explain the relationship between the base SI and the gain in expressed SI. The logarithm model (model 2 in Table 2-4) performed similarly to the linear model, but the latter was preferred for being more parsimonious. For the bedding treatment, parameters of the tested models were not significantly different from zero, meaning that there is not a significant change in expressed SI due to this treatment, therefore, no modification was needed for the SI equations including bedding. Consequently, for the treatments with vegetation control, the same model can be used regardless of being combined with bedding or not. The SI model can be then modified as follows:

$$HD_t = [SI_{25} + (\beta_0 + \beta_1\, SI_{25})]\left[\frac{1 - exp\left(-a_1 b_3^{z_1} b_4^{z_2}\, t\right)}{1 - exp\left(-a_1 b_3^{z_1} b_4^{z_2}\, (25)\right)}\right]^{a_2}$$

With $\hat{\beta}_0 = 9.38$ and $\hat{\beta}_1 = -0.39$.

The gain function added to the SI model ensures that dominant height estimated at the base age of 25 years reflects the gain in dominant height due to the treatment applied. Without this modification, all the dominant height curves would converge to the base SI at age 25, without reflecting the actual gain generated by the treatments (even when parameter b_3 and b_4 modify the function). The gain model was constructed combining the information of the vegetation control and the combined treatment plots; the R^2 for this model was 0.50 with a standard error of 4.5 m. Figure 2-8 shows the data and fitted line for the gain model.

Using the derived site index model, SI curves were generated (Figure 2-9). It is evident how greater responses compared to the control treatment are observed for lower quality sites (SI 15 m and 18 m), versus the responses for higher quality sites (SI 21 m and 24 m).

2.4. Discussion

Long-term studies bring opportunities to analyze how silvicultural treatments affect pine plantation growth, how different treatments interact, and what type of response is observed. In this study, a type C response was observed for the bedding treatment, consistent with previously reported studies (Zhao et al., 2009). Bedding has been acknowledged as a tillage treatment that improves rooting, soil moisture and nitrogen availability (Morris, 1988), increasing diameter and heights in pines, especially when these are established in wet, poorly drained sites (Gent et al., 1986; Pritchett, 1979). In this research, installations were distributed over somewhat poorly drained soils (Shiver et al., 1990), which explains the gains observed. These gains were not maintained through age 31, as bedding is expected to improve establishment conditions without adding any extra resources to the site, therefore, a type C response was expected.

The effects of vegetation control on slash pine height growth have been widely documented in the southern US (Zhao et al., 2009, 2007). Gains between 0.7 m and 1.5 m have been reported at young ages (2-7 years) when treatments removed competing vegetation (Creighton et al., 1987; Lauer and Glover, 1998), showing a peak gain at age ~11 and a subsequent slow decrease in growth over the control (Zhao et al., 2009). Similar results are reported in this study, with a maximum average gain in dominant height of 2.2 m achieved at age ~11 for the vegetation control treatment, and a 2.9 m gain for the combined treatment at the same age. The type B response observed for this treatment has been previously observed for slash pine. Jokela et al. (2000) observed a decline in average height gains between ages 5 and 8 for slash pine due to early growth benefits requiring additional inputs to sustain acceptable growth rates (fertilizer additions or more intensive understory competition control). In this study, dominant height gains were maintained due to the continuous control of competing vegetation until crown closure. Nevertheless, a decline was still observed at later ages, probably due to the lack of available resources on the site and intra-specific competition that limited growth even when no competing vegetation was present.

The CR dominant height model performed similarly to the PR model when comparing R^2_{adj} and RMSE. Mason and Milne (1999), when modeling basal area growth in *Pinus radiata* in New Zealand, also compared the PR model with a model that included treatment responses using parameter modifiers. In their study, modifying the parameters to represent the effects of site-preparation treatments resulted in reductions (although very small) of the model residual sum of squares. These authors found that adding an adjustment factor (as in the PR model) provided a better fit. In the research presented here, parameter modifiers showed a better fit when modeling the effect of bedding, but not the effect of vegetation control.

With either of the models fitted, a dynamic equation that leads to a SI model can be derived. For both models, an additional factor (or gain function) that accounts for gains in expressed SI due to the treatments, must be included. When this is not done, all the treatments converge to the same SI value at the base age of 25 years. Since treatment gains were shown to be related to site quality, including a gain function that depends on the base SI was found to be necessary when modeling dominant height for the vegetation control treatment. The results showed expressed SI was not changed significantly when only bedding was applied (0.5 m increase), but those treatments including vegetation control had an average increase in expressed SI varying from approximately -1.0 to 5.0 m, with the magnitude decreasing with increasing base SI.

Negative values of changes in SI due to the treatments are generated when higher values of dominant height for the control are observed compared to the treated plots. This might have been the result of differences between the site qualities of a control plot and its counterpart receiving the treatment. Zhao et al., (2016) reported that if the plot receiving the treatment has a somewhat higher or lower site quality than its counterpart plot not receiving the treatment, the calculated response may be greater or less than the true response. The study design of this long-term study controlled for this factor by allowing no more than 1.5 m difference in SI between the control and the treated plots. Nevertheless, this difference can still have a significant influence in the resulting gain in SI values, especially for the low-quality sites, as showed in this research.

Higher gains in low quality sites have been previously reported for slash pine. Oppenheimer et al. (1989) found for this species that responses to complete vegetation control could be affected by site quality, expecting lower gains for high quality sites which can support both understory and overstory vegetation. Logan & Shiver (2006) also confirmed this relationship for slash pine, reporting the variation in maximum gain differentiated by base SI, after applying treatments including bedding, vegetation control during site preparation, fertilization and mid-rotation release. These authors proposed a linear relationship between base SI and maximum expected response. In this research, a linear relationship was confirmed for base SI and expected gain at age 25. This inverse relationship has been confirmed for other pines in the southern United States (Zhao et al., 2016). According to Zhao et al., (2016) the lack of response on high quality sites might be a consequence of the treatment not providing limiting resources, or the trees being limited by other resources different that the ones provided by the treatment, which limit the effect of the treatment applied. These authors also mention plot to plot site variability, pest activity, weather events, and potential uneven treatment applications, as possible factors driving this relationship.

2.5. Conclusions

The control of competing vegetation by means of herbicide applications during site preparation activities for slash pine plantations had a long-lasting effect on dominant height still observed at age 31. The observed site index at age 25 was increased up to 5 m, with higher responses (relative to the control treatment) observed in low quality sites (i.e., base site index of 14 m). Higher quality sites (i.e., SI > 22 m) showed no significant increase in observed site index due to the treatment application. These findings imply that a more efficient application of herbicides can be done by targeting low quality sites which will show higher responses. On the other hand, bedding did not have a long-term effect on dominant height. Nevertheless, this treatment improved growth in dominant height at earlier stages (<20 years). The proposed SI model can be used to determine likely gains in dominant height due to treatment applications up to a rotation age of approximately 30 years and to evaluate whether this expected gain justifies the application costs.

2.6. Acknowledgements

This research was conducted with support from the Plantation Management Research Cooperative (PMRC), where resources have been invested in establishing and collecting data from this long-term research trial. Special thanks to the PMRC Field Crew who collected the data for this research over the course of three decades. In addition, this work was supported by the Warnell School of Forestry and Natural Resources at the University of Georgia, the J.L. Clutter Forest Biometrics Scholarship, and the McIntire-Stennis Project GEOZ0202-MS.

2.7. TABLES AND FIGURES

Table 2-1. Dominant height statistics per treatment and age, in meters.

Age (years)	Treatment	No. of plots	Mean (m)	SD (m)	Min (m)	Max (m)
5	Control	20	3.2	0.7	2.1	4.7
8		18	6.2	1.1	3.6	8.7
11		18	9.1	1.6	5.6	12.2
14		18	12.2	2.0	7.9	15.7
17		18	14.5	2.4	9.4	18.6
20		18	16.4	2.5	10.6	20.1
23		17	18.3	2.9	12.6	22.5
26		16	19.8	3.4	13.8	24.8
31		12	20.9	4.1	14.8	26.7
5	Bedding	21	3.9	0.7	3.1	5.2
8		19	7.1	1.0	6.0	9.2
11		19	10.1	1.3	8.3	13.1
14		19	13.2	1.5	10.9	16.4
17		19	15.5	2.0	12.6	19.5
20		19	17.1	2.4	13.5	21.5
23		18	18.9	2.9	14.5	24.2
26		17	20.2	3.3	15.0	25.4
31		11	20.9	4.2	15.2	27.0
5	Vegetation control	18	4.7	0.8	2.9	6.0
8		17	8.3	0.9	6.5	9.5
11		17	11.3	1.1	9.6	13.2
14		17	14.3	1.5	12.1	16.7
17		17	16.5	1.8	14.2	19.9
20		16	18.3	2.3	15.2	22.7
23		15	20.1	2.4	16.7	24.5
26		14	21.5	2.8	17.8	26.3
31		10	22.8	3.2	19.3	28.5
5	Bedding + Vegetation control	20	5.3	0.6	3.9	6.3
8		18	8.9	0.7	7.8	10.1
11		18	11.9	0.8	11.1	13.4
14		18	14.9	1.1	13.2	16.8
17		18	17.1	1.4	14.8	20.0
20		17	18.6	1.8	16.4	22.1
23		15	20.4	2.2	17.2	24.4
26		15	21.9	2.4	18.6	26.2
31		11	22.8	3.0	19.7	28.2

Table 2-2. Additional stand characteristics, averaged over all the installations (DBH: Diameter at breast height, BA: Basal area, TPH: trees per hectare).

Treatment	**Age (years)**	**DBH (cm)**	**BA (m^2/ha)**	**TPH**
Control	5	3.5	1.4	1,184
	11	11.0	11.2	1,109
	17	15.1	19.3	1,046
	23	17.3	24.4	1,017
	31	19.4	27.1	973
Bedding	5	4.9	2.8	1,258
	11	11.9	13.8	1,192
	17	15.6	22.6	1,141
	23	17.5	27.4	1,095
	31	19.5	28.6	1,042
Vegetation control	5	7.2	5.3	1,157
	11	14.3	18.6	1,131
	17	17.9	26.9	1,076
	23	20.3	31.5	992
	31	22.7	34.7	969
Bedding + Vegetation control	5	8.2	6.8	1,189
	11	14.8	20.3	1,146
	17	18.1	28.4	1,083
	23	20.0	32.0	1,005
	31	22.1	37.0	1,039
	5	7.2	5.3	1,157

Table 2-3. Base models and ADA formulations for the Chapman-Richards (CR) and Pienaar and Rheney (PR) models. Parameters and variables as defined in section 2.2.

Model	**Parameter related to site**	**Solution for $\boldsymbol{X}$ with initial values $(\boldsymbol{t_0}, \boldsymbol{DH_{t_0}})$**	**Dynamic Equation**
CR	$a_0 = X$	$X_0 = \frac{DH_{t_0}}{b_1^{z_1} b_2^{z_2} \times \left(1 - exp\left(-a_1 b_3^{z_1} b_4^{z_2}\, t_0\right)\right)^{a_2}}$	$DH_t = DH_{t_0}\left[\frac{1 - exp\left(-a_1 b_3^{z_1} b_4^{z_2}\, t\right)}{1 - exp\left(-a_1 b_3^{z_1} b_4^{z_2}\, t_0\right)}\right]^{a_2}$
CR	$a_1 = X$	$X_0 = -\,ln\left[1 - \left(\frac{DH_{t_0}}{a_0 b_1^{z_1} b_2^{z_2}}\right)^{1/a_2}\right] \Big/ b_3^{z_1} b_4^{z_2}\, t_0$	$DH_t = a_0 b_1^{z_1} b_2^{z_2}\left[1 - \left[1 - \left(\frac{DH_{t_0}}{a_0 b_1^{z_1} b_2^{z_2}}\right)^{1/a_2}\right]^{t/t_0}\right]^{a_2}$
PR	$a_0 = X$	$X_0 = \frac{DH_{t_0} - (b_0 Z_1 + b_1 Z_2) t_0 e^{-b_2 t_0}}{(1 - exp(-a_1 t_0))^{a_2}}$	$DH_t = X_0(1 - exp(-a_1 t))^{a_2} + (b_0 Z_1 + b_1 Z_2) t e^{-b_2 YST}$
PR	$a_1 = X$	$X_0 = -\,ln\left[1 - \left(\frac{DH_{t_0} - (b_0 Z_1 + b_1 Z_2) t_0 e^{-b_2 t_0}}{a_0}\right)^{1/a_2}\right] \Big/ t_0$	$DH_t = a_0(1 - exp(-X_0 t))^{a_2} + (b_0 Z_1 + b_1 Z_2) t e^{-b_2 YST}$

Table 2-4. Models tested for base SI vs Gain in SI.

ID	Formula
1	$G_{HDOM_i} = \beta_{0i} + \beta_{1i} SI_{25}$
2	$G_{HDOM_i} = \beta_{0i} + \beta_{1i} \log(SI_{25})$
3	$G_{HDOM_i} = \beta_{0i}\ exp\ (\beta_{1i}\ SI_{25})$

Where G_{HDOM} is the gain in expressed SI (m), a_{0i} and a_{1i} are the specific parameters for treatment i, and SI_{25} is the base site index, or dominant height at age 25 for the control treatment. The indicator i goes from 1 to 4 and indicating the treatments applied in the following order: control, bedding, vegetation control, and Bed + Veg. treatment.

Table 2-5. Estimated parameters for the two models evaluated, Chapman-Richards (CR), and Pienaar and Rheney (PR).

Model	Parameter	Estimated value	Standard error
CR	a_0	26.94	1.27
	a_1	0.07	0.01
	a_2	1.68	0.07
	b_1	0.94	0.03
	b_2	0.94	0.03
	b_3	1.16	0.04
	b_4	1.31	0.05
	β	0.88	
PR	a_0	24.62	0.85
	a_1	0.08	0.01
	a_2	1.86	0.15
	b_1	0.15	0.05
	b_2	0.42	0.09
	b_3	0.07	0.01

Table 2-6. Statistics of the models evaluated, Chapman-Richards (CR), and Pienaar and Rheney (PR). R^2_{adj}: adjusted coefficient of determination, RMSE: Root mean square error, AIC: Akaike information criteria.

Model	R^2_{adj}	**RMSE (m)**	**RMSE crossvalidation (m)**	**AIC**
CR	0.8893	2.039	2.085	875.3
PR	0.8905	2.028	2.086	867.9

Table 2-7. Estimated parameters for the dynamic site index model.

Parameter	**Estimated value**	**Standard error**
a_1	0.070	0.003
a_2	1.649	0.049
b_1	0.959	0.016
b_2	1.001	0.017
b_3	1.120	0.026
b_4	1.221	0.029
β	0.207	

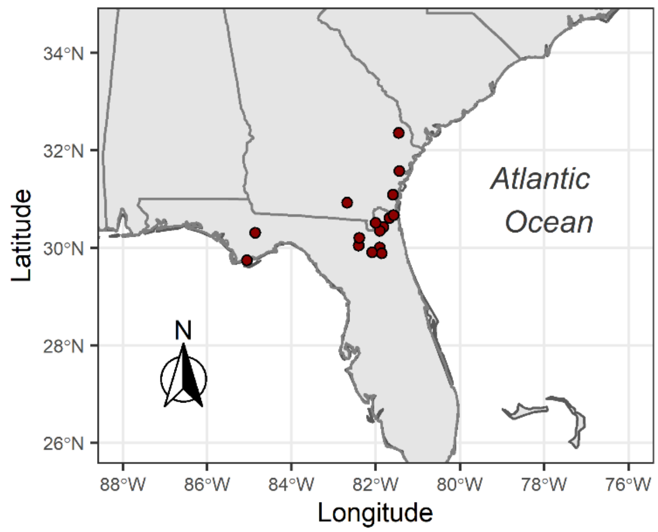

Figure 2-1. Location of the long-term slash pine study (each dot corresponds to one of the 16 installations used in this research).

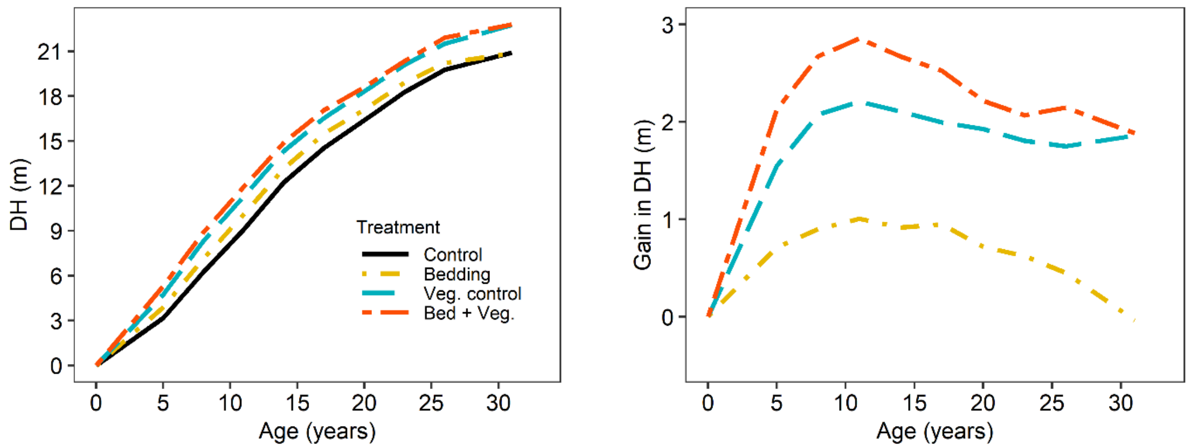

Figure 2-2. Average dominant height (DH) by treatment (left) and gain with respect to the control treatment (right).

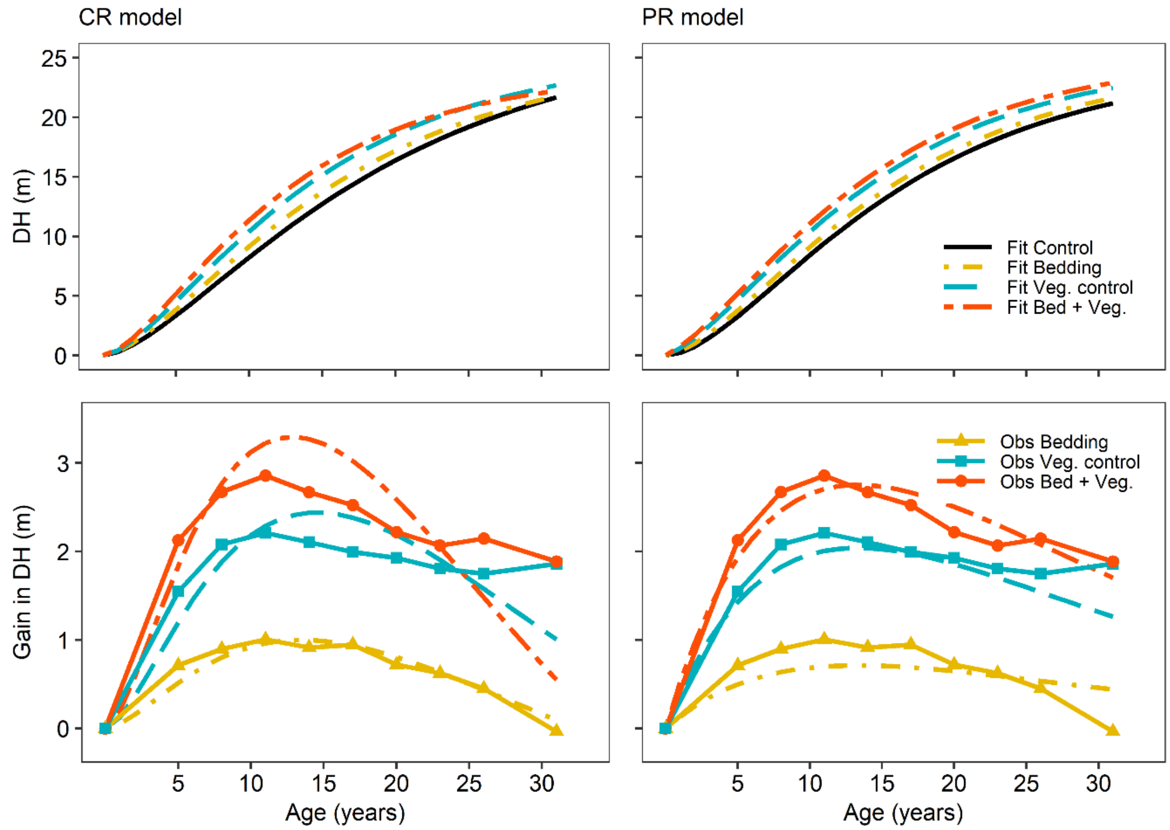

Figure 2-3. Average predicted dominant height (DH) and gain with the Chapman-Richards (CR) and Pienaar and Rheney (PR) models.

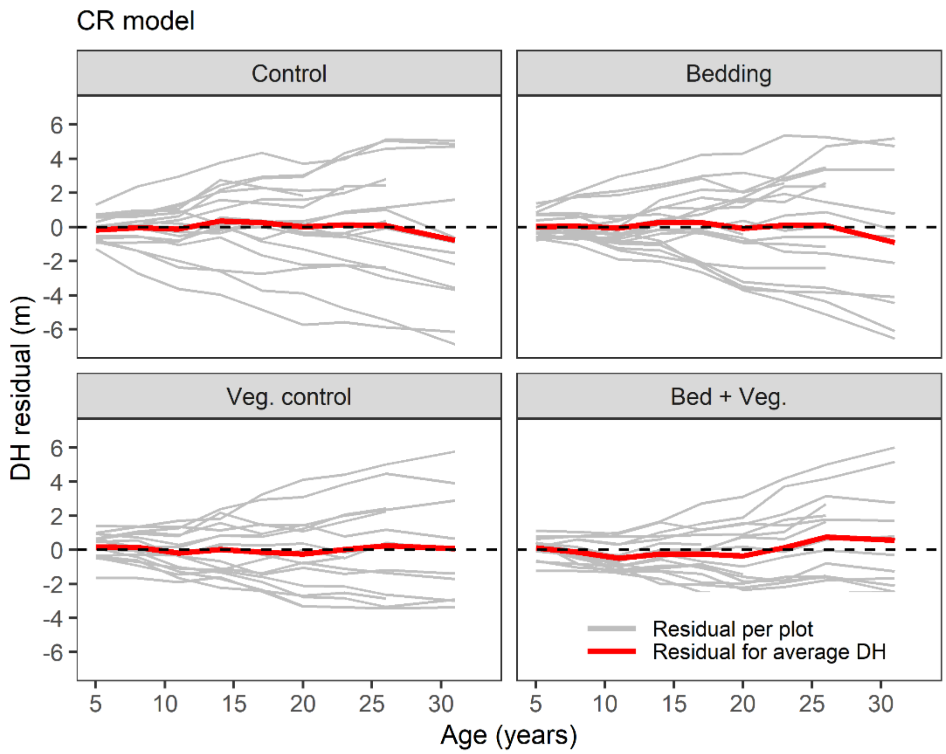

Figure 2-4. Residuals for the Chamman-Richards (CR) model calculated as 1. The observed value per plot minus the predicted value from the CR model (grey lines), and 2. The average dominant height (of all plots) at every age minus the predicted value from the CR model (red line).

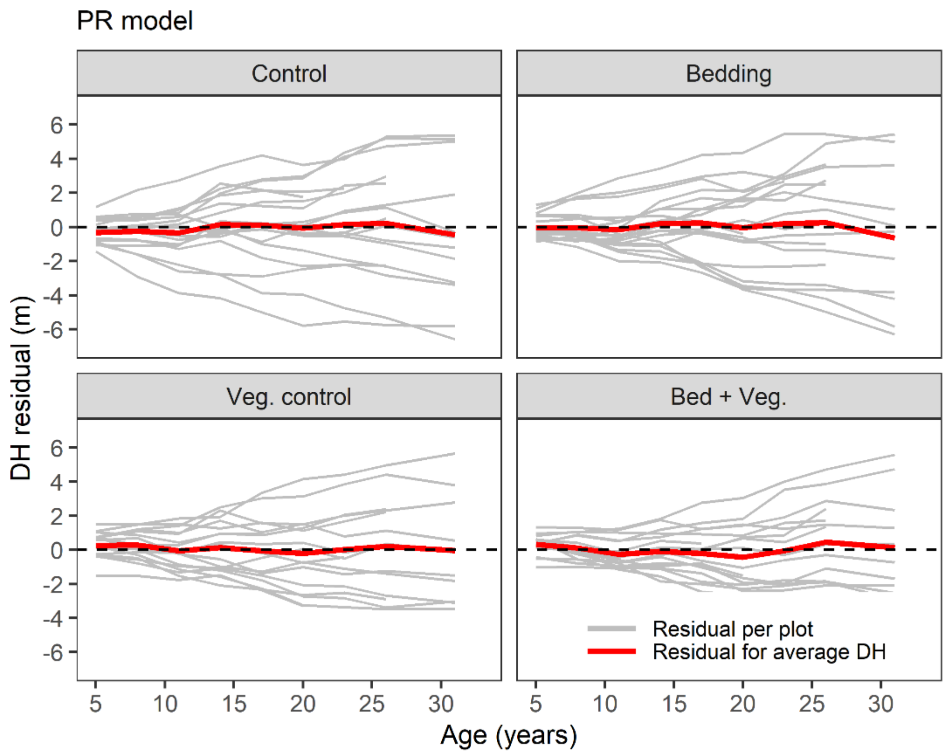

Figure 2-5. Residuals for the Pienaar and Rheney (PR) model calculated as 1. The observed value per plot minus the predicted value from the PR model (grey lines), and 2. The average dominant height (of all plots) at every age minus the predicted value from the PR model (red line).

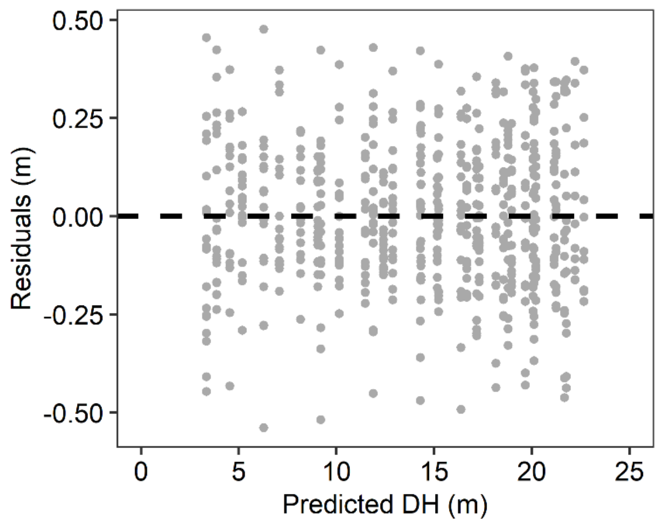

Figure 2-6. Residuals for the Chapman-Richards (CR) model. DH: dominant height.

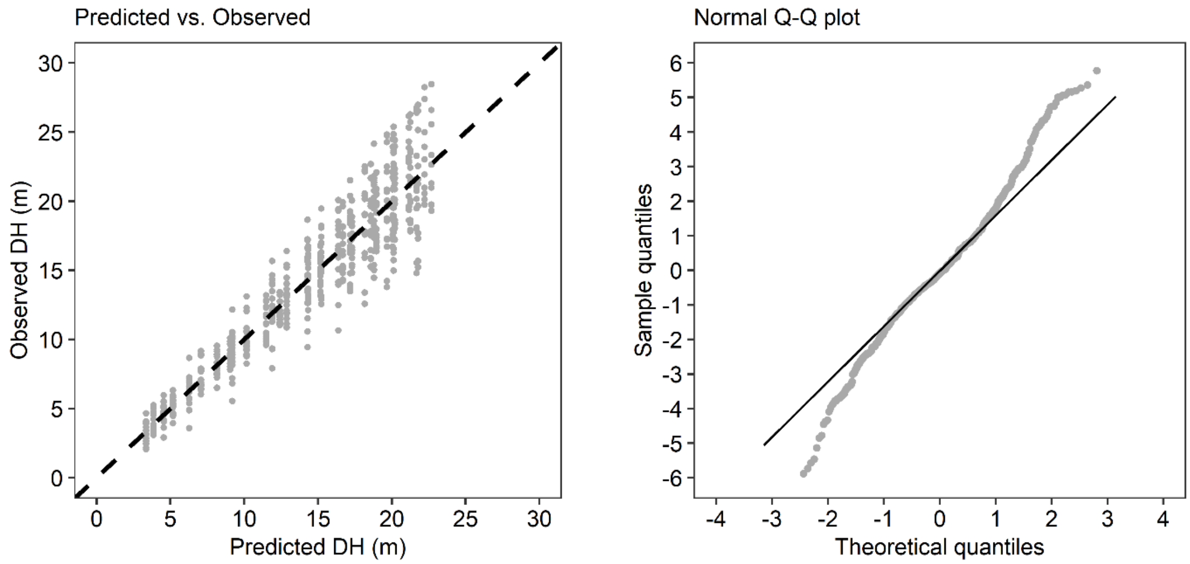

Figure 2-7. Diagnostic plots for the Chapman-Richards (CR) model. DH: Dominant height.

