Modeling Leaf Area as a Function of Environmental Variables to Predict Growth in Loblolly Pine Plantations

by

Stephen M. Kinane

(Under the Direction of Cristian R. Montes)

Abstract

Light interception is a key ecological indicator that is strongly related to forest productivity. Its application has increased over the last few decades with the advent of process based models. Evaluating light interception requires an assessment of the amount of leaves present in a given forest plantation. This assessment is done with indirect methods such as spectral indices derived from satellite data, that empirically relate different satellite bands with ground based leaf area measurements. These indices are not exempt from measurement error, due to the many factors affecting light reflection, refraction, and scattering in its path between the leaves and the satellite sensor. As a result, the assumption of error free predictors is violated, resulting in a biased estimator for leaf area. This is also known as an error-in-variable problem. To bridge this gap, a modeling framework is presented that account for stochastic deviations in the independent predictors, resulting in an improved model when compared to other published research.

Leaf area, as an indicator of site productivity is a function of climate and soil factors. Many of the relations between these environmental predictors are described through complex relations inside process based models, with assumptions like carbon partitioning fractions,

instantaneous carbon allocation and fixed respiration costs. These models trade simplicity for comprehensiveness, finding low acceptance from practitioners interested in operational applications. To bridge this gap, a semi-empirical model was developed to allow the description of foliage display in loblolly pine, as a function of environmental variables. The model was able to predict leaf area display, including a parameter responsible for foliage carrying capacity. The model was parameterized using plot level observations from two research studies across 24 locations across the southeast US.

Finally, a mechanistic model is presented that utilizes the foliage display model and relates it with a dominant height equation for loblolly pine. The model was fit as a system of simultaneous differential equations using local and global parameters. The model was able to correctly describe dominant height over age, bridging the gap between an empirical growth and yield approach and an important ecological indicator, estimated with an unbiased equation from satellite time series.

INDEX WORDS: Loblolly pine, remote sensing, Kalman filter, environmental variables, leaf area index, growth and yield

Modeling Leaf Area as a Function of Environmental Variables to Predict Growth in Loblolly Pine Plantations

by

Stephen M. Kinane

B.S., North Carolina State University, 2012

M.S., The University of Georgia, 2014

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

2020

Modeling Leaf Area as a Function of Environmental Variables to Predict Growth in Loblolly Pine Plantations

by

Stephen M. Kinane

Major Professor: Cristian R. Montes

Committee: Bronson P. Bullock
Rachel L. Cook
Deepak Mishra

Electronic Version Approved:

Ron Walcott
Dean of the Graduate School
The University of Georgia
December 2020

Dedication

To Claire

Acknowledgments

My gratitude to the Warnell School of Forestry and Natural Resources at the University of Georgia and to the Plantation Management Research Cooperative for providing the funding necessary to complete my degree is insurmountable. Many thanks to the researchers, staff, field crew, and cooperative members of the PMRC for providing the opportunity, means, and data used to complete this research. I am thankful to the awards committee for providing me the Clutter scholarship, a true honor.

A sincere thank you to my august advisor, Cristian, for taking a chance on me and expanding my island. A constant source of ideas, inspiration, and laughter, I appreciate his wisdom bestowed upon me. His support throughout my program is something I will always be thankful for. I would like to thank Dr. Bronson Bullock for igniting my interest in biometrics while an undergraduate at NC State and following me down to UGA to be an integral part of my education. Thank you to Drs. Deepak Mishra and Rachel Cook for their time, ideas, and resources to help me complete this research. I am also quite appreciative to my fellow graduate students and visiting researchers of the PMRC Biometrics Lab for their help, ideas, and jokes that made this program more fun than it should have been.

This could not have been completed without the unwavering love and support from my wife, Claire, who provided certainty when the analyses suggested otherwise. Thanks to our good dog, Shelby, for always greeting me no matter how late I came home. Many thanks to the boys back home who were always a phone call away when I needed it the most. An extended thank you and appreciation to my family, who provided a constant source of patience even though it wasn't always returned.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Understanding the relationships between stand productivity and the environment is essential for developing models that relate the two components. Stand productivity, the result of physiological processes, relates annual growth observed in plantations to the environment that supports it. One method of estimating stand productivity uses leaf area index, the dimensionless quantity representing the amount of one-sided leaf surface area over unit area of soil (Allen et al., 1998; Peduzzi et al., 2012; Waring, 1983). Leaf area index is an indicator of the ability of a stand to exchange material and energy with its environment (Grier and Running, 1977; Vose et al., 1994). Prior studies have supported the positive relationship between leaf biomass and productivity in loblolly pine stands (Teskey et al., 1987; Vose and Allen, 1988; Albaugh et al., 1998; Jokela and Martin, 2000). Using the dynamic properties of leaf area, it has been shown to respond to cultural treatments, including fertilization and weed control, increasing the leaf area a stand can support (Colbert et al., 1990; Vose and Allen, 1988; Dalla-Tea and Jokela, 1991; Jokela and Martin, 2000).

Forest management is traditionally an economically driven process, sometimes yielding to auxiliary constraints such as wildlife habitat or recreation, that requires long term projections of stand attributes to define current management decisions. Typically these projections require the modeling of a response to silvicultural activities and productivity is governed by a stand's estimated site index. Site index, the average height of the dominant and co-dominant trees at a base age, serves as the integrator of all environmental processes observed in a stand's rotation and their effect on dominant height, projected to the selected base age.

Inherent to this methodology is its inability to differentiate environmentally induced changes in productivity and the lasting effects it has on a stand's overall yield at rotation. Utilizing a biophysical variable, such as leaf area index, to be the underlying productivity driver of stand growth and yield systems can provide increased access to modeling scenarios in which a changing climate alters stand level processes, effectively reorienting the growth trajectory and associated variability (Vose and Swank, 1990; Sampson et al., 1997). The results of this research aim to: 1) explore methods to enhance leaf area index estimation accuracy, 2) understand the effects of environment on leaf area index across a stand's rotation, and 3) understand how leaf area index and light interception can be used to facilitate a dominant height growth model.

1.2 Literature Review

While theoretical limitations of leaf area index in loblolly pine in the southeastern United States have been predicted to be up to 5 m^2/m^2, field measurements have found that stands are routinely well below these limits (Jarvis and Leverenz, 1983; Gholz, 1986). Environmental factors, such as temperature, nutrient availability, water availability, and radiation, have been suggested to be the limiting factors keeping observed leaf area from its theoretical limits (Linder, 1987; Vose and Allen, 1988). Because leaf area indices for loblolly pine vary across the natural range, it has been hypothesized that climatic differences between the sites are the source of variability (Gholz et al., 1991; Benson et al., 1992).

Leaf area index dynamics have been proven to be influenced by several environmental and atmospheric conditions. These conditions not only affect the maximum leaf area index achieved in a stand, but also the seasonal variation observed (Vose et al., 1994). While temperature has been hypothesized to be influential on leaf area index (Gholz, 1986), calls for further research into the subject matter have been promoted due to mainly anecdotal evidence being reported (Vose et al., 1994). Model simulations have provided evidence for temperature influences on leaf area, showing that sites with lower nighttime temperatures

having lower maintenance respirations rates, increasing the available carbon for allocation to total leaf area (Cropper and Gholz, 1994). Temperature also affects the rate of leaf occurrence and growth, increasing up to an optimal temperature and decreasing thereafter (Monteith, 1977). Trees grown under warmer climates have been modeled to have greater leaf mass, leaf number, and leaf area (Way and Oren, 2010). In evergreen species, these responses were not as distinct as deciduous species, but trends continued to be positive as temperatures increased, suggesting that many temperate forests are operating below optimal temperatures (Way and Oren, 2010; Ryan, 2010).

Differences in leaf extension rates are shown by species in whether they are greater in day or night temperatures, how easily accessible nitrogen is, or during light or dark periods (McDonald et al., 1992). Additional research supported the idea that leaf area index is affected by temperature, showing that warmer sites had a higher change in annual foliage respiration relative to change in assimilation compared with cooler sites (Ryan et al., 1994).

Previous studies show a positive linear relationship between growth rate and radiation interception (Monteith, 1977; Jarvis and Leverenz, 1983). Leaf area reaches its optimum when the leaves at the bottom of the canopy can utilize radiation to, at minimum, maintain a balance between dry matter assimilation and respiration loss (Saeki, 1959). As leaf area increases, attenuation of photosynthetically active radiation increases, resulting in an increase in efficiency in conversion to stemwood (McCrady and Jokela, 1998). It has been shown that species with greater tolerances to shade are able to achieve higher peak leaf area indices (Vose et al., 1994). In other species, it has been shown that increasing levels of radiation increases the optimal leaf area index that maximizes growth (Black, 1963; Monsi and Saeki, 1953). Radiation is the driving force behind the seasonality seen in the natural range of loblolly pine, affecting the day length, temperature, evaporation, and humidity (Jarvis and Leverenz, 1983).

Increased availability of nutrients, mainly nitrogen, on limiting sites showed positive responses in leaf area (Fox et al., 2007; Jokela et al., 2004; Samuelson et al., 2008; Gholz

et al., 1991; Colbert et al., 1990; Campoe et al., 2013; Vose and Allen, 1988). This increase in leaf area index may result from the increased total number of fascicles per tree and mass per unit length of needles that results from increased nutrient availability, needle size, total number of needles, and decreased mortality (Gholz, 1986; Linder, 1987; Raison et al., 1992a). Other studies have shown that in the southeast United States, nutrient availability is the main driver of leaf area development (Fox et al., 2007). Increased leaf area may result from changes in crown structure that have been reported on resulting from increased availability to phosphorus at stand establishment (Jokela et al., 1989). Increasing availability of nutrients modifies the proportion of carbon allocation between above and below ground sinks, favoring aboveground processes (Albaugh et al., 1998; Vose and Allen, 1988; Cannell, 1989). Within year peak leaf area index, usually achieved in August, has been shown to maintain this positive relationship with stem volume increment and total biomass production across a range of nutrient and water availability (Albaugh et al., 1998).

Increased water availability has been shown to increase leaf area index in coniferous forests (Grier and Running, 1977). This relationship may be a result from increases in needle length in response to increased water availability (Raison et al., 1992b). In response to water stress, trees close stomata and senesce leaves, thus reducing the leaf area of a stand (Linder, 1987). The efficiency of energy conversion is influenced by water availability, decreasing under stress and increasing under availability until the optimum is reached (Linder, 1987). In areas with limiting water, irrigation treatments were shown to increase the growing period for carbon assimilation under optimal temperature conditions for eucalyptus, thus allowing for higher rates of leaf production and ultimately higher leaf area indices (Pereira et al., 1994). In the southeastern U.S., available water is not the most limiting resource, thus irrigation treatments have shown insignificant effect on peak leaf area (Campoe et al., 2013). Interactions between nutrient and water availability are displayed since adequate water availability is necessary for trees to utilize available soil nitrogen (Benson et al., 1992). These interactions have been shown to affect leaf area index in other species through effects on leaf size and

expansion (Pereira et al., 1994). These variables and their associated effects on leaf area all affect the amount of light intercepted by the canopy, which has been shown to have a strong linear relationship to total biomass production (Linder, 1987; Monteith, 1977; Vose and Allen, 1988). Cultural treatments, including fertilization, competition control, initial planting density, and thinning, can be used to manipulate the leaf area in each stand to promote increased growth rates. Timing is essential and the faster a stand reaches its peak leaf area, the more growth that can occur.

With the understanding of how these environmental variables affect leaf area and ultimately stand growth, process based models attempt to simulate the theoretical relationships between the stand and its environment. Process based models have been used for decades to promote understanding of the systems that drive growth (Running and Coughlan, 1988; Landsberg and Waring, 1997; Sands, 2003, 2004). Process based models, in comparison to empirically based models, are easily able to incorporate many site factors, including changes in site conditions, insect and disease effects, and risk (Sands, 2004). FOREST-BCG incorporated leaf area index as one of its primary independent variables, used to estimate the transpiration, photosynthesis, carbon allocation, respiration, and litterfall of the canopy structure (Running and Coughlan, 1988). In their research, Running and Coughlan (1988) found that leaf area index had different effects on the model in varying climates. One of the most popular processed based models utilizing the dynamics controlling leaf area development in loblolly pine stands is Physiological Principles in Predicting Growth (3-PG) (Landsberg and Waring, 1997). This model utilizes input variable modifiers to impose the ecophysiological constraints and established ecological constants on forest growth (Landsberg and Waring, 1997). In relation to leaf area, 3-PG utilizes environmental modifiers to impose constraints on leaf area development in stands, thus affecting overall projected productivity. Leaf area is also a function of allocation parameters in 3-PG, all which can be parameterized for specific species and situations. Other simulations, using process model BIOMASS, showed that

the long-term average productivity of loblolly pine was best predicted by leaf area index as compared with soil available water-holding capacity (Sampson and Allen, 1999).

Empirical methods, which utilize the statistical relationships between variables' behavior to model leaf area index development, have been developed for several agricultural species (Korzukhin et al., 1996; Chen et al., 2014; Schultz, 1992). To predict changes in leaf area of maize, researchers modeled leaf expansion and senescence were modeled as a function of species variety, leaf characteristics, planting density, and fertilization rates (Chen et al., 2014). A *Vitis* model utilized thermal time, canopy structure, and hedging effects to successfully simulate leaf area development throughout the growing season (Schultz, 1992). Tesemma et al. (2014) incorporated environmental variables in predicting the development of leaf area index for crops, pastures, and trees in southern Australia using empirical models. Process and empirical models can be appropriate methods for modeling different circumstances. While process models are less variable and can be used across a wide range of conditions, when predictions are made in a new condition, questions concerning the accuracy arise (Korzukhin et al., 1996). A criticism of empirical methods is the lack of explanation for a given process, only the statistical relationship between variables presented, thus extrapolation outside the conditions is limited (Korzukhin et al., 1996). It has been shown that loblolly pine leaf area can be fit to an empirical dynamic equation across a variety of conditions in soil fertility and water availability (Montes, 2012). The need exists for empirical leaf area equations to incorporate environmental variables to predict stand leaf area development in loblolly pine. With the understanding of the importance of leaf area in relating the productivity of a site, incorporation of leaf area into an empirical yield model may provide a better projection of stand growth with lower variation. Understanding of the environmental variables that affect leaf area on a site may aid in the prediction of the peak leaf a site can sustain, helping foresters to best choose silvicultural plans on a site-specific basis.

1.3 Rationale and Significance

Advances in the availability and ease of access to large collections of remotely sensed data have allowed for forest practitioners and researchers to gain a better understanding of how forested stands develop and change over time. Coupled with growth data obtained from long term, remeasured permanent sampling plots, the ability to understand the effects of climatic variations on forest productivity have greatly increased. The ability to estimate biophysical variables from remotely sensed data that provide insight into the health and structure of any given stand across the world is a powerful tool that can be used to model past, present, and future biological and economical scenarios.

The Landsat program has provided an outstanding amount of data since it's inception in the 1970s. The Landsat 5 and 7 platforms have provided essentially continuous high resolution, multi-spectral earth observation from 1984 to current (2020) on relatively short 8-16 day revisit period. The synchronicity of the specifications between these two platforms has allowed for cross-platform analysis of the data, providing almost forty years of imagery that can be used in long term growth and yield studies established by organizations, such as university research cooperatives. Provisions of the Landsat program include similarities in sensor specifications such that newer (Landsat 8) and future sensors can be used with models designated for Landsat 5 and/or 7 through tuning and reparameterization.

Using the remotely sensed data comes with the responsibility of understanding the assumptions and limitations to how it was observed, manipulated through processing techniques, and provided to the end user. Coupled with the knowledge of modeling assumptions in the statistical framework, using appropriate techniques to reduce or remove the effects of observation and modeling errors is essential to formulating unbiased models.

Using growth models to understand the complex dynamics of the biophysical variables under analysis provides a more cause and effect comprehension rather than correlation. Growth models provide the ability to test hypotheses on the effects of external forces on a system and to provide the long term trends that result from these scenarios outside of

what we have observed. Using leaf area index as the productivity driver of a forest growth system is not a new nor original idea. But with the massive sets of environmental variables, remotely sensed imagery, and long term forest growth and yield data now widely available, the ability to look at trends across a wide geographical range with high temporal and spatial resolution can provide new insights or reinforce currently held beliefs to our understanding of the uncertainty surrounding forest productivity.

1.4 Goals, Objectives, and Hypotheses

1.4.1 General Goal

To understand the effects of environmental variables on leaf area index and how that relates to forest productivity.

1.4.2 Study objectives

1. To provide an unbiased model that estimates leaf area index in loblolly pine plantations from satellite imagery.
    (a) Hypothesis: No relationship between vegetation indices derived from remotely sensed data and leaf area index will be identified.
    (b) Methodological objective: Account for observation and process errors in the modeling framework by using error invariable methods to reduce the effects of errors on estimated parameter values.
2. To determine the influence of environmental variables on loblolly pine leaf area index growth and development in the southeastern U.S.
    (a) Hypothesis: No effect of the environment will be present on leaf area index growth and development.

(b) Methodological objective: Construct a mechanistic model describing leaf area index growth and test the effects of environmental variables on the growth model.

3. To evaluate dominant height as a function of leaf area index and its interaction with the environment.

   (a) Hypothesis: No relationship between leaf area index and dominant height growth will be identified.

   (b) Methodological objective: Using a mechanistic model of leaf area index growth, imposed environmental modifiers, and monthly solar radiation data, calculate intercepted radiation and tests its effects on a dominant height growth model.

1.5 References

Albaugh, T. J., Allen, H. L., Dougherty, P. M., Kress, L. W., and King, J. S. (1998). Leaf area and above- and belowground growth responses of loblolly pine to nutrient and water additions. *Forest Science*, 44(2):317–328.

Allen, R. G., Pereira, L., Raes, D., and Smith, M. (1998). Crop evapotranspiration: Guidelines for computing crop requirements. *Irrigation and Drainage Paper No. 56, FAO*, (56):300.

Benson, M., Myers, B. J., and Raison, R. J. (1992). Dynamics of stem growth of Pinus radiata as affected by water and nitrogen supply. *Forest Ecology and Management*, 52(1-4):117–137.

Black, J. (1963). The interrelationship of solar radiation and leaf area index in determining the rate of dry matter production of swards of subterranean clover (Trifolium subterraneum L.). *Australian Journal of Agricultural Research*, 14(1):20–38.

Campoe, O. C., Stape, J. L., Albaugh, T. J., Lee Allen, H., Fox, T. R., Rubilar, R., and Binkley, D. (2013). Fertilization and irrigation effects on tree level aboveground net primary production, light interception and light use efficiency in a loblolly pine plantation. *Forest Ecology and Management*, 288:43–48.

Cannell, M. G. R. (1989). Physiological basis of wood production: a review. *Scandinavian Journal of Forest Research*, 4:459–490.

Chen, G., Zhang, J., Liu, P., and Dong, S. (2014). An empirical model for changes in the leaf area of maize. *Canadian Journal of Plant Science*, 94(4):749–757.

Colbert, S. R., Jokela, E. J., and 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. *Forest Science*, 36(4):995–1014.

Cropper, W. P. and Gholz, H. L. (1994). Evaluating Potential Response Mechanisms of a Forest Stand to Fertilization and Night Temperature : A Case Study Using Pinus Elliottii. *Ecological Bulletins*, (43):154–160.

Dalla-Tea, F. and Jokela, E. J. (1991). Needlefall, canopy light interception, and productivity of young intensively managed slash and loblolly pine stands. *Forest Science*, 37(5):1298–1313.

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

Gholz, A. H. L., Vogel, S. A., Cropper, W. P., Mckelvey, K., Ewel, K. C., Teskey, R. O., and Curran, P. J. (1991). Dynamics of canopy structure and light interception in pinus elliottii stands, north Florida. *Ecological Monographs*, 61(1):33–51.

Gholz, H. L. (1986). Crown and Canopy Structure in Relation to Productivity. In Fujimori, T. and Whitehead, D., editors, *Crown and Canopy Structure in Relation to Productivity*, pages 224–242. Forestry Products Research Institute, Ibaraki, Japan.

Grier, C. G. and Running, S. W. (1977). Leaf area of mature northwestern coniferous forests: Relation to site water balance. *Ecology*, 58(4):893–899.

Jarvis, P. G. and Leverenz, J. W. (1983). *Productivity of Temperate, Deciduous and Evergreen Forests*, chapter 8, pages 233–280. Springer Berlin Heidelberg, Berlin, Heidelberg.

Jokela, E. J., Dickens, E., Barnett, J., and Hubbard, W. (2004). Nutrient management of southern pines. In *Slash pine: still growing and growing! Proceedings of the slash pine symposium*, pages 27–35.

Jokela, E. J., Harding, R. B., and Nowak, C. A. (1989). Long-term effects of fertilization on stem form, growth relations, and yield estimates of Slash pine. *Forest Science*, 35(3):832–842.

Jokela, E. J. and Martin, T. A. (2000). Effects of ontogeny and soil nutrient supply on production, allocation, and leaf area efficiency in loblolly and slash pine stands. *Canadian Journal of Forest Research*, 30(10):1511–1524.

Korzukhin, M. D., Ter-Mikaelian, M. T., and Wagner, R. G. (1996). Process versus empirical models: which approach for forest ecosystem management? *Canadian Journal of Forest Research*, 26(5):879–887.

Landsberg, J. and Waring, R. (1997). A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. *Forest Ecology and Management*, 95(3):209–228.

Linder, S. (1987). Responses to water and nutrients in coniferous ecosystems. *Ecological Studies*, 61:180–202.

McCrady, R. and Jokela, E. (1998). Canopy dynamics, light interception, and radiation use efficiency of selected loblolly pine families. *Forest Science*, 44(1):64–72.

McDonald, A. J. S., Stadenberg, I., and Sands, R. (1992). Diurnal variation in extension growth of leaves of Salix viminalis. *Tree Physiology*, 11(2):123–132.

Monsi, M. and Saeki, T. (1953). On the Factor Light in Plant Communities and its Importance for Matter Production. *Japanese Journal of Botany*, 14:22–52.

Monteith, J. L. (1977). Climate and the Efficiency of Crop Production in Britain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 281(980):277–294.

Montes, C. R. (2012). *A Resource Driven Growth and Yield Model for Loblolly Pine Plantations*. PhD thesis, North Carolina State University.

Peduzzi, A., Wynne, R. H., Fox, T. R., Nelson, R. F., and Thomas, V. A. (2012). Estimating leaf area index in intensively managed pine plantations using airborne laser scanner data. *Forest Ecology and Management*, 270:54–65.

Pereira, J., Madeira, M. V., Linder, S., Tome, M., and Araujo, M. C. (1994). Biomass production with optimized nutrition in Eucalyptus globulus plantations. In *Eucalyptus for Biomass Production. Commission of the European Communities*, pages 13–30.

Raison, R. J., Khanna, P. K., Benson, M. L., Myers, B. J., McMurtrie, R. E., and Lang, A. R. G. (1992a). Dynamics of Pinus radiata foliage in relation to water and nitrogen stress: II. Needle loss and temporal changes in total foliage mass. *Forest Ecology and Management*, 52(1-4):159–178.

Raison, R. J., Myers, B. J., and Benson, M. L. (1992b). Dynamics of Pinus radiata foliage in relation to water and nitrogen stress: I. Needle production and properties. *Forest Ecology and Management*, 52(1-4):139–158.

Running, S. W. and Coughlan, J. C. (1988). A general model of forest ecosystem processes for regional applications I. Hydrologic balance, canopy gas exchange and primary production processes. *Ecological Modelling*, 42(2):125–154.

Ryan, M. G. (2010). Temperature and tree growth. *Tree Physiology*, 30(6):667–668.

Ryan, M. G., Linder, S., Vose, J. M., Hubbard, R. M., and Ryan, G. (1994). Dark Respiration of Pines. *Ecological Bulletins*, (43):50–63.

Saeki, T. (1959). Interrelationships Total between Leaf Amount, Light Distribution Photosynthesis in a Plant Community *. *Japanese Journal of Botany*, 14:22–52.

Sampson, D. and Allen, H. (1999). Regional influences of soil available water-holding capacity and climate, and leaf area index on simulated loblolly pine productivity. *Forest Ecology and Management*, 124(1):1–12.

Sampson, D. A., Vose, J. M., and Allen, H. L. (1997). A conceptual approach to stand management using leaf area index as the integral of site structure, physiological function, and resource supply. *Proceedings of the ninth biennial southern silvicultureal research conference*, 2:25–27.

Samuelson, L. J., Butnor, J., Maier, C., Stokes, T. A., Johnsen, K., and Kane, M. (2008). Growth and physiology of loblolly pine in response to long-term resource management: defining growth potential in the southern United States. *Canadian Journal of Forest Research*, 38(4):721–732.

Sands, P. (2004). Adaptation of 3-PG to novel species : guidelines for data collection and parameter assignment. Technical Report Technical Report 141, CRC for Sustainable Production Forestry and CSIRO Forestry and Forest Products, Hobart, Australia.

Sands, P. J. (2003). Process-based Models for Forest Management – Integrating determinants of growth into practical management systems. Technical Report Technical Report 126, Cooperative Research Centre for Sustainable Production Forestry.

Schultz, H. R. (1992). An empirical model for the simulation of leaf appearance and leaf area development of primary shoots of several grapevine (Vitis vinifera L.) canopy-systems. *Scientia Horticulturae*, 52(3):179–200.

Tesemma, Z. K., Wei, Y., Western, A. W., and Peel, M. C. (2014). Leaf area index variation for cropland, pasture and tree in response to climatic variation in the Goulburn-Broken catchment, Australia. *Journal of Hydrometeorology*, page 140401140559005.

Teskey, R. O., Bongarten, B. C., Cregg, B. M., Dougherty, P. M., and Hennessey, T. C. (1987). Physiology and genetics of tree growth response to moisture and temperature stress: an examination of the characteristics of loblolly pine (Pinus taeda L.). *Tree Physiology*, 3(1):41–61.

Vose, J. M. and Allen, H. L. (1988). Leaf area, stemwood growth, and nutrition relationships in loblolly pine. *Forest Science*, 34(3):547–563.

Vose, J. M., Dougherty, P. M., Long, J. N., Smith, F. W., Gholz, H. L., and Curran Vose, P. J. (1994). Factors influencing the amount and distribution of leaf area of pine stands. *Ecological Bulletins*, 43:102–114.

Vose, J. M. and Swank, W. T. (1990). A conceptual model of forest growth emphasizing stand leaf area. In Dixon, R. K., Meldahl, R. S., Ruark, G. A., and Warren, W. G., editors, *Process Modeling of Forest Growth Responses to Environmental Stress*, chapter 24, pages 278–287. Timber Press, Inc., Portland, OR.

Waring, R. (1983). Estimating forest growth and efficiency in relation to canopy leaf area. In MacFadyen, A. and Ford, E., editors, *Advances in Ecological Research*, chapter 13, pages 327 – 354. Academic Press.

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

Chapter 2

A Model to Estimate Leaf Area Index in Loblolly Pine Plantations in the Southeastern United States[1]

[1]Kinane, S.M., T.J. Albaugh, D. Mishra, and C.R. Montes. To be submitted to *Remote Sensing*.

Abstract

Vegetation indices calculated from remotely sensed satellite imagery are commonly used within empirically derived models to estimate leaf area index in loblolly pine plantations in the southeastern United States. The vegetation index data used to parameterize the models typically comes with observation errors, resulting in biased parameters. Using error-in-variable methods, we evaluated multiple vegetation indices, calculated errors associated with their observations, and corrected for them in the modeling process. We found that the normalized difference moisture index provided the best correlation with below canopy leaf area index measurements. A nonlinear model that accounts for the fertilization status of the stand was found to provide the best estimates of leaf area index. The analysis in this research provides a more extensive evaluation on common vegetation indices used to estimate leaf area index in loblolly pine plantations and a modeling framework that extends beyond the typical linear framework used.