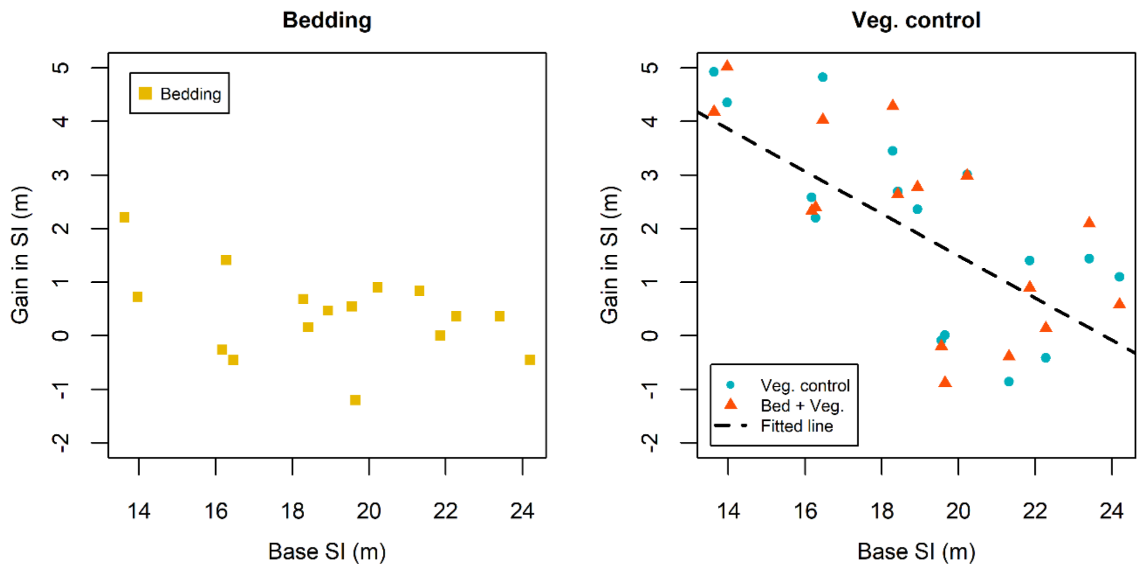

Figure 2-8. Relationship between base site index (SI) and gain in SI.

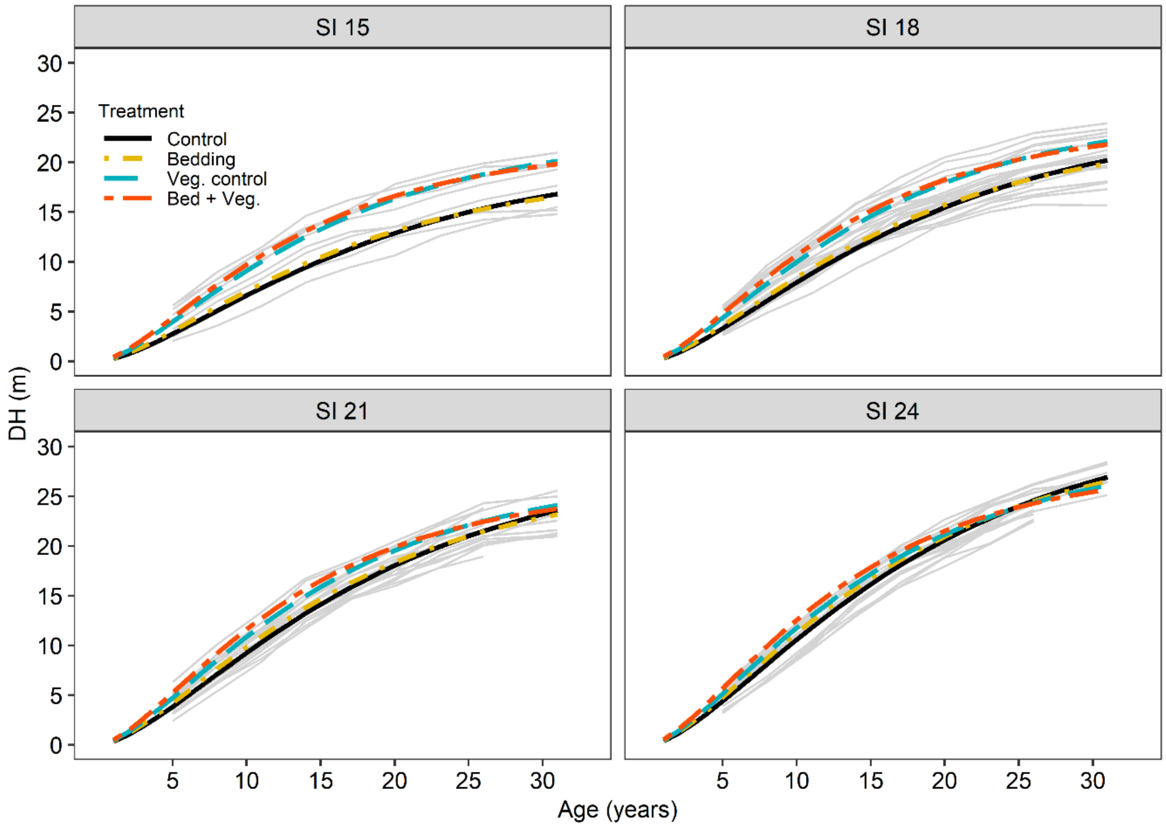

Figure 2-9. Site Index curves including treatment response. DH: dominant height.

2.8. References

Allen, H.L., Dougherty, P.M., Campbell, R.G., 1990. Manipulation of water and nutrients — Practice and opportunity in Southern U.S. pine forests. For. Ecol. Manage. 30, 437–453. https://doi.org/10.1016/0378-1127(90)90153-3

Amateis, R.L., Liu, J., Ducey, M.J., Allen, H.L., 2000. Modeling Response to Midrotation Nitrogen and Phosphorus Fertilization in Loblolly Pine Plantations. South. J. Appl. For. 24, 207–212. https://doi.org/10.1093/sjaf/24.4.207

Antón-Fernández, C., Burkhart, H.E., Strub, M., Amateis, R.L., 2011. Effects of initial spacing on height development of loblolly pine. For. Sci. 57, 201–211.

Bailey, R., Clutter, J., 1974. Base-Age Invariant Polymorphic Site Curves. For. Sci. 20, 155–159. https://doi.org/10.1093/forestscience/20.2.155

Barnett, J.P., Sheffield, R.M., 2004. Slash Pine: Characteristics, history, status and trends, in: Dickens, E.D.; Barnett, J.P.; Hubbard, W.G.; Jokela, E.L., Eds. Slash Pine: Still Growing and Growing! Proccedings of the Slash Pine Composium. pp. 1–6.

Carlson, C.A., Fox, T.R., Lee Allen, H., Albaugh, T.J., 2008. Modeling mid-rotation fertilizer responses using the age-shift approach. For. Ecol. Manage. 256, 256–262. https://doi.org/10.1016/j.foreco.2008.04.020

Cieszewski, C.J., Bailey, R.L., 2000. Generalized Algebraic Difference Approach: Theory Based Derivation of Dynamic Site Equations with Polymorphism and Variable Asymptotes. For. Sci. 46, 116–126. https://doi.org/https://doi.org/10.1093/forestscience/46.1.116

Colbert, S.R., Jokela, E.J., Neary, D.G., 1990. Effects of annual fertilization and sustained weed control on dry matter partitioning, leaf area, and growth efficiency of juvenile loblolly and slash pine. For. Sci. 36, 995–1014. https://doi.org/10.1093/forestscience/36.4.995

Creighton, J.L., Zutter, B.R., Glover, G.R., Gjerstad, D.H., 1987. Planted Pine Growth and Survival Responses to Herbaceous Vegetation Control, Treatment Duration, and Herbicide

Application Technique. South. J. Appl. For. 11, 223–227. https://doi.org/10.1093/sjaf/11.4.223

Diéguez-Aranda, U., Burkhart, H.E., Amateis, R.L., 2006. Dynamic site model for loblolly pine (Pinus taeda L.) plantations in the United States. For. Sci. 52, 262–272. https://doi.org/10.1093/forestscience/52.3.262

Diéguez-Aranda, U., Burkhart, H.E., Rodríguez-Soalleiro, R., 2005. Modeling dominant height growth of radiata pine (Pinus radiata D. Don) plantations in north-western Spain. For. Ecol. Manage. 215, 271–284. https://doi.org/10.1016/j.foreco.2005.05.015

Fang, Z., Bailey, R.L., 2001. Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. For. Sci. 47, 287–300. https://doi.org/10.1093/forestscience/47.3.287

Fox, T.R., Jokela, E.J., Allen, H.L., 2007. The development of pine plantation silviculture in the Southern United States. J. For. 105, 337–347. https://doi.org/10.1093/jof/105.7.337

Gent, J.A., Allen, H.L., Campbell, R.G., Wells, C.G., 1986. Magnitude, duration, and economic analysis of loblolly pine growth response following bedding and phosphorus fertilization. South. J. Appl. For. 10, 124–128. https://doi.org/10.1093/sjaf/10.3.124

Gyawali, N., Burkhart, H.E., 2015. General response functions to silvicultural treatments in loblolly pine plantations. Can. J. For. Res. 45, 252–265. https://doi.org/10.1139/cjfr-2014-0172

Hughes, J.H., Campbell, R.G., Duzan Sr, H.W., Dudley, C.S., 1979. Site index adjustment for intensive forest management treatments at North Carolina. For. Res. Tech. Report. Weyerhaeuser Co., New Bern, NC. 042-1404/79/24.

Hynynen, J., 1995. Predicting the growth response to thinning for Scots pine stands using individual-tree growth models. Silva Fenn. 29, 225–246. https://doi.org/10.14214/sf.a9210

Hynynen, J., Burkhart, H.E., Allen, H.L., 1998. Modeling tree growth in fertilized midrotation loblolly pine plantations. For. Ecol. Manage. 107, 213–229. https://doi.org/10.1016/S0378-1127(97)00334-4

Jokela, E.J., Martin, T.A., Vogel, J.G., 2010. Twenty-five years of intensive forest management with southern pines: Important lessons learned. J. For. 108, 338–347. https://doi.org/10.1093/jof/108.7.338

Jokela, E.J., Wilson, D.S., Allen, J.E., 2000. Early growth responses of slash and loblolly pine following fertilization and herbaceous weed control treatments at establishment. South. J. Appl. For. 24, 23–30. https://doi.org/10.1093/sjaf/24.1.23

Lauer, D.K., Glover, G.R., 1998. Early pine response to control of herbaceous and shrub vegetation in the flatwoods. South. J. Appl. For. 22, 201–208. https://doi.org/10.1093/sjaf/22.4.201

Logan, S.R., Shiver, B.D., 2006. Adjusting Slash Pine Growth and Yield for Silvicultural Treatments. Gen. Tech. Rep. SRS-92. Asheville, NC U.S. Dep. Agric. For. Serv. South. Res. Stn. 328–332.

Martin, S.W., Bailey, R.L., Jokela, E.J., 1999. Growth and yield predictions for lower coastal plain slash pine plantations fertilized at mid-rotation. South. J. Appl. For. 23, 39–45. https://doi.org/10.1093/sjaf/23.1.39

Mason, E.G., Milne, P.G., 1999. Effects of weed control, fertilization, and soil cultivation on the growth of Pinus radiata at midrotation in Canterbury, New Zealand. Can. J. For. Res. 29, 985–992. https://doi.org/10.1139/x99-087

Montes, C.R., 2001. A Silvicultural Decision Support System for Loblolly Pine Plantations. North Carolina State Unviersity.

Morris, L.A., 1988. Influence of Site Preparation on Soil Conditions Affecting Stabd Estaclishment and Tree Growth. South. J. Appl. For. 12, 170–178.

Oppenheimer, M.J., Shiver, B.D., Rheney, J.W., 1989. Ten-year growth response of midrotation slash pine plantations to control of competing vegetation. Can. J. For. Res. 19, 329–334. https://doi.org/10.1139/x89-049

Pienaar, L., Rheney, J., 1995. Modeling Stand Level Growth and Yield Response to Silvicultural Treatments. For. Sci. 41, 629–638. https://doi.org/10.1093/forestscience/41.3.629

Pritchett, W.L., 1979. Site preparation and fertilization of slash pine on a wet savanna soil. South. J. Appl. For. 3, 86–90.

Quicke, H., Glover, G., Glover, R., 1999. Loblolly pine growth response to herbaceous vegetation control at different planting densities. Can. J. For. Res. 29, 960–967. https://doi.org/10.1139/x99-069

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Salas, C., Stage, A.R., Robinson, A.P., 2008. Modeling effects of overstory density and competing vegetation on tree height growth. For. Sci. 54, 107–122.

Scolforo, H.F., Montes, C., Cook, R.L., Lee Allen, H., Albaugh, T.J., Rubilar, R., Campoe, O., 2020. A New Approach for Modeling Volume Response from Mid-Rotation Fertilization of Pinus taeda L. Plantations. Forests 11, 646. https://doi.org/10.3390/f11060646

Sharma, M., Burkhart, H.E., Amateis, R.L., 2002. Modeling the effect of density on the growth of loblolly pine trees. South. J. Appl. For. 26, 124–133. https://doi.org/10.1093/sjaf/26.3.124

Shiver, B.D., Rheney, J.W., Oppenheimer, M.J., 1990. Site-Preparation Method and Early Cultural Treatments Affect Growth of Flatwoods Slash Pine Plantations. South. J. Appl. For. 14, 183–188. https://doi.org/https://doi.org/10.1093/sjaf/14.4.183

Snowdon, P., 2002. Modeling type 1 and type 2 growth responses in plantations after application of fertilizer or other silvicultural treatments. For. Ecol. Manage. 163, 229–244. https://doi.org/10.1016/S0378-1127(01)00582-5

Snowdon, P., Waring, H.D., 1984. Long-term nature of growth responses obtained to fertilizer and weed control applied at planting and their consequences for forest management, in: Grey, D.., Shönau, A.P.., Schutz, C.. (Eds.), Proceedings of the IUFRO Symposium on Site and Productivity of Fast Growing Plantations. Forest Research Institute, Pretoria and Pietermaritzberg, South Africa, pp. 701–711.

Snowdon, P., Waring, H.D., 1981. Forest fertilisation and thinning practice, in: Proceedings of the IUFRO Regional Thinning Meeting, Economics and Techniques of Thinned Plantations in Australia and New Zealand. CSIRO Division of Forest Research, Camberra, pp. 159–162.

Socha, J., Tymińska-Czabańska, L., Bronisz, K., Zięba, S., Hawryło, P., 2021. Regional height growth models for Scots pine in Poland. Sci. Rep. 11, 1–14. https://doi.org/10.1038/s41598-021-89826-9

South, D.B., Miller, J.H., Kimberley, M.O., Vanderschaaf, C.L., 2006. Determining productivity gains from herbaceous vegetation management with "age-shift" calculations. Forestry 79, 43–56. https://doi.org/10.1093/forestry/cpi058

Tymińska-Czabańska, L., Hawryło, P., Socha, J., 2022. Assessment of the effect of stand density on the height growth of Scots pine using repeated ALS data. Int. J. Appl. Earth Obs. Geoinf. 108. https://doi.org/10.1016/j.jag.2022.102763

Zhao, D., Kane, M., Borders, B., Harrison, M., 2009. Long-term effects of site preparation treatments, complete competition control, and repeated fertilization on growth of slash pine plantations in the flatwoods of the southeastern United States. For. Sci. 55, 403–410. https://doi.org/10.1093/forestscience/55.5.403

Zhao, D., Kane, M., Borders, B., Harrison, M., 2008. Pine growth response to different site-preparation methods with or without post-plant herbaceous weed control on North Florida's Lower Coastal Plain. For. Ecol. Manage. 255, 2512–2523. https://doi.org/10.1016/j.foreco.2008.01.011

Zhao, D., Kane, M., Borders, B., Harrison, W.M., 2007. Slash pine site preparation study: Results through age 26. PMRC Tech. Rep. 2007-3 61.

Zhao, D., Kane, M., Teskey, R., Fox, T.R., Albaugh, T.J., Allen, H.L., Rubilar, R., 2016. Maximum response of loblolly pine plantations to silvicultural management in the southern United States. For. Ecol. Manage. 375, 105–111. https://doi.org/10.1016/j.foreco.2016.05.035

CHAPTER 3

MODELING SLASH PINE MORTALITY RATES INCORPORATING RESPONSES TO SILVICULTURAL TREATMENTS AND FUSIFORM RUST INFECTION RATES[1]

[1] Ramirez, L., Montes, C. R., & Bullock, B. P. 2023. *Forest Ecology and Management*, 532: 120832. Reprinted here with permission of publisher.

Abstract

The evaluation of stand survival at any age is an essential task that allows foresters to estimate stand dynamics and ultimately, the value of a forest stand. Therefore, any growth and yield system developed with the aim of predicting or projecting standing value requires precise estimates of the number of trees surviving at any point in time. When describing survival through modeling the mortality rate, differential equations using height increments rather than time increments have been shown to improve the overall fit of survival models. Using a long-term data set of slash pine plantations (*Pinus elliottii* Engelm.) in which silvicultural treatments (i.e., bedding and complete vegetation control) were applied during the establishment phase, and fusiform rust (*Cronartium quercuum* Berk.) infection rates were recorded, a survival/mortality model of this nature was constructed. Height increments were taken from a proposed dominant height model that included explicit treatment effects. In addition, the proportion of trees infected with fusiform rust at an age of 5 years was added as a predictor describing the mortality rate. Our results show that stand survival is better described by a model in which time increments are used rather than height increments, and although silvicultural treatments were essential for modeling dominant height, mortality was not greatly affected by these treatments and therefore, no additional parameter modifiers associated to these treatments were needed. On the other hand, the inclusion of average fusiform rust infections was essential to describe mortality rates in these stands, with higher infection rates associated to higher mortality rates.

3.1. Introduction

Forest stand dynamics and the associated financial value of a forest are largely determined by the stand survival at any point in time. It is therefore crucial that growth and yield systems developed to predict or project stand-level metrics (e.g., cubic volume, tons, or value), include a precise estimate of the number of trees surviving at any age. In forest plantations, density-dependent mortality is the most common process modeled. There are numerous challenges associated with modeling other stochastic factors such as diseases, pest infestations, and extreme weather events (Lee, 1971). In the southeastern United States, the stand-level relative rate of density dependent mortality has been modeled using differential equations since the 1980s (Bailey et al., 1985; Clutter and Jones, 1980; Pienaar and Shiver, 1981). This rate has been described as either a constant over time or as a function of current stand density, age and/or site quality. An extensive review of mortality functions was presented by Zhao et al., (2007a) concluding that no single equation was the best for all areas and management scenarios.

A different approach to model survival in which mortality rates are modeled with respect to dominant height instead of time, also using differential equations, was proposed by Garcia (2009). Garcia argued that using dominant height instead of age was more appropriate to describe mortality rates since dominant height is directly describing size, while age, although related to size, is not a direct measurement of size. One advantage of following this modeling approach is the ability to obtain a model that is independent of site quality (a quantity expected to reflect the site conditions over time), but that at the same time is able to accommodate changes in dominant height trajectories, providing a good fit even when the data is scarce (Garcia, 2010; Garcia et al., 2011; Tewari and Singh, 2018).

In the southeastern United States, pine plantations are intensively managed through silvicultural prescriptions aiming to reach the maximum production potential on a given site (Fox et al., 2007). Along with this intensive silviculture and management, growth and yield models have been constructed to accurately capture the effects of silvicultural treatment applications (Bailey and Burgan, 1989; Clutter and Jones, 1980; Gyawali and Burkhart, 2015; Martin et al., 1999; Pienaar and Rheney, 1995). Silvicultural treatment effects are usually incorporated directly into growth and yield models by adding a response value to a control (or base) model (Pienaar and Rheney, 1995) or modifying the parameters of the model (Mason and Milne, 1999).

Particularly for slash pine (*Pinus elliottii* Engelm.), the second most important planted species in the Southern United States (Barnett and Sheffield, 2004), several authors have proposed dominant height and survival models that include the effect of silvicultural treatments mainly by adding additional treatment factors or modifying the parameters of the model according to the treatment applied (Bailey et al., 1985; Bailey and Burgan, 1989; Pienaar and Rheney, 1995; Ramirez et al., 2022). In light of the methodology proposed by Garcia (2009), in which mortality is modeled with respect to dominant height instead of time, it is of interest to determine whether a mortality model for a forest stand in which silvicultural treatments were applied, requires additional modifications if the dominant height model already includes the treatment effects.

Survival models for slash pine plantations usually include fusiform rust (*Cronartium quercuum* Berk.) presence and/or infection rates as an important component determining the mortality rate (Bailey and Burgan, 1989; Devine and Clutter, 1985; Nance et al., 1981). Fusiform rust is one of the most relevant pathogens affecting slash pine plantations in the southeastern United States, and despite loss estimates of around 84 million dollars (2020 US dollars) per year (Susaeta, 2020), infection rates have not declined significantly for this important species (Randolph, 2016). Although mortality is not always directly caused by the cankers produced in an infected tree, several studies have shown evidence of higher mortality rates in infected trees compared to uninfected trees (Jones, 1972; Sluder, 1977), and infection rate at young ages (< 5 years) has been identified as a good predictor of future mortality and volume loss in slash pine plantations (Wells and Dinus, 1978).

The main objective of this paper was to develop a survival/mortality model for slash pine plantations including the effect of silvicultural treatments and the effects of fusiform rust infection on the mortality rate. The methodology proposed by Garcia (2009) involving dominant height increments was tested and model performance was compared with other mortality models that included the treatment effect explicitly and described the mortality rate using time increments. Our hypothesis regarding the different modeling techniques tested was that similar performance could be achieved when modeling mortality rates with respect to height without including additional explicit treatments effects and when modeling mortality rates with respect to time but including explicit treatment effects.

3.2. METHODS

3.2.1. DATA

A long-term slash pine study established in 1979 by the Plantation Management Research Cooperative (PMRC) at the University of Georgia in the southeastern United States was used to test the proposed hypothesis. The study was established on 16 different installations throughout the Lower Coastal Plain in northern Florida and southern Georgia (Figure 3-1). The main silvicultural treatments considered were bedding, consisting of a double pass with a bedding harrow during site preparation, complete competing vegetation control (using herbicides), and the combination of these two treatments. The vegetation control treatment included an herbicide application before site preparation (3% solution of Roundup) and repeated localized applications of Roundup or Garlon to remove most of the competing vegetation until crown closure (Zhao et al., 2009). These silvicultural treatments were replicated at least once on each one of the installations and dominant height, stand density measurements, and average fusiform rust infection rates per plot were available every 3 years starting from age 5 and up to age 31 for the longest series. In Table 3-1 a summary of the dominant height and number of trees per hectare for each of the treatments is presented. More details about this study are provided by Zhao et al., (2009), Zhao et al., (2007b) and Ramirez et al., (2022). Dominant height (H_D) trajectories are plotted in Figure 3-2 and trees per hectare (TPH) with respect to time and dominant height are plotted in Figure 3-3 and Figure 3-4, respectively.

3.2.2. MODELS

Dominant height model

A dominant height model was proposed to subsequently estimate dominant height increments and test the hypothesis related to using height increments versus time increments in the mortality model. The Gompertz model which has been used previously to model forest stand height trajectories (Medeiros et al., 2017; Zang et al., 2016), was chosen to describe the dominant height trajectories. The basic assumption of this model is that growth is proportional to size with a constant of proportionality μ, and that the effectiveness of the growth mechanism decays over time, generating an exponential decay (France and Thornley, 1984). In mathematical terms this can be described with the following system of equations:

$$\frac{dH_D}{dt} = \mu\, H_D \qquad \text{Eq. 3-1}$$

$$\frac{d\mu}{dt} = -K\mu \qquad \text{Eq. 3-2}$$

Where dH_D/dt is the change in dominant height (H_D) over a period of time (dt), t is time in years, and μ and K are parameters determining the dominant height trajectory. When Eq. 3-2 is solved as a separable differential equation and the value of μ is replaced back in Eq. 3-1, the following differential equation is obtained.

$$\frac{dH}{dt} = \mu_0 H e^{-Kt} \quad \text{Eq. 3-3}$$

Dominant height has been found to be positively affected by bedding and vegetation control (Ramirez et al., 2022). To incorporate silvicultural treatment responses into the Gompertz model (Eq. 3-3), parameter modifiers were proposed to be added to this model. Bedding is a silvicultural treatment that improves growth in the early stages of a stand by enhancing rooting along with improved soil moisture conditions and nitrogen availability (Morris and Lowery, 1988). To incorporate these effects into the model, a modifier was added to the parameter μ_0 which is the parameter directly associated with growth. For the vegetation control treatment, a modifier was added into the parameter K since this treatment does not directly improve growth, but reduces the limiting factors on site for the crop trees. Therefore, it was expected that this treatment would affect the decay rate (K) at which the growth rate μ_0 decreases. The dominant height model with the modifiers can be expressed as:

$$\frac{dH_D}{dt} = \mu_0 b_1^{Z_1} H_D e^{-K b_2^{Z_2} t} \quad \text{Eq. 3-4}$$

Where b_1 and b_2 are the parameter modifiers to be estimated, and Z_1 and Z_2 are dummy variables equal to 1 if bedding or vegetation control was applied, respectively, and zero otherwise.

Model from Eq. 3-4 was compared to the (null) model without treatment effects (Eq. 3-3) to evaluate the effectiveness of the proposed model to incorporate treatments effect into dominant height.

Survival model

Marked differences in survival trajectories were observed in the studied plots, with some of them experiencing high mortality while others experienced little to no mortality (Figure 3-3 and Figure 3-4). These differences were found to be strongly associated with the average percentage of fusiform rust infected trees at year 5. Overall, when this percentage was less than 15%, less mortality was recorded, while higher mortality rates were observed for plots in which the fusiform rust infection rate was higher than 15% (Figure 3-5). This was in line with the differential mortality rates reported by other authors assessing mortality in slash pine plantations in the southeastern United States (Jones, 1972; Sluder, 1977; Wells and Dinus, 1978). Therefore, all mortality models proposed in this research included average fusiform rust infection rate (varying from 0 to 1) as a predictor variable. In Table 3-2 the average fusiform rust infection proportion differentiated by treatment and by the cutoff value of 15% is presented. It is important to note that this classification of mortality trajectories for plots with less or more than 15% fusiform rust infection was done to visually inspect more clearly the different survival trajectories. Nevertheless, since the actual proportion of infected trees was used in the models tested, taking this infection rate as a continuous value, this cutoff value is not relevant for modeling purposes. The recorded fusiform infection rates for each plot during the whole period of study are presented in Figure 3-6 differentiated by silvicultural treatments.