2.1 Introduction

Foliage biomass is an ecological parameter often used as an indicator of forest productivity in commercial forest plantations (Waring, 1983; Shi and Cao, 1997). Its direct determination requires expensive destructive sampling methods that are difficult to measure and laborious to acquire which is why practitioners turned to indirect biomass estimation methods that rely on light reflectance as captured by remote sensing devices for its determination (Gower et al., 1999). The increased availability of surface reflectance products has allowed for the use of higher quality data in biophysical modeling due to the correction of atmospheric distortions, providing more consistent and less variable data (Claverie et al., 2015; Masek et al., 2006). Over the last two decades, use of remotely sensed data has increased as products become more readily available, allowing commercial forest plantation managers to use it as part of their decision making process (Flores, 2003; Fox et al., 2007; Oswalt et al., 2014; Parresol et al., 2017). Estimates of structure, productivity, disturbance, and phenology of loblolly pine (*Pinus taeda* L.) plantations are some of the applications for these data (Coleman et al., 1990; Flores et al., 2006; Iiames et al., 2008; Iverson et al., 1989; Sivanpillai et al., 2006; Stein et al., 2014; Yang et al., 2017). Nevertheless the application, most indirect methods aim at estimating leaf area index (LAI). LAI, the one-sided leaf surface area over a fixed ground area, is a biophysical parameter that has been used to understand productivity drivers in loblolly pine stands (Vose and Allen, 1988; Albaugh et al., 1998; Peduzzi et al., 2012; Waring, 1983). LAI serves as a proxy for the processes involving intercepted photosynthetically active radiation (PAR) and atmospheric gas and moisture exchange (Monteith, 1972; Vose and Swank, 1990; Mahowald et al., 2016). As an indicator for stand productivity, LAI has shown a positive correlation with stem volume increment, biomass production, site index, basal area, and stand density index (Albaugh et al., 1998; Vose and Allen, 1988; Teskey et al., 1987; Jokela and Martin, 2000; Gonzalez-Benecke et al., 2012; Dalla-Tea and Jokela, 1991). Similarly, LAI is responsive to changes in resource availability due to silvicultural inputs like fertilization (Albaugh et al., 2004; Jokela et al., 2004; Sampson et al., 2011).

Previous models to estimate LAI from passive remotely sensed data have utilized varying spectral, spatial, and temporal resolutions of satellite imagery and different modeling techniques to estimate in situ LAI measurements (Wenze Yang et al., 2006; Qi et al., 2000; Flores et al., 2006; Peduzzi et al., 2012; Shoemaker and Cropper, 2008). Estimation of LAI and other biophysical parameters from satellite imagery rely on the algebraic manipulation of spectral reflectance bands to form vegetation indices (VI) which are then used to model ecological processes. VIs, such as the normalized difference vegetation index (NDVI) and the simple ratio (SR), are used to separate vegetation dynamics from background processes allowing researchers to use remotely sensed data to model the underlying biological phenomena (Wang et al., 2005; Flores et al., 2006; Qu and Zhuang, 2018). To better understand the behavior of canopy spectral reflectance as sampled from satellite imagery, we need to start at the underlying structure of the individual leaf and the behavior of electromagnetic energy. Whether it is reflected, absorbed, or transmitted, the behavior of electromagnetic energy in a specific wavelength depends on the surface with which it interacts (Curran, 1980). At the individual leaf level, physiological properties, including cell structure, concentration of pigments, air, and water within the leaf determine the reflectance behavior (Knipling, 1970). When scaled to the canopy level, orientation and concentration of leaves, type and concentration of background matter, and the angle of illumination affect reflectance behavior due to the potential for shadows (Knipling, 1970). With these sources of potential variation in mind, measuring leaf and canopy reflectance to estimate biophysical parameters must take advantage of known spectral responses to remove noise. One common source of variation in spectral reflectance is a function of the atmospheric scattering which vary spatially and temporally (Masek et al., 2006; Roy et al., 2016). To reduce the effects of these perturbations, surface reflectance data in satellite imagery is created from top of atmosphere reflectance using radiative transfer models and information about atmospheric variables to correct for absorption and scattering of radiation as it passes through the atmosphere (Vermote et al., 1997). Furthermore, variations in sensor and solar angles can lead to directional effects in the

observed reflectance values (Roy et al., 2016). These effects can be reduced using ansiotropic functions, such as the bidirection reflectance distribution function (BRDF), to normalize all the pixels in a image to nadir reflectance. After outside influences of radiation are corrected, VIs such as NDVI work by associating known reflectance characteristics, such as high absorption of visible light and high reflectance of near infrared, to build ratio equations that relate in situ conditions to remotely sensed estimates (Knipling, 1970; Rouse et al., 1974). In this case, NDVI provides an estimate of the sample area chlorophyll concentration from healthy, green vegetation having high reflectance rates for near infrared and high absorption rates for green visible light.

Other considerations for developing a prediction model based on remotely sensed data include selection of a VI that best relates in situ conditions to the dependent variable. VIs commonly suffer from saturation effects in densely vegetated stands where light in the red bands is absorbed by chlorophyll resulting in losses of sensitivity in sites with LAIs greater than 2 (Huete et al., 1997; Huete and Justice, 1999; Mutanga and Skidmore, 2004). To avoid issues with saturation, (Gitelson et al., 1996) recommended the use of VIs constructed without bands in the red portion of the spectrum. Other considerations for selection of a VI include the ability to differentiate changes in leaf physical properties (i.e. changes in concentration of pigments due to stress) and changes in amount of leaf area (i.e. reduction of foliage biomass, increasing amount of background reflectance) (Knipling, 1970).

The most commonly used model for predicting loblolly pine LAI in the southeastern United States utilized a linear model framework and the simple ratio vegetation index (Flores et al., 2006). In spite of the large coefficient of determination found for the model in (Flores et al., 2006), some sources of error were omitted in model development that influence its practical application. These errors include data acquisition errors in the sensors and processing errors from geometric or radiometric rectifications which lead to biased predictions (Flores et al., 2006; Lunetta et al., 1991; Fernandes and G. Leblanc, 2005). Curran and Hay discussed the inadequacies of simple regression for remote sensing modeling resulting

from violation of the measurements without errors assumption, and when that assumption is violated, parameter estimates will be underestimated (Curran and Hay, 1986). To overcome this limitation, parametric and non-parametric techniques, including modified least squares, Thiel-Sen, and Simulation-Extrapolation (SIMEX), have been used to quantify and reduce measurement error effects (Fernandes and G. Leblanc, 2005; Cook and Stefanski, 1994)

SIMEX was chosen as the methodology to solve this errors-in-variables issue due to its heuristic based approach for reducing biased parameters (Ponzi et al., 2019). SIMEX permits flexibility in modeling because it does not require specifying an error structure model or a ranking of groups (Ponzi et al., 2019; Fernandes and G. Leblanc, 2005). SIMEX, a simulation-based, error invariant method of manipulating the estimated model parameters to reduce the effects of measurement errors, can reduce bias in models with measurement error-prone variables (Cook and Stefanski, 1994; Misumi et al., 2017; He et al., 2007). One method for obtaining initial estimates of the measurement variance is through the use of the Kalman filter. The Kalman filter, a recursive algorithm for data processing, provides mean and variance estimates of a known process under the influence of process and measurement errors through an iterative modeling process (Kalman, 1960; Maybeck, 1979; Montes, 2012).

The objective of this research was to develop an unbiased satellite-to-LAI calibration model using Landsat 5 and 7 images, and to test different spectral indices in their ability to predict Li-Cor 2000 Plant Canopy Analyzer LAI values. The novelty of this research was using error-in-variable methods to correct for the quantified measurement error associated with an observation on the model parameters.

2.2 Methods

2.2.1 Study Site

For this study, we used leaf area index measurements taken at the Southeast Tree Research and Education Site (SETRES) study. The study was established in 1992 on a loblolly pine plantation in Scotland County, North Carolina with an initial density of 1,667 trees per

hectare (2 m x 3 m spacing). A total of sixteen 50 m x 50 m treatment plots were installed to cover four replications of a 2 x 2 factorial of irrigation and fertilization (Albaugh et al., 1998). Individual plot LAI measurements were made using a Li-Cor 2000 Plant Canopy Analyzer from March 1992 to September 2004 for a total of 1673 samples (Albaugh et al., 1998; Li-Cor, 1991). The age range covered by the in situ LAI measurements was between 7 and 20 years old. Treatments applied to the study site included complete continuous mechanical and chemical competition control. The fertilization treatment applied was determined by foliar analysis to maintain predefined foliar concentrations of nitrogen, phosphorus, potassium, calcium, magnesium, and boron as needed (Albaugh et al., 1998). The irrigation treatment consisted of additional water added to the plot in excess of naturally occurring precipitation via plot level irrigation system to maintain targeted soil water content (Albaugh et al., 1998). Further details of the fertilization and irrigation treatments can be seen in Albaugh et al. (1998). There were four replicates of each treatment of the fertilization by irrigation factorial design.

2.2.2 Satellite Data

Landsat 5 Thematic Mapper and Landsat 7 Enhanced Thematic Mapper Plus (ETM+) imagery were selected for this study due to the temporal overlap between in situ measurements and image dates. Landsat Level-1 Tier 1 products were used to ensure the highest radiometric and geometric calibration across sensors for time-series analysis. The Landsat Level-1 Tier 1 products were atmospherically corrected to provide surface reflectance data, correcting for atmospheric distortions and allowing for the comparison with the in situ measurements (Masek et al., 2006). U.S. Geological Survey (USGS) Landsats 5 and 7 Surface Reflectance Level 1 Tier 1 data were queried in the Google Earth Engine from January 1, 1991 to December 31, 2005 (Gorelick et al., 2017). Nadir BRDF-adjusted reflectance values were calculated using the c-factor approach to provide temporal and geographic consistency across the thirteen years of observation (Roy et al., 2016). Data were further filtered to

remove cloud and shadow anomalies based on quality assurance bands and excessive haziness based on the atmospheric opacity band. Using individual polygons for the 16 study plots, the weighted average of the within plot pixel values was calculated and returned. Above canopy satellite observations were paired to the in situ LAI estimates by matching observation dates. To account for the below canopy in situ LAI measurements that did not have matching satellite imagery sample dates, the four most previous and the four following satellite observations were subset and fitted with a thin plate spline. The thin plate spline was weighted by the associated standard deviation of the plot level reflectance summarization. The interpolated VI estimates and associated interpolation standard errors were extracted from the thin plate spline model for the in situ observation date. Using the interpolated and matched values, multiple vegetation indices were calculated using the individual band values for the polygons (Table 2.1) and fit to the paired in situ LAI estimate in a simple linear model form.

2.2.3 LAI Estimation Model Building

The model used the annual peak leaf area index (pLAI) values for each plot (Figure 2.1). The pLAI values, which typically occurred in August or September of any given year in the study, were identified and subset at the individual plot and year level to provide a series of pLAI values for each of the 16 plots, for a total of 208 pLAI observations. The pLAI data were randomly split into individual training (n=156) and validation (n=52) sets. Using the selected VI and observed pLAI, several linear, segmented, and nonlinear models were fit to the training data using the maximum likelihood framework (Table 2.2). Two segmented regression models were fit to the data set (Models 2 and 3, Table 2.3). These segmented models included a linear and non-linear component, joined at a common point (θ) such that the line is continuous and smooth. Additional variables were included in the model form to test performance. The variance structure was assumed to be non-constant for all but two models, Models 0 and 2 (Table 2.2). Model selection was based on information criterion

(Akaike's and Bayesian) and minimizing root mean square error (RMSE). During the model testing and diagnostics procedures, one pLAI observation was found to be highly influential and was removed from the training pLAI data set (n=155).

2.2.4 Measurement Error Estimation

To estimate the amount of measurement error associated with the selected VI, the Kalman filter was used to step through the individual plot level observations to determine the amount of error associated with the measurements and selected model. The selected VI values were fit to a three-parameter logistic equation (Equation 2.1) to model the overall trend of the asymptotic VI development observed at the plot level (Montes, 2012):

$$V(t) \ = \frac{\alpha}{1 \ + \ e^{(\frac{\beta - t}{\gamma})}} \tag{2.1}$$

Where α is the asymptotic maximum value for VI, β is the location parameter for rate increase, and γ is the shape parameter of the relation, t is time, and V is the VI. The Kalman filter utilized the algebraic differenced form of the equation (Equation 2.2) to provide one step ahead estimates of VI for data with different starting points:

$$V(t+i) \ = \ V(t) \ * \ \frac{\frac{\alpha}{1 \ + \ e^{(\frac{\beta - t}{\gamma})}}}{\frac{\alpha}{1 \ + \ e^{(\frac{\beta - (t+i)}{\gamma})}}} \tag{2.2}$$

The generalized form of model was used to incorporate process error into the formula:

$$V(t_n) \ \sim \mathcal{N} \left(V(t) \ * \ \frac{\frac{\alpha}{1 \ + \ e^{(\frac{\beta - t_{n-1}}{\gamma})}}}{\frac{\alpha}{1 \ + \ e^{(\frac{\beta - t_n}{\gamma})}}}, \ \sigma^2_{proc} \right) \tag{2.3}$$

Using the generalized formulation (Equation 2.3), the VI for time n was projected using initial starting parameters, followed by an updated variance estimate. A new VI measurement was observed and input into the system. The observed measurement and projected estimate were assimilated, weighted by the inverse of their variances. The assimilated value was the new estimate of VI and the process was repeated by projecting it forward in time along

with a variance update until the next VI observed measurement time. The recursive process was repeated until the system state was optimal given a set of constraints, which in this case minimized the negative log-likelihood. Output from the Kalman filter included updated estimates of the equation parameters in addition to estimates of the process and observation error.

2.2.5 Bias Correction

To correct for the effects of measurement error on the parameter estimates from VI observations, an error invariable approach was used to account for the fact that observations in the analysis come with errors. The Kalman filter estimate of observation error associated with the selected VI time series for all 16 plots was used in conjunction with the interpolation errors in the SIMEX algorithm to correct for bias in the initial model parameters. SIMEX works with two main steps, a simulation step and an extrapolation step. Using the initial estimate of measurement variance and fitted parameters, SIMEX simulates increasing amounts of variance in the data and re-parameterizes the model. After several iterations of re-parameterizing the model with inflated variance, SIMEX extrapolates across the estimated (and biased) parameters to find the parameters with no variance (Cook and Stefanski, 1994; Misumi et al., 2017). To implement the SIMEX algorithm, the selected model and framework were refit with modified observations simulated to represent the inflated observation error estimate (Lederer and Seibold, 2013). To create these modified observations, synthetic noise ($U^*_{(\lambda_l)}$) was generated following $U^*_{(\lambda_l)} \sim IID\ N(0, \lambda_l \sigma^2)$ and was added to the observation (Lederer and Seibold, 2013). The values for λ_l ranged from 0 to 2 and were used to weight the amount of generated noise being added to the observation. A total of twenty steps between 0 and 2 were simulated 1000 times each for λ_l. After simulating the observation error inflation, the refit coefficients were averaged for each λ_l and plotted against the observation error values. A quadratic linear model was fit to correlate the new coefficient values to the λ_l.

The unbiased estimates were then extrapolated from the trend to the value of -1, as if the observations were measured completely without error.

To evaluate the robustness of the methodology and the predictive power of the model, the SIMEX parameter estimates were used to calculate mean square error (MSE) on the training data set and mean square prediction error (MSPR) on the validation data set. If MSE and MSPR comparisons are reasonable, the methodology is repeated with a combination of the two data sets.

To compare the proposed model to the current industry standard for LAI estimation from Landsat 5/7 imagery, the same procedures for handling in situ LAI measurements without exact satellite imagery matches were followed, Landsat 5 and 7 Top of Atmosphere (TOA) data were queried, matched, and interpolated to find corresponding SR values. Using the Flores et al. (2006) model, LAI was estimated and compared to observed LAI.

2.3 Results

2.3.1 Relationship between leaf area index and vegetation indices

A total of 1673 measured observations were used to test the relationship between in situ LAI and the individual vegetation indices. Of the 1673 observations, 1546 did not have exact day matches between in situ LAI measurements and satellite imagery (92%). The VI with the highest correlation with LAI was the normalized difference moisture index (NDMI), with a value of 0.765, indicating a strong positive relationship, followed by simple ratio 2 (SR2), -0.756, and the simple ratio (SR), 0.735 (Figure 2.2). When fitted to the 1673 observations, the relationship between NDMI and LAI had an adjusted R^2 of 0.6172 (Figure 2.3). For this study, we shifted NDMI by one to be in the positive domain. September (n=125) and August (n=54) had the highest frequency of pLAI values (Figure 2.1).

Little improvement was observed in model performance as order increased and form changed as compared to the simple linear model until the treatment variable was included (Table 2.3). The model that best predicted pLAI based on lowest AIC, BIC, and RMSE using

was Model 9 (Table 2.3, Figure 2.4). In addition to including NDMI, Model 9 included an indicator variable for FERT, referring to the optimum nutrition treatment that was applied to the fertilization and combination plots. With FERT in the model, a slight curvilinear relationship between NDMI and LAI was observed in the unfertilized plots (Figure 2.4). The fertilized plots maintained an overall linear relationship between the two variables, with higher expected LAI for a given NDMI as compared to the unfertilized plots.

The addition of FERT into the model improved performance as seen by reductions in AIC and RMSE (Table 2.3). Using Model 6 as an example, the addition of FERT (Model 7) reduced RMSE by 0.0877 units of LAI, decreasing the AIC from 221.74 to 173.42. A further reduction in AIC and RMSE occurred from the change in how FERT was incorporated into the model (Table 2.3). Model 9 used a nonlinear function of the VI to model the effect of FERT rather than as a constant, reducing the overall RMSE to 0.3947 and AIC to 163.15 (Table 2.3).

Estimated observation error for the individual treatments were combined following additive error propagation protocol to a final estimate of 0.0019 units NDMI (σ^2_{obs}). The SIMEX algorithm modified and unbiased the intercept and slope coefficient values to reduce the effects of measurement error in the NDMI model (Figure 2.5). The effects of increasing the amount of measurement error, or noise, on the individual parameters influenced how the parameters behaved, with well-defined trends in not only the mean value from the simulations, but also in the distribution of the parameter values in the simulations. For example, for β_0, as λ increases from 0 to 2, the mean estimates for β_0 increase from approximately -4.12 to -2.91 while standard deviations increased from approximately 0.04 to 0.10 (Figure 2.5).

The calculated training dataset MSPR was 0.1431, and smaller than the calculated MSE on the training data set (0.1651), indicating that the model provided reasonable predictions. The combined model including the training and validation data (n=207) provided the final model with unbiased parameters:

$$LAI \ = \frac{-4.520 + 4.930 * \text{NDMI} + \left(0.271 * \text{NDMI}^{1.816}\right) * \text{FERT}}{1 \ - 0.149 * \text{NDMI}} \tag{2.4}$$

With a model for the associated errors:

$$SD(LAI) \ = \ 0.156 * \text{NDMI}^{0.165} \tag{2.5}$$

Residual analysis for the model showed no heteroscedasticity and no obvious bias comparing predicted and observed LAI across all observed levels of LAI (Figures 2.6 and 2.7). The final MSE for the unbiased model was 0.1748 units of LAI (m^2/m^2) and RMSE was 0.4181 units of LAI (m^2/m^2).

Comparisons to the Flores et al. model showed an overall underprediction of LAI as compared to the observed, with a calculated RMSE of 1.47 (Flores et al., 2006) (Figure 2.7.

2.4 Discussion

NDMI was a better predictor of LAI in loblolly pine plantations in the southeastern United States when compared to more commonly used vegetation indices, such as NDVI and SR, supporting the findings by Blinn et al. (2019). Whereas literature has supported the use of NDMI to predict vegetation density and canopy moisture, Horler and Ahern (1986) hypothesized that the shortwave infrared band reflectance is more sensitive to shadows in the image, warranting further investigation of the various influences of canopy geometry on NDMI (Hardisky et al., 1983). NDMI, calculated by the difference ratio between near infrared (0.76-0.90 μm) and shortwave infrared (1.55-1.75 μm) bands, was a good predictor of live leaf biomass, total aboveground biomass, leaf moisture, and canopy moisture in other species (Tucker, 1980; Hardisky et al., 1983; Gao, 1996; Ahamed et al., 2011). Compared with the near infrared band's propensity to be reflected by the canopy, the shortwave infrared band wavelengths are absorbed by water, providing a measure of water (or moisture) content on the surface of interest (Wilson and Sader, 2002; Cibula et al., 1992). The shortwave infrared

band is sensitive to vegetation density in forests and to total biomass (biomass and necromass), which is what plant canopy analyzers, such as the Li-Cor 2000, are influenced by in their measurements of intercepted light (Horler and Ahern, 1986). It is important to make the distinction that while LAI was the biophysical variable under analysis in this research, limitations of the plant canopy analyzer include not separating out leaf area from additional plant matter, so plant area index may be the more appropriate designation. Optimum nutrition management influenced the behavior of the vegetation indices. Changes in the relationship between the VI and LAI when under optimum nutrition management may be a result of the increased photosynthetic capacity of the canopy, changes in leaf morphology, or allometric changes within stand (Maier et al., 2008; Gough et al., 2004; King et al., 1999). Physical or physiological changes, such as transpiration, chlorophyll concentrations, foliar elemental concentrations, radiation use efficiency, or leaf size and orientation may be the driving force behind the fertilization effect (Al-Abbas et al., 1974; Knipling, 1970; Martin and Jokela, 2004; Wightman et al., 2016; Gough et al., 2004). In other crops, a fertilization effect on plant spectral reflectance properties has been observed in the near infrared and middle infrared wavelengths, with increases in reflectance for the middle infrared and decreases for the near infrared when nitrogen fertilization had been applied (Walburg et al., 1982; Ahlrichs and Bauer, 1983; Hinzman et al., 1986). Using the study level average reflectance for plots receiving (1) and not receiving (0) fertilization treatments, our study showed similar results, with decreased reflectance of the middle infrared band (Band 5 in Landsat 5 and 7), and increased reflectance in in the near infrared (Band 4) (Figure 2.10).

Using error invariable methods to correct for observations with error is a necessary step to provide unbiased estimators for modeling remotely sensed data. The SIMEX algorithm provided a flexible method for estimating the effects of increasing amounts of error on the individual parameters for the selected model, allowing for the extrapolation of the trend to find the value of the unbiased parameter. Including the estimated error for the interpolated values due to temporal misregistration and the overall process error were two ways to account

for the several sources of error that may affect remotely sensed data (Curran and Hay, 1986). Overall, estimated errors from the Kalman filter and thin plate spline interpolations were relatively low, resulting in a slight change in model fit (Figure 2.8). Accounting for the errors in remotely sensed data provides the means to avoid violating the basic assumptions of linear modeling and reduces the influence of error prone variables.

While the data in this study come from a relatively small geographical area, the high temporal sampling provided a great number of LAI observations for the 16 plots. It is important to note that the treatments applied in this study are not considered operational for loblolly pine plantations in the southeastern U.S. Consequently, our results may not well represent the relationship between remotely sensed VIs and observed LAI measurements in operational settings. The long timeline of this study may also introduce errors due to orbit drift in Landsat 5 for which we did not explicitly manage (Zhang and Roy, 2016; Roy et al., 2020). The study plots were relatively small compared to the pixel size used by Landsat imagery. The 50 m by 50 m measurement plot size did not allow for adequate sampling of the 30 m Landsat pixels to reduce spatial autocorrelation between pixels, for which a 3 by 3 pixel sampling area is recommended (Curran and Hay, 1986). New sensors with higher resolution imagery will provide better sampling of the small study plots typically used for forestry research and should avoid this issue. While the weighted average of the pixels falling within the study plot were used to calculate the VIs, a further potential source of errors may come from interference in reflectance from areas surrounding the study plot under evaluation. Results from Cohrs et al. (2020) corroborate the general underprediction of the LAI model in Flores et al. (2006), thus the model presented in this research (Equation 2.4) should provide an update to sites relying on legacy data from Landsat 5 and 7 ETM+.

2.5 Conclusions

Overall we identified NDMI as the best VI for predicting LAI in loblolly pine plantations in the southeastern United States using Landsat 5 and 7 surface reflectance data. A model form

including an indicator variable for optimum nutrition management was the best framework for modeling the relationship between NDMI and LAI. SIMEX probed to be an effective methodology for reducing the effects of measurement error associated with the VI on the final parameter estimates.

2.6 Acknowledgements

We gratefully acknowledge support provided by the Plantation Management Research Cooperative at the University of Georgia, the Forest Productivity Cooperative and the Department of Forest Resources and Environmental Conservation at Virginia Polytechnic Institute and State University. Funding for this work was provided by the Warnell School of Forestry and Natural Resources, by the Virginia Agricultural Experiment Station and the McIntire-Stennis Program of the National Institute of Food and Agriculture, U.S. Department of Agriculture. We are also grateful for the resources and technical expertise from the Georgia Advanced Computing Resource Center, a partnership between the University of Georgia's Office of the Vice President for Research and Office of the Vice President for Information Technology.

2.7 References

Ahamed, T., Tian, L., Zhang, Y., and Ting, K. (2011). A review of remote sensing methods for biomass feedstock production. *Biomass and Bioenergy*, 35(7):2455–2469.

Ahlrichs, J. S. and Bauer, M. E. (1983). Relation of agronomic and multispectral reflectance characteristics of spring wheat canopies. *Agronomy Journal*, 75(6):987.

Albaugh, T. J., Allen, H. L., Dougherty, P. M., Kress, L. W., and King, J. S. (1998). Leaf area and above- and belowground growth responses of loblolly pine to nutrient and water additions. *Forest Science*, 44(2):317–328.

Albaugh, T. J., Lee Allen, H., Dougherty, P. M., and Johnsen, K. H. (2004). Long term growth responses of loblolly pine to optimal nutrient and water resource availability. *Forest Ecology and Management*, 192(1):3–19.

Al-Abbas, A. H., Barr, R., Hall, J. D., Crane, F. L., and Baumgardner, M. F. (1974). Spectra of normal and nutrient-deficient maize leaves. *Agronomy Journal*, 66(1):16–20.

Blinn, C., House, M., Wynne, R., Thomas, V., Fox, T., and Sumnall, M. (2019). Landsat 8 based leaf area index estimation in loblolly pine plantations. *Forests*, 10(3).

Chen, J. M. and Cihlar, J. (1996). Retrieving leaf area index of boreal conifer forests using landsat TM images. *Remote Sensing of Environment*, 55(2):153–162.

Cibula, W. G., Zetka, E. F., and Rickman, D. L. (1992). Response of thematic mapper bands to plant water stress. *International Journal of Remote Sensing*, 13(10):1869–1880.

Claverie, M., Vermote, E. F., Franch, B., and Masek, J. G. (2015). Evaluation of the Landsat-5 TM and Landsat-7 ETM+ surface reflectance products. *Remote Sensing of Environment*, 169(November):390–403.

Cohrs, C. W., Cook, R. L., Gray, J. M., and Albaugh, T. J. (2020). Sentinel-2 leaf area index estimation for pine plantations in the southeastern United States. *Remote Sensing*, 12(9):1406.

Coleman, T., Gudapati, L., and Derrington, J. (1990). Monitoring forest plantations using Landsat Thematic Mapper data. *Remote Sensing of Environment*, 33(3):211–221.

Cook, J. R. and Stefanski, L. A. (1994). Simulation-extrapolation estimation in parametric measurement error models. *Journal of the American Statistical Association*, 89(428):1314–1328.

Curran, P. (1980). Multispectral remote sensing of vegetation amount. *Progress in Physical Geography: Earth and Environment*, 4(3):315–341.

Curran, P. and Hay, A. (1986). The importance of measurement error for certain procedures in remote sensing at optical wavelengths. *Photogrammetric Engineering and Remote Sensing*, 52(2):229–241.

Dalla-Tea, F. and Jokela, E. J. (1991). Needlefall, canopy light interception, and productivity of young intensively managed slash and loblolly pine stands. *Forest Science*, 37(5):1298–1313.

Fernandes, R. and G. Leblanc, S. (2005). Parametric (modified least squares) and non-parametric (Theil–Sen) linear regressions for predicting biophysical parameters in the presence of measurement errors. *Remote Sensing of Environment*, 95(3):303–316.

Flores, F. J. (2003). *Using Remote Sensing Data To Estimate Leaf Area Index and Foliar Nitrogen of Loblolly Pine Plantations*. PhD thesis, North Carolina State University.

Flores, F. J., Allen, H. L., Cheshire, H. M., Davis, J. M., Fuentes, M., and Kelting, D. (2006). Using multispectral satellite imagery to estimate leaf area and response to silvicultural treatments in loblolly pine stands. *Canadian Journal of Forest Research*, 36(6):1587–1596.

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

Gao, B.-c. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3):257–266.

Gitelson, A. A., Kaufman, Y. J., and Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, 58(3):289–298.

Gitelson, A. A., Kaufman, Y. J., Stark, R., and Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment*, 80(1):76–87.

Gitelson, A. A., Viña, A., Arkebauer, T. J., Rundquist, D. C., Keydan, G., and Leavitt, B. (2003). Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophysical Research Letters*, 30(5):n/a–n/a.

Gonzalez-Benecke, C. A., Jokela, E. J., and Martin, T. A. (2012). Modeling the effects of stand development, site quality, and silviculture on leaf area index, litterfall, and forest floor accumulations in loblolly and slash pine plantations. *Forest Science*, 58(5):457–471.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202:18–27.

Gough, C. M., Seiler, J. R., and Maier, C. A. (2004). Short-term effects of fertilization on loblolly pine (Pinus taeda L.) physiology. *Plant, Cell and Environment*, 27(7):876–886.

Gower, S. T., Kucharik, C. J., and Norman, J. M. (1999). Direct and indirect estimation of leaf area index, fapar, and net primary production of terrestrial ecosystems. *Remote Sensing of Environment*, 70(1):29–51.

Haboudane, D. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 90(3):337–352.

Hardisky, M. A., Klemas, V., and Smart, R. M. (1983). The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of spartina alterniflora canopies. *Photogrammetric Engineering and Remote Sensing*, 49(1):77–83.

He, W., Yi, G. Y., and Xiong, J. (2007). Accelerated failure time models with covariates subject to measurement error. *Statistics in Medicine*, 26(26):4817–4832.

Hinzman, L., Bauer, M., and Daughtry, C. (1986). Effects of nitrogen fertilization on growth and reflectance characteristics of winter wheat. *Remote Sensing of Environment*, 19(1):47–61.

Horler, D. N. H. and Ahern, F. J. (1986). Forestry information content of thematic mapper data. *International Journal of Remote Sensing*, 7(3):405–428.

Huete, A. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3):295–309.

Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., and Ferreira, L. (2002). Overview of the radiometric and biophysical performance of the modis vegetation indices. *Remote Sensing of Environment*, 83(1):195 – 213. The Moderate Resolution Imaging Spectroradiometer (MODIS): a new generation of Land Surface Monitoring.

Huete, A., HuiQing Liu, and van Leeuwen, W. (1997). The use of vegetation indices in forested regions: issues of linearity and saturation. In *IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing - A Scientific Vision for Sustainable Development*, volume 4, pages 1966–1968. IEEE.

Huete, A. and Justice, C. (1999). MODIS Vegetation Index (MOD 13) Algorithm Theoretical Basis Document. Technical report.

Iiames, J., Congalton, R., Pilant, A., and Lewis, T. (2008). Leaf area index (lai) change detection analysis on loblolly pine (*Pinus taeda*) following complete understory removal. *Photogrammetric Engineering & Remote Sensing*, 74(11):1389–1400.

Iverson, L. R., Graham, R. L., and Cook, E. A. (1989). Applications of satellite remote sensing to forested ecosystems. *Landscape Ecology*, 3(2):131–143.

Jokela, E. J., Dickens, E., Barnett, J., and Hubbard, W. (2004). Nutrient management of southern pines. In *Slash pine: still growing and growing! Proceedings of the slash pine symposium*, pages 27–35.

Jokela, E. J. and Martin, T. A. (2000). Effects of ontogeny and soil nutrient supply on production, allocation, and leaf area efficiency in loblolly and slash pine stands. *Canadian Journal of Forest Research*, 30(10):1511–1524.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45.

Key, C. H. and Benson, N. C. (2006). Landscape assessment (la). In Lutes, D. C., Keane, R. E., Caratti, J. F., Key, C. H., Benson, N. C., Sutherland, S., and Gangi, L. J., editors, *FIREMON: Fire effects monitoring and inventory system.*, volume Gen. Tech. Rep. RMRS-GTR-164-CD., pages 1–55. US Department of Agriculture, Forest Service, Rocky Mountain Research Station., Fort Collins, CO.

King, J. S., Albaugh, T. J., Allen, H. L., and Kress, L. W. (1999). Stand-level allometry in Pinus taeda as affected by irrigation and fertilization. *Tree Physiology*, 19(12):769–778.

Knipling, E. B. (1970). Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sensing of Environment*, 1(3):155–159.

Lederer, W. and Seibold, H. (2013). *simex: SIMEX- And MCSIMEX-Algorithm for Measurement Error Models*. R package version 1.5.

Li-Cor, I. (1991). Lai-2000 plant canopy analyzer, instruction manual.

Lunetta, R. S., Congalton, R. G., Fenstermaker, L. K., Jensen, J. R., McGwire, K. C., and Tinney, L. R. (1991). Remote sensing and geographic information system data integration: Error sources and research issues. *Photogrammetric Engineering & Remote Sensing*, 57(6):677–687.

Mahowald, N., Lo, F., Zheng, Y., Harrison, L., Funk, C., Lombardozzi, D., and Goodale, C. (2016). Projections of leaf area index in earth system models. *Earth System Dynamics*, 7(1):211–229.

Maier, C. A., Palmroth, S., and Ward, E. (2008). Short-term effects of fertilization on photosynthesis and leaf morphology of field-grown loblolly pine following long-term exposure to elevated CO2 concentration. *Tree Physiology*, 28(4):597–606.

Martin, T. A. and Jokela, E. J. (2004). Developmental patterns and nutrition impact radiation use effieciency components in southern pine stands. *Ecological Applications*, 14(6):1839–1854.

Masek, J., Vermote, E., Saleous, N., Wolfe, R., Hall, F., Huemmrich, K., Gao, F., Kutler, J., and Lim, T.-K. (2006). A Landsat surface reflectance dataset for North America, 1990—2000. *IEEE Geoscience and Remote Sensing Letters*, 3(1):68–72.

Maybeck, P. S. (1979). *Stochastic Models, Estimation and Control: Volume 1*, volume I. Academic Press, Inc, New York, New York.

Misumi, M., Furukawa, K., Cologne, J. B., and Cullings, H. M. (2017). Simulation–extrapolation for bias correction with exposure uncertainty in radiation risk analysis utilizing grouped data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 67(1):275–289.

Monteith, J. L. (1972). Solar radiation and productivity in tropical ecosystems. *The Journal of Applied Ecology*, 9(3):747.

Montes, C. R. (2012). *A Resource Driven Growth and Yield Model for Loblolly Pine Plantations*. PhD thesis, North Carolina State University.

Mutanga, O. and Skidmore, A. K. (2004). Narrow band vegetation indices overcome the saturation problem in biomass estimation. *International Journal of Remote Sensing*, 25(19):3999–4014.

Oswalt, S. N., Smith, W. B., Miles, P. D., and Pugh, S. A. (2014). Forest Resources of the United States, 2012: a technical document supporting the Forest Service 2010 update of the RPA Assessment. Technical Report October.

Parresol, B., Scott, D., Zarnoch, S., Edwards, L., and Blake, J. (2017). Modeling forest site productivity using mapped geospatial attributes within a South Carolina Landscape, USA. *Forest Ecology and Management*, 406(June):196–207.

Peduzzi, A., Wynne, R. H., Fox, T. R., Nelson, R. F., and Thomas, V. A. (2012). Estimating leaf area index in intensively managed pine plantations using airborne laser scanner data. *Forest Ecology and Management*, 270:54–65.

Ponzi, E., Keller, L. F., Muff, S., and Hansen, T. (2019). The simulation extrapolation technique meets ecology and evolution: A general and intuitive method to account for measurement error. *Methods in Ecology and Evolution*, 10(10):1734–1748.

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., and Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, 48(2):119–126.

Qi, J., Kerr, Y., Moran, M., Weltz, M., Huete, A., Sorooshian, S., and Bryant, R. (2000). Leaf area index estimates using remotely sensed data and brdf models in a semiarid region. *Remote Sensing of Environment*, 73(1):18–30.

Qu, Y. and Zhuang, Q. (2018). Modeling leaf area index in North America using a process-based terrestrial ecosystem model. *Ecosphere*, 9(1):e02046.

Rouse, J., Hass, R., Deering, D., Shell, J., and Harlan, J. (1974). Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. Report, NASA.

Roy, D., Zhang, H., Ju, J., Gomez-Dans, J., Lewis, P., Schaaf, C., Sun, Q., Li, J., Huang, H., and Kovalskyy, V. (2016). A general method to normalize landsat reflectance data to nadir brdf adjusted reflectance. *Remote Sensing of Environment*, 176:255 – 271.

Roy, D. P., Li, Z., Zhang, H. K., and Huang, H. (2020). A conterminous United States analysis of the impact of Landsat 5 orbit drift on the temporal consistency of Landsat 5 Thematic Mapper data. *Remote Sensing of Environment*, 240(February):111701.

Sampson, D. A., Amatya, D. M., Lawson, C. D. B., and Skaggs, R. W. (2011). Leaf area index (lai) of loblolly pine and emergent vegetation following a harvest. *Advances in Forest Hydrology*, 54(6):2057–2066.

Shi, K. and Cao, Q. V. (1997). Predicted leaf area growth and foliage efficiency of loblolly pine plantations. *Forest Ecology and Management*, 95(2):109–115.

Shibayama, M., Salli, A., Häme, T., Iso-Iivari, L., Heino, S., Alanen, M., Morinaga, S., Inoue, Y., and Akiyama, T. (1999). Detecting phenophases of subarctic shrub canopies by using automated reflectance measurements. *Remote Sensing of Environment*, 67(2):160–180.

Shoemaker, D. and Cropper, W. (2008). Prediction of leaf area index for southern pine plantations from satellite imagery using regression and artificial neural networks. *Proceedings of the 6th Southern Forestry and Natural Resources GIS Conference*, (2008):139–160.

Sivanpillai, R., Smith, C. T., Srinivasan, R., Messina, M. G., and Wu, X. B. (2006). Estimation of managed loblolly pine stand age and density with Landsat ETM+ data. *Forest Ecology and Management*, 223(1-3):247–254.

Stein, B. R., Thomas, V. A., Lorentz, L. J., and Strahm, B. D. (2014). Predicting macronutrient concentrations from loblolly pine leaf reflectance across local and regional scales. *GIScience & Remote Sensing*, 51(3):269–287.

Teskey, R. O., Bongarten, B. C., Cregg, B. M., Dougherty, P. M., and Hennessey, T. C. (1987). Physiology and genetics of tree growth response to moisture and temperature stress: an examination of the characteristics of loblolly pine (Pinus taeda L.). *Tree Physiology*, 3(1):41–61.

Tucker, C. J. (1980). Remote sensing of leaf water content in the near infrared. *Remote Sensing of Environment*, 10(1):23–32.

Vermote, E., Tanre, D., Deuze, J., Herman, M., and Morcette, J.-J. (1997). Second simulation of the satellite signal in the solar spectrum, 6s: an overview. *IEEE Transactions on Geoscience and Remote Sensing*, 35(3):675–686.

Vose, J. M. and Allen, H. L. (1988). Leaf area, stemwood growth, and nutrition relationships in loblolly pine. *Forest Science*, 34(3):547–563.

Vose, J. M. and Swank, W. T. (1990). A conceptual model of forest growth emphasizing stand leaf area. In Dixon, R. K., Meldahl, R. S., Ruark, G. A., and Warren, W. G., editors, *Process Modeling of Forest Growth Responses to Environmental Stress*, chapter 24, pages 278–287. Timber Press, Inc., Portland, OR.

Walburg, G., Bauer, M. E., Daughtry, C. S. T., and Housley, T. L. (1982). Effects of nitrogen nutrition on the growth, yield, and reflectance characteristics of corn canopies. *Agronomy Journal*, 74(4):677.

Wang, F.-m., Huang, J.-f., Tang, Y.-l., and Wang, X.-z. (2007). New vegetation index and its application in estimating leaf area index of rice. *Rice Science*, 14(3):195–203.

Wang, Q., Adiku, S., Tenhunen, J., and Granier, A. (2005). On the relationship of NDVI with leaf area index in a deciduous forest site. *Remote Sensing of Environment*, 94(2):244–255.

Waring, R. (1983). Estimating forest growth and efficiency in relation to canopy leaf area. In MacFadyen, A. and Ford, E., editors, *Advances in Ecological Research*, chapter 13, pages 327 – 354. Academic Press.

Wenze Yang, Bin Tan, Dong Huang, Rautiainen, M., Shabanov, N. V., Wang, Y., Privette, J. L., Huemmrich, K. F., Fensholt, R., Sandholt, I., Weiss, M., Ahl, D. E., Gower, S. T., Nemani, R. R., Knyazikhin, Y., and Myneni, R. B. (2006). Modis leaf area index products: from validation to algorithm improvement. *IEEE Transactions on Geoscience and Remote Sensing*, 44(7):1885–1898.

Wightman, M., Martin, T., Gonzalez-Benecke, C., Jokela, E., Cropper, W., and Ward, E. (2016). Loblolly pine productivity and water relations in response to throughfall reduction and fertilizer application on a poorly drained site in northern Florida. *Forests*, 7(12):214.

Wilson, E. H. and Sader, S. A. (2002). Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sensing of Environment*, 80(3):385–396.

Yang, Y., Anderson, M. C., Gao, F., Hain, C. R., Semmens, K. A., Kustas, W. P., Noormets, A., Wynne, R. H., Thomas, V. A., and Sun, G. (2017). Daily Landsat-scale evapotranspiration estimation over a forested landscape in North Carolina, USA, using multi-satellite data fusion. *Hydrology and Earth System Sciences*, 21(2):1017–1037.

Zhang, H. K. and Roy, D. P. (2016). Landsat 5 Thematic Mapper reflectance and NDVI 27-year time series inconsistencies due to satellite orbit change. *Remote Sensing of Environment*, 186:217–233.

2.8 Tables and Figures

Table 2.1: Vegetation indices evaluated in the analysis. Near-infrared (NIR), band 4 in Landsat 5 and 7 ETM+; Shortwave-infrared (SWIR), band 5 in Landsat 5 and 7 ETM+; Shortwave-infrared 2 (SWIR2), band 7 in Landsat 5 and 7 ETM+

Vegetation Index	**Equation**	**Source**
Normalized Difference Moisture Index (NDVI)	$\frac{\text{NIR}-\text{Red}}{\text{NIR}+\text{Red}}$	(Rouse et al., 1974)
Simple Ratio (SR)	$\frac{\text{NIR}}{\text{Red}}$	(Chen and Cihlar, 1996)
Simple Ratio 2 (SR2)	$\frac{\text{SWIR}}{\text{NIR}}$	(Shibayama et al., 1999)
Enhanced Vegetation Index (EVI)	$2.5 * \left(\frac{\text{NIR}-\text{Red}}{\text{NIR}+6*\text{Red}-7.5*\text{Blue}+1}\right)$	(Huete et al., 2002)
Soil Adjusted Vegetation Index (SAVI)	$1.5 * \frac{\text{NIR}-\text{Red}}{\text{NIR}+\text{Red}+0.5}$	(Huete, 1988)
Modified Soil Adjusted Vegetation Index (MSAVI)	$\frac{2*\text{NIR}+1-\sqrt{(2*\text{NIR}+1)^2-8*(\text{NIR}-\text{Red})}}{2}$	(Qi et al., 1994)
Normalized Difference Moisture Index (NDMI)	$\frac{\text{NIR}-\text{SWIR}}{\text{NIR}+\text{SWIR}}$	(Hardisky et al., 1983)
Normalized Burn Index (NBI)	$\frac{\text{NIR}-\text{SWIR2}}{\text{NIR}+\text{SWIR2}}$	(Key and Benson, 2006)
Chlorophyll Index Green (CIG)	$\frac{\text{NIR}}{\text{Green}} - 1$	(Gitelson et al., 2003)
Modified Chlorophyll Absorption in Reflectance Index (MCARI)	$1.2 * (2.5 * (\text{NIR} - \text{Red}) - 1.3 * (\text{NIR} - \text{Green}))$	(Haboudane, 2004)
Modified Chlorophyll Absorption in Reflectance Index 2 (MCARI2)	$1.5 * \left(\frac{2.5*(\text{NIR}-\text{Red})-1.3*(\text{NIR}-\text{Green})}{\sqrt{(2*\text{NIR}+1)^2-(6*\text{NIR}-5*\sqrt{\text{Red}}-0.5)}}\right)$	(Haboudane, 2004)
Pan Normalized Difference Vegetation Index (PanNDVI)	$\frac{\text{NDVI}-(\text{Blue}+\text{Green}+\text{Red})}{\text{NDVI}+(\text{Blue}+\text{Green}+\text{Red})}$	(Wang et al., 2007)
Green Ratio Vegetation Index (GR)	$\frac{\text{NIR}}{\text{Green}}$	(Gitelson et al., 2002)

Table 2.2: Model forms and error structure evaluated using the normalized difference moisture index (NDMI) to predicted leaf area index (LAI). An indicator variable, FERT, is used to denote if the area under evaluation is under optimum nutrition management (1) or not (0).

Model Number	**Model Form**	**Error Model**
0	$y_i = \beta_0 + \beta_1 * \text{NDMI}$	ρ
1	$y_i = \beta_0 + \beta_1 * \text{NDMI}$	$\rho_1 * \text{NDMI}^{\rho_2}$
2	$y_i = \begin{cases} \beta_0 + \beta_1 X_i \text{ if } X_i \leq \theta \\ \gamma_0 X_i^{\gamma_1} \text{ if } X_i > \theta \end{cases}$	ρ
3	$y_i = \begin{cases} \beta_0 + \beta_1 X_i \text{ if } X_i \leq \theta \\ \gamma_0 X_i^{\gamma_1} \text{ if } X_i > \theta \end{cases}$	$\rho_1 * \text{NDMI}^{\rho_2}$
4	$y_i = \beta_0 + \beta_1 * \text{NDMI}^{\beta_2}$	$\rho_1 * \text{NDMI}^{\rho_2}$
5	$y_i = \beta_0 + \beta_1 * \text{NDMI} + \beta_2 * \text{NDMI}^2$	$\rho_1 * \text{NDMI}^{\rho_2}$
6	$y_i = \frac{\beta_0 + \beta_1 * \text{NDMI}}{1 + \beta_2 * \text{NDMI}}$	$\rho_1 * \text{NDMI}^{\rho_2}$
7	$y_i = \frac{\beta_0 + \beta_1 * \text{NDMI} + \beta_2 * \text{FERT}}{1 + \beta_3 * \text{NDMI}}$	$\rho_1 * \text{NDMI}^{\rho_2}$
8	$y_i = \beta_0 + \beta_1 * \text{NDMI} + \beta_2 * \text{NDMI}^2 + \beta_3 * \text{FERT}$	$\rho_1 * \text{NDMI}^{\rho_2}$
9	$y_i = \frac{\beta_0 + \beta_1 * \text{NDMI} + (\alpha_1 * \text{NDMI}^{\alpha_2}) * \text{FERT}}{1 + \beta_2 * \text{NDMI}}$	$\rho_1 * \text{NDMI}^{\rho_2}$

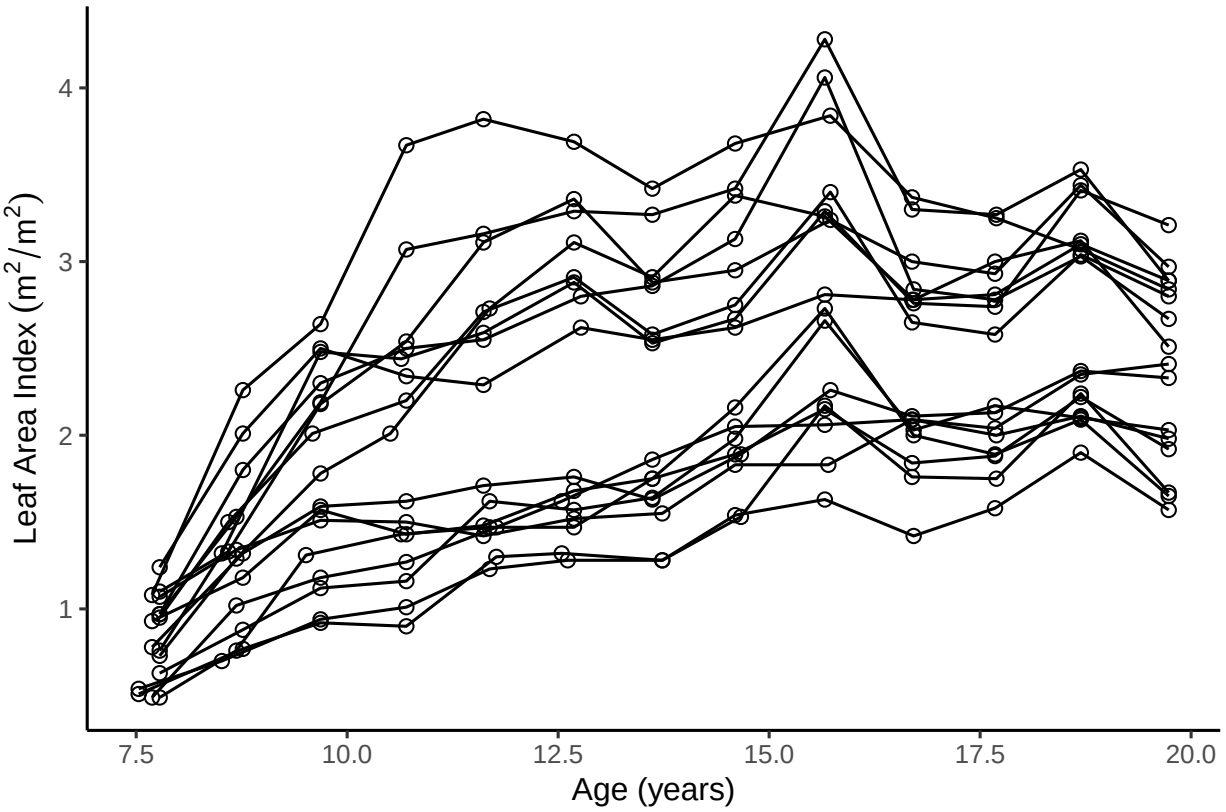


Figure 2.1: Peak leaf area index values for the sixteen plots. Peak leaf area index typically occurred in August or September of any given year.

Table 2.3: Associated fit statistics and evaluation criteria for the models evaluated using the training data (n=155). Negative log-likelihood (-loglik).

Model Number	**AIC**	**BIC**	**RMSE**	**Bias**	**-LogLik**	**Number of Parameters**
0	227.74	236.87	0.4948	2.05E-07	110.87	3
1	225.13	237.30	0.4954	-0.00117	108.57	4
2	233.94	249.16	0.4983	-0.00421	111.97	5
3	227.14	245.40	0.4993	-6.99E-05	107.57	6
4	223.92	239.14	0.4947	2.15E-05	106.96	5
5	223.60	238.81	0.4948	-1.74E-05	106.80	5
6	224.46	239.68	0.4950	-4.29E-05	107.23	5
7	176.74	195.00	0.4118	-0.0004	82.37	6
8	175.83	194.09	0.4136	7.34E-03	81.91	6
9	167.82	189.12	0.3988	8.28E-04	76.91	7

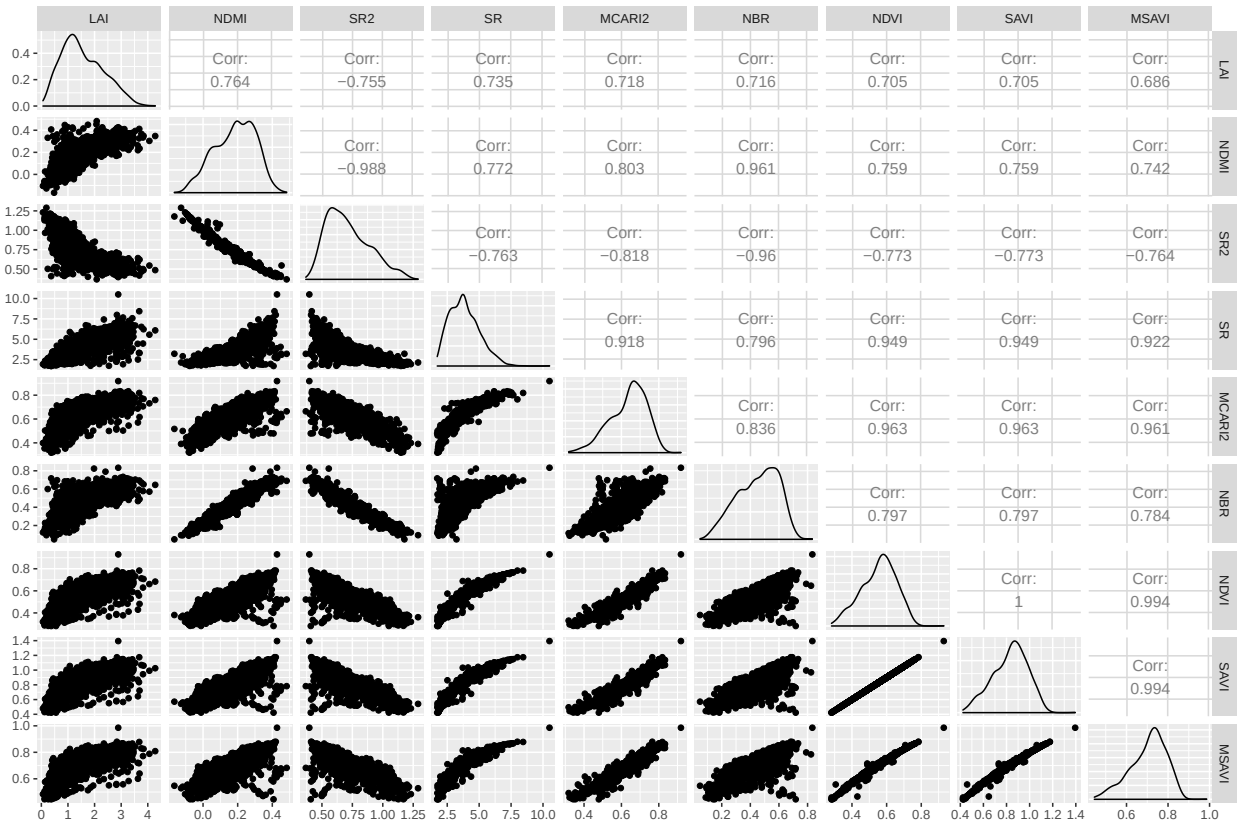


Figure 2.2: Correlation matrix for selected vegetation indices under analysis and their relationship to in situ leaf area index (LAI) for the SETRES data. Normalized difference moisture index (NDMI); simple ratio - 2 (SR2); SR (simple ratio); modified chlorophyll absorption in reflectance index 2 (MCARI2), normalized burn ratio (NBR); normalized difference vegetation index (NDVI); soil adjusted vegetation index (SAVI); modified soil adjusted vegetation index (MSAVI). Axis values are determined by the units for LAI (m^2/m^2) and for the vegetation indices. Where the variable on the x and y axis is equivalent, the distribution of values is returned.