Five different models found in the literature for slash pine were fitted first for the control plots, and the model with the best fit was then selected to be modified to include treatment effects and test the mentioned hypothesis related to the modeling approach. The models considered in this research are referenced in Table 3-3, where Eq. 3-10, Eq. 3-12, and Eq. 3-14 have been used previously in slash pine (Clutter and Jones, 1980; Devine and Clutter, 1985; Pienaar and Shiver, 1981) and Eq. 3-11 and Eq. 3-13 have been used for modeling mortality in other pine species (Zhao et al., 2007a). All these models were modified to include the effect of fusiform rust infestation by modeling the mortality rate (α) as a linear function of the average fusiform rust infection rate at year 5.

The model that had the best fit from Table 3-3 was selected to be modified following Garcia's (2009) approach. The different model variations are presented in a general form in Table 3-4, with all the variables as described before, θ representing the parameters from the base model in Table 3-3, and c_1, c_2 corresponding to the explicit treatment effects of bedding and vegetation control on the mortality rate, respectively. The inclusion of fusiform rust infection rates was maintained for these models.

Models where mortality was modeled with respect to dominant height instead of time (dN/dH_D instead of dN/dt) were obtained by combining the model in Eq. 3-15 with the best dominant height model from section 3.2.2 as follows:

$$\frac{dN}{dt} = f(N, t, \theta)\ \frac{dH_D}{dt} \qquad \text{Eq. 3-5}$$

$$\frac{dN}{dH_D} = f(N, t, \theta) \qquad \text{Eq. 3-6}$$

Where dH_D/dt in Eq. 3-5 is taken from the proposed dominant height model (Eq. 3-4). The models in Eq. 3-17 and Eq. 3-20 are similarly derived, although the time variable is completely replaced by the dominant height variable (H_D) for these models.

3.2.3. Parameters estimation

Parameters were estimated using the maximum likelihood framework in the Julia programming language (Bezanson et al., 2017). The procedure consisted of solving the differential equation numerically as an initial value problem and then finding the combination of parameters that maximized the likelihood of observing the data collected (given the assumed model), that is, as an inverse problem with unknown starting values. Therefore, in addition to the parameters for each one of the models in Table 3-3 and Table 3-4, the starting values for each equation (either dominant height at age 5 or stand density at age 5, for each plot), were defined as additional parameters to be estimated as part of the optimization procedure. Given the 72 plots used in this research, there were 72 dominant height values and 72 starting densities estimated using a dummy variable approach, similar to what was proposed by Cieszewski and Bailey (2000) for dominant height. An approximation of the standard error for all the parameters was obtained by calculating the Hessian matrix during the optimization process. For all models, the variable N (TPH) was scaled by dividing the values by a factor of 1,000 to facilitate the optimization process.

The following statistics were calculated to evaluate model fit:

- Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i - \widehat{Y_i}\right)^2}{n-p}}$$

- Mean Difference (MD):

$$MD = \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)}{n}$$

- Mean Absolute Difference (MAD):

$$MAD = \frac{\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|}{n}$$

And for model comparison:

- Akaike Information Criteria (AIC):

$$AIC = -2\ loglik + 2p$$

Where n is the total number of observations, p is the number of parameters in each model, Y_i is the observed value, $\widehat{Y_i}$ is the predicted value.

3.3. RESULTS

3.3.1. DOMINANT HEIGHT MODEL

Fit statistics for the two dominant height models evaluated are presented in Table 3-5. The proposed model with the addition of parameter modifiers representing treatments effects was effective at reducing the average error and bias. Therefore, model from Eq. 3-4 was chosen to test the hypothesis of using height increments for the mortality model in section 3.3.2 . Global parameter estimates for this model are presented in Table 3-6. The local parameters corresponding to the initial state of the variable, which is the dominant height (H_D) at age five ($t = 5$), are not presented.

Eq. 3-4 expressed as a difference equation (projection form), is presented in Eq. 3-7. The predicted dominant height trajectories and residuals are presented in Figure 3-7 and Figure 3-8, respectively.

$$H_{D_2} = H_{D_1}\, exp\left[\frac{\mu_0 b_1^{Z_1}}{K b_2^{Z_2}}\left(e^{-K b_2^{Z_2} t_1} - e^{-K b_2^{Z_2} t_2}\right)\right] \qquad \text{Eq. 3-7}$$

Given the importance of fusiform rust infection on the survival of the slash pine stands evaluated, it is natural to raise the question of the effect fusiform rust infection on dominant height. Nevertheless, this effect was not evaluated in this research due to previous research showing little or a non-significant reduction in height growth due to fusiform rust infection in slash pine (Jones, 1972; Nance et al., 1981; Sluder, 1977). The main documented effect of fusiform rust in slash pine has been the rust-associated mortality (RAM), which generates volume losses due to low stocking at the end of the rotation (Wells and Dinus, 1978). Economic losses are also usually associated with the low quality timber affected by stem cankers (Sluder, 1977) rather than with a reduction in growth. Burton et al., (1985) argued that although they did find significant differences in height growth at an early age (less than 5 years), over the long-term, losses associated with this lower growth rate are shadowed by RAM since young infected trees are the most likely to die before rotation age.

3.3.2. SURVIVAL MODELS

Fit statistics for the evaluated base survival models from Table 3-3 are presented in Table 3-7. The model with better performance was the model form from Eq. 3-12, which is the model proposed by Pienaar and Shiver (1981), and represents a reduced version of the model proposed by Clutter and Jones (1980) (Eq. 3-14 in Table 3-7) if parameter γ is equal to one ($\gamma = 1$). The full version of this model with $\gamma \neq 1$ generated a slightly better fit, with lower values of RMSE, MD and MAD. Nevertheless, this model had an additional parameter, and when using AIC as the model selection criteria, the model with the lowest (better) value was the model from Eq. 3-12.

Following what presented in Table 3-4, modifications were made to the model form from Eq. 3-12 to include treatments effects (i.e., bedding and vegetation control). The fit statistics for these models are presented in Table 3-8. The model with improved performance was model in Eq. 3-21, where the mortality rate was modeled with respect to time, and no additional treatment factors were used. When explicit treatment effects were added (Eq. 3-24), the RMSE and MAD were reduced slightly compared to model in Eq. 3-21, nevertheless, this was at the expense of two additional parameters, which it is not justified from a statistical point of view.

Parameter estimates for the final model (Eq. 3-21) are presented in Table 3-9. Only global parameters are presented. The trees per hectare (N) trajectories for this model are presented in Figure 3-9 along with the residuals in Figure 3-10. The residuals presented in Figure 3-10 are well distributed around zero across the predicted TPH values. In Figure 3-11 the histogram of these residuals and a Q-Q plot are presented and show no major deviations from a normal distribution, nevertheless some heavy tails are observed (Figure 3-11).

The final recommended model in its difference (projection) form is as follows:

$$N_2 = N_1 \exp\left[\frac{\alpha_0 + \alpha_1\, FR}{\delta + 1}\, \left(t_2^{\delta+1} - t_1^{\delta+1}\right)\right] \qquad \text{Eq. 3-8}$$

In the particular case of $t_1 = 0$ and $N_1 = N_0 =$ initial planting density, the prediction equation is as follows:

$$N_t = N_0 \exp\left[\frac{\alpha_0 + \alpha_1\, FR}{\delta + 1}\, t^{\delta+1}\right] \qquad \text{Eq. 3-9}$$

3.4. DISCUSSION

Stand survival in slash pine plantations including silvicultural treatment applications was best described when the mortality rate was modeled with respect to time increments rather than with respect to dominant height increments as proposed by Garcia (2009). Stankova and Diéguez-Aranda (2014) had attributed improvements in model fit when using Garcia's approach to the implicit inclusion of site quality factors when including dominant height into the model. Nevertheless, in all the models tested, the estimated local values per plot likely accounted for some of the site-specific variation. Garcia (2009) also mentioned his approach was particularly useful when dealing with scarce or low quality data that does not cover a wide enough range of growing conditions, which was not the case for this study where the installation locations were located throughout the slash pine range in southern Georgia and northern Florida.

The explicit treatment modifiers added to include the effect of bedding and or vegetation control as part of the survival function, were not successful in improving model fit significantly. Although some authors have found bedding to have a positive effect on pine plantations growth and survival when stands are located in poorly drained soils (Gent et al., 1986; Pritchett, 1979), we found that bedding did not have a significant effect on the mortality rate in this study, or more precisely, no effect was observed for the measurement period considered in this research, which started at year five. It is likely that the effect of bedding on mortality rates was more pronounced during the first years after planting and that no significant effect was observed in later years, when the overstory measurements commenced.

Regarding the vegetation control treatment, although growth gains have been reported for slash pine due to this treatment (Creighton et al., 1987; Lauer and Glover, 1998; Ramirez et al., 2022; Zhao et al., 2009), a less marked effect has been found for survival. Jokela et al., (2000) did not find a significant effect on slash pine survival rates at early ages (5 and 8 years) due to herbaceous weed control applications during site preparation, and Creighton et al. (1987) did report an improvement in survival rates due to vegetation control applications, although this was only when plantations were established on sites with a water deficit and high levels of competition. In contrast, the sites where these research plots were established were more likely to have excess water, explaining the non-significant effect of the vegetation control treatment on mortality observed in this research.

Fusiform infection rates have been found to be differential when combined with early silvicultural treatments. Burton et al., (1985) found for young stands (less than 5 years) that fusiform infection rates were higher when complete vegetation control and bedding were applied compared to a control. Although this was not an objective of our study, the fact that explicit treatment effects did not significantly improve model fit, suggest that in the long-term, the interaction between the silvicultural treatments and the fusiform rust infection rates might be less relevant.

Results from this study indicate that the most important variable to describe mortality in slash pine plantations was the fusiform rust infection rate. Wells and Dinus (1978) also found that the number of trees infected with fusiform rust at year five was a reliable predictor of rust-associated mortality at year 10. In our proposed model, the proportion of infected trees at year five was useful to model survival trajectories over time. The rate of infection at age five was maintained through the observed measurement period (Figure 3-6), with low infection rates continuing for sites in which low infection rates were observed at year five and vice versa. Therefore, it is believed that fusiform rust infection has an impact on survival for a sustained period of time beyond year five, but the first measurement taken at year five was a good proxy of the fusiform rust infection impact through the whole period evaluated. Performing an assessment of fusiform rust at year 5 is then recommended to use the proposed mortality model. Nevertheless, in the absence of this assessment, localized historical infection rates could be used to approximate the infection rate at year 5, which is preferable to ignore fusiform rust infection completely.

The use of differential equations in the construction of the proposed dominant height and survival models facilitated the inclusion of silvicultural treatments and fusiform rust infection effects. Although both models can be integrated analytically and can be fitted in this form using non-linear least squares or maximum likelihood, modifying the model in its differential form allowed us to better evaluate which terms should be modified according to what each parameter represented in the model.

3.5. Conclusions

We tested for the effect of silvicultural treatments (bedding and vegetation control) as they affect the rate of stand mortality in slash pine plantations. Treatment effects were incorporated either implicitly, through the use of dominant height increments which already captured treatment effects, or explicitly through treatment terms in the mortality model form. The use of dominant height increments and explicit parameter modifiers in the mortality model were not effective in improving model fit, implying bedding and vegetation control did not affect mortality rates at or beyond age 5 in the slash pine plantations evaluated in this study. On the other hand, knowledge of fusiform rust infection rates was essential to accurately describe mortality trajectories, with higher fusiform rust infection rates implying higher mortality. Mortality was modeled using differential equations in which the change in the number of trees per hectare at a given time was described by the current number of trees, a power function of age, and the observed average fusiform rust infection rate at age 5. Although fusiform rust has been previously identified by several authors as one of the main drivers of mortality in slash pine plantations, risk of infection remains high in the southeastern United States (Randolph, 2016; Weng et al., 2018). Genetic improvement has proven to be efficient at reducing infestation rates in other pines in the region (Randolph, 2016), suggesting the need to prioritize genetic improvement for rust resistance for the species to reduce the impact in overall value (Susaeta, 2020).

3.6. ACKNOWLEDGEMENTS

This research was conducted with support from the Plantation Management Research Cooperative (PMRC), where resources have been invested in establishing and collecting data from this long-term research trial. Special thanks to the PMRC Field Crew who collected the data for this research over the course of three decades. In addition, this work was supported by the Warnell School of Forestry and Natural Resources at the University of Georgia, the J.L. Clutter Forest Biometrics Scholarship, and the McIntire-Stennis Project GEOZ0202-MS.

3.7. TABLES AND FIGURES

Table 3-1. Dominant height (H_D) and trees per hectare (TPH), average and standard deviation values per plot.

Age (years)	Treatment	No. plots	Mean H_D(m)	SD H_D (m)	Mean TPH	SD TPH
5	Control	18	3.3	0.7	1181.9	157.2
8		18	6.2	1.1	1137.2	194.4
11		18	9.1	1.6	1108.9	194.4
14		18	12.2	2.0	1062.9	221.3
17		18	14.5	2.4	1046.3	218.5
20		18	16.4	2.5	1048.0	231.2
23		17	18.3	2.9	1017.2	241.2
26		16	19.8	3.4	990.8	250.5
31		12	20.9	4.1	972.8	244.4
5	Bedding	19	3.9	0.7	1252.2	127.8
8		19	7.1	1.0	1228.6	142.4
11		19	10.1	1.3	1191.7	132.5
14		19	13.2	1.5	1150.6	149.6
17		19	15.5	2.0	1141.0	151.2
20		19	17.1	2.4	1124.1	167.6
23		18	18.9	2.9	1095.0	190.0
26		17	20.2	3.3	1032.3	225.6
31		11	20.9	4.2	1042.3	272.2
5	Vegetation control (Chem)	17	4.8	0.7	1181.2	130.7
8		17	8.3	0.9	1144.2	171.1
11		17	11.3	1.1	1130.5	186.0
14		17	14.3	1.5	1083.0	242.0
17		17	16.5	1.8	1076.3	249.8
20		16	18.3	2.3	1032.3	251.0
23		15	20.1	2.4	992.1	259.8
26		14	21.5	2.8	964.4	256.1
31		10	22.8	3.2	968.7	276.5
5	Bedding + Vegetation control (Chem)	18	5.3	0.5	1200.9	139.4
8		18	8.9	0.7	1163.8	147.7
11		18	11.9	0.8	1146.5	163.7
14		18	14.9	1.1	1110.0	172.0
17		18	17.1	1.4	1083.0	191.0
20		17	18.6	1.8	1063.3	195.5
23		15	20.4	2.2	1005.3	246.5
26		15	21.9	2.4	963.3	243.2
31		11	22.8	3.0	1038.5	226.9

Table 3-2. Average fusiform rust infection rates at age 5 differentiated by treatment.

Treat	**No. plots**	**No. plots FR < 15%**	**No. plots FR ≥ 15%**	**Avg FR infection % (FR< 15%)**	**Avg FR infection % (FR ≥15%)**	**Avg FR infection % (all plots)**
Control	18	14	4	2.01	28.45	7.89
Bedding	19	14	5	2.01	22.34	7.36
Veg. control	17	14	3	3.56	38.30	9.69
Bed + Veg. control	18	14	4	2.75	36.68	10.29

Table 3-3. Base mortality model forms tested.

Equation	Model form	Model with Fusiform rust	Reference
Eq. 3-10	$\frac{dN}{dt} = \alpha\ N$	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\ N$	(Devine and Clutter, 1985)
Eq. 3-11	$\frac{dN}{dt} = \alpha\ Nt$	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\ Nt$	(Zhao et al., 2007a)
Eq. 3-12	$\frac{dN}{dt} = \alpha\ Nt^{\delta}$	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\ Nt^{\delta}$	(Pienaar and Shiver, 1981)
Eq. 3-13	$\frac{dN}{dt} = \alpha\ N^{\gamma}t$	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\ N^{\gamma}t$	(Zhao et al., 2007a)
Eq. 3-14	$\frac{dN}{dt} = \alpha\ N^{\gamma}t^{\delta}$	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\ N^{\gamma}t^{\delta}$	(Clutter and Jones, 1980)

N: trees per hectare at time t, t: time in years, FR: Fusiform rust infection rate per plot at year 5 (varying from 0 to 1), α, α_0, α_1 , γ, δ: parameters to be estimated.

Table 3-4. Mortality model modifications to include treatment effects.

Model form	Description	Equation
$\frac{dN}{dt} = f(N, t, \theta)$	Base model from Table 3-3	Eq. 3-15
$\frac{dN}{dH_D} = f(N, t, \theta)$	Modified model following Garcia (2009) using dH_D	Eq. 3-16
$\frac{dN}{dH_D} = f(N, H_D, \theta)$	Modified model following Garcia (2009) using dH_D and H_D (no t involved)	Eq. 3-17
$\frac{dN}{dt} = f(N, t, \theta) \cdot c_1^{Z_1} c_2^{Z_2}$	Base model with explicit treatments effect	Eq. 3-18
$\frac{dN}{dH_D} = f(N, t, \theta) \cdot c_1^{Z_1} c_2^{Z_2}$	Modified model following Garcia (2009) with explicit treatments effect	Eq. 3-19
$\frac{dN}{dH_D} = f(N, H_D, \theta) \cdot c_1^{Z_1} c_2^{Z_2}$	Modified model following Garcia (2009) with explicit treatments effect (no t involved)	Eq. 3-20

N: trees per hectare at time t, t: time in years, H_D: dominant height (meters), Z_1: dummy variable equal to 1 if bedding was applied and 0 otherwise, Z_2: dummy variable equal to 1 if vegetation control was applied and 0 otherwise, c_1, c_2, θ: parameters to be estimated.

Table 3-5. Fit statistics for dominant height model without treatments (Eq. 3-3) and with treatments (Eq. 3-4).

Model	RMSE (m)	MD (m)	MAD (m)	AIC
Eq. 3-3-Without treatments	0.7057	-0.0046	0.5469	1436.52
Eq. 3-4-With treatments	0.5724	0.0012	0.4463	1189.01

Table 3-6. Parameter estimates for the dominant height model (Eq. 3-4).

Parameter	**Estimated value**	**Standard error**
μ_0	0.3786	0.0093
K	0.1140	0.0015
b_1	0.9174	0.0102
b_2	1.1039	0.0068
$\ln(\sigma)$	-0.5579	0.0288

Table 3-7. Fit statistics[1] for base survival models (control only, no-treatment effects).

Eq.	Model form	RMSE (TPH)	MD (TPH)	MAD (TPH)	AIC	Log-like
Eq. 3-10	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\ N$	41.55	0.22	30.08	-533.18	-269.59
Eq. 3-11	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\ Nt$	52.79	0.31	35.60	-459.88	-232.94
Eq. 3-12	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\ Nt^{\delta}$	35.54	0.02	26.37	-579.00	-293.50
Eq. 3-13	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\ N^{\gamma}t$	57.58	-0.001	38.37	-431.31	-219.66
Eq. 3-14	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\ N^{\gamma}t^{\delta}$	35.46	0.003	26.23	-577.68	-293.84

[1] Fit statistics were scaled to the original value by multiplying by 1,000.

N: trees per hectare at time t, t: time in years, FR: Fusiform rust infection rate per plot at year 5 (varying from 0 to 1), α_0, α_1, γ, δ: parameters to be estimated.

Table 3-8. Fit statistics[1] for survival models including treatment effects.

Eq.	Model form	RMSE (TPH)	MD (TPH)	MAD (TPH)	AIC	Log-like
Eq. 3-21	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)Nt^{\delta}$	45.63	-0.01	32.59	-1997.18	-1002.59
Eq. 3-22	$\frac{dN}{dH_D} = (\alpha_0 + \alpha_1 * FR)Nt^{\delta}$	48.89	0.00	34.73	-1914.32	-961.16
Eq. 3-23	$\frac{dN}{dH_D} = (\alpha_0 + \alpha_1 * FR)NH_D^{\delta}$	48.56	-0.10	34.58	-1922.35	-965.18
Eq. 3-24	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)c_1^{Z_1}c_2^{Z_2}Nt^{\delta}$	45.53	0.01	32.57	-1995.77	-1003.88
Eq. 3-25	$\frac{dN}{dH_D} = (\alpha_0 + \alpha_1 * FR)c_1^{Z_1}c_2^{Z_2}Nt^{\delta}$	48.75	0.02	34.76	-1913.65	-962.82
Eq. 3-26	$\frac{dN}{dH_D} = (\alpha_0 + \alpha_1 * FR)c_1^{Z_1}c_2^{Z_2}NH_D^{\delta}$	48.49	-0.09	34.54	-1920.10	-966.05

[1] Fit statistics were scaled to the original value by multiplying by 1,000.

N: trees per hectare at time t, t: time in years, FR: Fusiform rust infection rate per plot at year 5 (varying from 0 to 1), H_D: dominant height (meters), Z_1: dummy variable equal to 1 if bedding was applied and 0 otherwise, Z_2: dummy variable equal to 1 if vegetation control was applied and 0 otherwise, α_0, α_1, c_1, c_2, δ: parameters to be estimated

Table 3-9. Parameter estimates for the final survival model including fusiform rust and treatment effects.

Parameter	Estimated value	Standard error estimate
α_0	-0.0091	0.0020
α_1	-0.2272	0.0472
δ	-0.4306	0.0798
$\ln(\sigma)$	-3.0871	0.0288

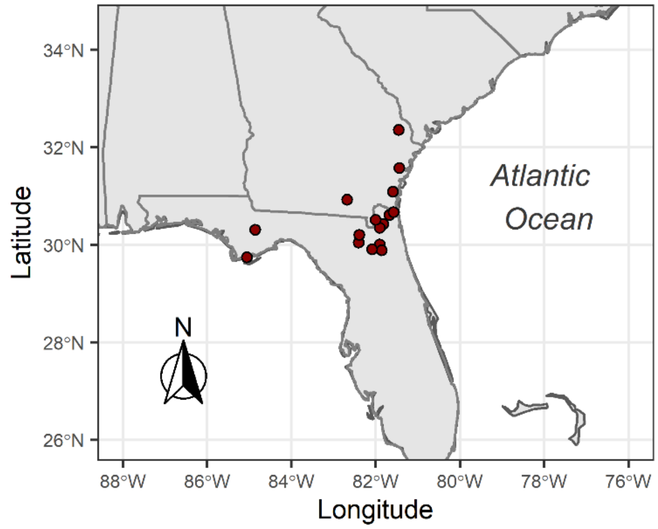

Figure 3-1. Location of the 16 installations of the study.

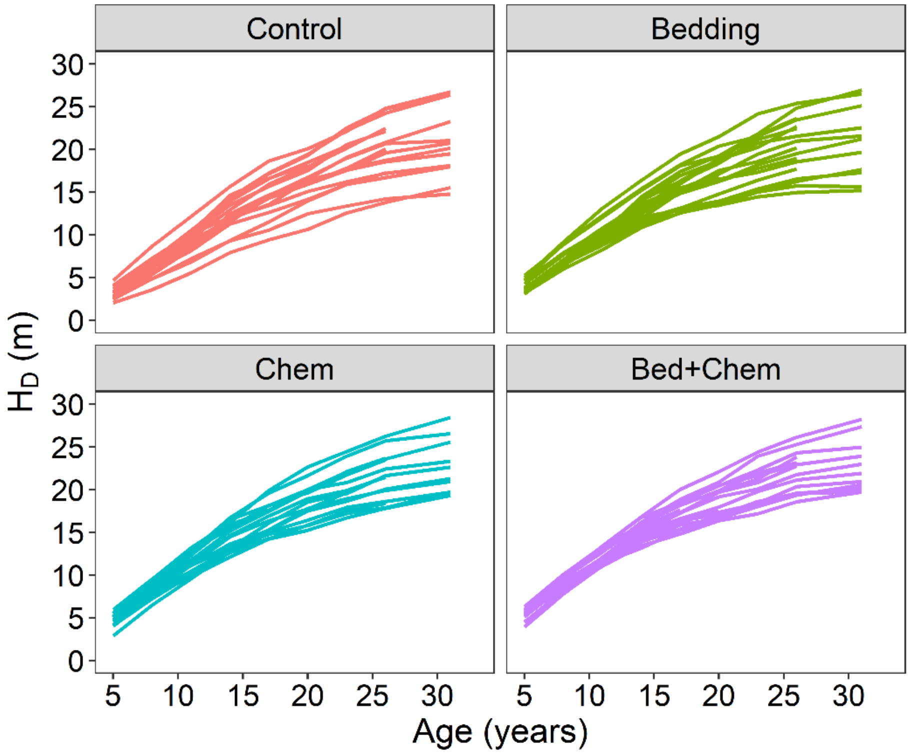

Figure 3-2. Dominant height trajectories for slash pine by treatment.

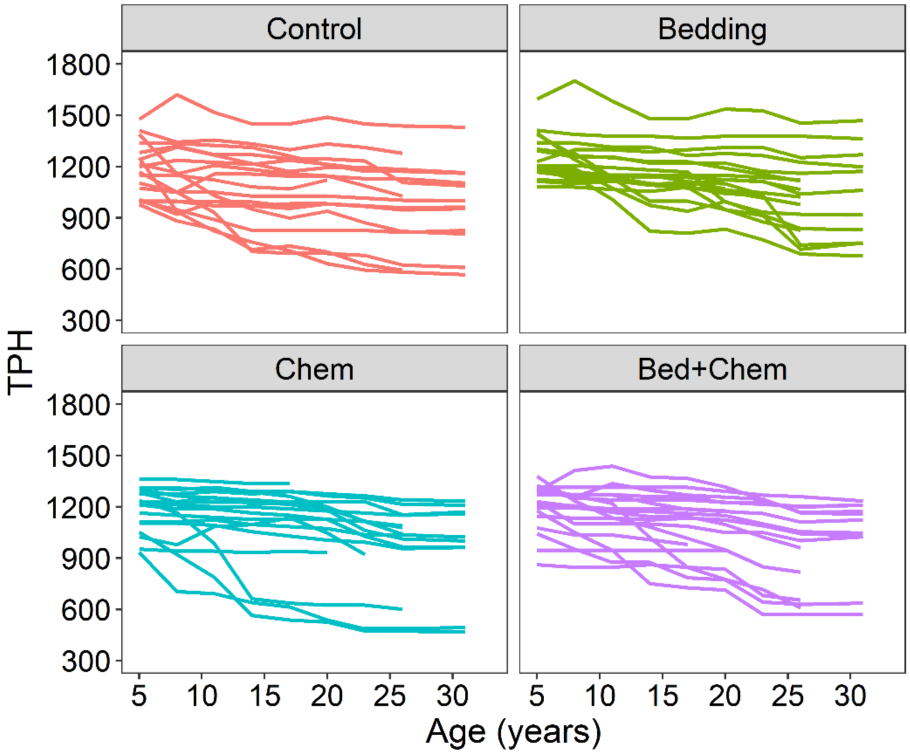

Figure 3-3. Trees per hectare (TPH) over time by treatment.

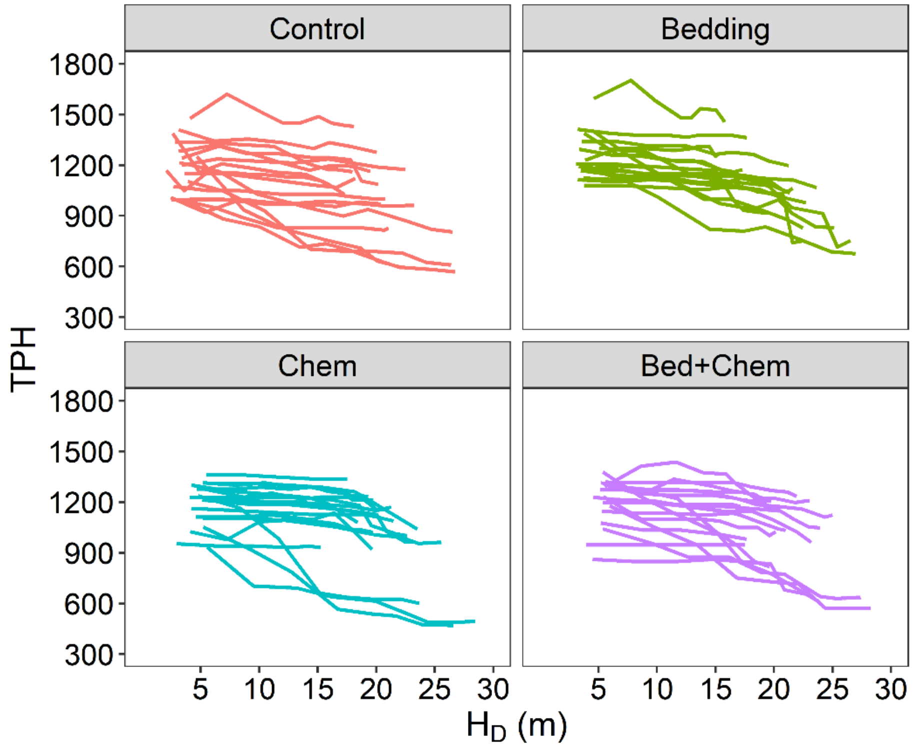

Figure 3-4. Trees per hectare (TPH) versus dominant height (H_D) by treatment.