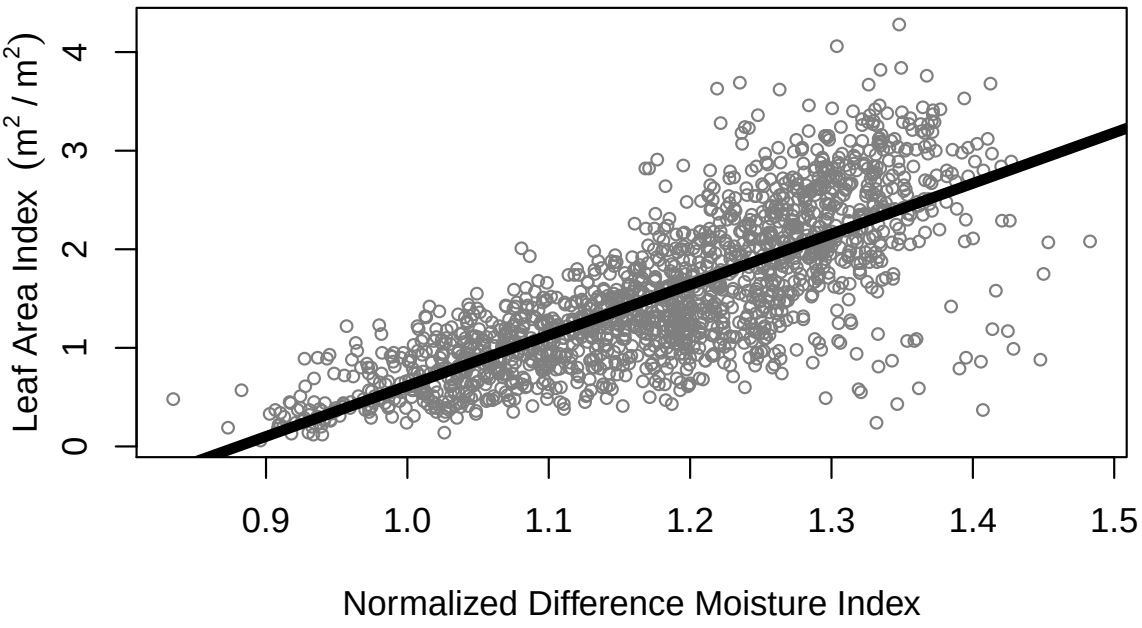


Figure 2.3: Relationship between the normalized difference moisture index and leaf area index data fit to a simple linear model for all observed measurement points (n=1673). Adjusted R^2 for this relationship was 0.5837 and root mean square error was 0.5049.

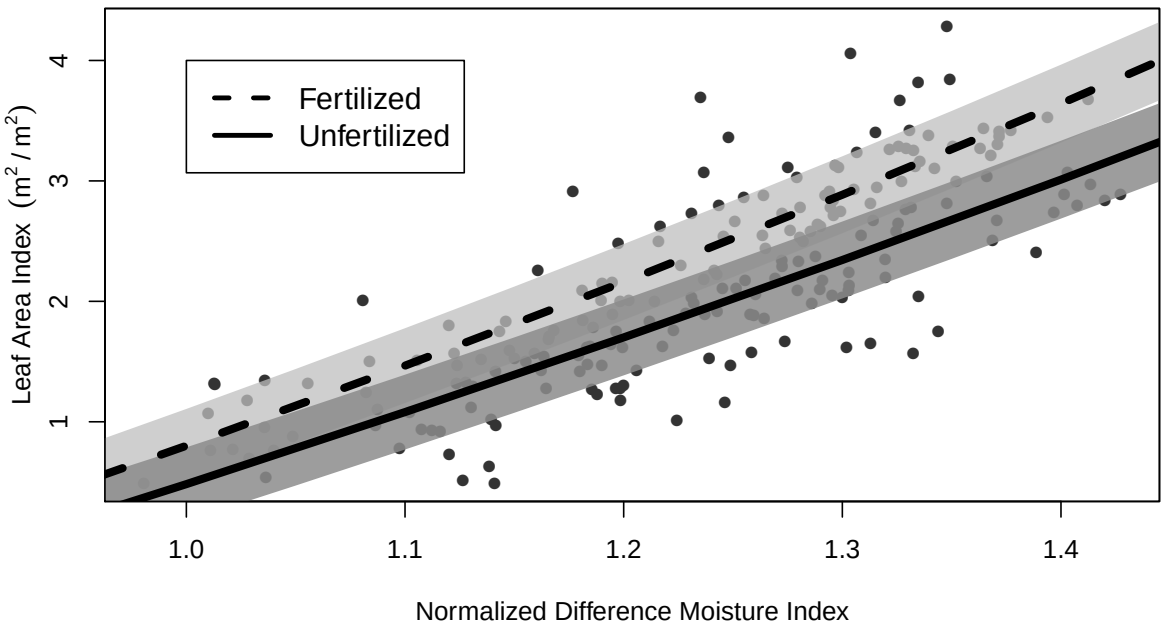


Figure 2.4: Model 10 fit using the unbiased parameters estimated using the SIMEX methodology.

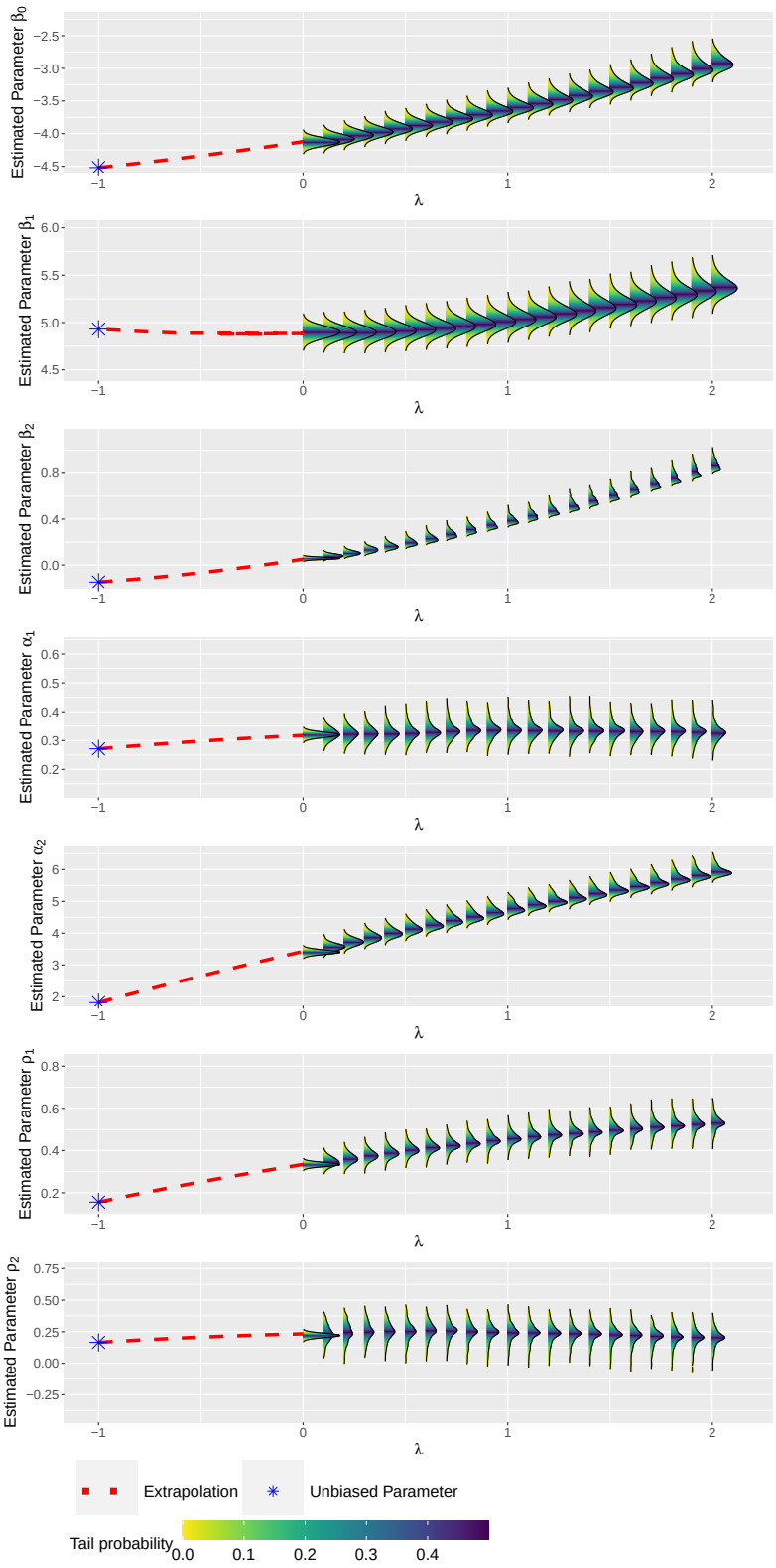


Figure 2.5: Results from the simulation-extrapolation (SIMEX) method for reducing the effects of measurement error on the parameter value. As error was increased from 0 to 2 (λ) in the 1000 simulations per step, a distribution was calculated. A quadratic linear model was used to fit the simulations and extrapolate to a λ level of -1, as if there was no measurement error.

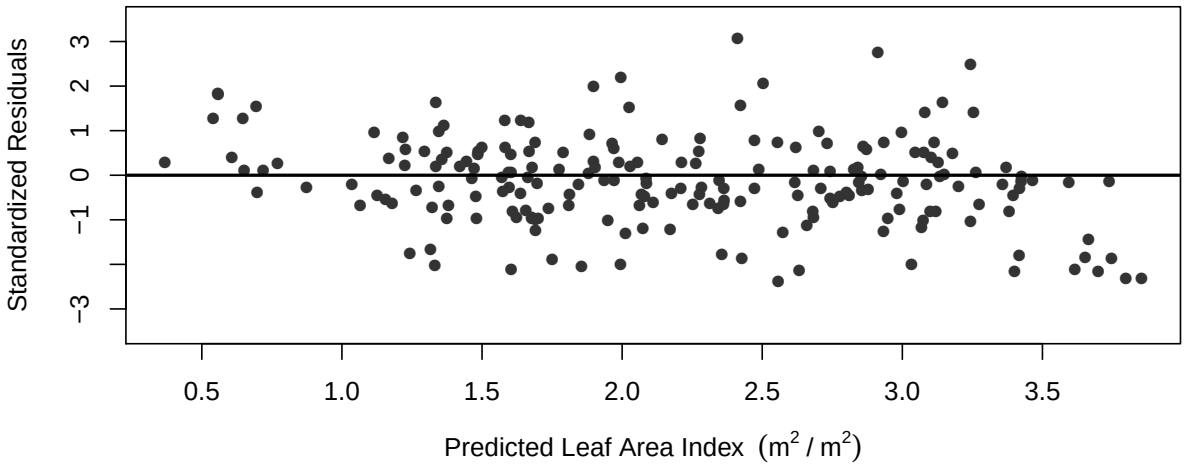


Figure 2.6: Studentized residuals for predictions.

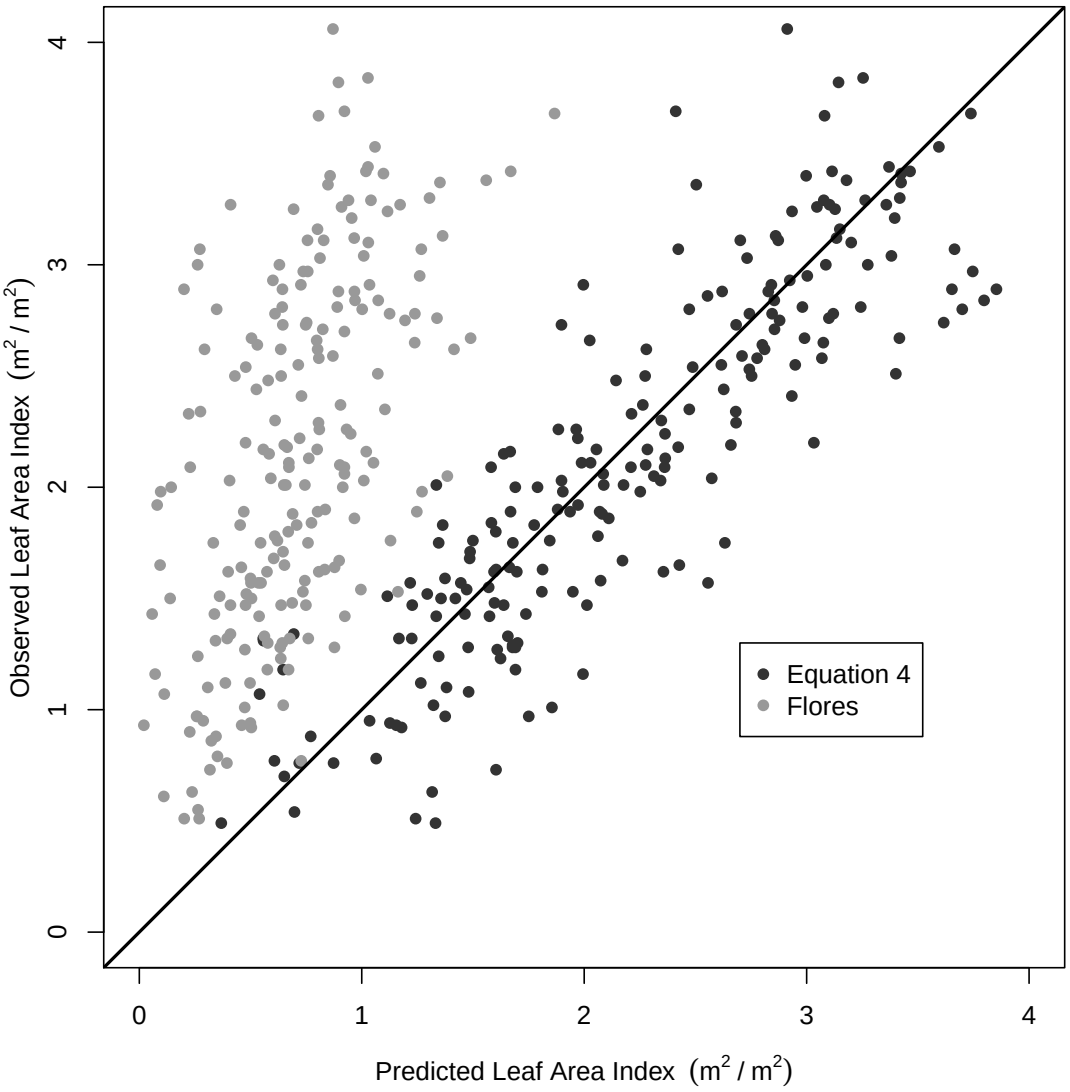


Figure 2.7: Predicted vs. measured leaf area index from the SIMEX bias corrected model.

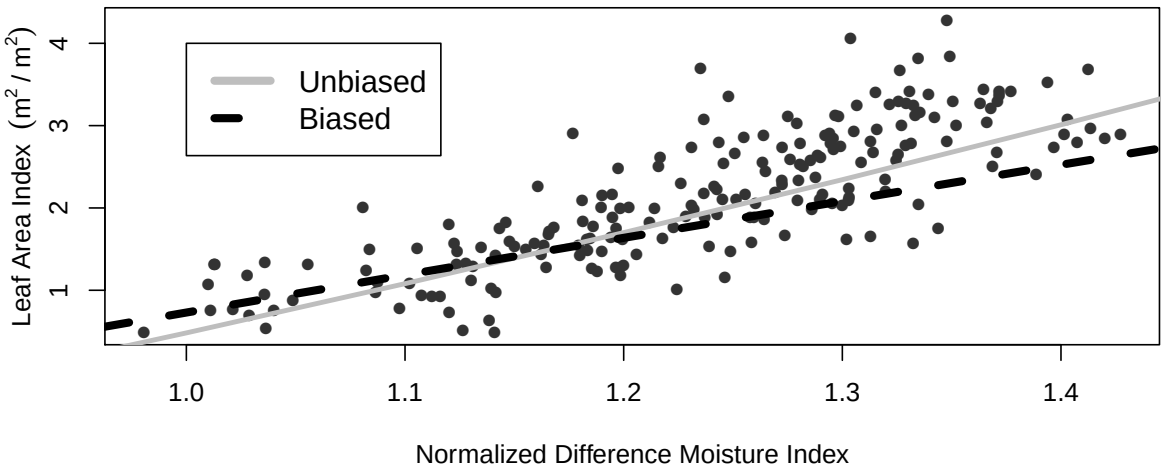


Figure 2.8: Comparison of the unbiased and biased model fits for the unfertilized (FERT = 0) model.

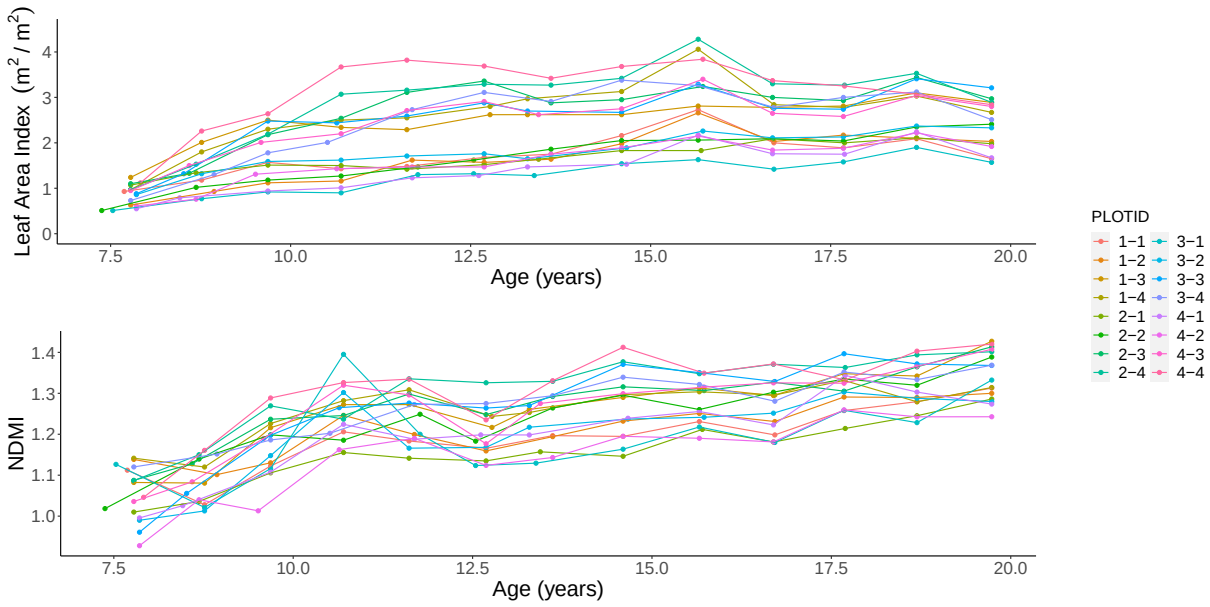


Figure 2.9: Trend comparison for LAI and NDMI for the SETRES data across the ages of observation. PLOTID indicates the (replication) - (treatment), where treatment 1 = control, 2 = irrigation only , 3 = fertilization only , and 4 = irrigation x fertilization.

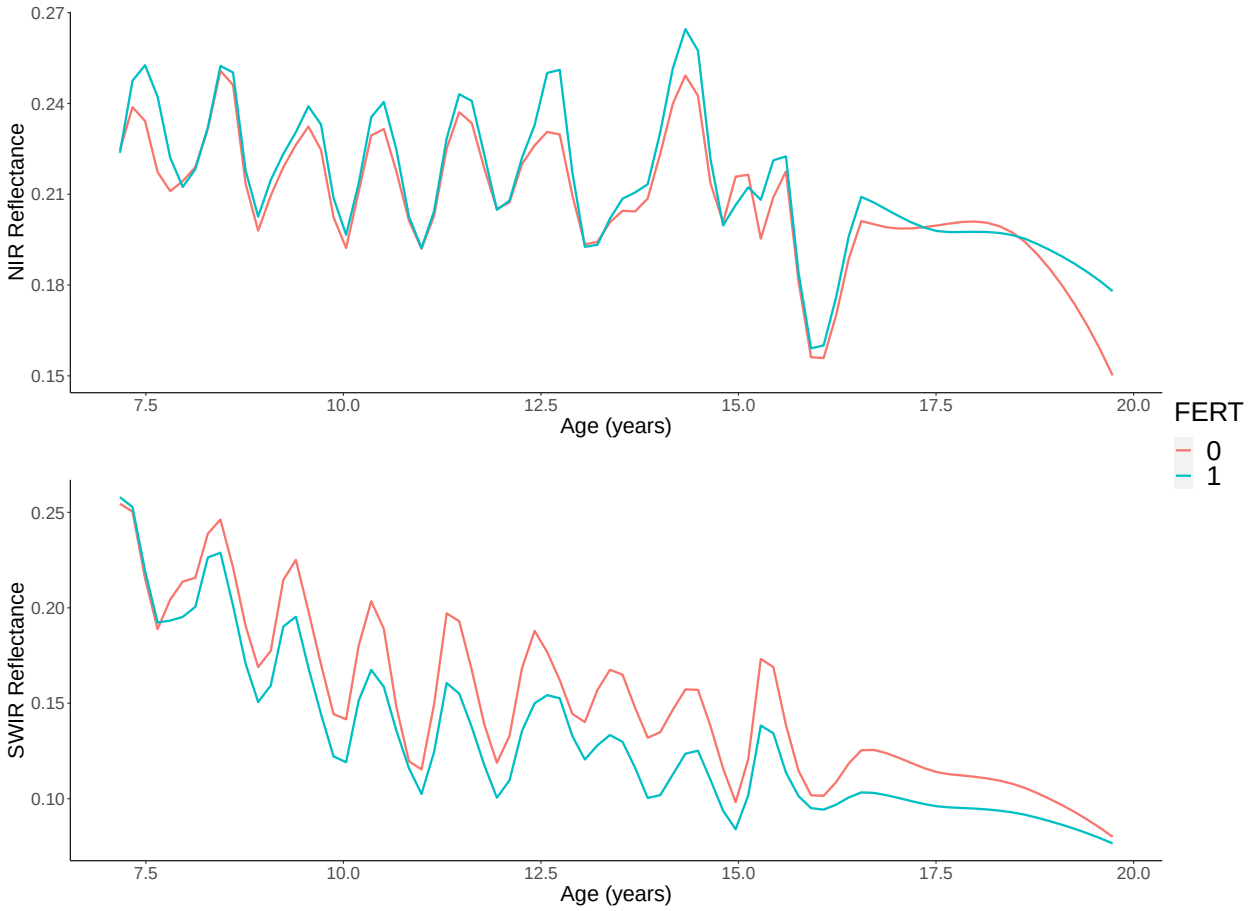


Figure 2.10: Changes in average Landsat 5 and 7 Band 4 (NIR) and Band 5 (SWIR) reflectance values by fertilization status.

Chapter 3

Influence of Environmental Variables on Leaf Area Index in Loblolly Pine Plantations[1]

[1]Kinane, S.M., M. Zapata, B.P. Bullock, R.L. Cook, D. Mishra, and C.R. Montes. To be submitted to *Agricultural and Forest Meteorology*.

Abstract

Variation in radiation interception is a considerable source of the variability observed in forest productivity. Leaf area index is an important biophysical variable that serves as the surface in which energy is exchanged in the canopy. To adequately capture the variability in radiation interception, we proposed a model that describes leaf area index growth in loblolly pine plantations in the southeastern United States using a delayed differential equation and periodic coefficients to enforce the seasonality of resource availability. Furthermore, productivity modifiers were incorporated to test the effect of different environmental variables on capturing additional variability observed in leaf area index time series. Monthly maximum temperature and monthly excess water were found to influence the leaf area index of stands and modifiers were included improve overall model performance.

3.1 Introduction

Variability in forest productivity has largely been attributed to the variation in radiation interception (Monteith, 1972; Grier and Running, 1977; Jarvis and Leverenz, 1983; Vose and Allen, 1988; Leverenz and Hinckley, 1990). Substantiating the variability in radiation interception is a multifaceted issue; with potential changes being driven by the environment (Vose et al., 1994; Dougherty et al., 1995), phenology (Vose et al., 1994), stand composition (Pretzsch, 2014), crown structure/architecture (Leverenz and Hinckley, 1990; Stenberg et al., 1994; McCrady and Jokela, 1998; Falster and Westoby, 2003; Niinemets et al., 2004; Binkley et al., 2010; Duursma et al., 2012), species (Chmura and Tjoelker, 2008), genotype (Jayawickrama et al., 1998; Chmura and Tjoelker, 2008), or silvicultural practices (Vose and Allen, 1988; Zhang et al., 1997; Will et al., 2005; Chmura and Tjoelker, 2008; Campoe et al., 2013). While other factors, such as light use efficiency and available photosynthetically active radiation, are necessary components in determining a forest's productivity (Martin and Jokela, 2004; Reich, 2012), leaf area index (LAI) is a popular subrogate for radiation interception due to its ecological significance as the surface that energy, carbon dioxide, and oxygen are exchanged between the environment (Grier and Running, 1977; Vose et al., 1994; Dougherty et al., 1995) and its ability to be estimated using readily available remotely sensed data (Badhwar and MacDonald, 1986; Spanner et al., 1990; Chen and Cihlar, 1996; Flores et al., 2006; Peduzzi et al., 2012; Savoy and Mackay, 2015; Blinn et al., 2019). LAI is the single-sided sum of the leaf surface area per unit ground area (m^2/m^2) and has been used in process models to assist in modeling energy fluxes, most successfully with physiological responses to climate (Running et al., 1989). Solar radiation is an essential source of energy for many biological processes and governs plant growth and development.

In regards to loblolly pine (*Pinus taeda* L.), the concentration of leaf area in a stand at any given time is a function of stand structure, nutrient availability, annual phenology, and environmental trends pertaining to the current year's leaf biomass and, further complicating the matter, the same trends' effects in the year prior due to the carryover of the newest age

class of needles (Dougherty et al., 1995). Structure of the stand has a fundamental influence on how the foliage is distributed within the canopy, affecting the probability that radiation passing through the canopy will intercepted (Russell et al., 1989). Among the structural factors affecting that probability besides the amount of leaf area include the grouping and orientation of foliage throughout the canopy (Russell et al., 1989). Simulations have shown that the non-uniform distribution of leaves within a canopy and with a range of leaf inclination angles can intercept more radiation and assimilate more carbon dioxide compared to a uniformly distributed canopy of the same leaf area (Russell et al., 1989). This increase in radiation interception is a result of more needles intercepting radiation as the solar incident angle changes throughout the day (Russell et al., 1989).

The phenology of loblolly pine needles is an environmentally driven process that represents a flux of nutrients within a given tree and corresponds to changes in growth rates on an intra-annual seasonal basis (Dougherty et al., 1995; Jayawickrama et al., 1998; Albaugh et al., 2012). The phenological process relies on genetics, temperature, photoperiod, water availability, and nutrient availability to stimulate the hormones that spur physiological changes (Dougherty et al., 1994; Bahuguna and Jagadish, 2015; Forrest and Miller-Rushing, 2010; Quesada et al., 2017). The annual phenological trend of loblolly pine needles has been characterized as a system of three stages: 1) a period of rapid needle accretion, followed by 2) a reduction in growth in which the needles reach their final length, and 3) a period of needle abscission (Sampson et al., 2003). Needle elongation follows the onset of bud break and shoot elongation, which can be temporally related to thermal units (Dougherty et al., 1994; Teskey et al., 1987; Russell et al., 1989), but not necessarily affected by fertilization status (Zhang et al., 1997). Multiple flushes of foliage occur throughout the growing seasons for loblolly pine, with later flushes attaining similar elongation rates but not overall lengths as first flushes (Dougherty et al., 1994). In other pine species, determination of final needle length was greater from droughts occurring earlier in the growing season (Dougherty et al., 1994) while overall leaf area levels were lower on sites with reduced water availability

(Teskey et al., 1987; Vose and Swank, 1990). Nutritional status has been shown to affect needle elongation rates and overall needle length (Zhang et al., 1997). The overall duration for needles is a function of the environment, nutrition status, and genetics (Vose and Allen, 1988; Dougherty et al., 1994; Zhang et al., 1997; Gholz et al., 1991; Hennessey et al., 1992). Fertilized stands have been shown to retain foliage longer as compared to unfertilized stands (Zhang et al., 1997). Needle abscission introduces a higher degree of variability into the system as compared to accretion and has been shown to be affected by site nutritional and water status, density, and environmental variables (Hennessey et al., 1992; Vose and Allen, 1988). Longer growing seasons, a function of the number of frost free days observed, provide more time for suitable growing conditions but may increase the respiration losses due to higher temperatures (Teskey et al., 1987). Understanding the environmental drivers of loblolly pine foliage provides a necessary understanding of the factors that promote or limit growth within and between years, and contribute to the variability observed in radiation interception. Environmental regulation of leaf area has been reported to be influenced by temperature, water availability, and site available nutrients (Waring et al., 1978; Vose and Allen, 1988; Hennessey et al., 1992; Dougherty et al., 1995; Zhang et al., 1997; Savoy and Mackay, 2015).

The goal of many forest management cultural treatments is to increase the availability of limiting resources on the site to improve conditions and promote growth. With a better understanding of the factors that affect loblolly pine productivity, we can improve the effectiveness of forest management goals and practices by providing tools for simulation purposes to understand the potential effects of a changing climate on the mechanisms that govern growth.

The objective of this research was to model the phenological development and trends of LAI in loblolly pine stands in the southeastern United States and to identify and model the effects of environmental variables on LAI development. The two main objectives of this research were to:

- Derive a model that accurately describes the development of LAI in a loblolly pine plantation
- Incorporate environmental variables to account for variations in LAI

3.2 Materials and Methods

3.2.1 Data

Data from the Plantation Management Research Cooperative's (PMRC) Coastal Plain Culture Density (CPCD) and South Atlantic Gulf Slopes (SAGS) studies were selected for this analysis. The CPCD study was installed in 1995 and 1996 as a split-plot design to study the effects of two cultural regimes (main plot) across a range of planting densities (subplot) in seventeen installations throughout the lower coastal plain of Florida, Georgia, and South Carolina (Zhao et al., 2011). The cultural regimes for the CPCD study included: 1) operational, consisting of bedding and banded chemical site preparation, herbaceous weed control after the first growing season, and fertilization at planting, before the 8^{th} growing season, and before the 12^{th} growing season and 2) intensive, consisting of bedding and broadcast chemical site preparation, tip moth control, complete vegetation control, and multiple fertilization treatments throughout the first twelve growing seasons (Zhao et al., 2011). The planting densities ranged from 741 to 4448 trees per hectare on study plots sized from 0.12 to 0.23 hectares to accommodate the minimum number of measurement trees. Further details for the CPCD study can be seen in Harrison and Kane (2008). Complementary to the CPCD, the SAGS study was installed in 1997 and 1998 at 23 installations throughout the Piedmont/Upper Coastal Plain physiographic region across five southeastern states (Zhao et al., 2010). Experimental design for the SAGS study was similar to the design of the CPCD study, testing the effects of cultural treatments (plot) across planting densities (subplot). For further details on the SAGS experimental design and treatments, see Zhao et al. (2008). A total of 24 sites, 11 from the CPCD study and 13 from the SAGS study were selected for analysis (Figure

3.1). Only plots under the intensive treatment were selected to remove potential sources of variability associated from differing treatment responses.

Due to the range of subplot size in the CPCD and SAGS studies, subplots from the 741 trees per hectare planting density (3.66 m x 3.66 m spacing) receiving intensive treatments were selected to maximize the subplot size (0.23 hectares) to increase the geographic footprint for remotely sensed data acquisition. All available surface reflectance scenes from the Landsat 5 TM and Landsat 7 ETM+ sensors were queried in the Google Earth Engine platform from site establishment to end of year 2019, masked for clouds, and exported as individual band values averaged for the subplot (Gorelick et al., 2017). Data were further filtered based upon cloud quality attributes, pixel quality attributes, and radiometric saturation. Using near-infrared (band 4) and short-wave infrared (band 5) bands, the normalized difference moisture index (NDMI) was calculated. Estimates of LAI were produced from an empirical model using the NDMI values. The number of available scene observations per study plot ranged from 334 to 1005 from plot establishment to December 31, 2019. Additionally, erroneous and outlying plots within the first five years of development were removed to reduce the amount of noise observed in the time series. The time series of individual plots were shortened to remove any observations that included post-thinning or increased mortality to insects. To provide increased temporal consistency and determine the underlying LAI development trend and seasonality, a thin plate spline (TPS) from the fields package was used to smooth the observed data (Savoy and Mackay, 2015; Nychka et al., 2017). LAI time series for each plot were fitted using the TPS function with the cost function set to 1.2. Using the fitted TPS model, a new time series for each plot was created on a 0.05 year time step for the age range observed at each plot.

Environmental variables for the individual study sites were sourced from the University of East Anglia's Climate Research Unit (CRU) (Harris et al., 2020). Using CRU's TS v. 4.04 NetCDF product, monthly values at the individual study plots were acquired from establishment to December 2019. Water storage capacity for the first 1.5 meters of soil

was estimated at the individual site using data from SSURGO spatial data base. Additional indices, including water deficit, water deficit index, excess water, and excess water index were calculated from the environmental variables in association with water storage capacity at the individual study plots. Water deficit for the site is calculated by subtracting the monthly precipitation from the monthly estimated potential evapotranspiration. Conversion of water deficit to water deficit index incorporates subtracting out the water storage capacity and summing the water deficits (Ojeda et al., 2018). Excess water and its index are calculated in similar fashion by subtracting the monthly potential evapotranspiration from the monthly precipitation.

3.2.2 Model Formulation

To provide more insight into the development trends of LAI and potential for influence by environmental variables, a growth model was developed to understand some of the underlying mechanisms (Zeide, 1993). Differential equations have been used to model growth, among many other things, as a rate of change of one variable in relation to another, primarily time (Fulford et al., 1997). These equations provide a mechanistic hypothesis for how systems interact and their parameterization can provide biological interpretation, which is paramount to understanding the complex dynamics that factor into tree growth (Zeide, 1993; Garcia, 2001). From observing many LAI time series for plantations across the southeast, several overarching trends were distinguished. Overall, LAI increases from stand initiation until some upper asymptote is reached and is maintained with some variations until a disturbance removes a significant portion of leaf area (e.g. thinning, harvest, pests or disease) (Switzer et al., 1966; Jarvis and Leverenz, 1983). To further the biological applicability and interpretation of our model, we proposed a delayed differential equation (DDE) to model the growth and development of LAI in loblolly pine plantations in the southeast United States. DDEs provide unique insight into biological systems by using the current state and a historical state to determine the future state. DDEs are a common way of modeling biolog-

ical systems that tend to rely on pasts states due to delays in development, such as sexual maturity, energy fluxes, overlapping populations, or predator/prey dynamics (Hutchinson, 1948; Wright, 1955; Nisbet and Gurney, 1976; Gopalsamy et al., 1990; Gopalsamy, 1992; Gallegos et al., 2008; Ghil et al., 2008; Keane et al., 2017). The common theme for use of DDEs is the understanding that feedback mechanisms that occur in nature typically come with some delay. DDEs have been used in forestry to model competition for light (Magal and Zhang, 2017, 2018) and forest planning (Garcia, 2001; Chaudhary and Pathak, 2017). In our case of modeling loblolly pine LAI, DDEs provide the ability to model the effect of prior and current leaf area states on future states, an important feature since loblolly pine canopies feature delays in physiological and phenological processes. Furthermore, historical trends provide influence on current canopy dynamics due to foliage in loblolly pines typically last for two years, providing up to two concurrent age classes of needles occupying the canopy at any given time, and a cause of great variability observed in canopies (Dougherty et al., 1995; Sampson et al., 2003). Additionally, the environmental regulation of LAI provides a scenario in which a delay between a LAI altering event and the resulting reaction in LAI may occur. DDEs are typically formulated as: $x_{t+1} = f(x_t, x_{t-\tau})$, where τ is the lag determining the location of the prior state. An effective DDE provides a glimpse into the dynamics of a given population, which in this case is leaf area, and can provide a cause and effect analysis into the factors determining growth (Kuang, 1993). The baseline model form proposed to represent the development of LAI is:

$$\frac{dL}{dt} = r * L_t * \left(1 - \frac{L_{t-\tau}}{K + A * \sin\left(\omega * t + \mu\right)}\right) - h * L_t \tag{3.1}$$

Where L is LAI, r is the growth rate parameter, L_t is the current state at time t, $L_{t-\tau}$ is the previous state τ units ago, K is the carrying capacity parameter, A is the amplitude, ω is period, μ is the phase shift, and h is another parameter to denote mortality. Equation 3.1 is split into three main components, 1) logistic equation ($r * L_t * (1 - \frac{L_{t-\tau}}{K})$), 2) seasonal component ($A * \sin\left(\omega * t + \mu\right)$), and 3) a mortality component ($h * L_t$). The logistic

equation provides the underlying trend of growth and development approaching an upper asymptote, typically referred to as a carrying capacity in the literature and denoted by K in the model. The r parameter provides a measure as to how quickly the trend reaches the upper asymptote. This particular implementation of the logistic model, sometimes referred to as the Hutchinson model, uses a prior state $L_{t-\tau}$ to provide reference and determine the trajectory of the current state L_t. As $L_{t-\tau}$ approaches K, the relative growth rate decreases, indicating that the carrying capacity has been met. In this case, we are proposing that our delay is discrete, as opposed to vanishing, neutral, or distributed delays that are used in other DDEs (Kuang, 1993).

Seasonal forcing was imposed through the use of a trigonometric function to provide the phenological representation of the leaf development and senescence annual cycle as a result of changing resource availability and reflected on the K parameter. With this formulation, K serves as the vertical shift in the periodic function. The final component, mortality, provides an accounting of additional biomass mortality that occurs outside of the seasonal leaf component.

Inclusion of Environmental Variables

Extendability of the DDE to relate to environmental conditions that govern growth can be implemented by establishing the model parameters as functions of environmental variables (Powers et al., 2003; Kramer, 1994). It has been well established that environmental variables, such as temperature and water, are important drivers of loblolly pine productivity and regulate the annual phenological processes that determine the annual growth (Albaugh et al., 2012; Dougherty et al., 1994). To incorporate these environmental variables as growth regulators into the system, several model forms were evaluated for the individual components, including the beta equation (Equation 3.2), the logistic equation (Equation 3.3), the double logistic equation (Equation 3.4), the Arrenhius equation (Equation 3.5), the Heaviside equation (Equation 3.6, and a simple linear equation (Equation 3.7).

$$f(env(t)) = \exp(\mu) * (env(t) - T_b)^{\alpha} * (T_c - env(t))^{\beta} \tag{3.2}$$

$$f(env(t)) = \alpha \left(\frac{1}{1 + \exp(-b * (env(t) - m))} \right) \tag{3.3}$$

$$f(env(t)) = \alpha \left(\frac{1}{(1 + \exp(-b * (env(t) - m)))(1 + \exp(d * (env(t) - n))} \right) \tag{3.4}$$

$$f(env(t)) = \alpha \left(\frac{\exp(\frac{-b}{env(t)})}{1 + c * \exp(\frac{-d}{env(t)})} \right) \tag{3.5}$$

$$\begin{aligned} f(env(t)) &= \alpha(env(t) - T_b) * H(env(t) - T_b) \\ \text{where} \quad H(env(t) - T_b) &= \begin{cases} 1 \text{ if } env(t) \geq T_b \\ 0 \text{ if } env(t) < T_b \end{cases} \end{aligned} \tag{3.6}$$

$$f(env(t)) = \beta_0 + \beta_1 * env(t) \tag{3.7}$$

3.2.3 Model Parameterization

The DDE model proposed was parameterized using maximum likelihood estimation. To account for differences between the individual plots, parameterization included global variables for the base model and modifier components and local variables for initial starting values at the individual plot level. The model relied on 33 parameters: eight for the LAI equation, one for the standard deviation for the maximum likelihood procedure, and 24 local variables describing the starting points for each time series. Using the environmental modifiers to calculate a productivity adjustment factor for the base model, the relationship was assumed to be interactive, so the base model and environmental models were formulated in a multiplicative fashion (e.g. $\frac{dL}{dt} = f(ENV) * g(L)$). The dede function in the deSolve package was used to solve the DDE model and a combination of the BFGS, Nelder-Mead, and CG

methods were used for optimization methods to maximize the log likelihood of the functions as implemented in the optimx package (Soetaert et al., 2010; Nash and Varadhan, 2011). Starting values for the local parameters were defined by the first observations of LAI at the individual plot level. For the models using local K values, starting values were the global K parameterized in the model fitting procedure. Global parameters starting values were selected through a trial and error approach to define the model form and starting values for the environmental modifiers were selected based on the reported literature response between the environment and loblolly pine (Nedlo et al., 2009; Albaugh et al., 2004; Teskey et al., 1987). Addition of the environmental variables and their subsequent model form were evaluated on reductions of root mean square error (RMSE) and Akaike's Information Criterion (AIC) to account for the additional parameters for the given model. Evaluations for parameter uncertainty estimates required the fixing of local variables to allow for the final standard error calculations. To evaluate model robustness, k-fold (k=9) cross validation was used. The 24 plots were randomly assigned to one of nine groups. For each of the nine groups, the plots in the group were withheld from the training data set, the model was fit on the training data set and then evaluated on the withheld group. The process was repeated for each group.

3.3 Results

3.3.1 Base Model Parameterization

Parameterization of the base model (Equation 3.1) resulted in an overall root mean square error (RMSE) for the 24 plots of 0.4427 and an AIC of 12496.7 (Table 3.1). Estimated standard errors for the individual parameter estimates indicate little variability with the exception of the K parameter. At the individual plot level, RMSE ranged from 0.2594 to 0.6158 LAI m^2/m^2. Inspection of the residuals showed little evidence of heteroscedacity (Figure 3.2) and comparison of the observations to predictions showed an adequate 1:1 relationship (Figure 3.3). Model predictions show a defined upper limit at approximately 3.72 m^2/m^2, where the base model limits the overall LAI production (Figure 3.3). Annual

seasonality was adequately captured by the trigonometric portion of the base model, with peak LAI occurring in September/October every year (Figure 3.4). The τ estimate indicated little importance of the prior state, with a value of 0.0002 (Table 3.1). Without a considerable lag effect, the relationship between the current and prior state maintained linearity (Figure 3.5). The K parameter showed an overall estimated carrying capacity for LAI at 4.7760 m^2/m^2, subject to the model formulated constraints and imposed seasonality. Systematic patterns were observed in the residuals as LAI increased, an artefact of the model reaching a global steady state with no outside forces imposing individuality at the plot level (Figure 3.2). The base level model was inadequate in capturing the seasonality and trend of LAI in the formative years of stand development, common to all observed plots (Figure 3.4). While the underlying development trend was captured with the inclusion of the local initial value at the individual plot level, the base model provided a gross simplification of the system. The amplitude of seasonality in the parameterized model showed large variations as compared to amplitude of seasonality in the observed plots and their subsequent smoothed spline time series.

3.3.2 Inclusion of Environmental Variables

Temperature was found to be the most influential on improving the LAI, with monthly maximum temperature having the greatest effect. The logistic, double logistic, and beta model forms proved to be the most flexible in parameterizing for inclusion of the environmental variables. Maximum monthly temperature in the beta form showed the greatest increase in model performance when included as a modifier to the base LAI model. RMSE decreased to 0.4299 and AIC decreased to 11891.71. Addition of the environmental modifier increased τ to 0.1148, or approximately 41 days, indication of past state dependency and a delayed reaction in physiological response to temperature (Table 3.2). Base level carrying capacity (K) decreased from 4.7760 to 4.5565 while the rate parameter (r) showed a slight increase from 1.0750 to 1.1109, a result of the modifier imposing changes to the base level productivity

at the individual sites (Table 3.2). Across the individual plots, inclusion of the temperature modifier produced a range of 0.2616 to 0.5846 RMSEs, indicating adverse effects on some plots and benefits to others. Fourteen of the 24 plots saw reductions in RMSE with the addition of the temperature modifier. Comparisons between the trend and the model showed an increase in the model's ability to react to environmental conditions, capturing some of the changes in the observed points and smooth trend (Figure 3.8). The resulting parameterized beta monthly maximum temperature modifier showed an asymmetrical modifier across the range of observed monthly maximum temperatures, displaying a slow modifier rate increase from the lower bound ($\approx$10°C) to a peak of approximately 30°C, and a sharp decline in modifier value following the peak as temperature increased to beyond 35°C (Figure 3.10). Increase of the τ parameter provided a relatively stable limit cycle relating the current and past states as LAI developed and reached its upper asymptote (Figure 3.9). Cycles around a central point, $\approx$ 3.1 LAI m^2/m^2 showed slight temperature induced deviations but overall stable process.

Evaluation of additional modifiers and their interactions with the beta monthly maximum temperature modifier focused primarily on water relations variables to reduce to collinearity between monthly maximum temperature and the other temperature variables under consideration. Examination of the water relations variables and their interactions with the monthly maximum temperature modifier found that monthly excess water in the double logistic model form to be the most influential for improving the model fit. Inclusion of excess water as an additional modifier had modest improvements in model fit as compared to monthly maximum temperature's effect on the base model, with the largest decrease of RMSE at the individual plot being 0.03 units of LAI. The updated parameter values for the base model and beta monthly maximum temperature modifier showed slight changes as compared to the temperature - modifier only model, and a slight overall decrease in RMSE to 0.4257 and AIC to 11684.19. The newly parameterized monthly maximum temperature modifier displayed a similar trend as the previous model (Figure 3.15, Table 3.3). The excess water modi-

fier displayed a symmetrical shape centered about approximately 175 mm of excess water (Figure 3.15). As excess water increased from 0 mm to the peak, the modifier increased from a baseline of approximately 0.70 to approximately 0.9. Following the peak, the modifier value began to decline to approximately 0.75 as the excess water value reached the end of its natural range of 350 mm observed in the data. When the two environmental modifiers' interactions were calculated, a defined area of increased productivity above the baseline model (modifier > 1) was observed at the peaks of temperature and excess water (Figure 3.16). A large area within the interaction plot indicates that a wide range of conditions can satisfy the conditions necessary for peak LAI growth.

3.3.3 Addition of Local Carrying Capacity

To test the effect of a local carrying capacity parameter compared with a global parameter, the final model was refitted with the addition of a local K parameter for each individual plot and dropping the global K, increasing the number of parameters by 23. RMSE across all observations decreased to 0.3802 accompanied by a decrease in AIC to 9383.04 (Table 3.3. Values for K ranged from 3.59 to 4.40, which had a weak positive relationship with estimated site index (Figure 3.23). As compared to the base model, the largest decrease in RMSE due to the local K parameter was a 0.29 reduction.

Model validation through k-fold (k=9) cross validation showed overall support for the proposed models (Table 3.5). Improvement was made with the inclusion of environmental variables as compared to the base model, while the addition of a second environmental variable showed minimal declines in RMSE in training and validation data sets (Table 3.5). Inclusion of the local parameter K showed a considerable decline in RMSE for both the training and validation data sets, indicating the importance of local features on LAI growth and development and its overall interaction with environmental variables.

3.4 Discussion

Dependency of the future state based on the current and prior states, denoted by τ, showed varying conditions based on whether additional information, such as temperature, was provided in terms of a productivity modifier. In the base model, τ proved to be quite small, an indication that, as modeled, LAI followed the development and seasonal trend defined by the periodic and rate parameters and didn't require further information from prior states. It has been shown that at sufficiently small τ, the Hutchinson model is equivalent to the traditional logistic, providing a simplified model (Gopalsamy, 1992). As additional variability in the modeled system was introduced by the environmental modifiers, a dependence on those prior states proved influential, indicating that the state at time $t - \tau$ is necessary in determining the future state. These larger lags produce a feedback loop, which can be visualized in Figures 3.9 and 3.14, where oscillations about the equilibrium point of the solution for the individual parameterized equations occurs as t increases and the population (LAI) reaches steady state (Gopalsamy, 1992). Fluctuations in the oscillations about the equilibrium point indicate the influence of the environment, but overall LAI shows to be a relatively stable process, or one that is able to return to its equilibrium (May, 1973).

Differences in the observed equilibrium point in the two models incorporating environmental effects and their individual parameterized K values shows a defined difference, with the equilibrium point far below the K. This difference is likely due to the fluctuations in K, which is representative of the carrying capacity but now defined as the vertical shift in our periodic coefficient, a representation of the changing resource availability and additionally influenced by the environmental variables, which has been reported to be the case in other population studies (May, 1973; Roff, 1974). Stability and persistence of a given population has been shown to have direct correlations with the heterogeneity of the population's environment, with the probability of persistence decreasing with increased environmental variability (Roff, 1974; Levins, 1969).

While the base model overall has adequate fit and conceptually represents LAI growth, the steady state reached does not provide an accurate understanding of the observed LAI dynamics, as once net growth for a period is zero, we see no changes to the dynamics, which is unlikely in any ecosystem. The mechanisms proposed to model LAI development were not able to fully capture the observed dynamics, especially in the early stages prior to reaching steady state (Figures 3.4, 3.8, and 3.11). Large changes between years in the seasonality, peak, and minimums observed were glossed over by the proposed model form, with poor representation of those early stand dynamics. From the remote sensing and LAI estimation standpoint, this portion of the individual time series provided great difficulty in providing a smooth seasonal trend due to a large number of erratic and outlying estimates from the Landsat 5 and 7 ETM+ imagery. Potential sources of this variability may be from pixel level saturation of neighboring plots, large amounts of competing vegetation present in the stand prior to treatment, among others. The oscillations observed between the prior and current state for the second, third, and fourth models about their equilibrium indicated a stable process with τ sufficiently large enough to provide adequate feedback (Figures 3.9, 3.14, 3.20) (Gopalsamy, 1992).

Monthly maximum temperature provided the greatest increase in model performance for capturing the dynamics of loblolly pine LAI. Similar to the results presented in Savoy and Mackay (2015), many of the sites evaluated in this study were highly influenced by the temperature modifier. Increased rates at the cellular level at higher temperatures have been reported to provide increases in photosynthesis and respiration, leading to greater increase of carbon assimilation rates (Way and Oren, 2010). More specifically, Way and Oren (2010) found that evergreens increased leaf mass, among other allometric changes, when temperatures were increased. This change may be driven by greater leaf elongation rates from increases in division and expansion at the cellular level (Ryan, 2010).

Large decreases observed in the time series data at individual plots may have been caused by interaction effects of environmental stress with disease or pests, clearly not captured in

the models proposed here. While these biotic effects can cause large variations in observed leaf area, further modeling developments can be used to incorporate the probability of these interactions affecting the system given environmental inputs and the magnitude of their potential effects. Providing insight on the conditions that cause fluctuations of LAI outside of the base model, shown here with maximum temperature and excess water, allow for a more systematic understanding of forest productivity and provide tools for management planning to account for potential disturbances to the population (Gopalsamy, 1992).

Further analysis into periodic coefficients may provide additional insight into the seasonality of growth, the effects of the environment on the fluctuations of growth and its effects on individual model parameters (Gopalsamy et al., 1990; Lisena, 2007). Additionally, periodic coefficients may provide the necessary switching of the modifier to correctly adjust the growth given the forced seasonality of the model, i.e. promote growth when conditions are adequate and restrict declines in growth when the seasonality begins its annual decline. Potential improvements in model and computation efficiency may be had from implementing neural or universal ordinary differential equations to handle the various data on different time steps to provide a better adapting system (Rackauckas et al., 2020). Recent advances in neural ODEs may provide additional insight to the noisy environment observed within. Examining the lag component as a neutral DDE, where τ is used to define historical derivatives as opposed to past states may increase the ability to adequately capture the changes imposed by environmental modifiers. Formulation of the model imposed seasonality based on the sine component, a product allowing for the capture of seasonal changes in LAI throughout the year and the observation series for the plots. Further investigation into more parsimonious model forms may provide accurate mechanistic options into the foundations of LAI development. Following the methodology used by Powers et al. (2003), we incorporated our environmental variable modifier on the entire state rather than individual parameters. While future work will investigate the environmental influences individually on the parameters, earlier work has shown that in the logistic model form (Levins, 1969), the current state will follow a weighted

harmonic mean of K. Further inquiries into the effect of additional delays (n$\geq$2) and combinations of positive and negative feedback may provide a more thorough investigation into environmental induced dynamics of the LAI system on an annual and lifespan basis (Keane et al., 2017).

A sensitivity analysis into the size of time step selected in the splined LAI and environmental variables used for the modeling purposes should be undertaken to verify that the underlying system being modeled isn't affected by the timing and amount of observations (Powers et al., 2003). Splining the data and providing a discrete time step between interpolated observations improved the model fitting procedure by ensuring adequate observations across the observation period. LAI observations proved to quite noisy and provide a suitable medium for future improvements by incorporating stochastic methods to handle the erratic time series observed. Additionally, using annual peaks and minimums may provide an alternative to capture the true amplitude of the seasonality component. While Powers et al. (2003) recommended a backwards selection approach to inclusion of environmental variables, computational demand of the DDE and the large amount of data used for model fitting prohibited that methodology from being completed considering reasonable time constraints.

Interpretability of the formulated model allows for our increased understanding of the dynamics that govern forest productivity and how they interact with resource availability, which is observed in annual seasonal cycles and interactions with the local environment. Here we defined the carrying capacity of our system as a periodic coefficient to represent the changes in available resources throughout the year. When imposed, environmental modifiers retard annual growth of the system and we see an overall decline in the carrying capacity, which may affect other physiological changes associated with forest growth. When defining the carrying capacity of a forested stand, using shorter-term phenomena that pair resource availability and consumption, such as leaf area index, as opposed to cumulative basal area yield or stand density index that is typically reported, may provide a better biological repre-

sentation of a dynamical system. Further investigations into the regional effects of different environmental variables may provide more insight into the factors limiting growth across the loblolly pine spatial extent.

3.5 Conclusions

Overall it was shown that a mechanistic derived model can describe aspects of the growth and development of LAI in loblolly pine plantations and provide additional insight into the seasonal and environmental dynamics that affect the current state. We were able to provide improvements in a base loblolly pine LAI growth model by incorporating environmental variables for 24 sites across a wide geographic range in the southeastern United States. Monthly maximum temperature and monthly water deficit were shown to be influential in the development and long term trends of loblolly pine LAI in the southeastern United States.

By using the environmental modifiers to adjust the system, we are accounting for the changes in resource availability due to the stochastic environment, resulting in changes of the carrying capacity. As resources and conditions becoming limiting to the biological processes, we are able to model the effects on how much leaf area a stand can support at any given time. Periodicity was forced by the inclusion of a periodic coefficient for the K parameter, resulting in forced oscillations on the basis of resources are limited seasonally, affecting overall carrying capacity and driving the observed patterns.