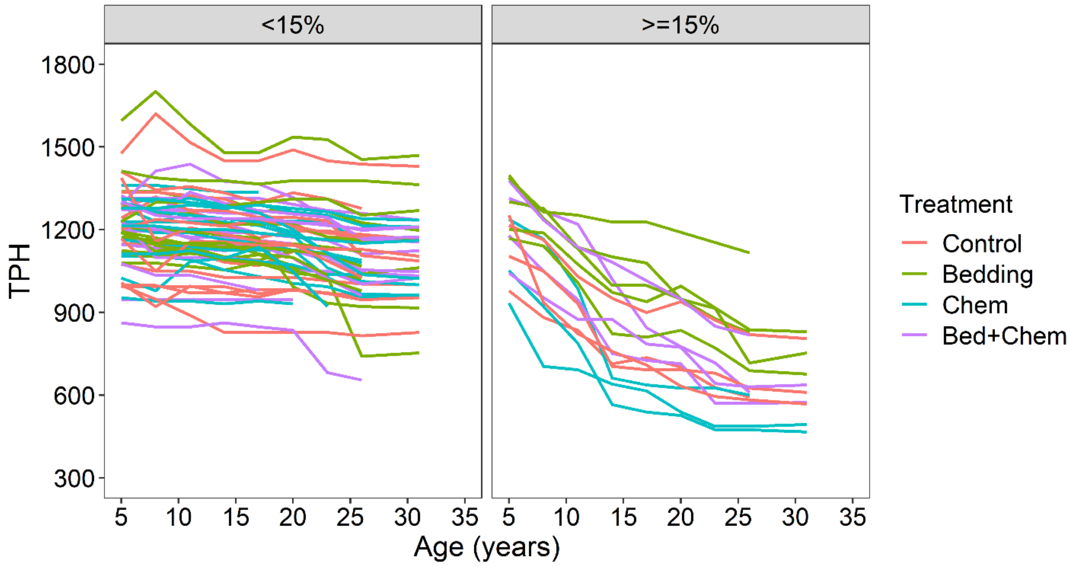

Figure 3-5. Trees per hectare (TPH) over time for different groups of fusiform rust infection rate (less than 15% or greater than or equal to 15%) at age 5.

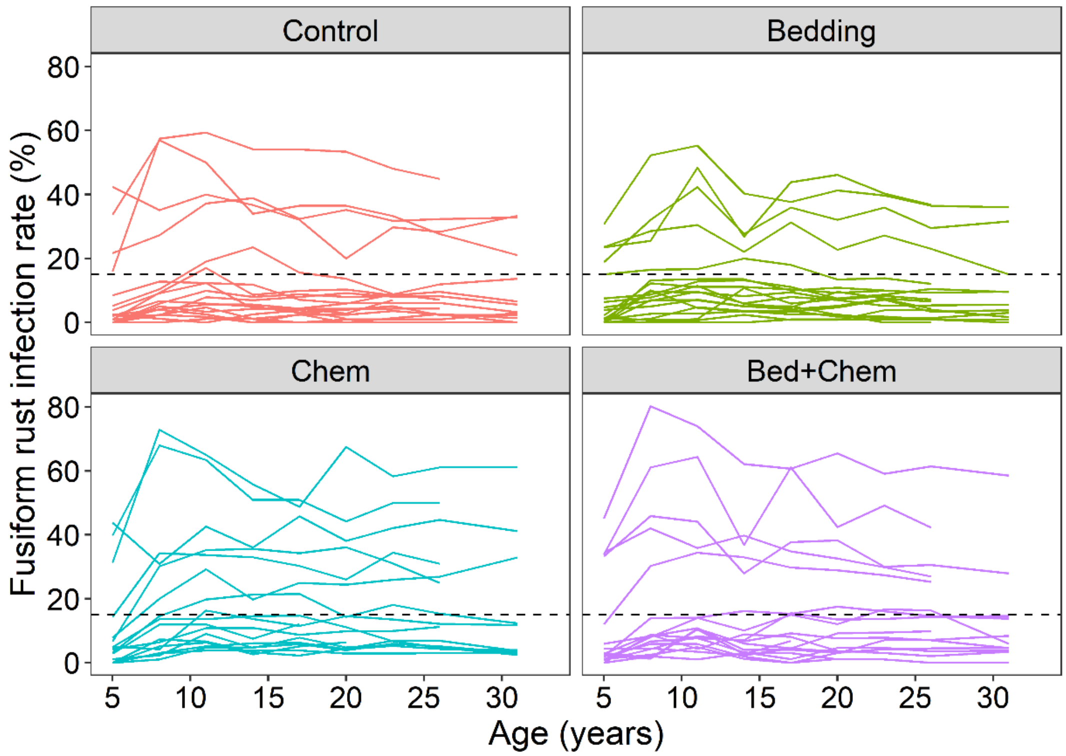

Figure 3-6. Fusiform rust average infection rate per plot, trajectory for the duration of the study. The dashed line represents the 15% infection rate threshold.

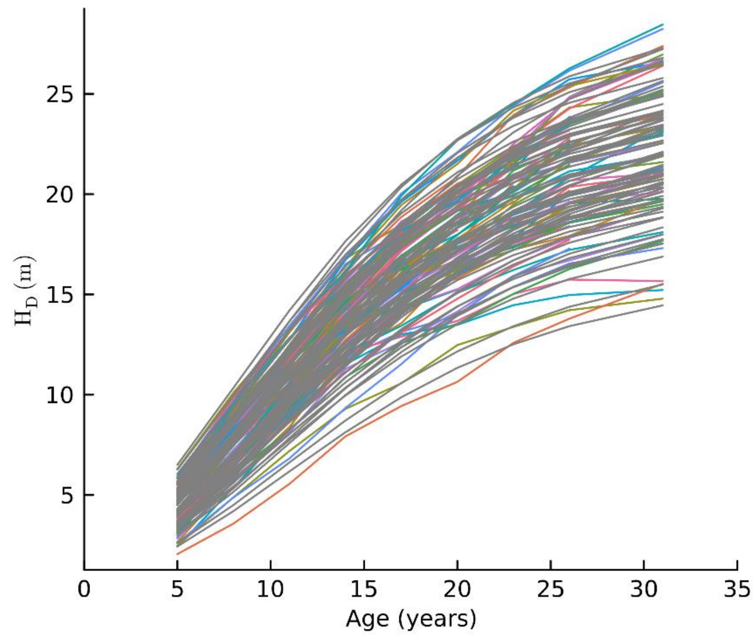

Figure 3-7. Estimated dominant height trajectories.

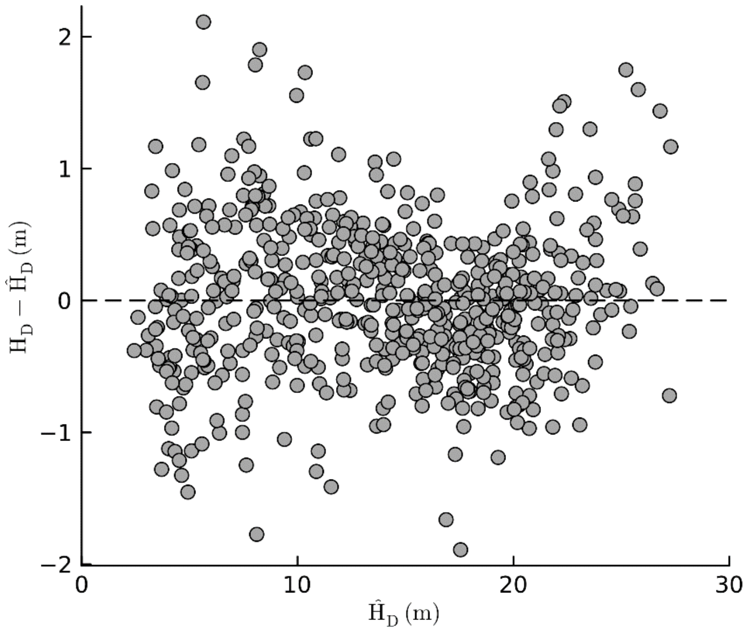

Figure 3-8. Dominant height residuals vs predicted dominant height values.

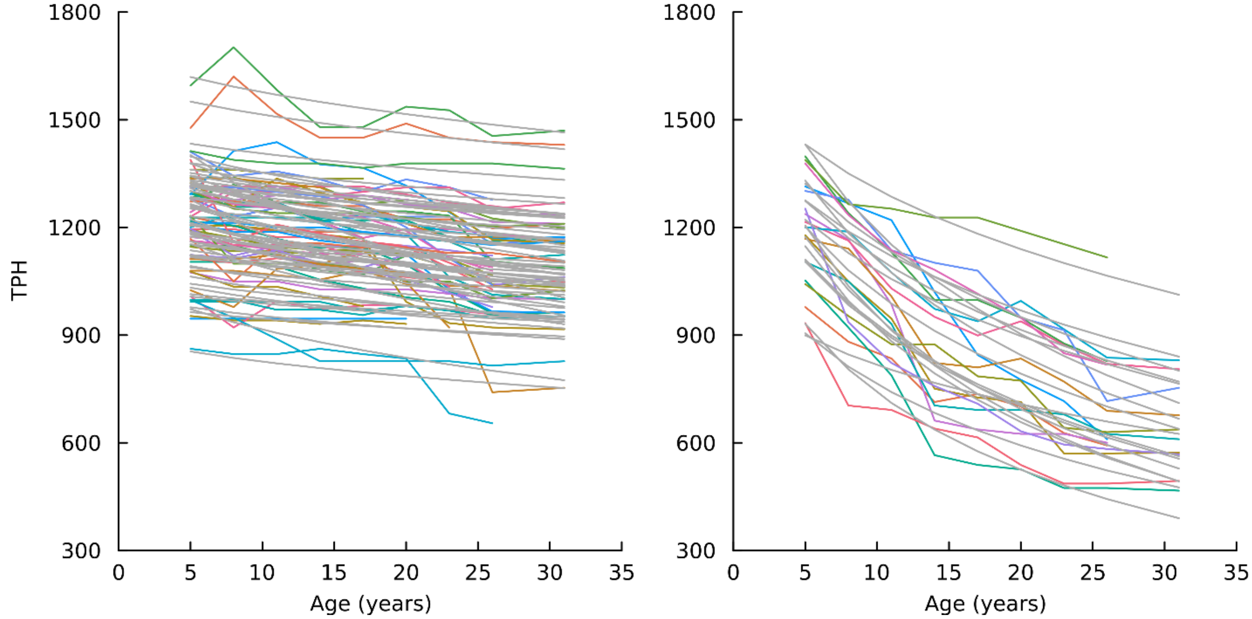

Figure 3-9. Trajectories of TPH estimated with Eq. 3-21 for plots with less than 15% (left), and higher than 15% (right) fusiform rust infection rates.

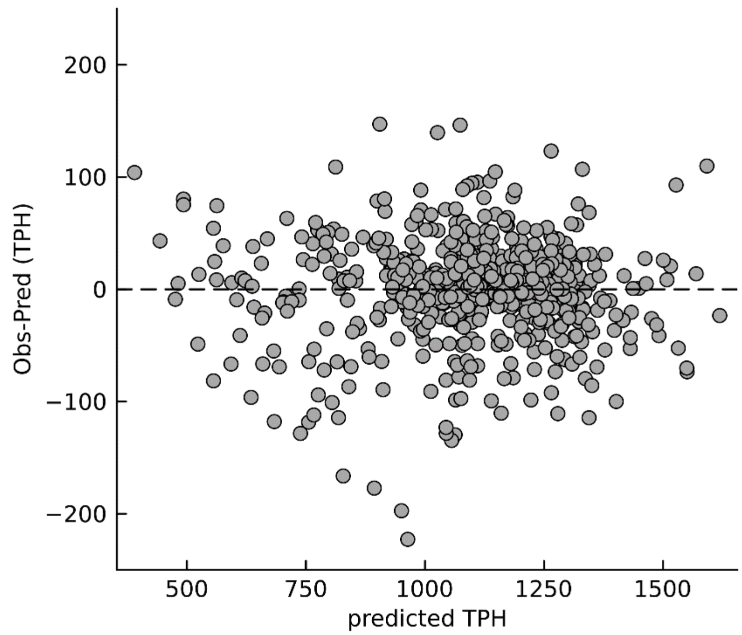

Figure 3-10. Residuals for the recommended mortality model (Eq. 3-21).

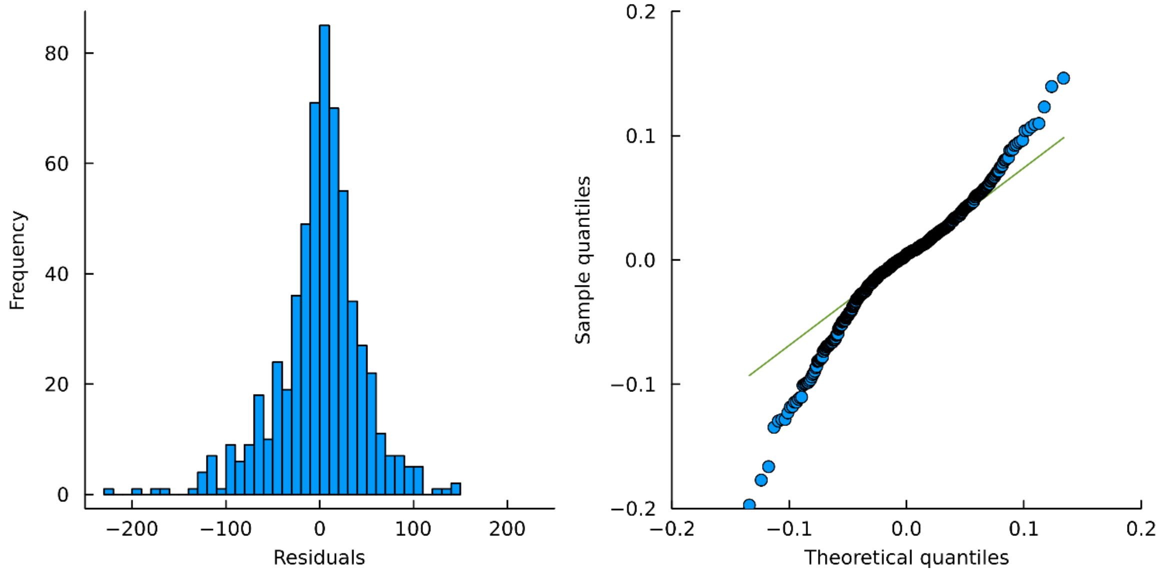

Figure 3-11. Residuals histrogram and Q-Q plot for the recommended mortality model including treatments effect (Eq. 3-21).

3.8. References

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

Bailey, R.L., Burgan, T.M., 1989. Fertilized midrotation-aged slash pine plantations-stand structure and yield prediction models. South. J. Appl. For. 13, 76–80.

Barnett, J.P., Sheffield, R.M., 2004. Slash pine: characteristics, history, status and trends. In: Dickens, E.D.; Barnett, J.P.; Hubbard, W.G.; Jokela, E.L., eds. Slash pine: still growing and growing! Proc. slash pine Symp. Gen. Tech. Rep. SRS-76. Asheville, NC U.S. Dep. Agric. For. Serv. South. Res. Station. 1–6.

Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B., 2017. Julia : A fresh approach to numerical computing. SIAM Rev. 59, 65–98. https://doi.org/10.1137/141000671. pdf

Burton, J.D., Shoulders, E., Snow, G.A., 1985. Incidence and impact of fusiform rust vary with silviculture in slash pine plantations. For. Sci. 31, 671–680.

Cieszewski, C.J., Bailey, R.L., 2000. Generalized algebraic difference approach: theory based derivation of dynamic site equations with polymorphism and variable asymptotes. For. Sci. 46, 116–126.

Clutter, J.L., Jones, J.E.P., 1980. Prediction of growth after thinning in old-field slash pine plantations. Dep. Agric. For. Serv. Southeast. For. Exp. Stn. Vol. 217.

Creighton, J.L., Zutter, B.R., Glover, G.R., Gjerstad, D.H., 1987. Planted pine growth and survival responses to herbaceous vegetation control, treatment duration, and herbicide application technique. South. J. Appl. For. 11, 223–227. https://doi.org/10.1093/sjaf/11.4.223

Devine, O.J., Clutter, J.L., 1985. Prediction of survival in slash pine plantations infected with fusiform rust. For. Sci. 31, 88–94.

Fox, T.R., Jokela, E.J., Allen, H.L., 2007. The development of pine plantation silviculture in the Southern United States. J. For. 105, 337–347. https://doi.org/10.1093/jof/105.7.337

France, J., Thornley, J.H., 1984. Mathematical models in agriculture. Butterworths.

Garcia, O., 2009. A simple and effective forest stand mortality model. Int. J. Math. Comput. For. Nat. Sci. 1, 1–9. https://doi.org/10.1136/bmj.317.7152.199

Garcia, O., 2010. A parsimonious dynamic stand model for interior spruce in British Columbia. For. Sci. 57, 265–280.

Garcia, O., Burkhart, H.E., Amateis, R.L., 2011. A biologically-consistent stand growth model for loblolly pine in the Piedmont physiographic region, USA. For. Ecol. Manage. 262, 2035–2041. https://doi.org/10.1016/j.foreco.2011.08.047

Gent, J.A., Allen, H.L., Campbell, R.G., Wells, C.G., 1986. Magnitude, duration, and economic analysis of loblolly pine growth response following bedding and phosphorus fertilization. South. J. Appl. For. 10, 124–128. https://doi.org/10.1093/sjaf/10.3.124

Gyawali, N., Burkhart, H.E., 2015. General response functions to silvicultural treatments in loblolly pine plantations. Can. J. For. Res. 45, 252–265. https://doi.org/10.1139/cjfr-2014-0172

Jokela, E.J., Wilson, D.S., Allen, J.E., 2000. Early growth responses of slash and loblolly pine following fertilization and herbaceous weed control treatments at establishment. South. J. Appl. For. 24, 23–30. https://doi.org/10.1093/sjaf/24.1.23

Jones, J.E.P., 1972. Fusiform rust affects planted slash pine. J. For. 70, 350–352.

Lauer, D.K., Glover, G.R., 1998. Early pine response to control of herbaceous and shrub vegetation in the flatwoods. South. J. Appl. For. 22, 201–208. https://doi.org/10.1093/sjaf/22.4.201

Lee, Y., 1971. Predicting mortality for even-aged stands of lodgepole pine. For. Chron. 47, 29–32. https://doi.org/10.5558/tfc47029-1

Martin, S.W., Bailey, R.L., Jokela, E.J., 1999. Growth and yield predictions for lower coastal plain slash pine plantations fertilized at mid-rotation. South. J. Appl. For. 23, 39–45. https://doi.org/10.1093/sjaf/23.1.39

Mason, E.G., Milne, P.G., 1999. Effects of weed control, fertilization, and soil cultivation on the growth of Pinus radiata at midrotation in Canterbury, New Zealand. Can. J. For. Res. 29, 985–992. https://doi.org/10.1139/x99-087

Medeiros, R.A., De Paiva, H.N., Soares, A.A.V., Da Cruz, J.P., Leite, H.G., 2017. Thinning from below: Effects on height of dominant trees and diameter distribution in Eucalyptus stands. J. Trop. For. Sci. 29, 238–247.

Morris, L.A., Lowery, R.F., 1988. Influence of site preparation on soil conditions affecting stand establishment and tree growth. South. J. Appl. For. 12, 170–178.

Nance, W.L., Froelich, R.C., Shoulders, E., 1981. Effects of fusiform rust on survival and structure of Mississippi and Louisiana slash pine plantations. US Dep. Agric. For. Serv. South. For. Exp. Stn. Vol. 232.

Pienaar, L., Rheney, J., 1995. Modeling stand level growth and yield response to silvicultural treatments. For. Sci. 41, 629–638. https://doi.org/10.1093/forestscience/41.3.629

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

Pritchett, W.L., 1979. Site preparation and fertilization of slash pine on a wet savanna soil. South. J. Appl. For. 3, 86–90.

Ramirez, L., Montes, C.R., Bullock, B.P., 2022. Long-term term effect of bedding and vegetation control on dominant height of slash pine plantations in the southeastern United States. For. Ecol. Manage. 522, 120479. https://doi.org/10.1016/j.foreco.2022.120479

Randolph, K.C., 2016. Status of fusiform rust incidence in slash and loblolly pine plantations in the southeastern United States. Gen. Tech. Rep. South. Res. Station. USDA For. Serv. 212, 53–59.

Sluder, E.R., 1977. Fusiform rust in loblolly and slash pine plantations on high-hazard sites in Georgia. For. Serv. US Dep. Agric. Southeast. For. Exp. Stn. Vol. 160.

Susaeta, A., 2020. Implications of future risk of fusiform rust on optimal forest management of even-aged slash pine plantations. For. Policy Econ. 116, 102183. https://doi.org/10.1016/j.forpol.2020.102183

Tewari, V.P., Singh, B., 2018. A first-approximation simple dynamic growth model for forest teak plantations in Gujarat state of India. South. For. 80, 59–65. https://doi.org/10.2989/20702620.2016.1277644

Stankova, T.V., Diéguez-Aranda, U., 2014. Derivation and analysis of new stand-level mortality models based on existing growth equations. Ecol. Res. 29, 319–330. https://doi.org/10.1007/s11284-014-1126-5

Wells, O.O., Dinus, R.J., 1978. Early infection as a predictor of mortality associated with fusiform rust of southern pines. J. For. 76, 8–12.

Weng, Y.H., Coble, D.W., Grogan, J., Stovall, J.P., 2018. Temporal trends in fusiform rust infections and their relationships with stand structure in pine plantations in East Texas. J. For. 116, 420–428. https://doi.org/10.1093/jofore/fvy035

Zang, H., Lei, X., Ma, W., Zeng, W., 2016. Spatial heterogeneity of climate change effects on dominant height of larch plantations in Northern and Northeastern China. Forests 7. https://doi.org/10.3390/f7070151

Zhao, D., Borders, B., Wang, M., Kane, M., 2007a. Modeling mortality of second-rotation loblolly pine plantations in the Piedmont/Upper Coastal Plain and Lower Coastal Plain of the southern United States. For. Ecol. Manage. 252, 132–143. https://doi.org/10.1016/j.foreco.2007.06.030

Zhao, D., Kane, M., Borders, B., Harrison, M., 2009. Long-term effects of site preparation treatments, complete competition control, and repeated fertilization on growth of slash pine plantations in the flatwoods of the southeastern United States. For. Sci. 55, 403–410. https://doi.org/10.1093/forestscience/55.5.403

Zhao, D., Kane, M., Borders, B., Harrison, W.M., 2007b. Slash pine site preparation study: Results through age 26. Plantation Management Research Cooperative (PMRC). Tech. Rep. 2007-3 61.

CHAPTER 4

A GROWTH AND YIELD SYSTEM OF DIFFERENTIAL EQUATIONS FOR SLASH PINE PLANTATIONS INCLUDING RESPONSE TO SILVICULTURAL TREATMENTS[1]

[1] Ramirez, L., Montes, C. R. & Bullock, B. P. Accepted by *Proceedings of the 22nd Biennial Southern Silvicultural Research Conference*. Reprinted here with permission of publisher.

ABSTRACT

Slash pine is the second most important commercial species in the southeastern United States, it is usually established in poorly drained flatwoods where it outperforms other common commercial pine species. Modeling slash pine growth and how it responds to silvicultural treatments is of interest to forest managers wanting to maximize their investments in the region. In this research, a growth and yield (G&Y) system of differential equations is proposed to model slash pine growth including the effect of silvicultural treatments (i.e., bedding, and vegetation control). Data for this model came from a long-term study (30 years) established by the Plantation Management Research Cooperative (PMRC) across Georgia, Florida, and South Carolina. The model system describes the trajectory of three state variables: dominant height, survival/mortality, and basal area. Treatments effects were incorporated into the dominant height and basal area models by using parameter modifiers and dummy variables associated with each one of the treatments. Survival was not affected by the studied treatments, but the presence of fusiform rust was found to be essential to determine the stand density trajectories for the evaluated stands. The three models were estimated simultaneously using maximum likelihood and the variance/covariance was modeled within the system. The G&Y system was validated with a slash pine independent dataset.

4.1. INTRODUCTION

Growth models in forestry, when not entirely empirical, have been constructed by trying to understand the underlying biological processes that generate net growth (Pienaar and Turnbull, 1972; Zeide, 1993). When constructing these models, differential equations have been useful to describe net growth as the interaction between two contrasting forces, one generating growth and one limiting it (Zeide, 1993). Some of the most commonly used models in forestry as the

Schumacher and the Chapman-Richards model (Pienaar and Turnbull, 1972), when expressed in the differential form, rely on growth theories based on assumptions about how organisms grow (e.g., growth proportional to surface area) and how this growth is limited (e.g., maximum carrying capacity). Therefore, parameters in these models have a biological interpretation which helps researchers to understand the underlying processes in a simplified way. This fact also facilitates the modification or construction of these models when the growth dynamic is changed by factors like silvicultural treatments or external pathogens.

Although many of the models used in forestry were originally developed from differential equations, the integrated form is usually used when fitting these models to data. Common statistical software used in forestry have facilitated model fit with routines and packages that can accommodate these integrated models (e.g., nls in *R*). Nevertheless, these software also have the ability to solve differential equations numerically, which combined with parameter estimation techniques like maximum likelihood, allows users to fit models directly using the differential form. When the equation can be integrated analytically, there is no difference between either approach used to fit the model. Nevertheless, when it is not possible to get a closed form of the model, using the differential equation is the only possibility. This in turn opens the door to new opportunities like adding stochastic components (e.g., random variation) making use of stochastic differential equations (Garcia, 1979).

Different strategies have been used to include silvicultural treatment response in forest growth models. These usually consist of adding additional factors to the (integrated) model that account for the treatment response (Pienaar and Rheney, 1995), modifying the parameters of the model according to the treatment (Mason and Milne, 1999), adding additional covariables related to the treatment (Hynynen, 1995), or using an age-shift approach in which the basic structure of the model is not modified (Snowdon, 2002). Using parameters modifiers is a flexible methodology that allows for the inclusion of multiple treatments and their interaction (Ramirez et al., 2022). Using models in the differential equation form facilitates the way treatments responses are incorporated into the growth model since the parameters that should be modified become more intuitive. Thus, treatments that aim to increase growth rate (e.g., fertilization), can be modeled by modifying parameters that describe growth in the model, while parameters that describe the decline or limiting factors within the model can be modified if a treatment targets the reduction of limiting factors (e.g., vegetation control). This strategy was used on this study to model the response to bedding and vegetation control for slash pine plantations in the southeastern United States. Growth responses were evaluated on three state variables, dominant height, survival, and basal area.

The growth and yield (G&Y) system constructed in this research incorporates the effect of bedding and vegetation control on dominant height and basal area growth. Parameters modifiers were tested to include this response. Survival was not affected by the treatments evaluated, but the presence of fusiform rust was evaluated as a potential predictor to help describe mortality for the evaluated stands. Fusiform rust is an endemic disease that started to significantly affect plantations in the southeastern United States in the 1960s (Randolph, 2016). Slash pines have been particularly affected by this pathogen given that the fungi generating the disease (*Cronartium quercuum* Berk.) completes part of its reproductive cycle in these pines (Phelps, 1978). Fusiform rust is currently the most common pathogen affecting slash pine in the southeastern United States, being associated with higher mortality rates and loses in product quality (Bailey and Burgan, 1989; Devine and Clutter, 1985; Nance et al., 1981; Susaeta, 2020).