3.6 References

Albaugh, T. J., Allen, H. L., Stape, J. L., Fox, T. R., Rubilar, R. A., and Price, J. W. (2012). Intra-annual nutrient flux in Pinus taeda. *Tree Physiology*, 32(10):1237–1258.

Albaugh, T. J., Lee Allen, H., Dougherty, P. M., and Johnsen, K. H. (2004). Long term growth responses of loblolly pine to optimal nutrient and water resource availability. *Forest Ecology and Management*, 192(1):3–19.

Badhwar, G. D. and MacDonald, R. B. (1986). Satellite-derived leaf-area-index and vegetation maps as input to global carbon cycle models-a hierarchical approach. *International Journal of Remote Sensing*, 7(2):265–281.

Bahuguna, R. N. and Jagadish, K. S. (2015). Temperature regulation of plant phenological development. *Environmental and Experimental Botany*, 111:83–90.

Binkley, D., Stape, J. L., Bauerle, W. L., and Ryan, M. G. (2010). Explaining growth of individual trees: Light interception and efficiency of light use by Eucalyptus at four sites in Brazil. *Forest Ecology and Management*, 259(9):1704–1713.

Blinn, C., House, M., Wynne, R., Thomas, V., Fox, T., and Sumnall, M. (2019). Landsat 8 based leaf area index estimation in loblolly pine plantations. *Forests*, 10(3):222.

Campoe, O. C., Stape, J. L., Albaugh, T. J., Lee Allen, H., Fox, T. R., Rubilar, R., and Binkley, D. (2013). Fertilization and irrigation effects on tree level aboveground net primary production, light interception and light use efficiency in a loblolly pine plantation. *Forest Ecology and Management*, 288:43–48.

Chaudhary, M. and Pathak, R. (2017). A dynamical approach to the legal and illegal logging of forestry population and conservation using taxation. *Advances in Difference Equations*, 2017(1):385.

Chen, J. M. and Cihlar, J. (1996). Retrieving leaf area index of boreal conifer forests using landsat TM images. *Remote Sensing of Environment*, 55(2):153–162.

Chmura, D. J. and Tjoelker, M. G. (2008). Leaf traits in relation to crown development, light interception and growth of elite families of loblolly and slash pine. *Tree Physiology*, 28(5):729–742.

Dougherty, P. M., Hennessey, T. C., Zarnoch, S. J., Stenberg, P. T., Holeman, R. T., and Wittwer, R. F. (1995). Effects of stand development and weather on monthly leaf biomass dynamics of a loblolly pine (Pinus taeda L.) stand. *Forest Ecology and Management*, 72(2-3):213–227.

Dougherty, P. M., Whitehead, D., and Vose, J. M. (1994). Environmental influences on the phenology of pine. *Ecological Bulletins*, 43:64–75.

Duursma, R. A., Falster, D. S., Valladares, F., Sterck, F. J., Pearcy, R. W., Lusk, C. H., Sendall, K. M., Nordenstahl, M., Houter, N. C., Atwell, B. J., Kelly, N., Kelly, J. W. G., Liberloo, M., Tissue, D. T., Medlyn, B. E., and Ellsworth, D. S. (2012). Light interception efficiency explained by two simple variables: a test using a diversity of small- to medium-sized woody plants. *New Phytologist*, 193(2):397–408.

Falster, D. S. and Westoby, M. (2003). Leaf size and angle vary widely across species: What consequences for light interception? *New Phytologist*, 158(3):509–525.

Flores, F. J., Allen, H. L., Cheshire, H. M., Davis, J. M., Fuentes, M., and Kelting, D. (2006). Using multispectral satellite imagery to estimate leaf area and response to silvicultural treatments in loblolly pine stands. *Canadian Journal of Forest Research*, 36(6):1587–1596.

Forrest, J. and Miller-Rushing, A. J. (2010). Toward a synthetic understanding of the role of phenology in ecology and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1555):3101–3112.

Fulford, G., Forrester, P., and Jones, A. (1997). *Modelling with Differential and Difference Equations*. Australian Mathematical Society Lecture Series. Cambridge University Press.

Gallegos, A., Plummer, T., Uminsky, D., Vega, C., Wickman, C., and Zawoiski, M. (2008). A mathematical model of a crocodilian population using delay-differential equations. *Journal of Mathematical Biology*, 57(5):737–754.

Garcia, O. (2001). Functional differential equations in sustainable forest harvesting. *Journal of Forest Planning*, 6(2):49–63.

Ghil, M., Zaliapin, I., and Thompson, S. (2008). A delay differential model of ENSO variability: Parametric instability and the distribution of extremes. *Nonlinear Processes in Geophysics*, 15(3):417–433.

Gholz, A. H. L., Vogel, S. A., Cropper, W. P., Mckelvey, K., Ewel, K. C., Teskey, R. O., and Curran, P. J. (1991). Dynamics of canopy structure and light interception in pinus elliottii stands, north Florida. *Ecological Monographs*, 61(1):33–51.

Gopalsamy, K. (1992). *Stability and Oscillations in Delay Differential Equations of Population Dynamics*. Springer Netherlands, Dordrecht.

Gopalsamy, K., Kulenović, M. R., and Ladas, G. (1990). Environmental periodicity and time delays in a "food-limited" population model. *Journal of Mathematical Analysis and Applications*, 147(2):545–555.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202:18–27.

Grier, C. G. and Running, S. W. (1977). Leaf area of mature northwestern coniferous forests: Relation to site water balance. *Ecology*, 58(4):893–899.

Harris, I., Osborn, T. J., Jones, P., and Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, 7(1):109.

Harrison, M. and Kane, M. (2008). PMRC Coastal plain culture/density study: age 12 analysis. *PMRC Technical Report*, 1:74.

Hennessey, T., Dougherty, P., Cregg, B., and Wittwer, R. (1992). Annual variation in needle fall of a loblolly pine stand in relation to climate and stand density. *Forest Ecology and Management*, 51(4):329–338.

Hutchinson, G. E. (1948). Circular causal systems in ecology. *Annals of the New York Academy of Sciences*, 50(4 Teleological):221–246.

Jarvis, P. G. and Leverenz, J. W. (1983). *Productivity of Temperate, Deciduous and Evergreen Forests*, chapter 8, pages 233–280. Springer Berlin Heidelberg, Berlin, Heidelberg.

Jayawickrama, K. J., Mckeand, S. E., and Jett, J. B. (1998). Phenological variation in height and diameter growth in provenances and families of loblolly pine. *New Forests*, 16(1):11–25.

Keane, A., Krauskopf, B., and Postlethwaite, C. M. (2017). Climate models with delay differential equations. *Chaos*, 27(11):114309.

Kramer, K. (1994). Selecting a model to predict the onset of growth of fagus sylvatica. *The Journal of Applied Ecology*, 31(1):172.

Kuang, Y. (1993). *Delay differential equations: with applications in population dynamics*, page 398. Academic Press.

Leverenz, J. W. and Hinckley, T. M. (1990). Shoot structure, leaf area index and productivity of evergreen conifer stands. *Tree Physiology*, 6(2):135–149.

Levins, R. (1969). The effect of random variations of different types on population growth. *Proceedings of the National Academy of Sciences*, 62(4):1061–1065.

Lisena, B. (2007). Periodic solutions of logistic equations with time delay. *Applied Mathematics Letters*, 20(10):1070–1074.

Magal, P. and Zhang, Z. (2017). Competition for light in forest population dynamics: From computer simulator to mathematical model. *Journal of Theoretical Biology*, 419(February):290–304.

Magal, P. and Zhang, Z. (2018). Numerical simulations of a population dynamic model describing parasite destruction in a wild type pine forest. *Ecological Complexity*, 34:147–160.

Martin, T. A. and Jokela, E. J. (2004). Developmental patterns and nutrition impact radiation use effieciency components in southern pine stands. *Ecological Applications*, 14(6):1839–1854.

May, R. M. (1973). Stability and complexity in model ecosystems. *Monographs in population biology*, 6:1–235.

McCrady, R. and Jokela, E. (1998). Canopy dynamics, light interception, and radiation use efficiency of selected loblolly pine families. *Forest Science*, 44(1):64–72.

Monteith, J. L. (1972). Solar radiation and productivity in tropical ecosystems. *The Journal of Applied Ecology*, 9(3):747.

Nash, J. C. and Varadhan, R. (2011). Unifying optimization algorithms to aid software system users: optimx for R. *Journal of Statistical Software*, 43(9):1–14.

Nedlo, J. E., Martin, T. A., Vose, J. M., and Teskey, R. O. (2009). Growing season temperatures limit growth of loblolly pine (Pinus taeda L.) seedlings across a wide geographic transect. *Trees - Structure and Function*, 23(4):751–759.

Niinemets, Ü., Cescatti, A., and Christian, R. (2004). Constraints on light interception efficiency due to shoot architecture in broad-leaved Nothofagus species. *Tree Physiology*, 24(6):617–630.

Nisbet, R. and Gurney, W. (1976). Population dynamics in a periodically varying environment. *Journal of Theoretical Biology*, 56(2):459 – 475.

Nychka, D., Furrer, R., Paige, J., and Sain, S. (2017). fields: Tools for spatial data. R package version 10.3.

Ojeda, H., Rubilar, R. A., Montes, C., Cancino, J., and Espinosa, M. (2018). Leaf area and growth of Chilean radiata pine plantations after thinning across a water stress gradient. *New Zealand Journal of Forestry Science*, 48(1):10.

Peduzzi, A., Wynne, R. H., Fox, T. R., Nelson, R. F., and Thomas, V. A. (2012). Estimating leaf area index in intensively managed pine plantations using airborne laser scanner data. *Forest Ecology and Management*, 270:54–65.

Powers, S. J., Brain, P., and Barlow, P. W. (2003). First-order differential equation models with estimable parameters as functions of environmental variables and their application to a study of vascular development in young hybrid aspen stems. *Journal of Theoretical Biology*, 222(2):219–232.

Pretzsch, H. (2014). Canopy space filling and tree crown morphology in mixed-species stands compared with monocultures. *Forest Ecology and Management*, 327:251–264.

Quesada, T., Parisi, L. M., Huber, D. A., Gezan, S. A., Martin, T. A., Davis, J. M., and Peter, G. F. (2017). Genetic control of growth and shoot phenology in juvenile loblolly pine (Pinus taeda L.) clonal trials. *Tree Genetics & Genomes*, 13(3):65.

Rackauckas, C., Ma, Y., Martensen, J., Warner, C., Zubov, K., Supekar, R., Skinner, D., Ramadhan, A., and Edelman, A. (2020). Universal differential equations for scientific machine learning.

Reich, P. B. (2012). Key canopy traits drive forest productivity. *Proceedings of the Royal Society B: Biological Sciences*, 279(1736):2128–2134.

Roff, A. D. A. (1974). Spatial heterogeneity and the persistence of populations. *Oecologia*, 15(3):245–258.

Running, S. W., Nemani, R. R., Peterson, D. L., Band, L. E., Potts, D. F., Pierce, L. L., and Spanner, M. A. (1989). Mapping regional forest evapotranspiration and photosynthesis by coupling satellite data with ecosystem simulation. *Ecology*, 70(4):1090–1101.

Russell, G., Jarvis, P., and Monteith, J. (1989). *Absorption of radiation by canopies and stand growth*, page 21–40. Society for Experimental Biology Seminar Series. Cambridge University Press.

Ryan, M. G. (2010). Temperature and tree growth. *Tree Physiology*, 30(6):667–668.

Sampson, D. A., Albaugh, T. J., Johnsen, K. H., Allen, H. L., and Zarnoch, S. J. (2003). Monthly leaf area index estimates from point-in-time measurements and needle phenology for Pinus taeda. *Canadian Journal of Forest Research*, 33(12):2477–2490.

Savoy, P. and Mackay, D. S. (2015). Modeling the seasonal dynamics of leaf area index based on environmental constraints to canopy development. *Agricultural and Forest Meteorology*, 200:46–56.

Soetaert, K., Petzoldt, T., and Setzer, R. W. (2010). Solving differential equations in R: Package deSolve. *Journal of Statistical Software*, 33(9):1–25.

Spanner, M. A., Pierce, L. L., Peterson, D. L., and Running, S. W. (1990). Remote sensing of temperate coniferous forest leaf area index the influence of canopy closure, understory vegetation and background reflectance. *International Journal of Remote Sensing*, 11(1):95–111.

Stenberg, P., Kuuluvainen, T., Kellomäki, S., Grace, J. C., Jokela, E. J., and Gholz, H. L. (1994). Crown Structure, Light Interception and Productivity of Pine Trees and Stands. *Ecological Bulletins*, (43):pp. 20–34.

Switzer, G. L., Nelson, L. E., and Smith, W. H. (1966). The characterization of dry matter and nitrogen accumulation by loblolly pine (pinus taeda l.). *Soil Science Society of America Journal*, 30(1):114–119.

Teskey, R. O., Bongarten, B. C., Cregg, B. M., Dougherty, P. M., and Hennessey, T. C. (1987). Physiology and genetics of tree growth response to moisture and temperature stress: an examination of the characteristics of loblolly pine (Pinus taeda L.). *Tree Physiology*, 3(1):41–61.

Vose, J. M. and Allen, H. L. (1988). Leaf area, stemwood growth, and nutrition relationships in loblolly pine. *Forest Science*, 34(3):547–563.

Vose, J. M., Dougherty, P. M., Long, J. N., Smith, F. W., Gholz, H. L., and Curran Vose, P. J. (1994). Factors influencing the amount and distribution of leaf area of pine stands. *Ecological Bulletins*, 43:102–114.

Vose, J. M. and Swank, W. T. (1990). A conceptual model of forest growth emphasizing stand leaf area. In Dixon, R. K., Meldahl, R. S., Ruark, G. A., and Warren, W. G., editors, *Process Modeling of Forest Growth Responses to Environmental Stress*, chapter 24, pages 278–287. Timber Press, Inc., Portland, OR.

Waring, R. H., Emmingham, W. H., Gholz, H. L., and Grier, C. C. (1978). Variation in maximum leaf area of coniferous forests in Oregon and its ecological significance. *Forest Science*, 24(1):131–140.

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

Will, R. E., Narahari, N. V., Shiver, B. D., and Teskey, R. O. (2005). Effects of planting density on canopy dynamics and stem growth for intensively managed loblolly pine stands. *Forest Ecology and Management*, 205(1-3):29–41.

Wright, E. M. (1955). A non-linear difference-differential equation. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 1955(194):66–87.

Zeide, B. (1993). Analysis of growth equations. *Forest Science*, 39(3):594–616.

Zhang, S., Allen, H. L., and Dougherty, P. M. (1997). Shoot and foliage growth phenology of loblolly pine trees as affected by nitrogen fertilization. *Canadian Journal of Forest Research*, 27(9):1420–1426.

Zhao, D., Kane, M., and Borders, B. E. (2010). Development and applications of the relative spacing model for loblolly pine plantations. *Forest Ecology and Management*, 259(10):1922–1929.

Zhao, D., Kane, M., and Borders, B. E. (2011). Growth responses to planting density and management intensity in loblolly pine plantations in the southeastern USA Lower Coastal Plain. *Annals of Forest Science*, 68(3):625–635.

Zhao, D., Kane, M., and Harrison, W. M. (2008). SAGS culture/density study: results through age 10. *PMRC Technical Report*, 3:33.

3.7 Tables and Figures

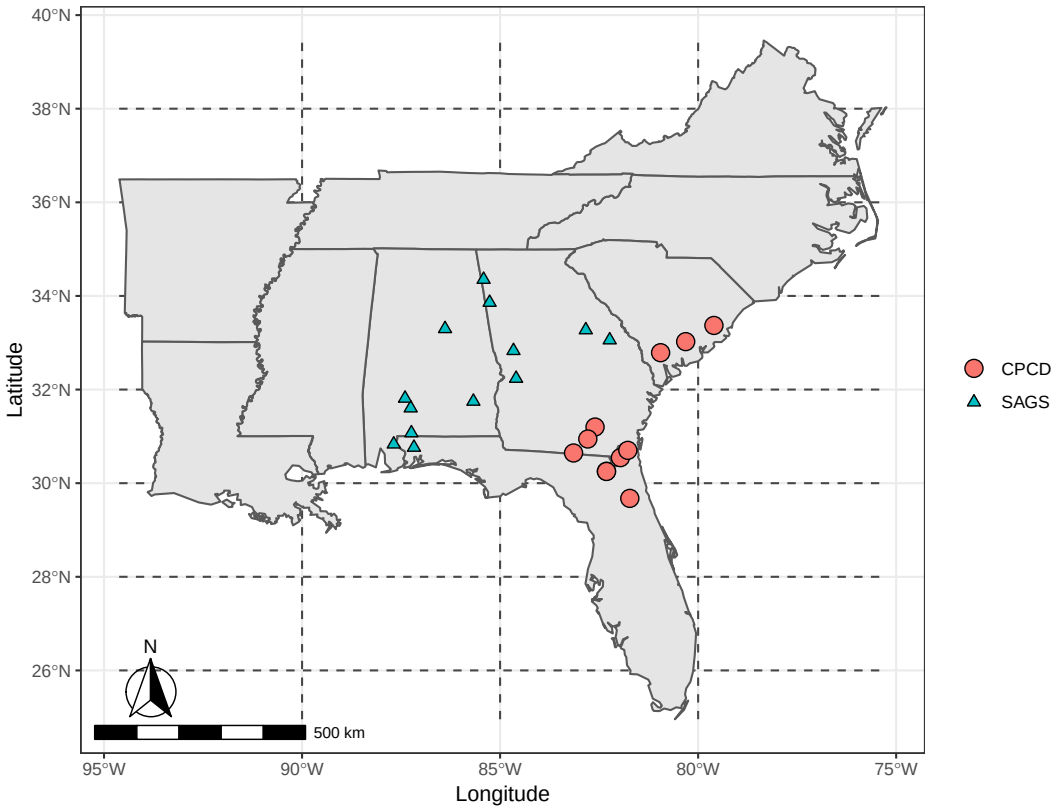


Figure 3.1: Site locations for the Plantation Management Research Cooperative's Coastal Plain Culture Density (CPCD) and South Atlantic Gulf Slopes (SAGS) installations used in analysis.

Table 3.1: Global parameter estimates and associated uncertainty for the base leaf area index model (Equation 3.1) fitted to 24 plots across the southeastern United States. Root mean square error (RMSE) was calculated across all observations.

Parameter	Estimate	Standard Error	Model Form	Environmental Variables
r	1.0750	0.0014		
τ	0.0002	0.0012		
K	4.7760	0.0016		
A	3.1199	0.0036	Base	N/A
ω	6.2933	0.0012		
μ	−1.8875	0.0016		
h	0.0961	0.0004		
RMSE	0.4427			
AIC	12 496.7			

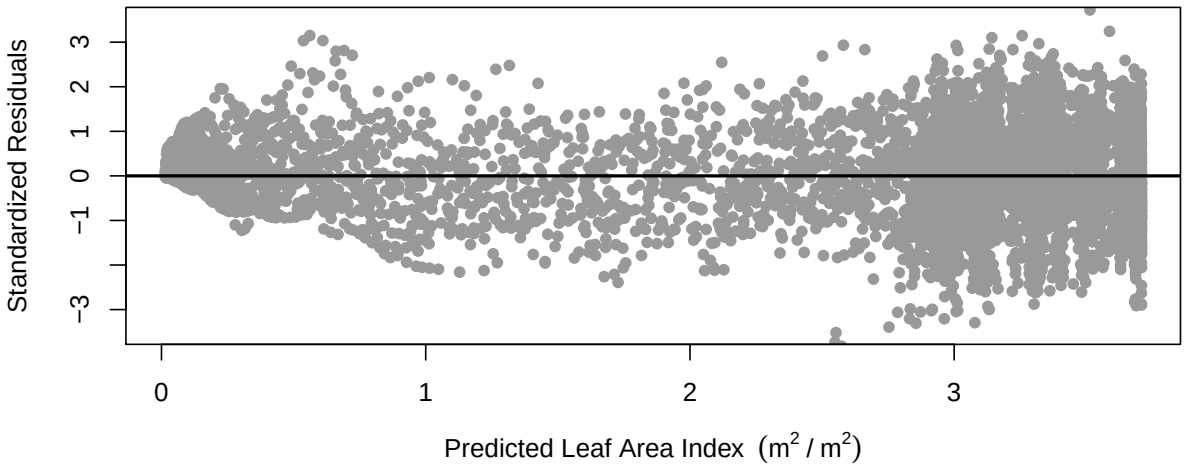


Figure 3.2: Studentized residuals for the base model.

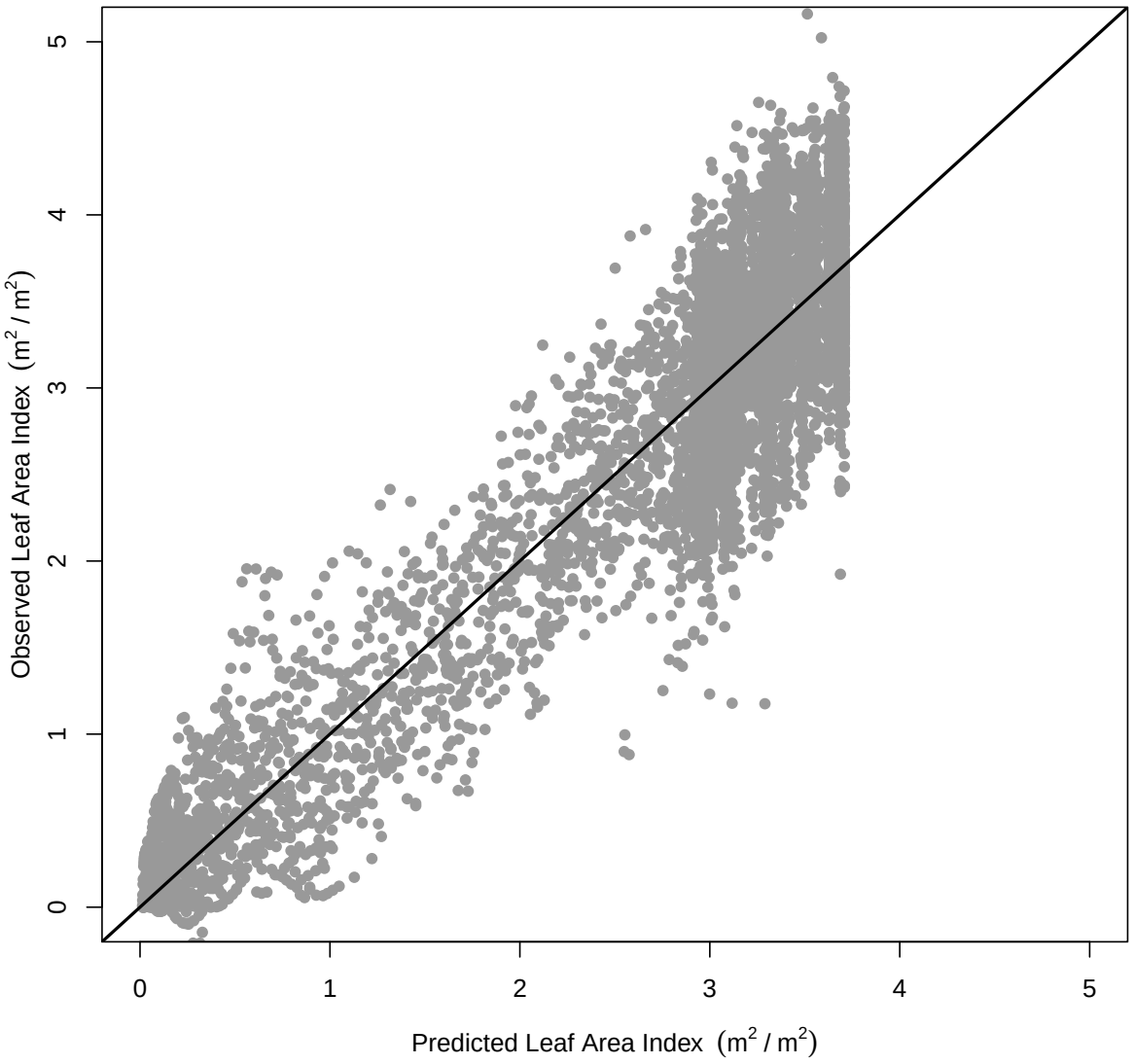


Figure 3.3: Observations vs predictions for base model. The abrupt termination of predicted leaf area index values is due to the model reaching the estimated carrying capacity.

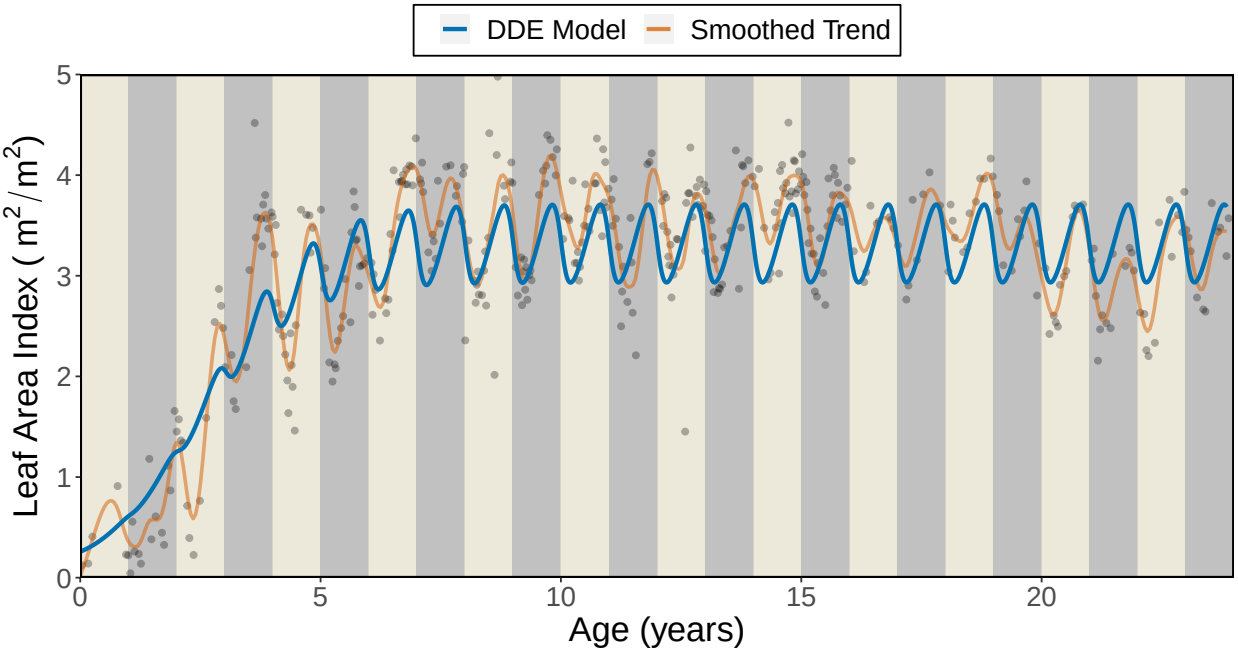


Figure 3.4: Results for the base model compared to the smoothed trend and observations. Root mean square error for this observed plot was 0.3398.

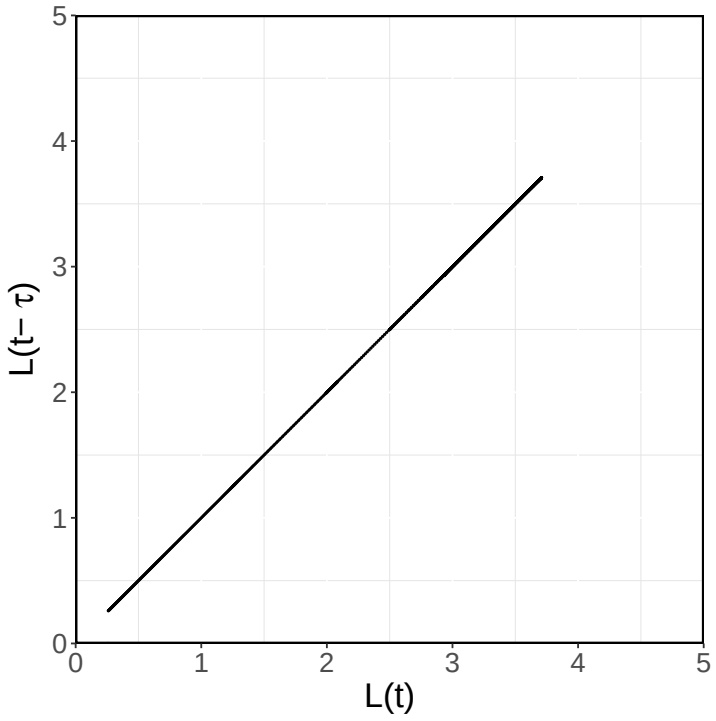


Figure 3.5: Results for the base model lag history trend. The x-axis shows the current state L(t) and its relationship with the historical state on the y-axis defined by τ. State units are LAI m^2/m^2.

Table 3.2: Global parameter estimates and associated uncertainty for the base leaf area index model (Equation 3.1) with monthly maximum temperature (°C) as a beta modifier fitted to 24 plots across the southeastern United States. Root mean square error (RMSE) was calculated across all observations.

Parameter	Estimate	Standard Error	Model Form	Environmental Variables
r	1.1109	0.0006		
τ	0.1148	0.0004		
K	4.5565	0.0010		
A	3.0997	0.0005	Base	N/A
ω	6.2948	0.0012		
μ	−1.8663	0.0009		
h	0.0885	0.0006		
μ_2	−5.7533	0.0009		
T_b	1.4255	0.0005		Monthly
α	1.4832	0.0006	Beta	Maximum
T_c	36.7020	0.0006		Temperature
β	0.3782	0.0011		
RMSE	0.4299			
AIC	11 891.71			

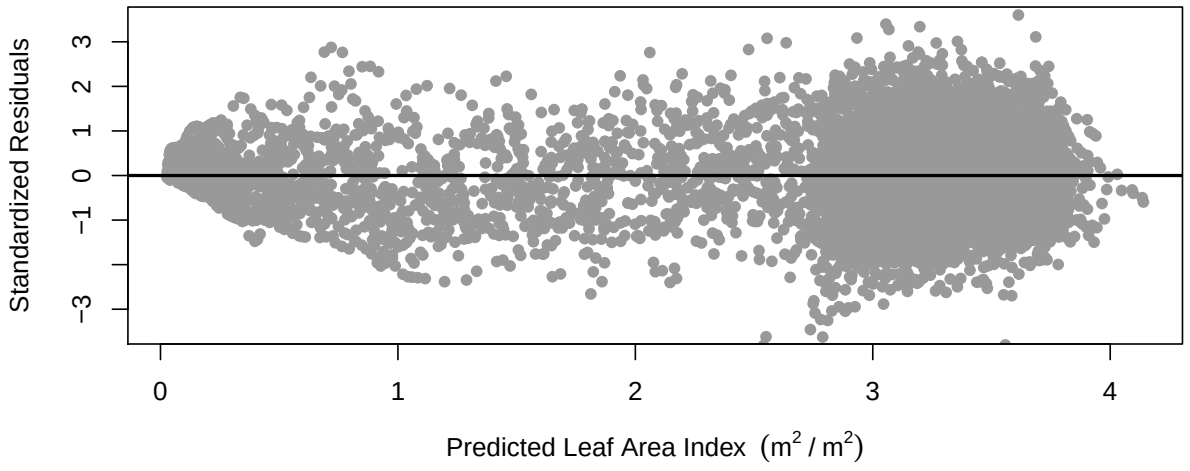


Figure 3.6: Standardized residuals for the model parameterized in Table 3.2 using monthly maximum temperature (°C) in beta form as a modifier.

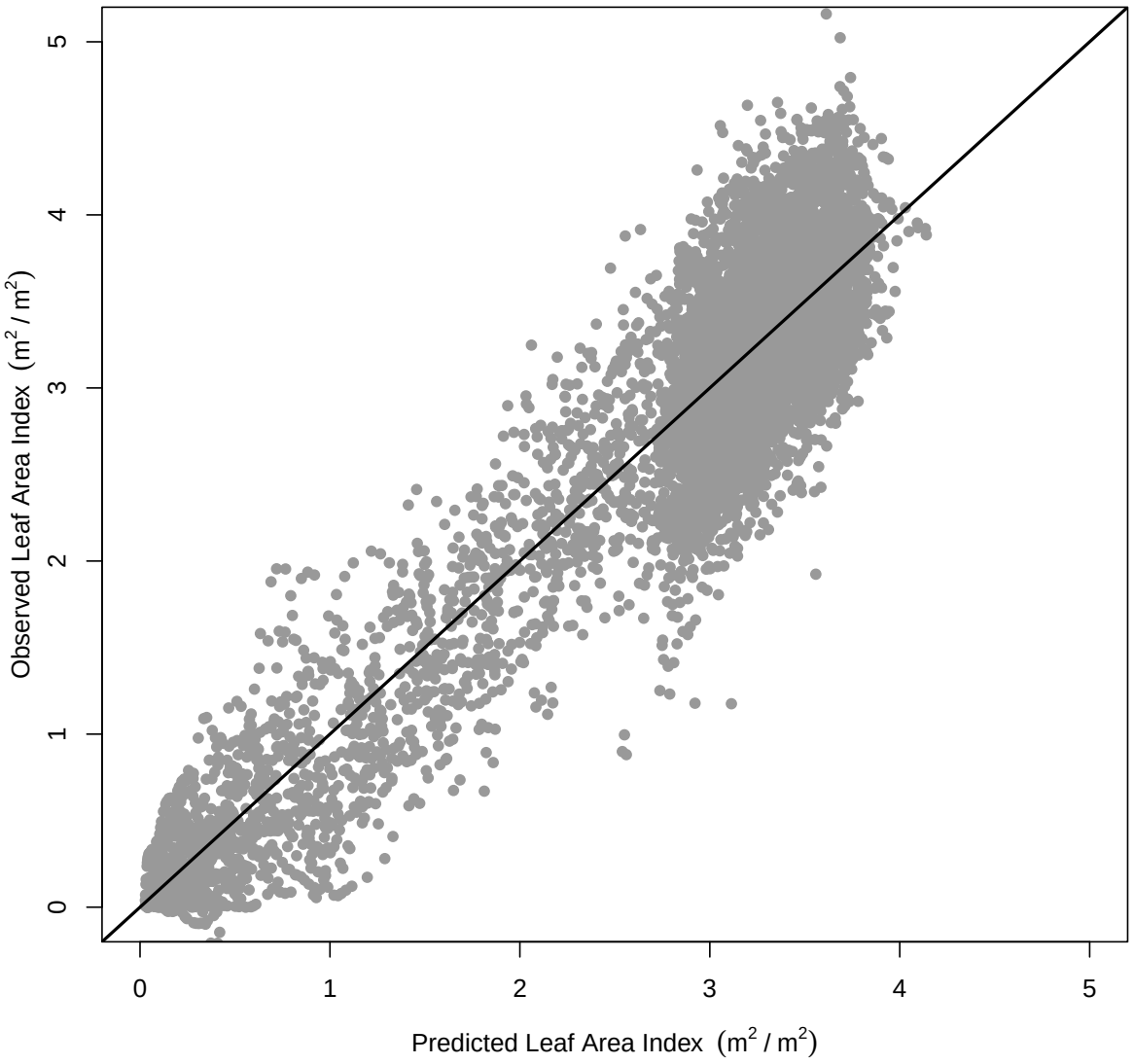


Figure 3.7: Observations vs. predictions for the model parameterized in Table 3.2 using monthly maximum temperature (°C) in beta form as a modifier.

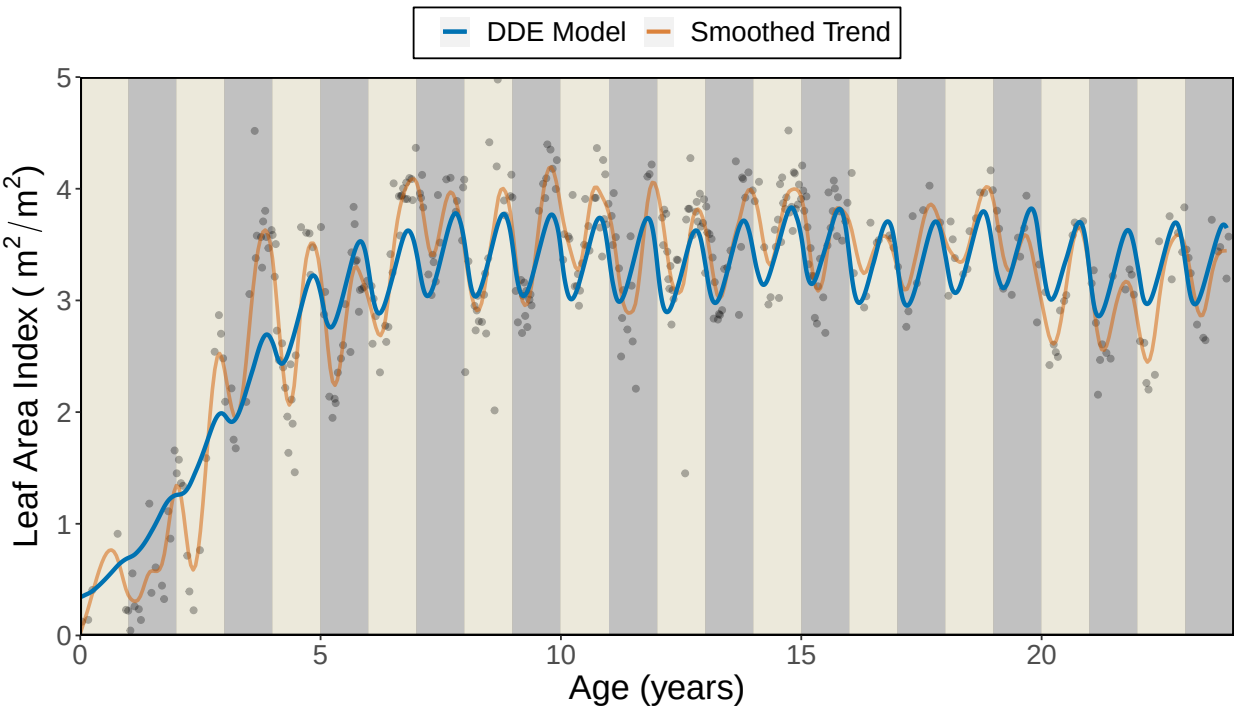


Figure 3.8: Results for the model parameterized in Table 3.2 using monthly maximum temperature (°C) in beta form as a modifier compared to the smoothed trend and observations. Root mean square error for this plot was 0.3171.

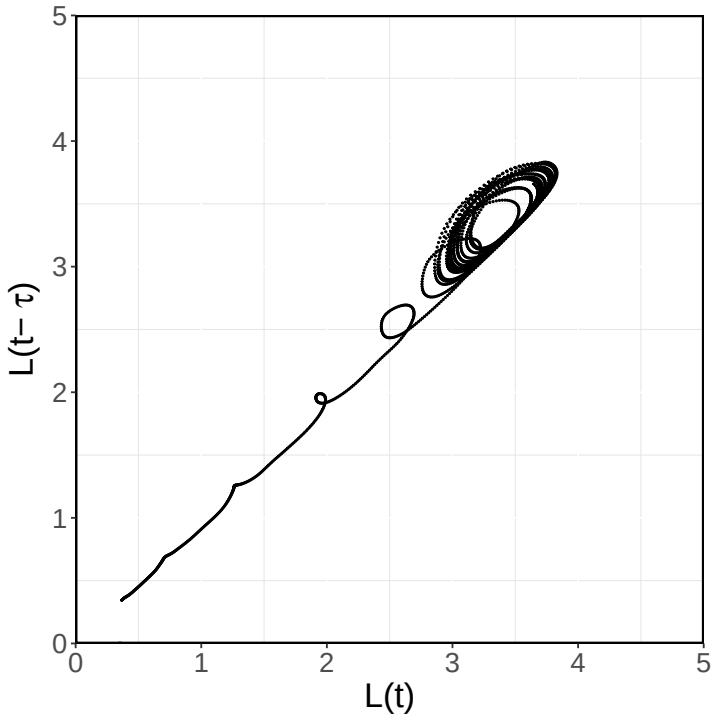


Figure 3.9: Results for the maximum monthly temperature model lag history trend. The x-axis shows the current state L(t) and its relationship with the historical state on the y-axis defined by τ. State units are LAI m^2/m^2.

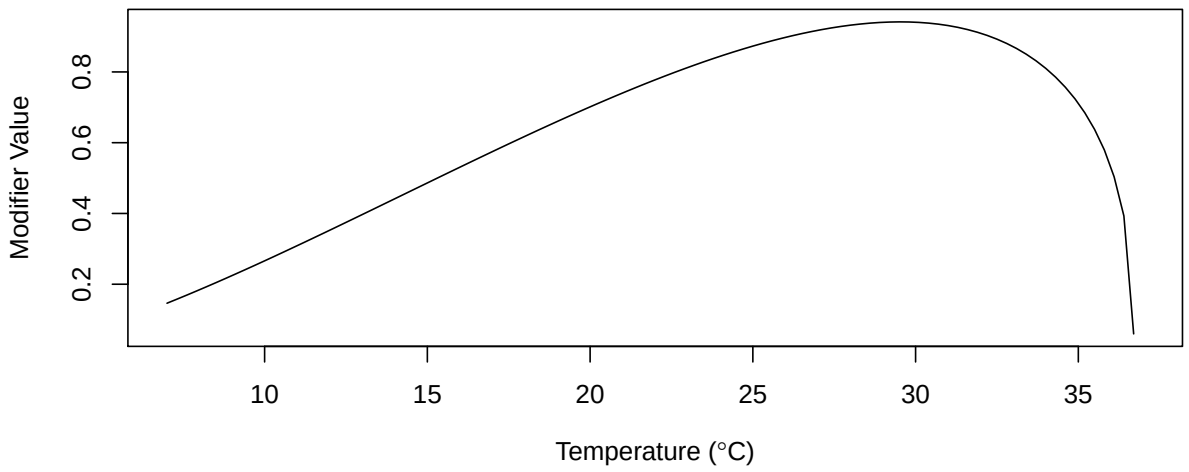


Figure 3.10: Evaluation of the parameterized monthly maximum temperature beta modifier

Table 3.3: Global parameter estimates and associated uncertainty for the base leaf area index model (Equation 3.1) fitted to 24 plots across the southeastern United States with the monthly maximum temperature (°C) included as a beta function and monthly excess water (mm) included as a double logistic function. Root mean square error (RMSE) was calculated across all observations.

Parameter	Estimate	SE	Model Form	Environmental Variable
r	1.1175	0.0060	Base	N/A
τ	0.1148	0.0011		
K	4.5535	0.0073		
A	3.1052	0.0004		
ω	6.2973	0.0061		
μ	−1.8661	0.0457		
h	0.088 48	0.0024		
μ_2	−5.5757	0.0876	Beta	Monthly Maximum Temperature
T_b	1.4248	0.1920		
α	1.5219	0.0185		
T_c	37.018 63	0.2095		
β	0.3799	0.0045		
α	1.5082	0.0007	Double Logistic	Monthly Excess Water
β	0.0067	0.0004		
δ	0.0072	0.0004		
T_b	−0.0005	0.000 04		
T_c	320.0497	0.0359		
RMSE	0.4257			
AIC	11 684.19			

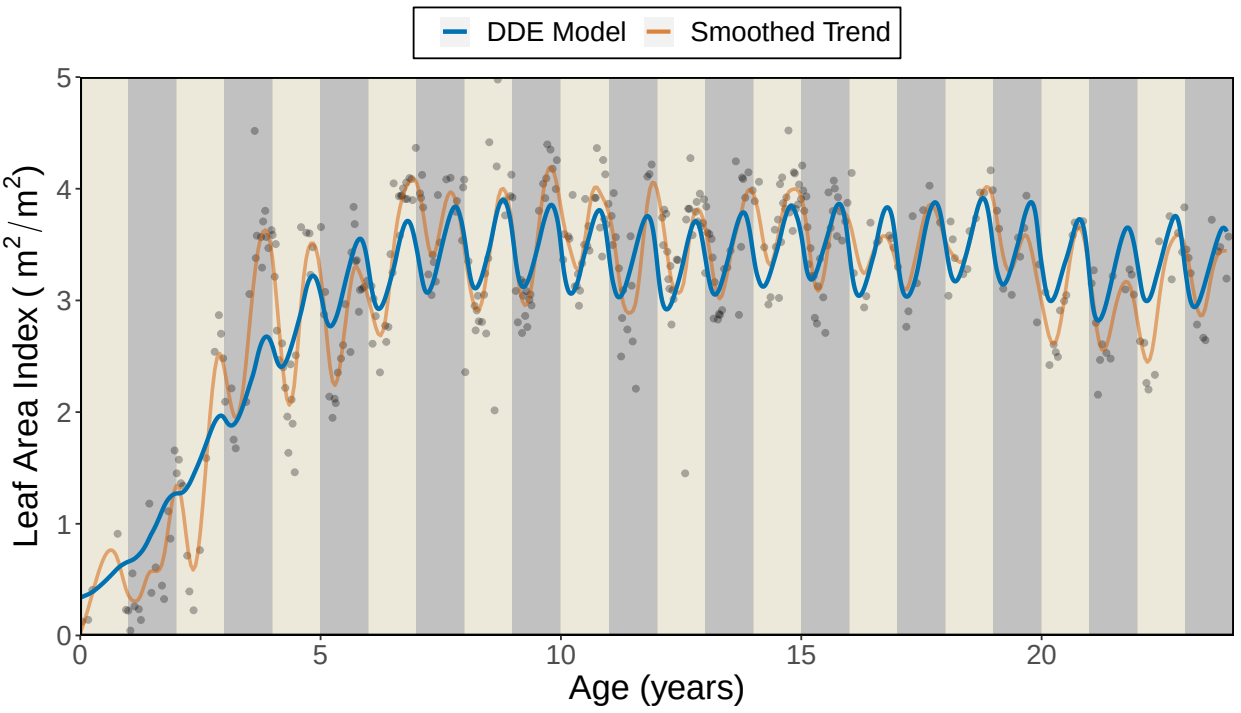


Figure 3.11: Results for the model parameterized in Table 3.3 using monthly maximum temperature (°C) in beta form and monthly excess water in double logistic form as modifiers compared to the smoothed trend and observations. Root mean square error for this plot was 0.3096.

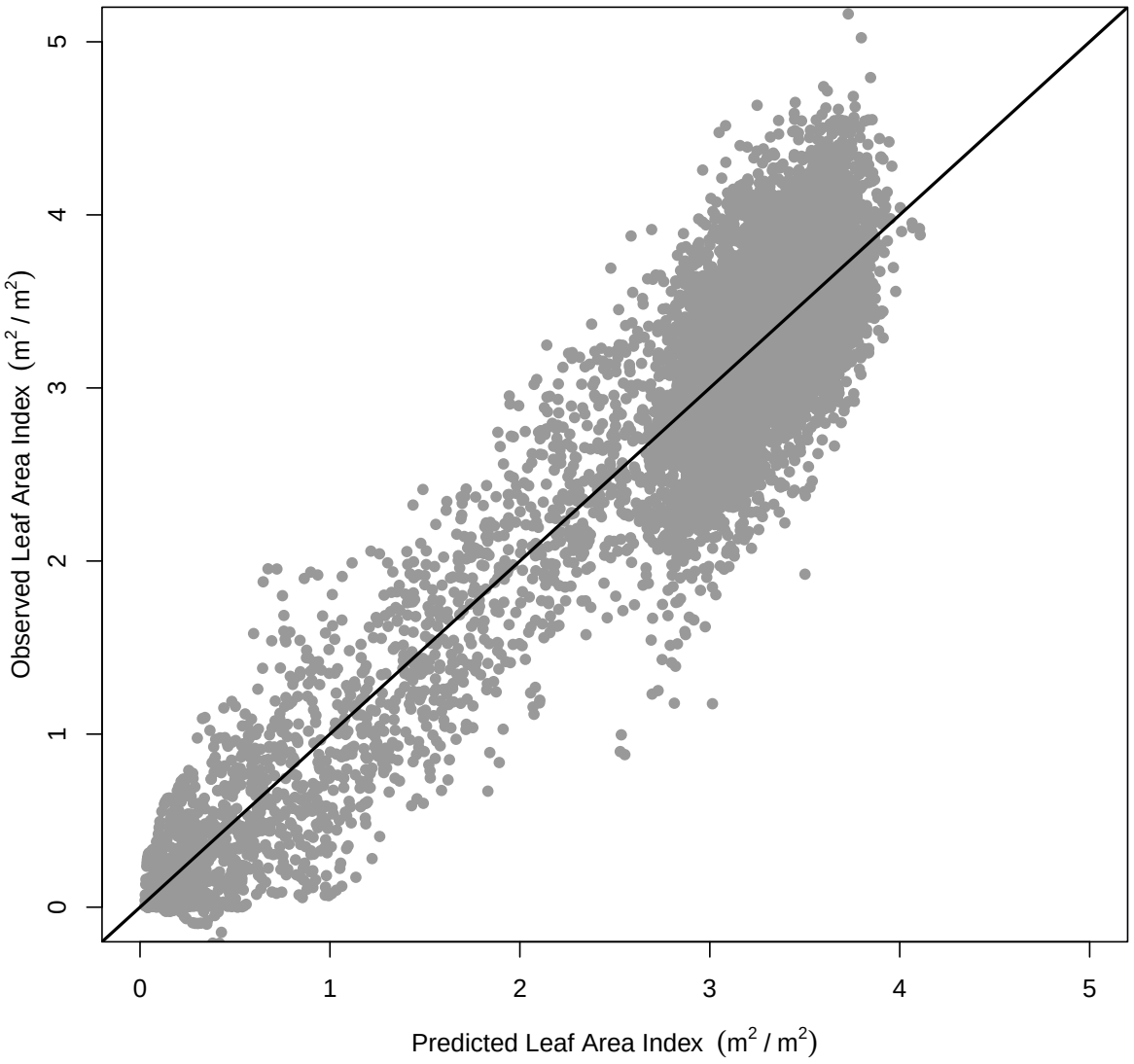


Figure 3.12: Observations vs. predictions for the model parameterized in Table 3.3 using monthly maximum temperature (°C) in logistic form and monthly excess water (mm) in double logistic form as modifiers.

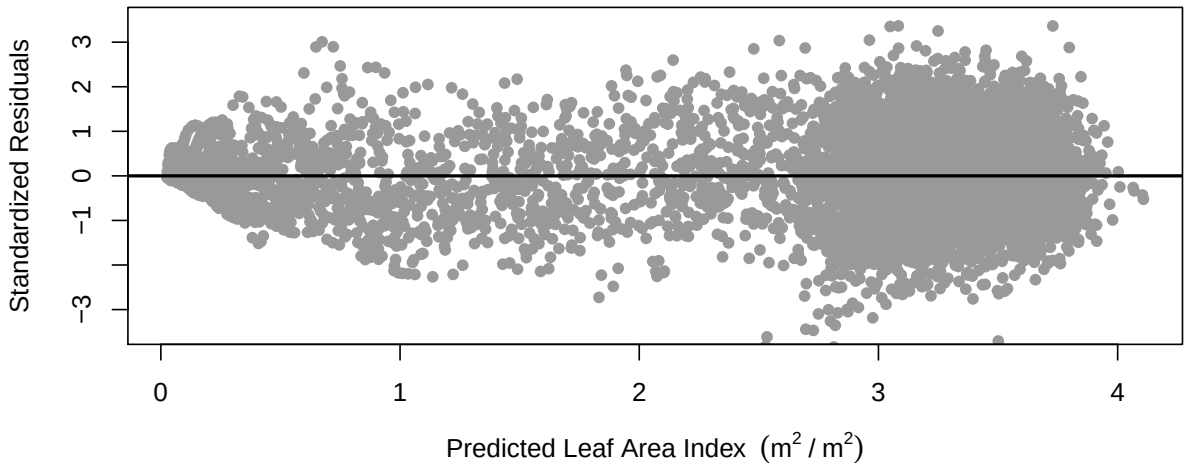


Figure 3.13: Results for the model parameterized in Table 3.3 using monthly maximum temperature (°C) in beta form and monthly excess water in double logistic form as modifiers compared to the smoothed trend and observations.

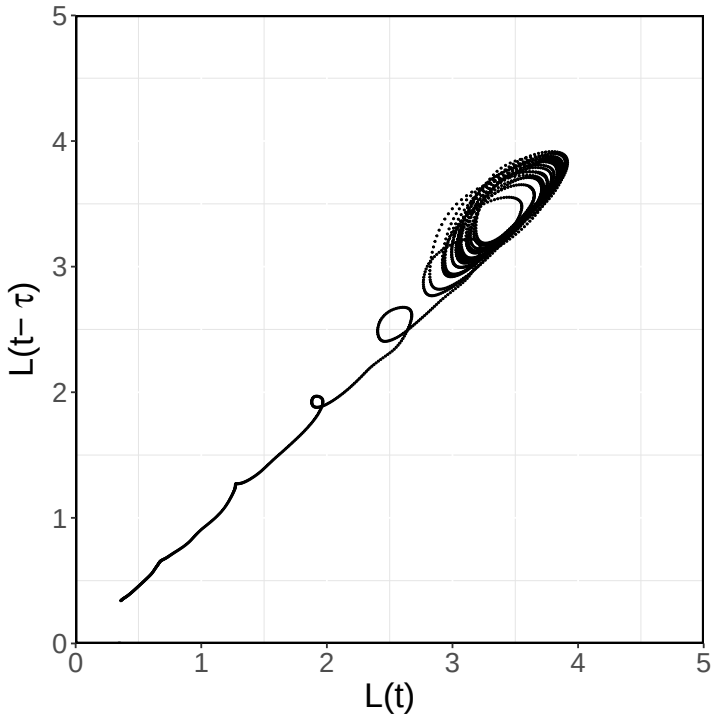
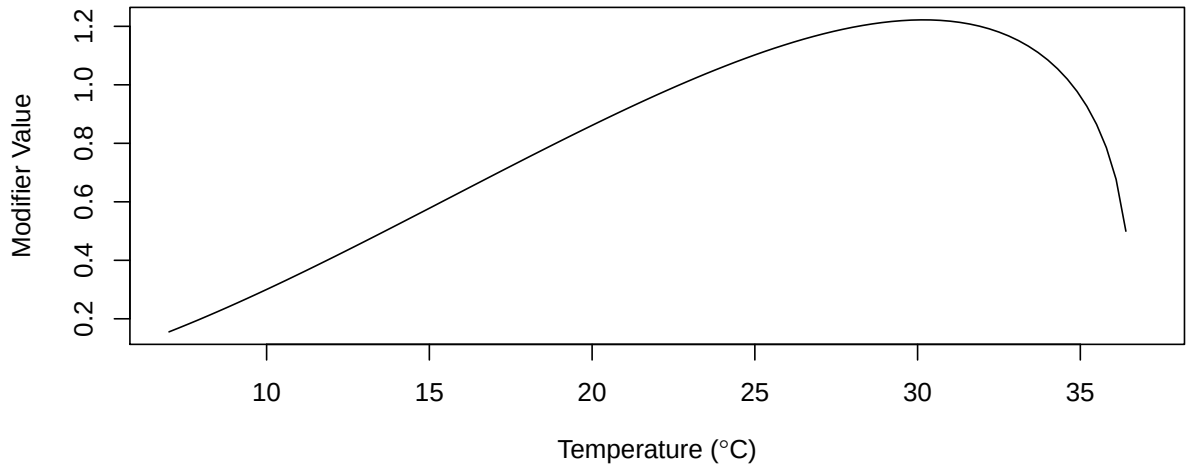


Figure 3.14: Results for the maximum monthly temperature and excess water interaction model lag history trend. The x-axis shows the current state L(t) and its relationship with the historical state on the y-axis defined by τ. State units are LAI m^2/m^2.