The three models composing the system were first fitted independently before fitting them simultaneously using maximum likelihood. All models were fitted using the differential equation form where the initial value for solving the differential equation was taken as an additional parameter to be estimated on each one of the models (i.e., local parameters per plot). These models were fitted using the Julia programming language (Bezanson et al., 2017). The G&Y system was tested against independent data coming from a different slash pine study stablished in the southeastern United States.

4.2. METHODS

4.2.1. DATA

Data from the three variables of interest were available from a slash pine site preparation trial established in 1979 by the Plantation Management Research Cooperative (PMRC) in the southeastern United States. Several treatments including bedding, vegetation control, and fertilization, were part of the mentioned study. Nevertheless, for this research, the focus was on bedding, competing vegetation control and the combination of these two treatments. The bedding treatment consisted of a double pass with a bedding harrow during site preparation and the vegetation control treatment included an herbicide application before site preparation (3% solution of Roundup) and repeated localized applications of Roundup or Garlon to remove most of the competing vegetation until crown closure (Zhao et al., 2009).

Measurements of diameter at breast height, dominant height, stand density, and average fusiform rust infection rates per plot were available every 3 years from age 5 to age 31 for the longest series (some plots were lost at earlier ages due to external factors). In Table 4-1 a summary of the dominant height, number of trees per hectare, basal area, and average fusiform rust infection rate for each one of the treatments and different ages is presented. More details about the study are provided by Zhao et al., (2009), Zhao et al., (2007) and Ramirez et al., (2022). Dominant height, survival, and basal area trajectories observed are plotted in Figure 4-1.

4.2.2. Dominant height model

A model form that had previously been used to model dominant height with the inclusion of bedding and vegetation control treatment effects was used (Ramirez et al., 2023). This model is based on the Gompertz equation (Eq. 4-1), and with the inclusion of the two silvicultural treatments of interest has the form of Eq. 4-2. For the construction of the growth and yield system, parameters of the model (Eq. 4-2) were not taken directly from the Ramirez et al. (2023) publication, but were estimated following the methodology explained in section 4.2.5.

$$\frac{dH}{dt} = \mu_0 H e^{-Kt} \qquad \text{Eq. 4-1}$$

$$\frac{dH}{dt} = \mu_0 b_1^{Z_1} H e^{-K b_2^{Z_2} t} \qquad \text{Eq. 4-2}$$

In these models, H is dominant height (m), t is time in years, μ_0, K, b_1, and b_2 are parameters to be estimated, and Z_1 and Z_2 are dummy variables equal to 1 if bedding or vegetation control was applied, respectively, and zero otherwise.

4.2.3. Mortality model

Mortality rates in slash pine plantations in the southeastern United States have been demonstrated to be influenced by the presence of fusiform rust infection (Jones, 1972; Sluder, 1977; Wells and Dinus, 1978). Recently, Ramirez et al. (2023) developed a mortality/survival model in which the rate of mortality depended on the proportion of fusiform rust infection at year 5. This model form (Eq. 4-3) was used in this research as part of the growth and yield system.

$$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\, N t^{\delta} \qquad \text{Eq. 4-3}$$

In this model N is trees per hectare, t is time in years, α_0 , α_1 , and δ are parameters to be estimated, and FR is the proportion of trees affected by fusiform rust on a given plot. Similar to the dominant height component, parameters for this model were estimated as part of the system and were not taken directly from Ramirez et al. (2023). The effect of bedding and vegetation control on survival of slash pine plantations was found to be non-significant by Ramirez et al. (2023), and therefore, this model was not further modified to include treatments effects.

4.2.4. Basal area model

Model from Eq. 4-4 was proposed to describe basal area trajectories. This model is based on the Korf growth model (Korf, 1939; Zarnovican, 1979) although it was modified to include dominant height and mortality as covariables, similar to the basal area model proposed by Pienaar and Shiver (1986) for slash pine plantations.

$$\frac{dB}{dt} = [\beta_1 + \beta_2 H + \beta_3 N] \frac{B}{t^{\beta_4}} \qquad \text{Eq. 4-4}$$

In this model, B is basal area (m^2/ha), H, N, and t are as defined before, and $\beta_1..\beta_4$ are parameters to be estimated. An analysis of variance (ANOVA) showed that there were significant differences in basal area for the treatments evaluated at every age. Therefore, the model from Eq. 4-4 was modified similarly to the dominant height model to include treatment effects as follows:

$$\frac{dB}{dt} = [\beta_1 + \beta_2 H + \beta_3 N] \frac{B}{t^{\beta_4}} \; d_1^{Z_1} d_2^{Z_2} \qquad \text{Eq. 4-5}$$

Where Z_1 and Z_2 are the same dummy variables related to the treatments and defined for the dominant height model, and d_1 and d_2 are additional parameters to be estimated.

In order to determine if the proposed modifications to the dominant height and basal area models were effective to include treatment effects, two G&Y systems were proposed and compared in this chapter, one with no treatment effects (System 1) and another (System 2) with the treatment effects in both components (dominant height and basal area). These systems are described in Table 4-2. The system with better fit statistics was chosen to model the three variables of interest.

4.2.5. Parameter estimation

Parameters for the two systems proposed were estimated using maximum likelihood in the Julia programing language (Bezanson et al., 2017). Thus, when fitting the system, the differential equations are solved analytically and the predicted values are compared to the observed data. The estimated (best) parameters are the ones that maximize the likelihood value assuming errors are normally distributed with zero mean. The parameter describing the variance of the normal distribution assumed for the errors is also estimated together with the parameters of the models. In addition to this parameter and to the parameters for each one of the models from Table 4-2 (i.e., global parameters), the initial values that define the trajectories (in any of the three variables) for each one of the plots (i.e., local parameters), were defined as additional parameters to be estimated as part of the optimization procedure.

Each one of the models from Table 4-2 was first fitted individually assuming a normal distribution of the errors (i.e., $H: \epsilon_i \sim N(0, \sigma_H)$, $N: \epsilon_i \sim N(0, \sigma_N)$, $B: \epsilon_i \sim N(0, \sigma_B)$). Then, the G&Y system was simultaneously estimated by fixing the local parameters previously estimated per plot. That is, when simultaneous estimation was done, only the global parameters were estimated, this was necessary to simplify the estimation procedure. For the simultaneous estimation, a multinormal distribution (MN) was used in which the errors are described by the following variance/covariance matrix:

$$\vec{\epsilon} \sim MN(\vec{0}, \Sigma), \quad \Sigma = \begin{bmatrix} \sigma_H^2 & \sigma_{HxN} & \sigma_{HxB} \\ .. & \sigma_N^2 & \sigma_{NxB} \\ .. & .. & \sigma_B^2 \end{bmatrix}$$

Where, $\vec{\epsilon}$ is the vector of errors (observed-predicted), Σ is the variance/covariance matrix (symmetrical), σ_X^2 is the variance of the X variable (H, N or B), and σ_{XxY} is the covariance between variables X and Y. During the estimation procedure, matrix Σ was rewritten in terms of the correlation between the variables as follows:

$$\Sigma = \begin{bmatrix} \sigma_H^2 & \sigma_{HxN} & \sigma_{HxB} \\ .. & \sigma_N^2 & \sigma_{NxB} \\ .. & .. & \sigma_B^2 \end{bmatrix} = \begin{bmatrix} \sigma_H^2 & \sigma_H \sigma_N \rho_{HxN} & \sigma_H \sigma_{BA} \rho_{HxB} \\ .. & \sigma_N^2 & \sigma_N \sigma_{BA} \rho_{NxB} \\ .. & .. & \sigma_B^2 \end{bmatrix}$$

Thus, the additional parameters estimated for this matrix were the correlation (and not directly the covariance) between the errors of the three variables of the system. The off-diagonal values of the Σ matrix were first assumed to be zero (i.e., zero correlation between the errors of the three state variables), to calibrate the initial values of the optimization and then, these values were estimated together with all the global parameters. An approximation of the standard error for the parameters was calculated using the Hessian matrix. For all models, the variable N was scaled by dividing the values by a factor of 1,000 to facilitate the optimization process. The following statistics were calculated to evaluate models fit.

- Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n-p}}$$

- Mean Difference (MD) or bias:

$$MD = \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)}{n}$$

- Mean Absolute Difference (MAD):

$$MAD = \frac{\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|}{n}$$

- AIC: for model comparison

$$AIC = -2\ loglik + 2p$$

Where n is the total number of observations, p is the number of parameters in each model, Y_i is the observed value, $\widehat{Y}_i$ is the estimated value.

4.2.6. Validation dataset

The proposed model was validated with an independent dataset available from another PMRC study. This study was established in 1976 on existing slash pine plantations located in south Georgia and north Florida (Oppenheimer et al., 1989). These plantations ranged in age from 9 to 15 years when the study was established and the first measurement was taken. Vegetation control was part of this study, but since this treatment was applied at later ages and not during the establishment phase (as done for the main slash pine dataset), only the control plots were selected for validation. More details about this study are described by Oppenheimer et al., (1989). In Table 4-3 a summary of the variables of interest for this dataset is presented. When compared to the main dataset, dominant height and basal area values are within the range of the original data. The main differences between the two datasets are the range of stand density (TPH) and fusiform rust infection values, which are higher for the validation dataset.

4.3. Results

Results comparing the two growth and yield systems tested are presented in Table 4-4. The AIC value for System 1 (no treatments) was 1,536.86, while for System 2 it was 778.36, indicating that the system with treatment factors performed better than the system without treatment effects (lower AIC values indicate better fit). From Table 4-4 it can be seen how there is significant gain in precision (reduction of error and bias) when using treatment modifiers. The average error (RMSE) for the dominant height model was reduced from 0.71 to 0.57 m (20% reduction), while the error for the basal area model was reduced from 1.88 m^2/ha to 1.25 m^2/ha (34% reduction). The same percentual reduction was obtained in bias (MD) for these two variables. The values for mortality are not shown due to this model did not change between the systems. Ramirez et al., (2023) showed that mortality was not affected by bedding and vegetation control, but that the inclusion of fusiform rust was crucial to describe the survival trajectories. An additional comparison was then made for the mortality component by fitting the same model but excluding the fusiform rust information (i.e., $dN/dt = \alpha N t^{\delta}$). When fusiform rust was not included into the model, the average error (RMSE) increased to 77.53 TPH and the bias increased to -0.43 TPH. This means that when including fusiform rust, the RMSE was reduced in a 41% and the bias in a 98%. Given these results, System 2 was selected as the best G&Y system and therefore, the results presented in this section refer to this system.

Parameter estimates and fit statistics for each one of the components of System 2 are presented in Table 4-5 and Table 4-6, respectively. The estimated values in Table 4-5 correspond to the values obtained with the simultaneous estimation of the global parameters using the multivariate normal in which the correlation between the variables was estimated and the local values per plot were taken from the independent estimation. When the correlation was assumed to be zero, parameter estimates were not different for the dominant height model but differed for the mortality and basal area models. When comparing the estimated values of the standard error of the parameters, no differences were found for the dominant height model and opposite trends were found for the other two variables, with an increase in the standard error for the mortality parameters, but a decrease in the standard error for the basal area parameters when including correlation. Despite these differences, the fit statistics were not significantly different when comparing the systems with and without correlation (Table 4-6).

Although no improvement was obtained in terms of model fit when accounting for correlation, accounting for this value during the estimation procedure is statistically sound and therefore, the estimated parameter values from Table 4-5 were used to plot the predicted trajectories for the three variables of interest in Figure 4-2. The residuals for these models are presented in Figure 4-3. Overall, a good fit was obtained and the residuals did not show any trend.

The robustness of the G&Y system proposed was evaluated from two different perspectives. First, by using the same dataset but modifying how predictions were done, and then by using the independent dataset described in section 4.2.6. The first assessment consisted in using only the estimated global parameters and utilizing the observed values as starting points of the differential equations to predict the whole trajectories forward. Different starting points were tested. For example, the first observation (at age 5) was used first as the starting point to predict the whole series until year 31, then, the second observation (at age 8) was used to predict the values from this age forward (until age 31), and this was repeated until the second to last observation was used to predict the last observation.

It is well known in forestry that accuracy and precision decreases when the projection interval increases (Wang et al., 2020). This fact was particularly important for the G&Y system proposed here given that the differential equations proposed are highly influenced by the starting point of the state variables. This was checked through the first assessment and it is shown in Figure 4-4. In this figure, the average length of the projection interval was plotted against the RMSE and the average bias (MD). Higher errors were obtained when the first observation (at year 5) was used to predict all the points forward (until year 31). This error was reduced when the projection interval was reduced. In Figure 4-4, the red dot indicates the error obtained when the projection was made using the closest previous observation to predict the current observation (e.g., observation at age 5 used to predict observation at age 8, observation at age 8 used to predict observation at age 11, etc.). This meant a projection length of 3 years for most of the observations (except for the last pair of observations which was from age 26 to age 31). Although not all possible combinations of projection lengths were used for evaluating the model, it is clear that this G&Y system has better performance with short predictions. The RMSE was consistently reduced when the projection length was reduced. No clear trend was observed for bias. Overall, when evaluating the magnitude of the errors, basal area was the most sensitive variable, while mortality was the least affected variable.

The second part of the robustness assessment consisted in using the validation dataset to evaluate the fit of the G&Y system proposed. For this data, the shortest prediction interval was used, which means predictions were made for an interval of two years. The fit statistics when using the validation dataset and how these compare to the fit statistics from the main dataset (or training dataset), are presented in Table 4-7. To fairly compare these values, when using the main dataset, the most recent observations were used for prediction. Although the average error (RMSE) and bias (MD and MAD) are higher for the validation dataset (as expected), these values are still small and acceptable for these types of systems. Relative to the fit statistics of the main dataset, errors increased in a 21% for the dominant height and basal area variables, and in a 38% for the mortality component. The observed versus predicted values for each one of the variables of the system for each one of the datasets are presented in Figure 4-5, while the residual plots are presented in Figure 4-6. In general, more negative bias (overestimation) is present for the validation dataset, whereas the model looks unbiased for the main dataset.

From the G&Y system presented, the dominant height and survival/mortality model can be integrated analytically. These two models are presented in Eq. 4-6 and Eq. 4-7. The basal area model cannot be integrated analytically and therefore, the differential equation form needs to be used when using the model. Therefore, a routine in *R* (R Core Team, 2018) was included to use this model with the estimated parameters for a simple example in Appendix A, and a more complete code was added in Appendix B using Julia to run three different examples.

$$H_2 = H_1\, exp\left[\frac{\mu_0 b_1^{Z_1}}{K b_2^{Z_2}}\left(e^{-Kb_2^{Z_2}t_1} - e^{-Kb_2^{Z_2}t_2}\right)\right] \qquad \text{Eq. 4-6}$$

$$N_2 = N_1\, exp\left[\frac{\alpha_0 + \alpha_1\, FR}{\delta + 1}\left(t_2^{\delta+1} - t_1^{\delta+1}\right)\right] \qquad \text{Eq. 4-7}$$

4.4. DISCUSSION

Bedding and vegetation control affect dominant height and basal area growth in slash pine plantations (Pienaar and Rheney, 1995). In this research, this effect was successfully incorporated into a growth and yield system by using parameter modifiers. This technique has been commonly used in forestry models (e.g., Hynynen et al., 1998) and has the advantage of incorporating treatments interaction and using both control and treatment plots during the fitting procedure, avoiding assigning all the variability of the data to the treatment effect which is the case when treatment effects are incorporated as an additive function to a base model (Gyawali and Burkhart, 2015; Logan and Shiver, 2006; Mason and Milne, 1999; Pienaar and Rheney, 1995). Both average error and bias were reduced when treatment modifiers were added to the base G&Y system, accounting for the treatment effects. These treatments, nevertheless, did not affect survival/mortality as shown in other research by Ramirez et al., (2023), and therefore, the G&Y system proposed did not include parameter modifiers for this component of the system, although it did include fusiform rust infection rates which showed to improve model fit, reducing both error and bias significantly.

When constructing G&Y systems for forest stands, is common to find variables that appear in both the left and right side of the equations, generating an interdependent system of equations that requires simultaneous estimation to get unbiased estimates of the parameters of the system (Borders and Bailey, 1986; Goelz and Burk, 1996). In addition, the correlation between the errors of the variables should be addressed. Although unbiased estimates of the parameters can be obtained when correlation is present, the standard error of the parameters is not correct when correlation exists and it is not taken into account. The parameters of the G&Y system proposed were estimated simultaneously using maximum likelihood and the variance/covariance matrix was estimated for the system, addressing these issues.

Several authors have modelled before the response to silvicultural treatments in G&Y systems, although without making use of simultaneous estimation (Bailey and Burgan, 1989; McTague, 2009; Pienaar and Rheney, 1995). When including simultaneous estimation, most of the authors have focused on the baseline system, in which silvicultural treatments are not considered (Borders and Bailey, 1986; Gallagher et al., 2019; Murphy and Sternitzke, 1979; Pienaar and Harrison, 1989; Sullivan and Clutter, 1972). Fewer authors have included both response to silvicultural treatments and simultaneous estimation (Fang et al., 2001; Martin et al., 1999). To the knowledge of the authors this is the first G&Y system for slash pine plantations where silvicultural treatments were included using a system of differential equations and simultaneous estimation was used. Nevertheless, accounting for the correlation between the errors of the different components of the system did not improve model fit and although the estimated value for the standard error changed for the mortality and basal area models, this change was not unidirectional.

The basal area component of the G&Y system proposed is a differential equation that cannot be solved analytically. Although the model looks fairly simple, the inclusion of dominant height and basal area as predictors makes this task difficult. Therefore, solving the differential equation numerically while estimating the parameters is the only approach for these types of models. In addition, the use of differential equations generally helps to construct the models when they need to be modified to incorporate treatment effects or the influence of pathogens like fusiform rust infection. This is because it is intuitive to determine what parameters should be modified according to the treatment being considered.

An additional factor that must be considered when using differential equations is how the trajectories of the different solutions change when changing the initial value for which the equation is being solved. For the models proposed, a big influence of these values was found. This is because the trajectory of the variable being modeled will highly differ with small variations on the initial value. For these models, the initial starting point defines the asymptote for each one of the curves. For this reason, when using the observed values instead of the predicted values as initial points to solve the differential equation, the errors increased considerably. Nevertheless, when points closer to the asymptote (later ages) and shorter projection intervals were used, the errors decreased. This is not the case for all the differential equations. For example, in the models propose by Garcia (1983) a change in the starting value (which is usually assumed as zero), does not influence the asymptote. For these models, other parameters generate the same phenomenon. Therefore, an understanding of the model being used and the influence of the different parameters of the model on the growth trajectory is necessary to properly use these models.

Although accessing an independent dataset with both a control and the two treatments evaluated in this study was not possible, the validation dataset allowed to evaluate model performance when used for the control scenario. Although error and bias increased, the magnitude of these values is acceptable for a G&Y system like the one presented here. The system proposed was not validated with an independent dataset for the treatments evaluated in this research, but the effectiveness of the G&Y system to incorporate treatment effects was assessed when comparing the model fit to the base G&Y system which did not have any parameter modifiers accounting for the treatments.

4.5. Conclusions

A growth and yield system for slash pine including dominant height, mortality, and basal area was fitted for slash pine plantations. The effect of bedding and vegetation control was successfully incorporated into dominant height and basal area components by using dummy variables. The average error and bias were reduced when these modifiers were added, compared to base models without these factors. Fusiform rust infection rates were crucial to describe mortality rates. These rates were modeled as a function of the initial infection rate per plot. Dominant height and mortality were used as predictor variables for the basal area model proposed, adding correlation within the system which was addressed by estimating the variance/covariance matrix of the system. The G&Y system proposed was shown to have lower errors for short projection intervals, and therefore its use is recommended for these intervals. When using an independent validation dataset to validate the model for the control plots, the error obtained increased but was maintained within acceptable ranges.

4.6. Acknowledgments

This research was conducted with support from the Plantation Management Research Cooperative (PMRC), where resources have been invested in establishing and collecting data from this long-term research trial. Special thanks to the PMRC Field Crew who collected the data for this research over the course of three decades. In addition, this work was supported by the Warnell School of Forestry and Natural Resources at the University of Georgia, the J.L. Clutter Forest Biometrics Scholarship, and the McIntire-Stennis Project GEOZ020205-MS.

4.7. TABLES AND FIGURES

Table 4-1. Dominant height (HD), trees per hectare (TPH), basal area (BA), and fusiform rust (FR) infection percentage, average values for the main dataset.

Age (years)	Treatment	No. plots	Mean HD (m)	Mean TPH	Mean BA (m^2/ha)	Mean FR Infection (%)
5	Control	18	3.26	1181.90	1.49	7.89
8		18	6.24	1137.16	6.02	14.02
11		18	9.08	1108.89	11.17	16.10
14		18	12.24	1062.92	15.66	13.70
17		18	14.54	1046.32	19.28	12.73
20		18	16.39	1047.97	22.25	11.87
23		17	18.29	1017.20	24.36	12.57
26		16	19.76	990.78	25.80	12.31
31		12	20.90	972.77	27.10	10.22
5	Bedding	19	3.87	1252.16	2.70	7.36
8		19	7.14	1228.63	8.30	12.63
11		19	10.09	1191.71	13.83	15.21
14		19	13.16	1150.63	18.94	12.03
17		19	15.49	1141.01	22.65	12.98
20		19	17.11	1124.11	25.02	12.25
23		18	18.91	1095.03	27.43	12.71
26		17	20.21	1032.31	27.86	11.26
31		11	20.87	1042.34	28.64	10.61
5	Chem	17	4.79	1181.24	5.50	9.69
8		17	8.32	1144.19	12.54	19.33
11		17	11.29	1130.53	18.56	22.18
14		17	14.35	1083.02	22.89	19.08
17		17	16.54	1076.34	26.92	19.68
20		16	18.32	1032.31	29.02	19.54
23		15	20.09	992.12	31.45	21.10
26		14	21.51	964.36	33.40	21.54
31		10	22.76	968.73	34.71	17.55
5	Bed+Chem	18	5.31	1200.97	6.84	10.29
8		18	8.92	1163.78	14.30	18.64
11		18	11.94	1146.49	20.25	19.77
14		18	14.91	1109.99	25.06	15.16
17		18	17.07	1082.96	28.39	17.29
20		17	18.61	1063.26	30.78	16.92
23		15	20.35	1005.29	32.00	18.61
26		15	21.91	963.30	33.36	17.73
31		11	22.79	1038.52	36.97	13.23

Table 4-2. Growth and yield systems tested. System 1(without treatment effects), system 2 (with treatment effects in dominant height and basal area).

Variable	System 1		System 2	
Dominant height	$\frac{dH}{dt} = \mu_0 H e^{-Kt}$	Eq. 4-1	$\frac{dH}{dt} = \mu_0 b_1^{Z_1} H e^{-K b_2^{Z_2} t}$	Eq. 4-2
Survival	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\, N t^{\delta}$	Eq. 4-3	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\, N t^{\delta}$	Eq. 4-3
Basal area	$\frac{dB}{dt} = [\beta_1 + \beta_2 H + \beta_3 N] \frac{B}{t^{\beta_4}}$	Eq. 4-4	$\frac{dB}{dt} = [\beta_1 + \beta_2 H + \beta_3 N] \frac{B}{t^{\beta_4}} d_1^{Z_1} d_2^{Z_2}$	Eq. 4-5

Table 4-3. Dominant height (HD), trees per hectare (TPH), basal area (BA), and fusiform rust (FR) infection percentage, average values for the validation dataset.

Age (years)	**Treatment**	**No. plots**	**Mean DH (m)**	**Mean TPH**	**Mean BA (m^2/ha)**	**Mean FR infection (%)**
10	Control	19	8.59	1415.00	12.54	24.77
12		21	9.91	1357.42	14.52	27.78
14		21	11.72	1301.10	16.74	24.77
16		21	13.28	1251.22	19.20	22.53
18		21	14.81	1217.18	21.35	88.82
20		19	15.90	1192.74	23.70	26.89
22		20	16.67	1116.92	24.54	22.04
24		17	17.82	1035.72	25.68	21.31
26		2	19.09	1090.20	26.35	0.00

Table 4-4. Fit statistics for dominant height and basal area with the two systems proposed.

Variable	**G&Y System**	**RMSE**	**MD**	**MAD**
Dominant height (m)	1. No treatments	0.71	0.03	0.55
	2. With treatments	0.57	0.01	0.45
Basal area (m^2ha^{-1})	1. No treatments	1.88	0.18	1.48
	2. With treatments	1.25	0.04	0.94

Table 4-5. Parameter estimates for the best growth and yield system for slash pine.

Variable	Model form	Parameter	Estimated value	Standard error
Dominant height (m)	$\frac{dH}{dt} = \mu_0 b_1^{Z_1} H e^{-K b_2^{Z_2} t}$	μ_0	0.3737	0.0042
		K	0.1128	0.0010
		b_1	0.9178	0.0020
		b_2	1.1057	0.0019
Mortality (TPH)	$\frac{dN}{dt} = (\alpha_0 + \alpha_1 * FR)\, N t^{\delta}$	α_0	-0.0151	0.0026
		α_1	-0.3596	0.0571
		δ	-0.6292	0.0674
Basal area (m^2ha^{-1})	$\frac{dB}{dt} = [\beta_1 + \beta_2 H + \beta_3 N] \frac{B}{t^{\beta_4}} d_1^{Z_1} d_2^{Z_2}$	β_1	10.4090	0.3921
		β_2	-0.2776	0.0097
		β_3	-1.4977	0.0649
		β_4	1.5512	0.0183
		d_1	0.8747	0.0024
		d_2	0.7449	0.0021
Variance/ Covariance Matrix (Σ)	$\vec{\epsilon} \sim MN(\vec{0}, \Sigma)$	$\ln(\sigma_H)$	-0.5556	0.0290
		$\ln(\sigma_N)$	-3.0782	0.0293
	$\Sigma = \begin{bmatrix} \sigma_H^2 & \sigma_H\sigma_N\rho_{HxN} & \sigma_H\sigma_{BA}\rho_{HxB} \\ .. & \sigma_N^2 & \sigma_N\sigma_{BA}\rho_{NxB} \\ .. & .. & \sigma_B^2 \end{bmatrix}$	$\ln(\sigma_B)$	0.2253	0.0294
		ρ_{HxN}	-0.0054	0.0418
		ρ_{HxB}	0.4408	0.0340
		ρ_{NxB}	0.4172	0.0354

Table 4-6. Fit statistics for the three variables of the best growth and yield system for slash pine.

Variable	Fit statistic	Value with correlation	Value without correlation
Dominant height (m)	RMSE	0.57	0.57
	MD	0.01	0.00
	MAD	0.45	0.45
Mortality (TPH)	RMSE	46.04	45.67
	MD	0.83	-0.14
	MAD	32.85	32.64
Basal area (m^2ha^{-1})	RMSE	1.25	1.24
	MD	0.04	-0.01
	MAD	0.94	0.94

Table 4-7. Fit statistics for the validation dataset (just control) compared to the statistics with the data used to fit the model.

Variable	Data	RMSE	MD	MAD
Dominant height (m)	Training	0.60	0.07	0.49
	Validation	0.73	-0.25	0.57
Mortality (TPH)	Training	44.91	-0.34	29.48
	Validation	62.06	-2.58	40.81
Basal area (m^2/ha)	Training	1.09	0.08	0.86
	Validation	1.32	-0.59	0.99

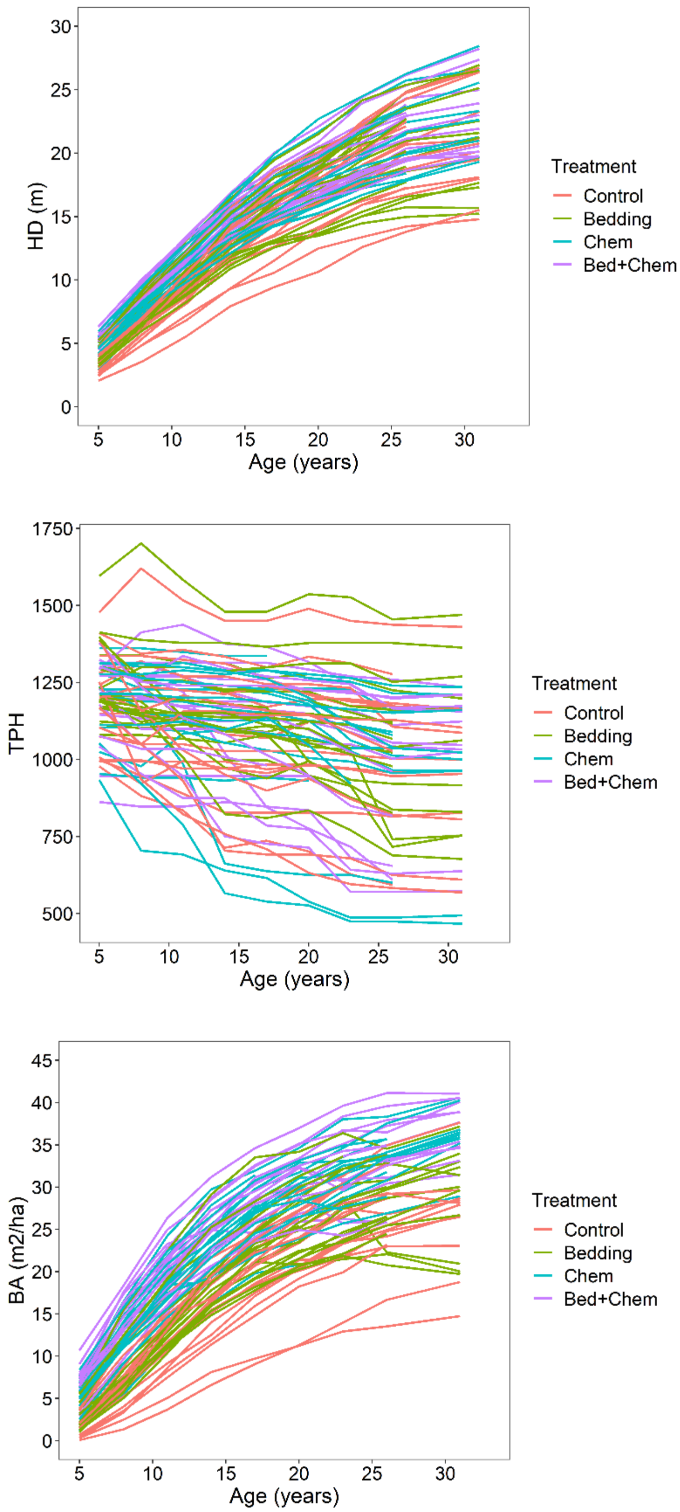

Figure 4-1. Dominant height (HD), survival (TPH), and basal area (BA) trajectories for the 30-yr study.

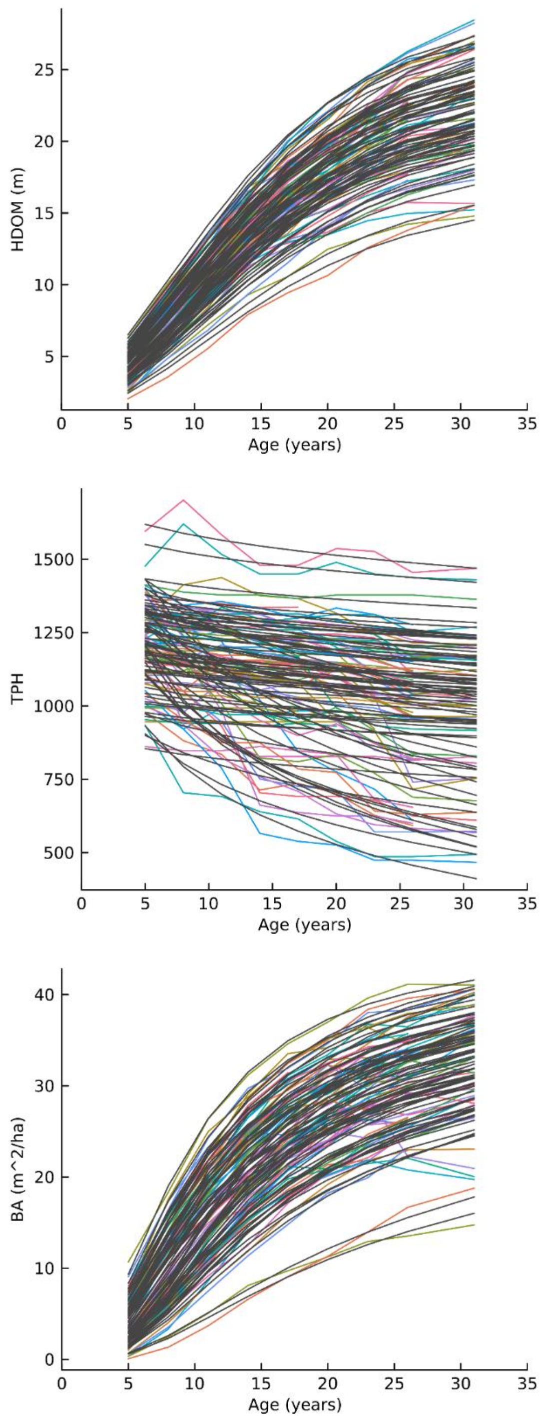

Figure 4-2. Dominant height, survival (mortality), and basal area predicted curves using the best growth and yield system for slash pine.

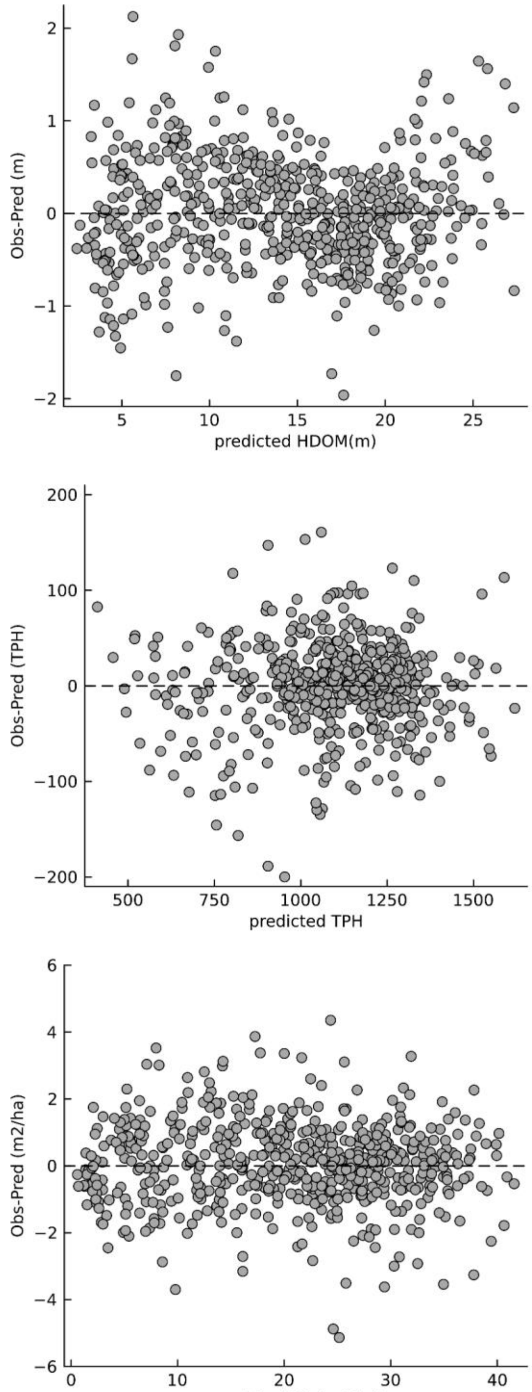

Figure 4-3. Dominant height, survival (mortality) and basal area residuals using the best growth and yield system for slash pine.

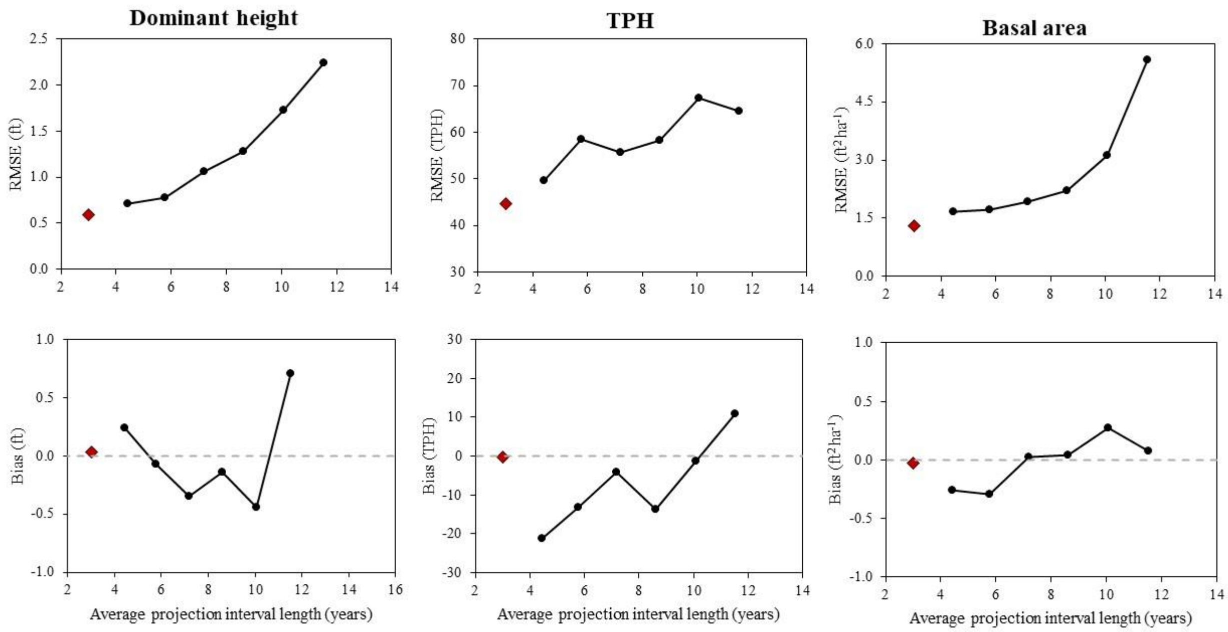

Figure 4-4. RMS and bias when projecting starting from a different age. The red dot refers to predictions using the immediate previous point to project forward.

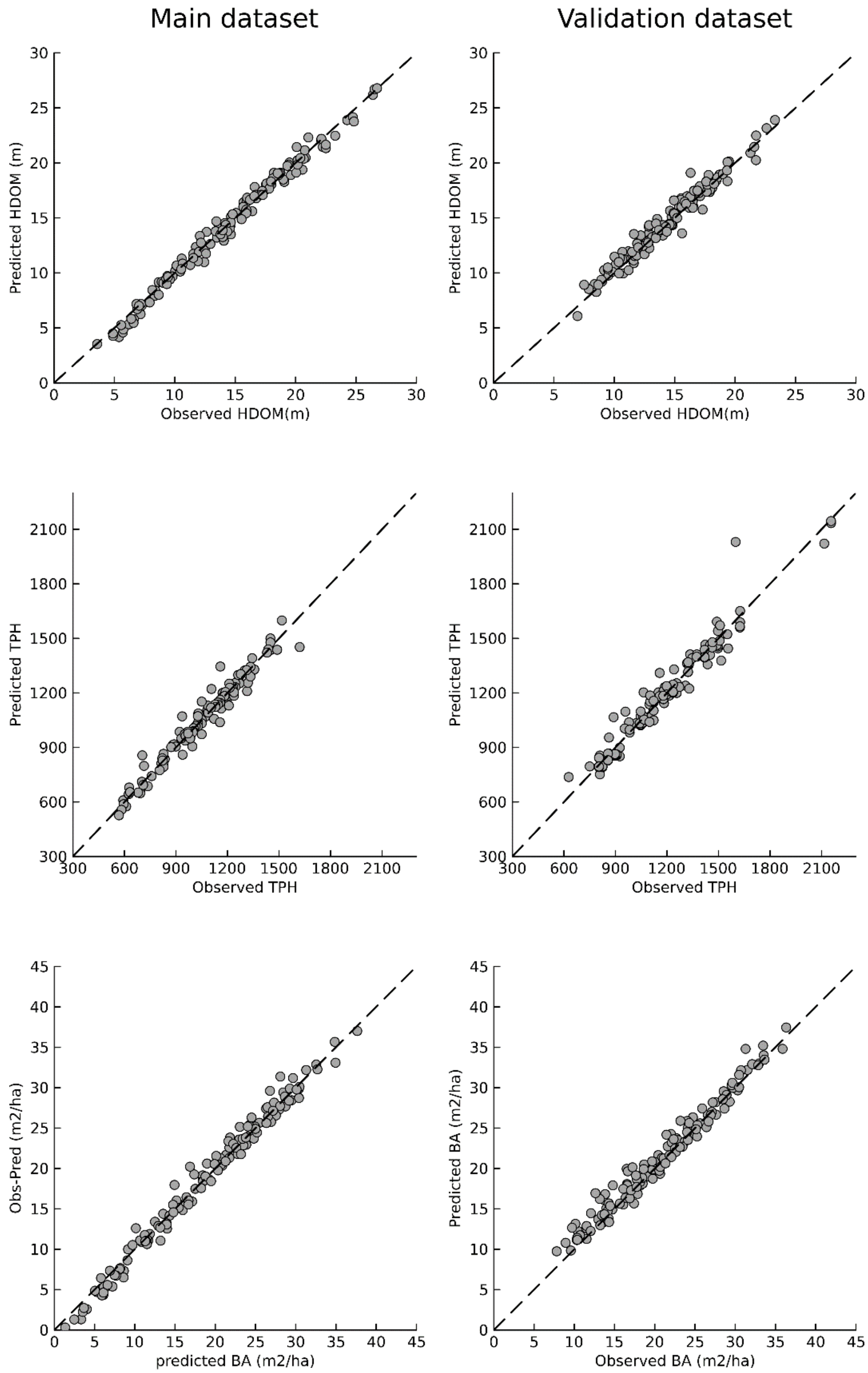

Figure 4-5. Observed vs predicted for the main and validation dataset when predicting values using the immediately previous observation as starting point.

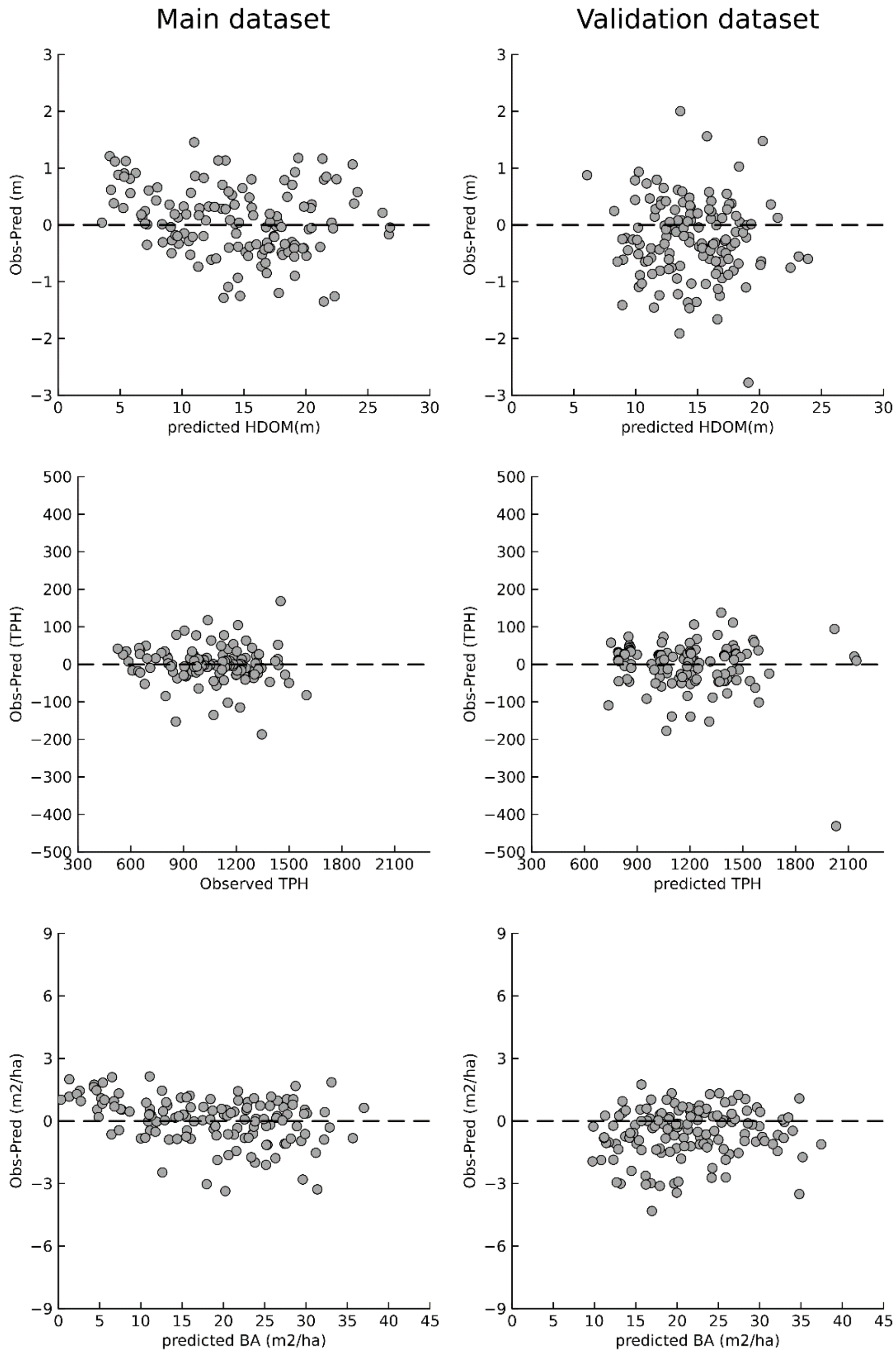

Figure 4-6. Residuals plot for the main and validation dataset when predicting values using the immediately previous observation as starting point.

4.8. REFERENCES

Bailey, R.L., Burgan, T.M., 1989. Fertilized midrotation-aged slash pine plantations-stand structure and yield prediction models. South. J. Appl. For. 13, 76–80.

Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B., 2017. Julia : A fresh approach to numerical computing. SIAM Rev. 59, 65–98. https://doi.org/10.1137/141000671. pdf

Borders, B.E., Bailey, R.L., 1986. A compatible system of growth and yield equations for slash pine fitted with restricted three-stage least squares. For. Sci. 32, 185–201. https://doi.org/10.1093/forestscience/32.1.185

Devine, O.J., Clutter, J.L., 1985. Prediction of survival in slash pine plantations infected with fusiform rust. For. Sci. 31, 88–94.

Fang, Z., Bailey, R.L., Shiver, B.D., 2001. A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects. For. Sci. 47, 550–562. https://doi.org/10.1093/forestscience/47.4.550

Gallagher, D.A., Bullock, B.P., Montes, C.R., Kane, M.B., 2019. Whole stand volume and green weight equations for loblolly pine in the Western Gulf Region of the United States through Age 15. For. Sci. 65, 420–428. https://doi.org/10.1093/forsci/fxy068

Garcia, O., 1983. A Stochastic Differential Equation Model for the Height Growth of Forest Stands. Biometrics 39, 1059. https://doi.org/10.2307/2531339

Garcia, O., 1979. Modelling stand development with stochastic differential equations. Mensurat. Manag. Plan. Exot. For. Plant. New Zeal. For. Serv. FRI Symp. 315–333.

Goelz, J.C.., Burk, T.., 1996. Measurement error causes bias in site index equations. Can. J. For. Res. 26, 1585–1593. https://doi.org/10.16309/j.cnki.issn.1007-1776.2003.03.004

Gyawali, N., Burkhart, H.E., 2015. General response functions to silvicultural treatments in loblolly pine plantations. Can. J. For. Res. 45, 252–265. https://doi.org/10.1139/cjfr-2014-0172

Hynynen, J., 1995. Predicting the growth response to thinning for Scots pine stands using individual-tree growth models. Silva Fenn. https://doi.org/10.14214/sf.a9210

Hynynen, J., Burkhart, H.E., Allen, H.L., 1998. Modeling tree growth in fertilized midrotation loblolly pine plantations. For. Ecol. Manage. 107, 213–229. https://doi.org/10.1016/S0378-1127(97)00334-4

Jones, J.E.P., 1972. Fusiform rust affects planted slash pine. J. For. 70, 350–352.

Korf, V., 1939. A mathematical definition of stand volume growth law. Lesn. práce 18.

Logan, S.R., Shiver, B.D., 2006. Adjusting Slash Pine Growth and Yield for Silvicultural Treatments. Gen. Tech. Rep. SRS-92. Asheville, NC U.S. Dep. Agric. For. Serv. South. Res. Stn. 328–332.

Martin, S.W., Bailey, R.L., Jokela, E.J., 1999. Growth and yield predictions for lower coastal plain slash pine plantations fertilized at mid-rotation. South. J. Appl. For. 23, 39–45. https://doi.org/10.1093/sjaf/23.1.39

Mason, E.G., Milne, P.G., 1999. Effects of weed control, fertilization, and soil cultivation on the growth of Pinus radiata at midrotation in Canterbury, New Zealand. Can. J. For. Res. 29, 985–992. https://doi.org/10.1139/x99-087

McTague, J.P., 2009. Modelling the Response of Loblolly Pine to Juvenile Fertilization. Open For. Sci. J. 1, 80–88. https://doi.org/10.2174/1874398600801010080

Murphy, P.A., Sternitzke, H.S., 1979. Growth and Yield Estimation for Loblolly Pine in the West Gulf. Res. Pap. SO-154. New Orleans, LA U.S. Dep. Agric. For. Serv. South. For. Exp. Station. 12 p. 154.

Nance, W.L., Froelich, R.C., Shoulders, E., 1981. Effects of fusiform rust on survival and structure of Mississippi and Louisiana slash pine plantations. US Dep. Agric. For. Serv. South. For. Exp. Stn. Vol. 232.

Oppenheimer, M.J., Shiver, B.D., Rheney, J.W., 1989. Effect of site factors on growth response of 9 to 15-year-old slash pine plantations following understory vegetation removal. Technical Report 1989-3. University of Georgia.

Phelps, W. R. (1978). Fusiform rust of southern pines (Vol. 26). US Department of Agriculture, Forest Service.

Pienaar, L., Rheney, J., 1995. Modeling stand level growth and yield response to silvicultural treatments. For. Sci. 41, 629–638. https://doi.org/10.1093/forestscience/41.3.629

Pienaar, L. V., Harrison, W.M., 1989. Simultaneous growth and yield prediction equations for pinus elliottii plantations in zululand. South African For. J. 149, 48–53. https://doi.org/10.1080/00382167.1989.9628992

Pienaar, L. V., Shiver, B.D., 1986. Basal Area Prediction and Projection Equations for Pine Plantations. For. Sci. 32, 626–633.

Pienaar, L. V, Turnbull, K.J., 1972. The Chapman-Richards Generalization of Von Bertalanffy's Growth Model for Basal Area Growth and Yield in Even Aged Stands. For. Sci. 19, 2–22.

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Ramirez, L., Montes, C.R., Bullock, B.P., 2023. Modeling slash pine mortality rates incorporating responses to silvicultural treatments and fusiform rust infection rates. For. Ecol. Manage. 532, 120832. https://doi.org/10.1016/j.foreco.2023.120832

Ramirez, L., Montes, C.R., Bullock, B.P., 2022. Long-term term effect of bedding and vegetation control on dominant height of slash pine plantations in the southeastern United States. For. Ecol. Manage. 522, 120479. https://doi.org/10.1016/j.foreco.2022.120479

Randolph, K. C. (2016). Status of fusiform rust incidence in slash and loblolly pine plantations in the southeastern United States. Gen. Tech. Rep. South. Res. Station. USDA For. Serv, 212, 53-59.

Sluder, E.R., 1977. Fusiform rust in loblolly and slash pine plantations on high-hazard sites in Georgia. For. Serv. US Dep. Agric. Southeast. For. Exp. Stn. Vol. 160.

Snowdon, P., 2002. Modeling type 1 and type 2 growth responses in plantations after application of fertilizer or other silvicultural treatments. For. Ecol. Manage. 163, 229–244. https://doi.org/10.1016/S0378-1127(01)00582-5

Sullivan, A.D., Clutter, J.L., 1972. A simultaneous growth and yield model for loblolly pine. For. Sci. 18, 76–86.

Susaeta, A., 2020. Implications of future risk of fusiform rust on optimal forest management of even-aged slash pine plantations. For. Policy Econ. 116, 102183. https://doi.org/10.1016/j.forpol.2020.102183

Wang, M., Montes, C.R., Bullock, B.P., Zhao, D., 2020. An empirical examination of dominant height projection accuracy using difference equation models. For. Sci. 66, 267–274. https://doi.org/10.1093/forsci/fxz079

Wells, O.O., Dinus, R.J., 1978. Early infection as a predictor of mortality associated with fusiform rust of southern pines. J. For. 76, 8–12.

Zarnovican, R., 1979. Fonction de la croissance de Korf. For. Chron. 55, 194–197.

Zeide, B., 1993. Analysis of Growth Equations. For. Sci. 39, 594–616. https://doi.org/10.1093/forestscience/39.3.594

Zhao, D., Kane, M., Borders, B., Harrison, M., 2009. Long-term effects of site preparation treatments, complete competition control, and repeated fertilization on growth of slash pine plantations in the flatwoods of the southeastern United States. For. Sci. 55, 403–410. https://doi.org/10.1093/forestscience/55.5.403

Zhao, D., Kane, M., Borders, B., Harrison, W.M., 2007. Slash pine site preparation study: Results through age 26. Plantation Management Research Cooperative (PMRC). Tech. Rep. 2007-3 61.

CHAPTER 5

USING STOCHASTIC DIFFERENTIAL EQUATIONS FOR MODELING FOREST GROWTH [1]

[1] Ramirez, L., Montes, C. R. To be submitted to *Forestry*.

ABSTRACT

Forest growth is a process influenced by stochastic factors such as environmental variations that should be considered when creating growth and yield models. Naturally, data collected from inventories also include observations and sampling errors that add noise to forestry data. One way in which this stochasticity can be incorporated when modeling forest growth, is making use of stochastic differential equations (SDEs). This approach has been incorporated into forestry since the 1980s, nevertheless, it has never been a widespread approach due to the complex concepts associated with it. The advantages and limitations of using SDEs for estimating parameters in growth models in forestry are described in this chapter. This framework is compared to the simpler alternative of using regular differential equations, together with modeling variance using predetermined functions when using maximum likelihood. The two approaches were compared using two different datasets, a loblolly pine dataset, publicly available, and a 30-year-old slash pine dataset available from a trial study established in the southeastern United States in the 1970s. When comparing the two approaches, it was found that modeling variance with a simpler function produced comparable results to the SDEs approach, without the complexity of the SDEs.

5.1. INTRODUCTION

When modeling growth and yield in forestry systems it is often desired that the models are able to capture the random environment in which the systems are embedded (Sandland and McGilchrist, 1979). Two main sources of uncertainty are usually distinguished in the forestry literature, the uncertainty linked to the growth process itself, often considered a stochastic process influenced by environmental factors, and the observation error, derived from measurement and sampling errors (Garcia, 1983). Including a measure of uncertainty when modeling forest growth provides managers with additional information about the expected yield and the potential variations of the predictions, providing confidence intervals around the predictions that can be useful when considering management regimes (Fox et al., 2001).

One common approach to include uncertainty into forest growth and yield systems (without differentiating between process and observation error) is to add a variance model during the estimation process. This refers to specifying for each one of the state variables, the variance of its errors, and how these errors are correlated to the errors of the other state variables. One clear example of this approach was followed by Gregoire (1987), who proposed four different variance/covariance matrices according to the expected error and correlations in the data. Following this approach, these matrices can be incorporated into regression techniques, which traditionally in forestry have included least squares, two or three stage least squares (2SLS-3SLS) (Borders and Bailey, 1986; Pienaar and Harrison, 1989), and mixed effect models (Fang et al., 2001; Gallagher et al., 2019). Along the same line, LeMay (1990) proposed a different approach, named multistage least squares (MSLS) to fit simultaneous, contemporaneously correlated systems of equations with both serial correlated and heteroscedastic error terms. This author proposed different variance/covariance matrices varying according to the combination of presence/absence of autocorrelation and heteroscedasticity.

Another approach to incorporate uncertainty into growth and yield models in forestry consists in using stochastic differential equations (SDEs). Garcia (1979) was the first author proposing this application in forestry. In Garcia's approach an additional factor representing random variation is added as a Wiener or Brownian motion process, representing the process error, and an additional constant term is added representing the observation error. Garcia has extensively worked on estimation procedures using this framework, either using maximum likelihood (Garcia, 1983), or reducing the problem to a least square problem (Garcia, 2019). Although some authors have followed this approach (Donnet et al., 2010; Orrego et al., 2021; Rupšys, 2019; Zhang and Borders, 2001), the use of SDEs in forestry is generally perceived as complex and hard to implement (Burkhart and Gregoire, 1994).

The main objective of this chapter is to explain SDEs when incorporating uncertainty into a growth and yield system in forestry and to compare it with the alternative option of modeling variance directly when using maximum likelihood. The theory of SDEs applied to growth models in forestry is first explained. Then, a case study using a literature example is presented in which parameters are estimated using the approach defined by Garcia (2019) and the alternative defined here. These two methodologies are later compared using the available dominant height data for slash pine presented in chapters 3 and 4 and some final remarks are made regarding the appropriateness of using SDEs in forestry and the separation of the errors into observation and process errors.

5.2. METHODS

5.2.1. THEORY OF STOCHASTIC DIFFERENTIAL EQUATIONS (SDES)[2]

In general, an SDE is defined by its deterministic component (e.g., Eq. 5-1), also called the drift, and its stochastic component, called the diffusion factor. The latter is usually defined as a multiple of White noise (W_t) and in the simplest case, an SDE has the form of Eq. 5-2.

$$\frac{dY_t}{dt} = \alpha Y_t \qquad \text{Eq. 5-1}$$

$$\frac{dY_t}{dt} = \alpha Y_t + \beta W_t \qquad \text{Eq. 5-2}$$

The term βW_t refers to the random error term. In this formulation β is a constant. White noise (W_t) is often expressed as $W_t = dB_t/dt$. That is, the "derivative" of a Brownian

[2] The concepts explained in this section are based on the explanations presented by Dobrow (2016b, 2016a).

motion (B_t). Brownian motion is a stochastic process that can describe the motion of a particle that diffuses randomly along a line. At each point *t*, the particle's position is normally distributed with variance t (Figure 5-1). Therefore, as t increases, the particle's position is more diffuse (Dobrow, 2016a).

Although White noise is often expressed as $W_t = dB_t/dt$, Brownian motion derivatives do not exist and therefore, this expression is better written as $W_t dt = dB_t$, and in a more rigorous language it is more appropriate to say that Brownian motion is integrated White noise. Thus, Eq. 5-2 can be rewritten as in Eq. 5-3, or equivalently, as in Eq. 5-4.

$$dY_t = \alpha Y_t dt + \beta W_t dt = \alpha Y_t dt + \beta \, dB_t \qquad \text{Eq. 5-3}$$

$$\int_0^t dY = \alpha \int_0^t Y_t \, dt + \beta \int_0^t dB_t \qquad \text{Eq. 5-4}$$

For solving the SDE defined by Eq. 5-4, stochastic calculus rules need to be used. Note that the second integral in the right hand side of Eq. 5-4 is not defined as an integral with respect to time (dt), as it is usually done in regular calculus, but it is defined as an integral with respect to Brownian motion (dB_t), a new concept particular to stochastic calculus. The solution to an SDE is a stochastic process for which the expected value is driven by the deterministic component of the SDE, and the variance can be calculated by solving the stochastic integral.

As it happens with regular differential equations, several techniques exist to solve SDEs. One common approach is to use Ito's Lemma, which can be interpreted as the stochastic calculus counterpart of the chain rule. Although not evident, this means that the stochastic component of the solution to an SDE is affected by the form of the deterministic component. For example, for the so called Langevin equations (in physics) described in Eq. 5-5, although σ is a constant, it can be shown that the expected value and variance of the solution to this SDE, have the form of Eq.

5-6 and Eq. 5-7, respectively. That is, the resulting variance is not constant, and increases with time with an asymptote at $\sigma^2/2r$.

$$dY_t = r(\mu - Y_t) + \sigma\, dB_t \qquad \text{Eq. 5-5}$$

$$E(Y_t) = \mu + (Y_0 - \mu)e^{-rt} \qquad \text{Eq. 5-6}$$

$$Var(Y_t) = \frac{\sigma^2}{2r}(1 - e^{-2rt}) \qquad \text{Eq. 5-7}$$

Not all SDEs have an analytical solution for which the expected value and variance functions can be derived. In fact, only very few (and simple) SDEs can be solved analytically. Numerical methods need to be applied to get one solution of the SDE for most cases (Nygaard et al., 2000). The example presented in Eq. 5-5 is relevant since this is the type of equations that Garcia utilizes for deriving its applications of SDEs to forestry data (Garcia, 2019, 1983). The model this author proposes has one variation to incorporate a wider range of models using the same approach. Garcia first introduced his approach for a dominant height model using the transformed variable H^c instead of H (Garcia, 1983). Thus, parameter c becomes a value between 0.3 and 1 for height-age curves that must be estimated together with the other parameters of the model (Garcia, 1983). When using this transformation, Eq. 5-5 can be rewritten as in Eq. 5-8, and the same solution from Eq. 5-6 and Eq. 5-7 can be applied to the transformed variable H^c.

$$dH^c = b(a^c - H^c)dt + \sigma\, dB_t \qquad \text{Eq. 5-8}$$

Using Eq. 5-8 and the theory of SDEs, Garcia developed a framework based on maximum likelihood where the parameters are estimated assuming errors are normally distributed with variance as in Eq. 5-7. Among the modifications proposed by Garcia are the addition of observation errors and the inclusion of correlation between errors at different points in time. Thus, Garcia's model had the form from Eq. 5-9.

$$h_i^c = a^c - (a^c - H_0^c)\exp\{-b(t_i - t_0)\} + \delta_i + \epsilon_i \qquad \text{Eq. 5-9}$$

Where δ_i represents the process error, normally distributed with variance as in Eq. 5-7 (with the proper change of variables and parameters), and ϵ_i represents the observation error, assumed to be normally distributed, $\epsilon_i \sim N(0, \eta^2)$. Note that this approach assumes the process and observation errors are additive. In addition, Garcia's approach is based on using pairs of observations for finding the parameters, that is, solving the differential equation from t_0 to t_i. The alternative would be to solve the equation for a given t.

Garcia later developed a simplification of its approach by converting the problem to a least square problem (Garcia, 2019), which facilitates its implementation and the parameter estimation. The proposal was based on the same type of equations described before and *R* code was provided to follow the approach. An example of how to use this approach is explained in Garcia's paper using the '*Loblolly'* dataset which is a preloaded dataset with observations of height for loblolly pine trees included in the software *R* (R Core Team, 2018). The same dataset was used in this chapter to reproduce Garcia's approach (using the *R* code provided by Garcia) and compare it with the proposed simpler approach.

5.2.2. Proposed approach

Although Garcia provided *R* code for implementing his proposed methodology, it is clear that users wanting to follow his approach need a deep understanding of SDEs to be able to properly use this tool. Therefore, it was of interest to compare the approach proposed by Garcia with a simpler approach in which differential equations and maximum likelihood are still used but errors are normally distributed with a given variance function determined by the user (either constant, or any function of time). The two approaches are compared in Table 5-1 and the proposed approach is described in the following pseudo code with an example to solve it presented in Appendix C using the programming language Julia[3] (Bezanson et al., 2017).

Pseudocode:

1. Define the function describing the differential equation (the deterministic component)
2. Define the initial values for the parameters (θ).
3. Solve the differential equation (numerically) using the initial value of the parameters and the defined function. Save the values of the solution for those times in which you have observations.
4. Calculate the errors (ϵ_i) in the prediction by comparing the observed values with the predicted values from step 3.
5. These errors are assumed to be normally distributed with mean zero and a given variance function for which parameters need to be estimated:

$$\epsilon_i \sim N(0, var(\epsilon_i))$$

Where $var(\epsilon_i)$ could be a constant ($var(\epsilon_i) = \sigma^2$), or a function of t (e.g., $var(\epsilon_i) = \sigma_0 + \sigma_t t$)

6. Calculate the loglikelihood value using the errors from step 4 and the Normal distribution from step 5. Sum up these values.
7. Using an optimization algorithm, find the parameters ($\hat{\theta}$) that minimize the negative loglikelihood (negative value from step 6).

[3] Although the example provided was written in Julia, this does not mean that the proposed approach can only be applied in this programming language. An equivalent routine can be written in any programming language that can solve differential equations numerically and has optimization algorithms to find the parameters.

In the next section both methodologies are compared using first the '*Loblolly'* dataset used by Garcia in 2019, and then using the slash pine dominant height data from chapters 2-4. Comparisons were based on the value of the estimated parameters and the fit statistics below.

- Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2}{n-p}}$$

- Mean Difference (MD):

$$MD = \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)}{n}$$

- Mean Absolute Difference (MAD):

$$MAD = \frac{\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|}{n}$$

Where n is the total number of observations, p is the number of parameters in each model, Y_i is the observed value, $\widehat{Y}_i$ is the predicted value.

When using the proposed approach, two different models were tested. These models differed in the assumption made about the variance function, one being constant (representing observation error) and the other having the form of Eq. 5-7 (representing process error). These models are described below:

- Model 1 (M_1): constant variance

$$\begin{aligned} dH^c &= b(a^c - H^c)dt \\ \epsilon_i &= H^c_{obs} - \widehat{H}^c_{pred} \\ var(\epsilon_i) &= \sigma^2 \end{aligned}$$

- Model 2 (M_2): variance function in the form of Eq. 5-7, simulating process error defined by Garcia.

$$
\begin{aligned}
dH^c &= b(a^c - H^c)dt \\
\epsilon_i &= H^c_{obs} - \widehat{H}^c_{pred} \\
var(\epsilon_i) &= \frac{\sigma^2}{2b}(1 - e^{-2bt_i})
\end{aligned}
$$

5.3. RESULTS

5.3.1. LOBLOLLY DATASET

Garcia initially presented one example of his approach using only one series of observations from the *Loblolly* dataset (Seed = 301, Example 2 in Garcia, 2019). The first comparison made in this chapter was made using the same subset and fixing the value for the transforming variable c by using the value reported by Garcia ($\hat{c} = 0.5024$). This guaranteed that the exact same data was being used. The parameters reported by Garcia were used to calculate the fit statistics described in section 5.2 and these values were compared with fit statistics from the two alternative models in which the proposed approach was used. The results for this comparison are presented in Table 5-2.

When using Model 1 (M_1), the estimated parameters for the dominant height model matched those reported by Garcia, and also did the parameter describing the variance. In Garcia's example, the variance associated to the process error was estimated as zero. Therefore, using the simpler approach presented here generated the same results as Garcia's approach for this example, being the proposed approach, a less complex one. When comparing Garcia's approach to Model 2 (M_2), where a more complex variance function was assumed, similar values for the parameters were obtained, and a slightly lower RMSE and MAD were obtained, although a higher MD was obtained.

To keep exploring the differences between Garcia's and the proposed approach, the data from all the observations of the *Loblolly* dataset (Example 3 in Garcia's paper) were then used to compare the different approaches. Again, the value of the transformation variable c was fixed to $\hat{c} = 0.49182$ as reported by Garcia. The results are presented in Table 5-3. Similar to what happened with the single plot, the two models proposed generated parameter estimates that do not differ significantly from Garcia's estimates.

5.3.2. Slash pine dataset

One last comparison was made by fitting the same dominant height model to the slash pine dataset described in chapters 2-4. Since no value of the transformation variable c was available for this dataset, this value had to be estimated to compare the two approaches. Nevertheless, when using the proposed approach, there were numerical difficulties when estimating this parameter, reaching non-convergence from the optimization algorithm. Therefore, to be able to compare both methodologies, a value of $c = 0.5$ was fixed. Note that although with the proposed methodology finding this parameter was not possible (for this specific dataset), the main advantage of using the proposed methodology is that it does not require transforming the data and that this transformation was made in this example to directly compare the results with Garcia's approach. Results from this dataset are presented in Table 5-4. The results are aligned with what obtained with the *Loblolly* dataset. With Garcia's approach the observation process was estimated as zero, and with the simplest model with constant variance (M_1), the RMSE was the same as with Garcia's, but the bias (MD) and absolute bias (MAD) are significantly smaller.

5.4. Discussion

When using the *Loblolly* dataset with one single plot, the two models proposed (M_1 and M_2) yielded very similar results compared to Garcia's results, raising the question of the necessity of using complex variance models when a constant variance could be used instead. To closely evaluate the differences between assuming a constant variance and a more complex variance function (M_1 vs. M_2), the estimated variances for these two models were plotted in Figure 5-2, and the different likelihood values in the log scale, obtained after using each one of these variances during the estimation procedure, were plotted against each other and presented in Figure 5-3. Given the differences in the variance assumed by each model (Figure 5-2), it is natural that the biggest differences in likelihood are given for the observations at smaller ages (i.e., 3, 5), the major differences were observed at age 3 (Figure 5-3). Nonetheless, overall, the likelihood values are very similar (close to the 1:1 line), regardless of the variance function used, which explains why the estimated parameters and fit statistics are non-significantly different between these two models.

A third alternative model was tested in which the variance was modeled as the sum of the two variances from M_1 and M_2, simulating what Garcia proposes when having process and observation errors as additive. For this model, the resulting parameters where highly influenced by the initial value of the parameters associated to each one of the variance functions. Thus, when the initial value for the parameter associated to the 'process error' was zero, the algorithm converged to zero for this parameter, and when the parameter associated to the 'observation error' (constant variance) was set equal to zero, the estimated 'observation error' was zero. In several applications of Garcia's approach this conclusion of having either process or observation errors as zero is common (Garcia, 2019, 2005, 1983; Orrego et al., 2021), which can be an indication that Garcia's optimization algorithm is unable to separate the two sources of errors, at least when they are defined to be additive.

When using the complete *Loblolly* dataset, results were on the same line. With the proposed models, parameter estimates, and fit statistics were very similar to what obtained with Garcia's approach. With this complete dataset, different to the single plot case, Garcia found the estimated observation error was zero, one more example of an application where the two sources of error cannot be separated. From the two models proposed, model M_1, with constant variance, produced the same RMSE as with Garcia's approach, with a much lower bias (MD), and slightly lower absolute bias (MAD).

When testing the different methodologies for the slash pine dataset, results were similar to the *Loblolly* dataset. In general, Garcia's approach generated higher (negative) bias (overestimation) when using several plots in both the examples tested, compared to almost unbiased results when using the proposed approach with either model M_1 or M_2. Although the estimation procedure proposed here does not have a derived maximum likelihood function that can be expressed in a closed form, as does Garcia's approach, the effectiveness of the proposed methodology was demonstrated with two different datasets and its use is recommended when using differential equations with either a constant or a custom variance function.

The examples presented in section 5.3 showed in practical terms how a simpler approach could be used instead of SDEs to model forest growth with similar (or better) results. Theoretically, another disadvantage of Garcia's approach is that only models that have the form from Eq. 5-5 can be used. Although having the transforming parameter c adds flexibility and growth curves of different shapes like the Bertalanffy-Richards, Logistic, and Gompertz can be approximated, this is still a limited range of models. Another drawback of Garcia's approach is that the variance function is assumed to have a predetermined form defined by Eq. 5-7, which has the general form presented in Figure 5-4, and when fitting this model to a single plot, it is impossible to check if the variance of the process being evaluated follows the pattern imposed by this SDE.

When using SDEs, even small variations to the deterministic component generate a different variance function (or variance/covariance matrix), which completely changes the maximum likelihood function or the least squares estimation. Moreover, if the stochastic component is defined as dependent on time or the current state of the variable being modeled, the SDE quickly evolves to an equation that cannot be solved analytically and Garcia's approach cannot be used. For these type of models, other parameter estimation approaches such as the method of moments, or Bayesian approaches need to be used (Nygaard et al., 2000). To illustrate how the complexity of the SDE quickly increases, we can define the dominant height model presented in Chapter 3, equation Eq. 3-3 as an SDE. The simplest approach to convert this model to an SDE would be to add white noise as in Eq. 5-3. Then, the SDE would have the following form:

$$dH = \mu_0 H e^{-Kt} + \sigma\, dB_t \qquad \text{Eq. 5-10}$$

Using the theory for solving SDEs, a solution for this equation would be given by Eq. 5-11, a stochastic process for which the expected value is as in Eq. 5-12, and the variance is defined by Eq. 5-13. Nevertheless, different to the previous example used by Garcia, this variance function cannot be defined in a closed expression since the stochastic integral in Eq. 5-13 cannot be solved analytically and therefore, numerical methods need to be used to estimate this variance. This implies that a close form of the maximum likelihood function cannot be derived, and Garcia's approach cannot be used.

$$H = H_0 \exp\left(-\frac{\mu_0}{K} e^{-Kt}\right) + \sigma \exp\left(-\frac{\mu_0}{K} e^{-Kt}\right) \int_0^t \exp\left(-\frac{\mu_0}{K} e^{-Kt}\right) dB_t \qquad \text{Eq. 5-11}$$

$$E(H) = H_0 \exp\left(-\frac{\mu_0}{K} e^{-Kt}\right) \qquad \text{Eq. 5-12}$$

$$Var(H) = \sigma \exp\left(-\frac{\mu_0}{K} e^{-Kt}\right) \int_0^t \exp\left(\frac{\mu_0}{K} e^{-Kt}\right) dB_t \qquad \text{Eq. 5-13}$$

When Eq. 5-10 is solved (numerically), it can be shown that the variance from Eq. 5-13 has the general form shown in Figure 5-5, a different form from the variance of the simplest case represented in Figure 5-4. This theoretical example highlights other limitations of using Garcia's approach. Overall, although Garcia's proposal was an innovative and elegant approach on how to apply SDEs to forestry data, not only its difficulty, but the limited set of functions that can be used, and the restricted variance that is obtained from the specific SDE being used, are limitations that support the use of other simpler and flexible approaches to estimate parameters for G&Y systems including uncertainty.

5.5. CONCLUSIONS

The appropriateness of using stochastic differential equations (SDEs) for modeling growth and incorporating uncertainty into forestry systems was evaluated. The strong mathematical background required to implement this approach had been recognized in the literature before, but in addition, other limitations of this approach were explained through different practical and theoretical examples. The limited set of functions that can be used with this approach along with the restricted form of the variance function that is derived from the SDEs were explained. Although the use of SDEs can be extended to include other function forms, the complexity of the models quickly increases to the point where only numerical solutions to the SDEs can be used, precluding the derivation of analytical forms for the likelihood function to be minimized when using maximum likelihood for estimating the parameters of a model. As an alternative to Garcia's approach, a simpler approach based on differential equations and maximum likelihood was presented, showing the same or better fit than Garcia's approach. An example of how to implement this approach was attached as an appendix.

5.6. TABLES AND FIGURES

Table 5-1. Comparison of Garcia's approach and proposed approach to estimate parameters using differential equations.

Garcia's approach	**Proposed approach**
Uses stochastic differential equations (requires understanding of stochastic calculus)	Uses regular (deterministic) differential equations (requires basic knowledge of differential equations)
Applies only to linear equations that can be expressed in the form of Eq. 5-5	Applies to any equation, even non-linear ones
The differential equation is solved for consecutive pairs of observations	The differential equation is solved for the whole time series
Autocorrelation is incorporated	Autocorrelation is ignored
Both observation and process errors are proposed	A single error is proposed
Uses the variance function derived after solving the SDE for a specific group of curves. This function cannot be modified to assume other forms without mathematically deriving the proper function	Can use any variance function proposed by the user, including constant or a custom function similar to the one proposed by Garcia
Can be applied using maximum likelihood or using non-linear least squares	Can be applied using maximum likelihood
Can incorporate local values per plot	Can incorporate local values per plot

Table 5-2. Fit statistics for the Loblolly 301 data (fixing $\hat{c} = 0.5024$), using Garcia's approach and models proposed (M_1 and M_2).

Approach	Parameter estimates			Fit statistics		
	$\hat{a}$	$\hat{b}$	$\hat{\sigma}$	RMSE (ft)	MD (ft)	MAD (ft)
Garcia (2019)	72.550	0.097	$\hat{\sigma}_{obs} = 0.049$ $\hat{\sigma}_{proc} = 0.000$	0.581	0.005	0.506
M_1: Constant variance	72.540	0.097	$\hat{\sigma} = 0.049$	0.581	0.002	0.506
M_2: Variance function	72.677	0.096	$\hat{\sigma} = 0.0235$	0.576	0.010	0.504

Table 5-3. Fit statistics for the whole *Loblolly* data (14 plots, fixing $\hat{c} = 0.49182$), using Garcia's approach and models M_1 and M_2.

Approach	**Parameter estimates**			**Fit statistics**		
	$\hat{a}$	$\hat{b}$	$\hat{\sigma}$	**RMSE (ft)**	**MD (ft)**	**MAD (ft)**
Garcia (2019)[1]	$\hat{a}_1 = 74.77$, ⋮ $\hat{a}_{14} = 70.44$	0.095	$\hat{\sigma}_{obs} = 0.000$ $\hat{\sigma}_{proc} = 0.034$	0.850	-0.143	0.694
M_1: Constant variance	$\hat{a}_1 = 74.23$ ⋮ $\hat{a}_{14} = 69.22$	0.095	$\hat{\sigma} = 0.070$	0.850	0.005	0.673
M_2: Variance function	$\hat{a}_1 = 74.34$ ⋮ $\hat{a}_{14} = 68.69$	0.096	$\hat{\sigma} = 0.035$	0.891	-0.002	0.680

[1] Note that the estimated values for the local parameter (a) for Garcia's model do not match the values reported directly in the supplementary material of Garcia (2019). The difference lies in the order in which these parameters were estimated, in the appendix presented by Garcia, the value of $\hat{a}_1$ does not correspond to the parameter of the first plot (Seed=301), but to plot 13 (Seed =329). This was generated by the order in which the factor levels of the variable Seed are ordered by default for that dataset (not in increasing order).

Table 5-4. Fit statistics for the slash pine dataset (dominant height) using Garcia's approach (c fixed at $c = 0.5$) and the proposed approach with models M_1 and M_2.

Approach	Parameter estimates			Fit statistics		
	$\hat{\boldsymbol{a}}$	$\hat{\boldsymbol{b}}$	$\hat{\boldsymbol{\sigma}}$	**RMSE (m)**	**MD (m)**	**MAD (m)**
Garcia (2019)	$\hat{a}_1 = 26.25$, ⋮ $\hat{a}_{72} = 23.64$	0.107	$\hat{\sigma}_{obs} = 0.000$ $\hat{\sigma}_{proc} = 0.075$	0.809	-0.120	0.643
M_1: Constant variance	$\hat{a}_1 = 26.09$ ⋮ $\hat{a}_{72} = 24.02$	0.105	$\hat{\sigma} = 0.120$	0.787	0.014	0.603
M_2: Variance function	$\hat{a}_1 = 25.89$ ⋮ $\hat{a}_{72} = 23.82$	0.106	$\hat{\sigma} = 0.061$	0.823	0.002	0.624

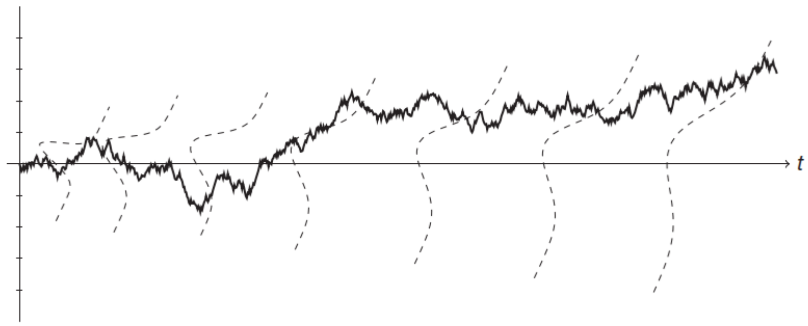

Figure 5-1. Brownian motion example, taken from (Dobrow, 2016a), Figure 8.1. Superimposed are normal density curves with mean 0 and variance t.

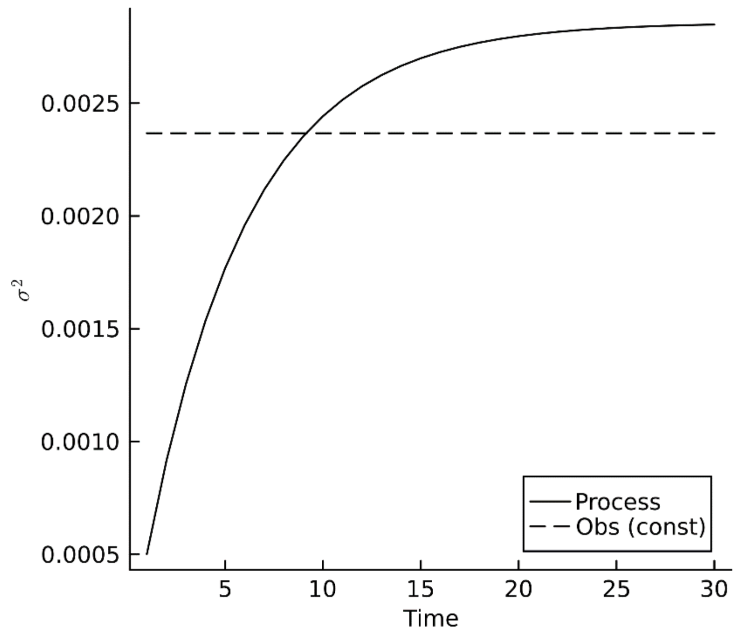

Figure 5-2. Comparison between assuming variance as a constant and as an increasing function of time.

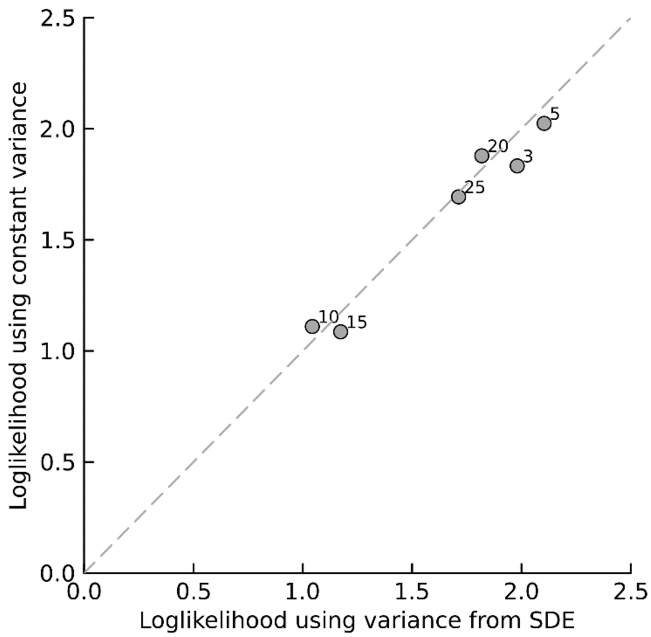

Figure 5-3. Loglikelihood values calculated using a constant value (M_1) or an increasing variance function (M_2). The labels indicate the age of the observation being evaluated.

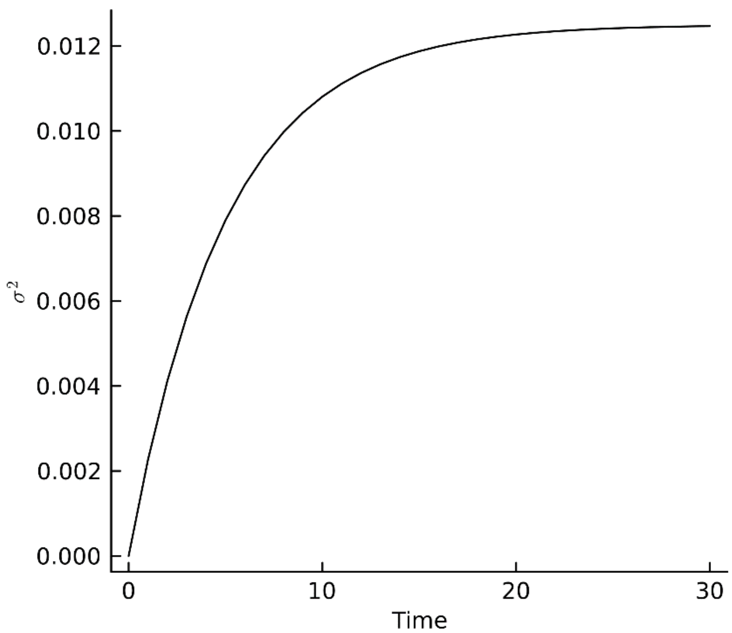

Figure 5-4. Variance from an SDE of the form from Eq. 5-5, using Eq. 5-7 with $r = 0.1$ and $\sigma = 0.05$.

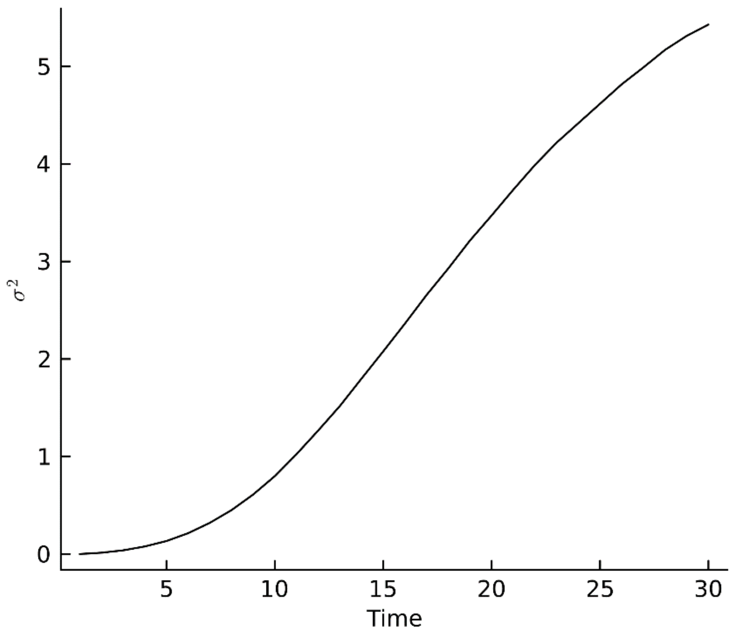

Figure 5-5.Variance form for the model from Eq. 5-13, with $\sigma = 0.1$, $\mu_0 = 0.4$, and $K = 0.12$.

5.7. REFERENCES

Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B., 2017. Julia : A fresh approach to numerical computing. SIAM Rev. 59, 65–98. https://doi.org/10.1137/141000671. pdf

Borders, B.E., Bailey, R.L., 1986. A compatible system of growth and yield equations for slash pine fitted with restricted three-stage least squares. For. Sci. 32, 185–201. https://doi.org/10.1093/forestscience/32.1.185

Burkhart, H.E., Gregoire, T.G., 1994. 11 Forest biometrics. Handb. Stat. 12, 377–407. https://doi.org/10.1016/S0169-7161(05)80013-6

Dobrow, R.P., 2016a. Brownian motion, in: Introduction to Stochastic Processes with R. John Wiley & Sons, Inc., pp. 320–371.

Dobrow, R.P., 2016b. A gentle introduction to stochastic calculus, in: Introduction to Stochastic Processes with R. John Wiley & Sons, Inc., pp. 372–396.

Donnet, S., Foulley, J.L., Samson, A., 2010. Bayesian analysis of growth curves using mixed models defined by stochastic differential equations. Biometrics 66, 733–741. https://doi.org/10.1111/j.1541-0420.2009.01342.x

Fang, Z., Bailey, R.L., Shiver, B.D., 2001. A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects. For. Sci. 47, 550–562. https://doi.org/10.1093/forestscience/47.4.550

Fox, J.C., Ades, P.K., Bi, H., 2001. Stochastic structure and individual-tree growth models. For. Ecol. Manage. 154, 261–276. https://doi.org/10.1016/S0378-1127(00)00632-0

Gallagher, D.A., Bullock, B.P., Montes, C.R., Kane, M.B., 2019. Whole stand volume and green weight equations for loblolly pine in the Western Gulf Region of the United States through Age 15. For. Sci. 65, 420–428. https://doi.org/10.1093/forsci/fxy068

Garcia, O., 2019. Estimating reducible stochastic differential equations by conversion to a least-squares problem. Comput. Stat. 34, 23–46. https://doi.org/10.1007/s00180-018-0837-4

Garcia, O., 2005. Comparing and combining stem analysis and permanent sample plot data in site index models. For. Sci. 51, 277–283.

Garcia, O., 1983. A Stochastic Differential Equation Model for the Height Growth of Forest Stands. Biometrics 39, 1059. https://doi.org/10.2307/2531339

Garcia, O., 1979. Modelling stand development with stochastic differential equations. Mensurat. Manag. Plan. Exot. For. Plant. New Zeal. For. Serv. FRI Symp. 315–333.

Gregoire, T.G., 1987. Generalized error structure for forestry yield models. For. Sci. 33, 423–444.

LeMay, V.M., 1990. MSLS: a linear least squares technique for fitting a simultaneous system of equations with a generalized error structure. Can. J. For. Res. 20, 1830–1839. https://doi.org/10.1139/x90-246

Nygaard, J., Henrik, N., Young, P.C., 2000. Parameter estimation in SDE: An overview. Annu. Rev. Control 24, 83–94.

Orrego, S., Montes, C., Restrepo, H.I., Bullock, B.P., Zapata, M., 2021. Modeling height growth for teak plantations in Colombia using the reducible stochastic differential equation approach. J. For. Res. 32, 1035–1045. https://doi.org/10.1007/s11676-020-01174-y

Pienaar, L. V., Harrison, W.M., 1989. Simultaneous growth and yield prediction equations for pinus elliottii plantations in zululand. South African For. J. 149, 48–53. https://doi.org/10.1080/00382167.1989.9628992

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Rupšys, P., 2019. Modeling dynamics of structural components of forest stands based on trivariate stochastic differential equation. Forests 10. https://doi.org/10.3390/f10060506

Sandland, R.L., McGilchrist, C.A., 1979. Stochastic Growth Curve Analysis Author. Biometrics 35, 255–271.

Zhang, Y., Borders, B.E., 2001. An iterative state-space growth and yield modeling approach for unthinned loblolly pine plantations. For. Ecol. Manage. 146, 89–98. https://doi.org/10.1016/S0378-1127(00)00448-5

CHAPTER 6

OVERALL CONCLUSIONS

The effect of bedding and vegetation control on slash pine growth was evaluated and modeled in this dissertation. In Chapter 2, a long-lasting effect of the vegetation control treatment on dominant height was observed, with estimated increments in site index (base age 25) of about 5 m, and some residual effect still present at year 31. Higher responses (relative to the control) were observed in low quality sites (i.e., base site index of 14 m), while higher quality sites (i.e., SI > 22 m) showed no significant increase in observed site index due to the treatment application. On the other hand, growth responses due to bedding were no longer observed for stands older than 20 years. This treatment improved dominant height for younger ages only. Results from this chapter imply that a more efficient application of herbicides can be done by targeting low quality sites which will show higher responses.

The treatments evaluated did not show any significant differences on mortality/survival. Although this component was greatly affected by the percentage of trees affected by fusiform rust within a plot. In Chapter 3, this is explained and a mortality/survival model including average fusiform rust infection rates at age 5 was proposed to describe stand mortality. Although the use of differential equations in which height increments are used instead of time increments had been found in the literature as a better alternative to model survival mortality, this was not the case for the models and data used for this research.

Results from chapters 2 and 3 were used to construct the final growth and yield model proposed for slash pine plantations. This system included the stand-level characteristics of dominant height, mortality, and basal area. The effect of bedding and vegetation control was successfully incorporated into the dominant height and basal area components by using dummy variables. The average error and bias of the models were reduced when parameter modifiers were added, compared to base models without these factors. Fusiform rust infection rates were crucial to describe mortality rates. These rates were modeled as a function of the initial infection rate per plot as presented in Chapter 3. Dominant height and mortality were used as predictor variables for the basal area model proposed. When having these variables as predictors, correlation between the errors of the variables is present. This correlation was addressed by estimating the variance-covariance matrix of the system along with the parameters of the system. The G&Y system proposed was shown to have lower errors for short projection intervals, and therefore its use is recommended for these intervals. When using an independent validation dataset to validate the model for the control plots, the error obtained increased but was maintained within acceptable ranges.

Finally, the appropriateness of using of stochastic differential equations (SDEs) for modeling growth and incorporating uncertainty into forestry systems was evaluated. The strong mathematical background required to implement this approach had been recognized in the literature before, but in addition, other limitations of this approach were explained through different practical and theoretical examples. The limited set of functions that can be used with this approach along with the restricted form of the variance function that is derived from the SDEs were explained. Although the use of SDEs can be extended to include other function forms, the complexity of the models quickly increases to the point where only numerical solutions to the SDEs can be used, precluding the derivation of analytical forms for the likelihood function to be minimized when using maximum likelihood for estimating the parameters of a model. As an alternative to using SDEs, a simpler approach based on differential equations and maximum likelihood was presented, showing the same or better fit than an approach using SDEs. An example of how to implement this approach was attached as an appendix.

APPENDIX A: CHAPTER 4 R CODE TO USE G&Y SYSTEM FOR A SINGLE PREDICTION

```r
# Loading the packages
library(deSolve)

#---------Estimated parameters----------------
#DH parms
u0 <- 0.3737; K <- 0.1128; b1 <- 0.9178; b2 <- 1.1057
#N parms
a0 <- -0.0151; a1 <- -0.3596; d <- -0.6292;
#BA parma
B1 <- 10.4090; B2 <- -0.2776; B3 <- -1.4977 ; B4 <- 1.5512; d1 <- 0.8747;
d2 <- 0.7449

parms_sys <- c(u0, K, b1, b2, a0, a1, d, B1, B2, B3, B4, d1, d2)
#--------Initial conditions-------------------
hini  <- 3
Nini  <- 1100/1000 #Needs to be divided by 1000
BAini <- 1.5
t1    <- 5
t2    <- 8
FR    <- 0.1
Z1    <- 0.0
Z2    <- 0.0
#--------Define system (differential equations)------
System <- function(t, u, p) {
  h0  <- u[1]
  N0  <- u[2]
  BA0 <- u[3]

  dH <- (h0 * p[1] * p[3]^Z1) * exp(-p[2] * p[4]^Z2 * t)
  dN <- N0 * (p[5] + p[6] * FR) * t^p[7]
  dB <- (p[8] + p[9] * h0 + p[10] * N0)*BA0 / t^p[11]) * p[12]^Z1 * p[13]^Z2

  return(list(c(dH, dN, dB)))
}
#-----Solve the system numerically--------------------
y0    <- c(hini, Nini, BAini)
times <- c(t1,t2)
sol   <- ode(y = y0, times = times, func = System, parms = parms_sys)

#---- extract the values of interest-----------------
HDOM_pred <- sol[2,2]
N_pred    <- sol[2,3] * 1000 #Needs to back transform the variable
BA_pred   <- sol[2,4]
```

APPENDIX B: CHAPTER 4 JULIA CODE TO USE G&Y SYSTEM FOR A SINGLE OR MULTIPLE PREDICTIONS AND WITH DIFFERENT PLOTS

```
#Loading the packages
using DifferentialEquations
using Plots
#---------Estimated parameters----------------
#DH parms
u0 = 0.3737;K  = 0.1128;b1 = 0.9178;b2 = 1.1057
#N parms
a0 = -0.0151;a1 = -0.3596;d  = -0.6292;
#BA parma
B1 =10.4090;B2 = -0.2776;B3 = -1.4977 ;B4 = 1.5512;d1 = 0.8747; d2 = 0.7449

parms_sys = [u0;K;b1;b2;a0;a1;d;B1;B2;B3;B4;d1;d2]

#--------Initial conditions-------------------
hini  = 3
Nini  = 1100/1000 #This variable need to be scaled
BAini = 1.5
t1    = 5
t2    = 8
FR    = 0.1
Z1    = 0.0
Z2    = 0.0

#--------Define system (differential equations)------
function System(du,u,p,t)
   h0 =  u[1]
   N0 =  u[2]
   BA0 = u[3]

 du[1]   =(h0.*(p[1]).*(p[3].^Z1)).*exp.(-p[2].*(p[4].^Z2).*t)
 du[2]   = N0.*(p[5].+p[6].*FR).*(t.^p[7])
 du[3]   =(p[8].+p[9].*h0.+p[10].*N0).*(BA0./(t^p[11])).*(p[12].^Z1).*(p[13].^Z2)
    nothing
end
```

```
#-----1. Predicting one plot, one single time --------------------
#--- Solve the system numerically---------------------
prob_sol = ODEProblem(System,vcat(hini,Nini,BAini),(t1, t2), parms_sys)
sol      = solve(prob_sol,Tsit5(),saveat = t2);

#---- Extract the values of interest-----------------
HDOM_pred  = transpose(hcat(sol.u...))[2,1];
N_pred     = transpose(hcat(sol.u...))[2,2].*1000; #scale back this variable
BA_pred    = transpose(hcat(sol.u...))[2,3];

#----2. Predicting one plot, different times--------------------

ages = collect(5:30)
t1   = minimum(ages)
t2   = maximum(ages)
prob_sol = ODEProblem(System,vcat(hini,Nini,BAini),(t1, t2), parms_sys)
sol      = solve(prob_sol,Tsit5(),saveat = ages);

#---- Extract the values of interest-----------------
HDOM_pred  = transpose(hcat(sol.u...))[:,1];
N_pred     = transpose(hcat(sol.u...))[:,2].*1000; #scale back this variable
BA_pred    = transpose(hcat(sol.u...))[:,3];

plot(ages, HDOM_pred)
plot(ages, N_pred)
plot(ages, BA_pred)


#----3. Predicting different plots, several different times--------------------
#--------Initial conditions-------------------
hini  = [3,2.5,3.5] #Assuming 3 plots
Nini  = [1100,1000,1150]./1000
BAini = [1.5,1.3,1.6]
FR    = [0.1,0.0,0.5]
Z1    = [0.0,1.0,1.0]
Z2    = [1.0,0.0,1.0]

nplot = length(hini)
ages = collect(5:30)
t1   = minimum(ages)
t2   = maximum(ages)

#Necessary to take advantage of system of DE
h1  = 1
hn  = nplot
```

```
N1  = nplot+1
Nn  = 2*nplot
BA1 = 2*nplot+1
BAn = 3*nplot;

#----------------Define system for vectors--------------------------------
function System2(du,u,p,t)
    h0 =  u[h1:hn]
    N0 =  u[N1:Nn]
    BA0 = u[BA1:BAn]

    du[h1:hn]   .=(h0.*(p[1]).*(p[3].^Z1)).*exp.(-p[2].*(p[4].^Z2).*t)
    du[N1:Nn]   .= N0.*(p[5].+p[6].*FR).*(t.^p[7])
    du[BA1:BAn].=(p[8].+p[9].*h0.+p[10].*N0).*(BA0./(t^p[11])).*(p[12].^Z1).*(p[13].^Z2)

    nothing
end

#------------------Plots (predicted lines)-----------------------------
prob_sol2 = ODEProblem{true}(System2,vcat(hini,Nini,BAini),(t1, t2), parms_sys)
sol2      = solve(prob_sol2,Tsit5(),saveat = ages);

sol_DH    = transpose(hcat(sol2.u...))[:,1:nplot];
sol_N     = transpose(hcat(sol2.u...))[:,(nplot+1):2*nplot].*1000;
sol_BA    = transpose(hcat(sol2.u...))[:,(2*nplot+1):3*nplot];

plot(sol2.t, sol_DH)
plot(sol2.t, sol_N)
plot(sol2.t, sol_BA)
```

Appendix C: Chapter 5 Julia code to estimate parameters of a DE

C1. Code to directly compare with Garcia's approach

```
#Loading the packages
using DifferentialEquations
using Plots
using Optim
using ForwardDiff
using DataFrames
using Distributions
using CSV
using BenchmarkTools
using StaticArrays
using LinearAlgebra
using StatsBase
using NLSolversBase
using Preferences
using DelimitedFiles
using Plots.Measures
#----------------Reading data----------------------------
#Need to export the Loblolly dataset beforehand from R
data_lob = DataFrame(CSV.File("Data/data_lob.csv"))
plot(data_lob.age,data_lob.height,group  = data_lob.Seed)

#---------------Definig variables as constant---------------
nplot = length(unique(data_lob.Seed))
AGE   = data_lob.age
nAge  = length(unique(AGE))
unique_AGES = sort(unique(AGE))
HLob = data_lob.height
inits = fill(0,nplot)
c_fix = 0.49182 #all sites, from Garcia (2019)

#---------------------Defining DDEEs---------------------------------------
function f(du,u,p,t)
    du .= p[nplot+1]*(abs.(p[1:nplot].+0im).^c_fix .-(u))
    nothing
end
#This function can be modified; it was just created to match Garcia's
#see the Alternative option at the end
```

```julia
#the abs.(..+0im) is a trick to avoid the optimization algorithm to stop
#if it finds complex numbers during the search
#This differential equation is being defined for the transformed variable H^c

#-----------------Solving DE numerically-----------------------------------
function DEsol_l(Age,parms)

    time_int      = (0,maximum(Age)) #Time interval to solve the DE
    u0            = inits            #Initial values to solve the DE

    prob          = ODEProblem{true}(f, u0,time_int, parms)
    sol           = solve(prob, Tsit5(),saveat = unique_AGES)
    sol_v         = vec(transpose(Array(sol)))
    return sol_v

end

#---------------------Log-lik function-------------------------------------------
function loglik(time, vari,θ)

    sig           = θ[nplot+2]
    hpred         = DEsol_l(time,θ)
    var_trans     = vari.^c_fix
    residual      = var_trans .- hpred
    sigma_fx      = exp(sig) #Constant variance, defined as sd

    result        = sum(logpdf.(Normal.(0,sigma_fx ),skipmissing(residual)))
    return -result

end
#Variance was defined as constant here, it can be modified to be
#any function. The exp guarantees positive numbers.


#Function to create initial values for the optimization
function get_par_inits(pars)
    inits = combine(groupby(data_lob, :Seed), :height => maximum)
    inits = inits[!,2]
    inits = [inits;pars]
end
par_inits = get_par_inits([0.1,log(0.1)]); #initial values for the optim
loglik(AGE,HLob,par_inits) #testing if everything is working
```

```
#----------------- Optimization--------------------------
func_twice = TwiceDifferentiable(b -> loglik(AGE, HLob, b), par_inits;
autodiff=:forward); #Function to be minimized, using automatic differentiation
@time opt = optimize(func_twice, par_inits,Newton(),Optim.Options(iterations =
1000000))
#Results
opt.minimum
res = Optim.minimizer(opt)
res[(nplot+1):(nplot+2)]
exp(res[nplot+2])

#Plot solution
inits_trans = inits.^c_fix
prob_sol = ODEProblem{true}(f,inits_trans,(0, 31), res)
sol      = solve(prob_sol,Tsit5(),saveat=unique_AGES);
plot(sol, legend=false)

data_lob.Hpred_trans = vec(transpose(Array(sol)))
data_lob.Hpred_orig  = data_lob.Hpred_trans.^(1.0/c_fix)


plot_res =  plot(data_lob.age,data_lob.height,
             group  = data_lob.Seed,
             seriestype=:scatter,
             xlabel = "Time (years)", ylabel = "HD (ft)",legend =
false,xlims     = (0, 30))
plot!(data_lob.age,data_lob.Hpred_orig,group  = data_lob.Seed,linewidth =
1,color="grey")

#---------------- Fit statistics---------------------------------------
#RMSE
sqrt(mean(skipmissing((HLob.-data_lob.Hpred_orig).^2)))

#Average bias
mean(skipmissing((HLob.-data_lob.Hpred_orig)))

#Average absolute bias
mean(abs.(skipmissing((HLob.-data_lob.Hpred_orig))))
```

```
#----------Alternative with variance function-------------------------------
#---------------------Log-lik function-------------
function loglik(time, vari,θ)

    r             = θ[nplot+1]
    sig           = θ[nplot+2]
    hpred         = DEsol_l(time,θ)
    var_trans     = vari.^c_fix
    residual      = var_trans .- hpred
    sigma_fx      = ((sig^2)/(2*r)).*(1.0.-exp.(-2*r*time)) #this is var not sd

    result        =sum(logpdf.(Normal.(0,sqrt.(sigma_fx)),skipmissing(residual)))

    return -result

end


#Function to create initial values
function get_par_inits(pars)
    inits = combine(groupby(data_lob, :Seed), :height => maximum)
    inits = inits[!,2]
    inits = inits
    inits = [inits;pars]
end
par_inits = get_par_inits([0.09,0.05]);
loglik(AGE,HLob,par_inits)


#----------------- Optimization---------------------------
func_twice = TwiceDifferentiable(b -> loglik(AGE, HLob, b), par_inits;
autodiff=:forward);
@time opt = optimize(func_twice, par_inits,Newton(),Optim.Options(iterations =
1000000))


#Results
opt.minimum
res = Optim.minimizer(opt)
res[(nplot+1):(nplot+2)]

var_teo =(((res[nplot+2])^2)/(2*res[nplot+1])).*(1.0.-exp.(-
2*res[nplot+1]*unique_AGES)) #checking the variance form
plot(unique_AGES, var_teo, ylims=(0.002,0.007))
```

```
#Plot solution
inits_trans = inits.^c_fix
prob_sol = ODEProblem{true}(f,inits_trans,(0, 31), res)
sol      = solve(prob_sol,Tsit5(),saveat=unique_AGES);
plot(sol, legend=false)

data_lob.Hpred_trans = vec(transpose(Array(sol)))
data_lob.Hpred_orig  = data_lob.Hpred_trans.^(1.0/c_fix)


plot_res =  plot(data_lob.age,data_lob.height,
             group  = data_lob.Seed,
             seriestype=:scatter,
             color =data_lob.Seed,
             xlabel = "Time (years)", ylabel = "HD (ft)",legend =
false,xlims     = (0, 30))
plot!(data_lob.age,data_lob.Hpred_orig,group  = data_lob.Seed,linewidth = 1,color
=data_lob.Seed)

#---------------- Fit statistics---------------------------------------
#RMSE
sqrt(mean(skipmissing((HLob.-data_lob.Hpred_orig).^2)))
#Average bias
mean(skipmissing((HLob.-data_lob.Hpred_orig)))
#Average absolute bias
mean(abs.(skipmissing((HLob.-data_lob.Hpred_orig))))
```

C2. CODE WITH PROPOSED METHODOLOGY AND SLASH PINE DATASET

This code cannot be run with publicly available data but is another example of how to use the proposed approach using a model different that the one used by Garcia. The Gompertz model from chapters 3 and 4, with a constant variance is exemplified here (without treatments).

This dataset has 72 different plots and local parameters corresponding to the initial state of the system for which the differential equation is solved are estimated. There are between 7 and 9 observations per plot (unbalanced dataset). The minimum age is 5 and the maximum is 31. If another dataset is used, change these values accordingly when plotting the solution. Each plot has a unique identifier called ‘myPlot’.

```
#Loading the packages
using DifferentialEquations
using Plots
using Optim
using ForwardDiff
using DataFrames
using Distributions
using CSV
using BenchmarkTools
using StaticArrays
using LinearAlgebra
using StatsBase
using NLSolversBase
using Preferences
using DelimitedFiles

#-----------Reading the data------------------------
data_s = DataFrame(CSV.File("Data/data_trt.csv"))
data_s = sort!(data_s,[:myPlot,:AGE])

#Definig variables as constant (better for optimization)
const nplot = length(unique(data_s.myPlot))
const AGE   = data_s.AGE
const nAge  = length(unique(AGE))
const unique_AGES = sort(unique(AGE))
const unique_plot = sort(unique(data_s.myPlot)) #Needs to be sorted by myplot
```

```
#Expanding dataset to be used in matrix form
#This is necessary to work efficiently with the unbalanced data
df_full      = DataFrame(myPlot  = repeat(unique_plot,inner = nAge),
                         AGE   = repeat(unique_AGES,outer = nplot))
data_sfullNA = sort!(leftjoin(df_full,data_s, on =[:myPlot,:AGE]),[:myPlot,:AGE])
const HDOM_NA = data_sfullNA.HDOM
#----------- Required functions----------------------
#Defining the Gompertz differential equation
function f(du,u,p,t)
    du.=(u.*(p[1])).*exp.(-p[2].*t)
    nothing
end

#Solving DE numerically
function DEsol_l(Age,parms)
    time_int      = (minimum(Age),maximum(Age))
    u0            = @view parms[1:nplot] #@view is used to speed up the optim

    prob          = ODEProblem{true}(f, u0,time_int, @view parms[(nplot+1):(nplot+2)])
    sol           = solve(prob, Tsit5(),saveat = unique_AGES)
    sol_v         = vec(transpose(Array(sol)))
    return sol_v

end

#Log-lik function
function loglik(time, var,θ)

    sigma         = exp(θ[nplot+3]) #constant variance
    hpred         = DEsol_l(time,θ)
    residual      = var .- hpred
    result        = sum(logpdf.(Normal(0, sigma),skipmissing(residual)))

    return -result

end

#Function to create initial values
function get_inits(pars)
    inits = combine(groupby(data_s, :myPlot), :HDOM => minimum)
    inits = inits[!,2]
    inits = [inits;pars]
end
inits = get_inits([0.4,0.12,0]);
loglik(AGE,HDOM_NA,inits)
```

```
#---------------Optimization--------------------------------------
@time result = optimize(b -> loglik(AGE, HDOM_NA, b),
inits,Optim.Options(iterations = 1000000))
#Results
result.minimum
res = Optim.minimizer(result)
res[(nplot+1):(nplot+3)]

#Plot solution
prob_sol = ODEProblem{true}(f,res[1:nplot],(5, 31), res[(nplot+1):(nplot+3)])
sol      = solve(prob_sol,Tsit5(),saveat=unique_AGES);

plot_res =  plot(data_s.AGE,data_s.HDOM,
             group  = data_s.myPlot,
             xlabel = "Time (years)", ylabel = "HD (m)",
              legend =false, xlims     = (0, 35))
plot!(sol, linewidth = 1,  xlabel    = "Age (years)",color="grey" ,
      xlims     = (0, 35))

#---------------- Fit statistics--------------------------------------
#RMSE
sqrt(mean(skipmissing((HDOM_NA.-vec(transpose(Array(sol)))).^2)))
#Average bias
mean(skipmissing((HDOM_NA.-vec(transpose(Array(sol))))))
#Average absolute bias
mean(abs.(skipmissing((HDOM_NA.-vec(transpose(Array(sol)))))))

#----------------- Optimization with AD to get- Hessian and standard error---
func_twice = TwiceDifferentiable(b -> loglik(AGE, HDOM_NA, b), inits;
autodiff=:forward);
@time opt = optimize(func_twice, inits,Newton(),Optim.Options(iterations =
1000000))

#Solution
Optim.minimizer(opt)
opt.minimum

#Parameters
parameters = Optim.minimizer(opt); #All parameters, including local
parameters[(nplot+1):(nplot+3)] #Want to focus on the global parameters

#Hessian-standard error
numerical_hessian = hessian!(func_twice,parameters);
var_cov_matrix = inv(numerical_hessian);
```

```
β    = parameters;
temp = diag(var_cov_matrix)
SE   = sqrt.(temp)
SE[(nplot+1):(nplot+3)] #Standard error of the global parameters
```