A)



B)

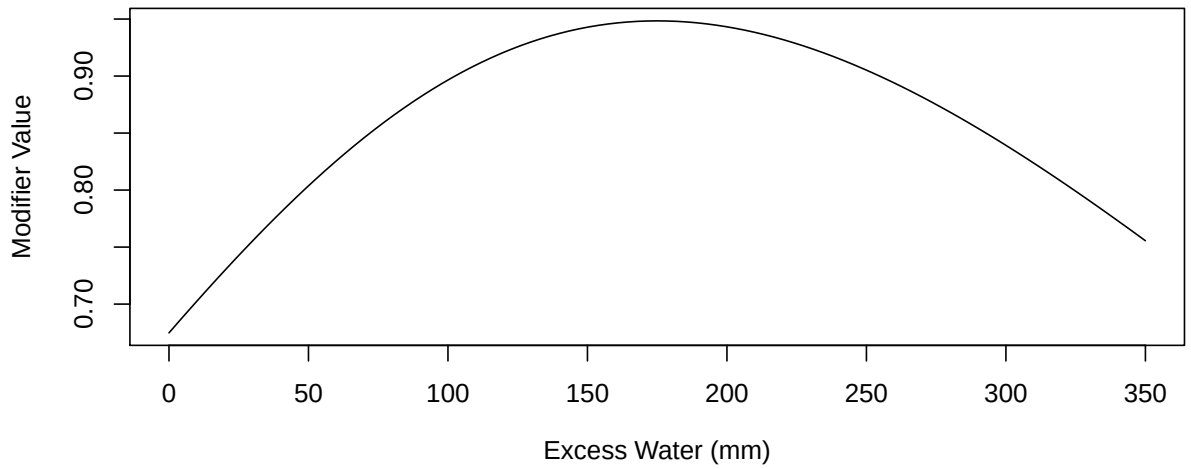


Figure 3.15: Evaluation of the parameterized modifiers for A) monthly maximum temperature (°C) in beta form and B) monthly excess water (mm) double logistic form across the ranges of observed values for the individual variable.

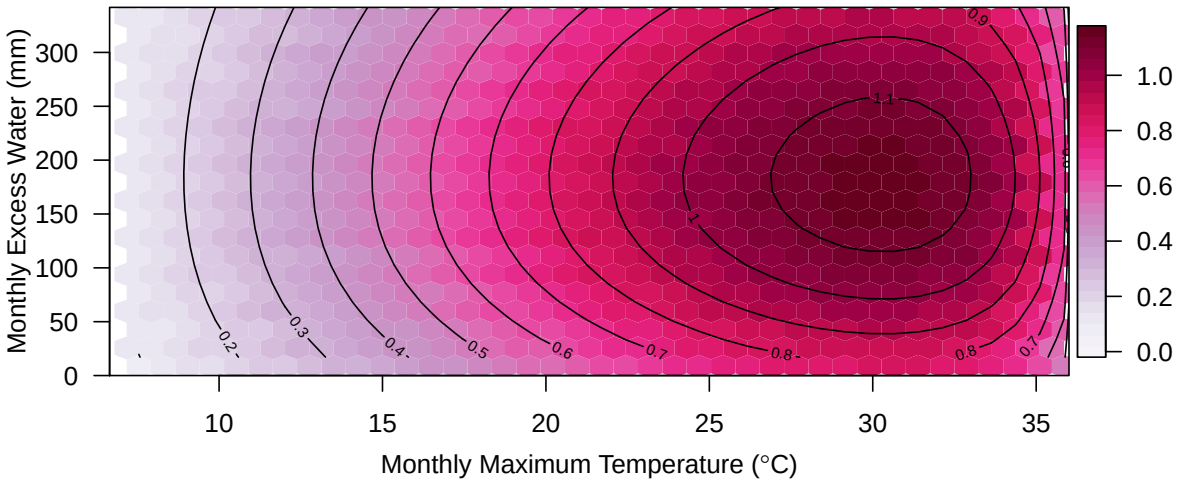


Figure 3.16: Resulting interaction modifier between the monthly maximum temperature beta modifier and the monthly excess water double logistic modifier across the ranges of observed values for the 24 study plot across the southeastern United States.

Table 3.4: Global parameter estimates and associated uncertainty for the leaf area index model (Equation 3.1) fitted to 24 plots across the southeastern United States with the environmental variables included as a beta and double logistic function, and a local carrying capacity parameter (K). Root mean square error (RMSE) was calculated across all observations.

Parameter	Estimate	SE	Model Form	Environmental Variable
r	1.1968	0.0181	Base	N/A
τ	0.1143	0.0080		
A	2.6254	0.0237		
ω	6.2877	0.0009		
μ	−1.7420	0.0062		
h	0.0647	0.0041		
μ_2	−6.0359	0.0324	Beta	Monthly Maximum Temperature
T_b	1.4592	0.0093		
α	1.5984	0.0081		
T_c	44.69	0.0197		
β	0.3419	0.0103		
α	1.5984	0.0081	Double Logistic	Monthly Excess Water
β	0.0063	0.0004		
δ	0.0071	0.0006		
T_b	−0.0004	0.000 04		
T_c	276.3619	0.0166		
RMSE	0.3802			
AIC	9383.04			

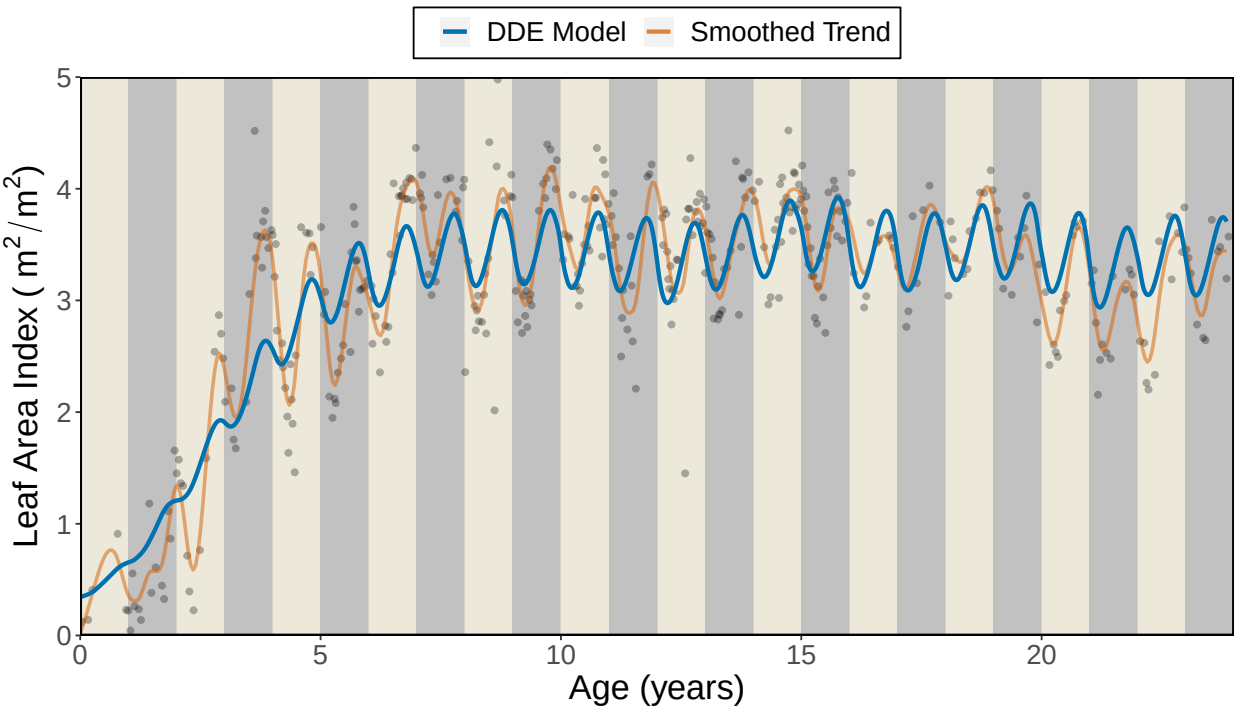


Figure 3.17: Results for the model parameterized in Table 3.3 using monthly maximum temperature (°C) in beta form, monthly excess water in double logistic form as modifiers and a local LAI carrying capacity parameter compared to the smoothed trend and observations. Root mean square error for this plot was 0.3096.

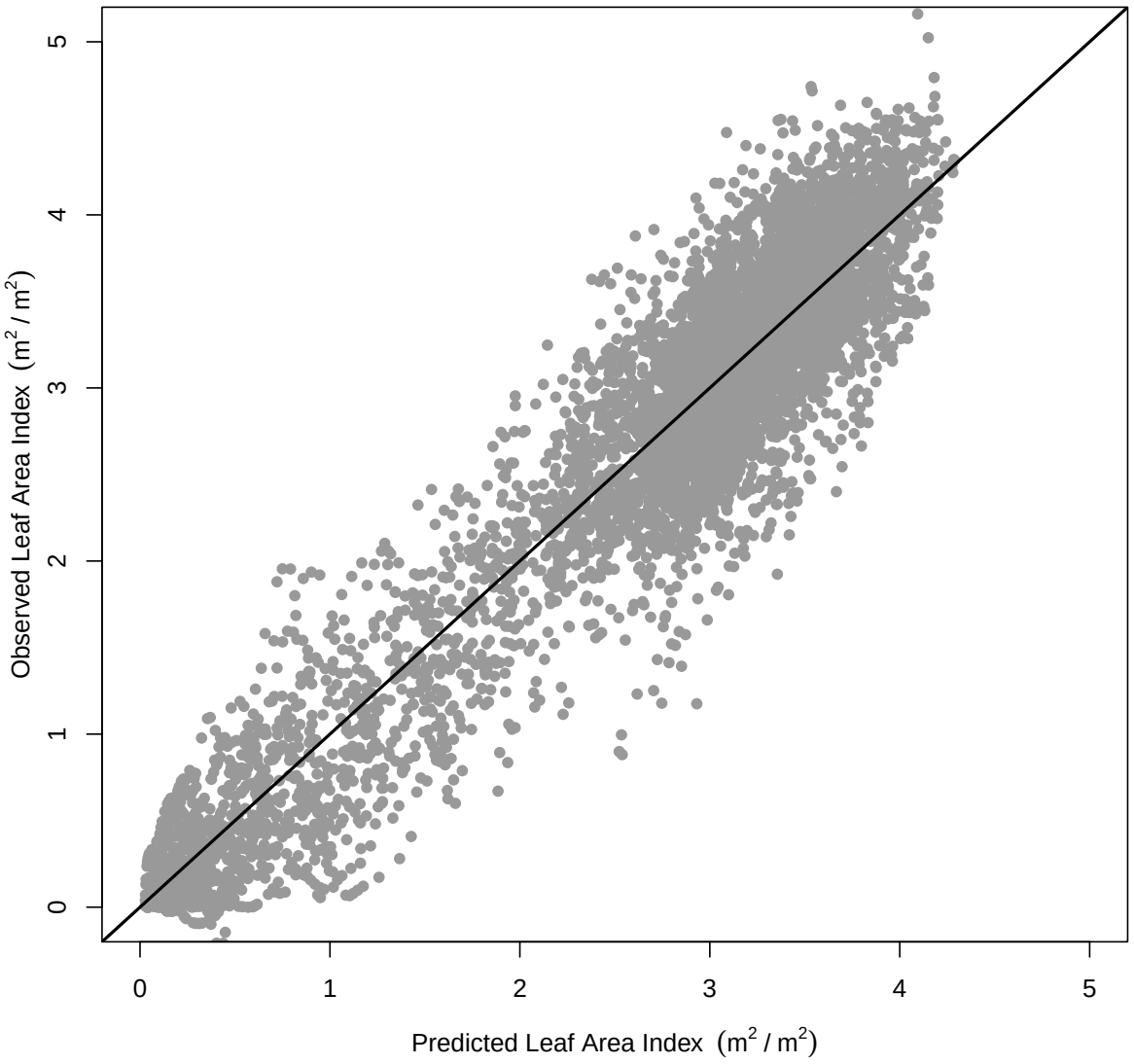


Figure 3.18: Observations vs. predictions for the model parameterized in Table 3.3 using monthly maximum temperature (°C) in logistic form, monthly excess water (mm) in double logistic form as modifiers and a local LAI carrying capacity parameter.

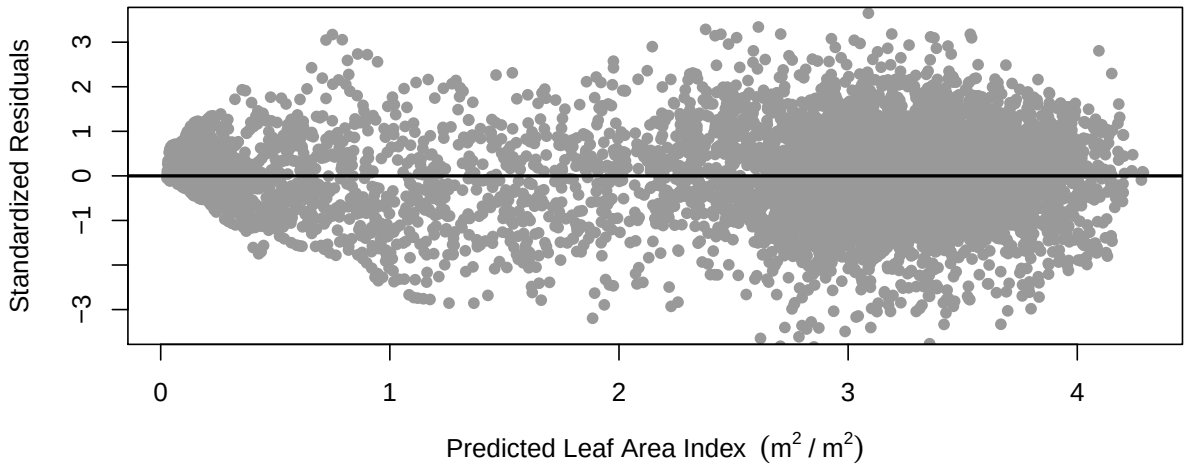


Figure 3.19: Results for the model parameterized in Table 3.3 using monthly maximum temperature (°C) in beta form and monthly excess water in double logistic form as modifiers and a local LAI carrying capacity parameter compared to the smoothed trend and observations.

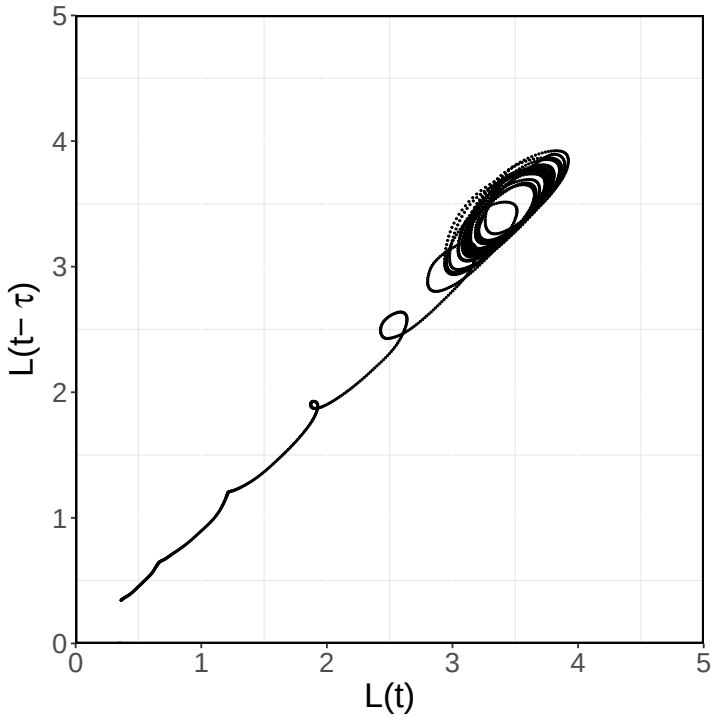


Figure 3.20: Results for the maximum monthly temperature and excess water interaction model with local LAI carrying capacity parameter lag history trend. The x-axis shows the current state L(t) and its relationship with the historical state on the y-axis defined by τ. State units are LAI m^2/m^2.

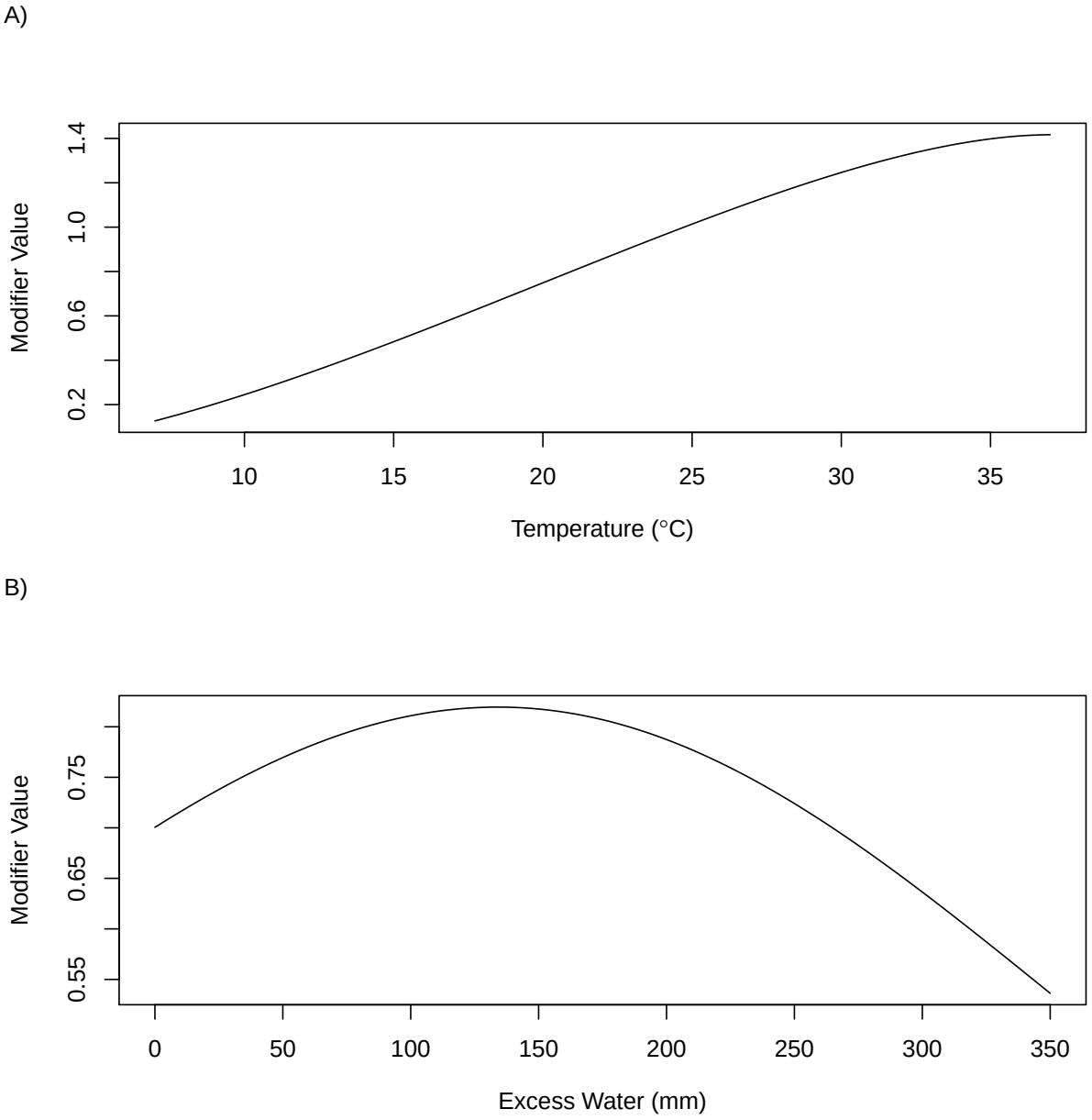


Figure 3.21: Evaluation of the parameterized modifiers for A) monthly maximum temperature (°C) in beta form and B) monthly excess water (mm) double logistic form across the ranges of observed values for the individual variable for the model fitted with a local LAI carrying capacity parameter.

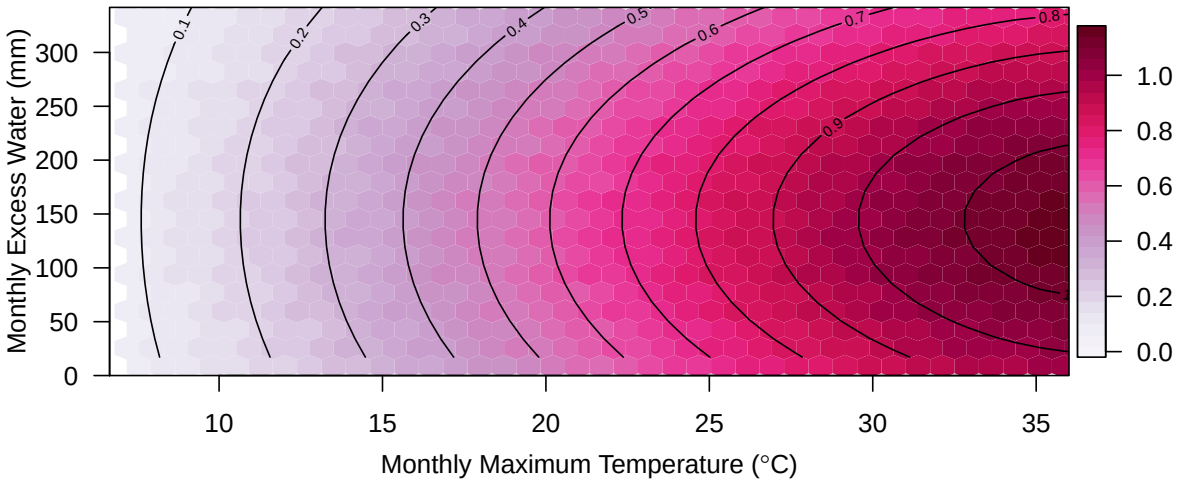


Figure 3.22: Resulting interaction modifier between the monthly maximum temperature beta modifier and the monthly excess water double logistic modifier across the ranges of observed values for the 24 study plot across the southeastern United States fitted with a local LAI carrying capacity parameter.

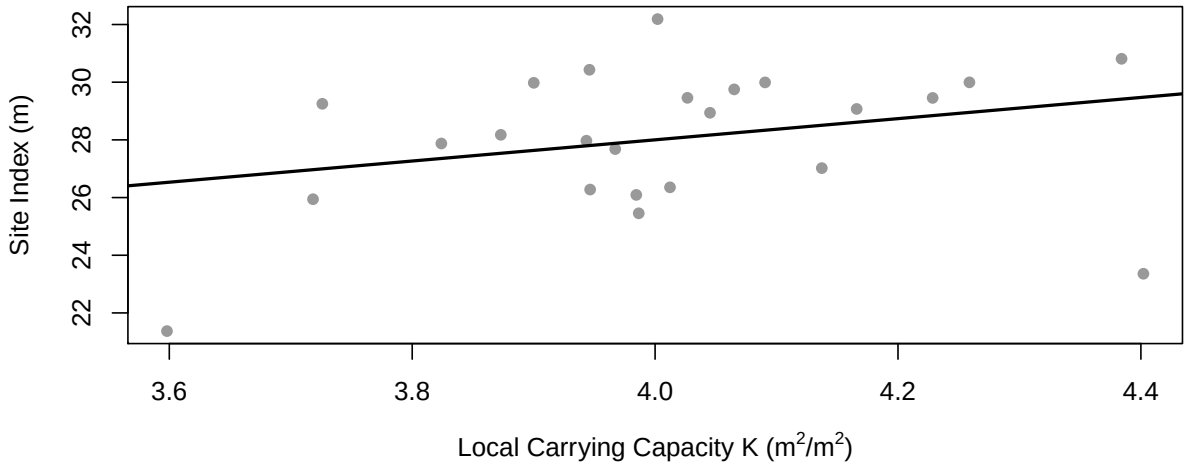


Figure 3.23: Comparison between the fitted local K LAI carrying capacity parameter and estimated site index (base age = 25) for the 24 plots.

Table 3.5: Root mean square error (RMSE) results from the k-fold (k=9) cross validations performed on the four models. Base indicates base model, Env1 indicates the inclusion of monthly maximum temperature as a beta function, Env2 indicates monthly excess water as a double logistic function, and Local K indicates the use of a local leaf area index carrying capacity parameter.

Model	Training			Validation		
	Min	Max	Avg	Min	Max	Avg
Base	0.4270	0.4544	0.4422	0.3161	0.5508	0.4367
Base + Env1	0.4056	0.4489	0.4285	0.2885	0.5210	0.4283
Base + Env1 + Env2	0.4082	0.4419	0.4265	0.3209	0.5801	0.4319
Base + Env1 + Env2 + Local K	0.3658	0.3886	0.3796	0.3237	0.4709	0.3819

Chapter 4

A Bio-physically Driven Dominant Height Model for Loblolly Pine Plantations[1]

[1]Kinane, S.M., M. Zapata, B.P. Bullock, R.L. Cook, D. Mishra, and C.R. Montes. To be submitted to *Forests*.

Abstract

Dominant height models are commonly used to serve as the productivity indicator in growth and yield systems when defined at a base age. Many sources of variation can influence dominant height estimates, leading to over or under estimates in long term projections. One of the main sources of variability in productivity of loblolly pine plantations is the variation in light interception. To account for that source of variation, the objective of this research was to model dominant height as a function of leaf area index. Results showed that incorporating estimated annual radiation intercepted using a leaf area index growth model improved the performance of a dominant height growth model. By adjusting the asymptotic parameter of a dominant height model as a function of predicted peak leaf area index and estimated annual intercepted light, annual dominant height growth was regulated. Sites with higher carrying capacities for leaf area index were shown to have higher dominants heights than sites with lower carrying capacities.

4.1 Introduction

Traditionally, forest growth for loblolly pine (*Pinus taeda* L.) in the southeastern United States has been modeled through empirically derived and regionally parameterized stand level equations to estimate attributes such as height, density, volume, and basal area (Landsberg and Sands, 2011). These models are formulated using the observed relationships between time and the independent variable to explain forest growth. Dominant height is the primary attribute for establishing stand level productivity at an individual site in these systems, defined at a base age and typically referred to as site index. These site index driven models of forest growth rely on the relative growth rate of dominant height over time to govern growth rates of the remaining equations due to the observations of correlation between dominant height and site productivity and its reported invariance to stand density. A variety of empirical modeling approaches have been used to predict the dominant height - age relationship of stands, ranging from three parameter von Bertalanffy equations to more complex generalized algebraic difference approaches. The definition of dominant height varies regionally and results can influence model predictions and overall assumptions of model responses. The calculated site index serves as the integration of all biotic and abiotic forces acting upon the stand throughout its lifetime rather than explicit individual inputs.

In contrast to the empirically derived models, processed-based or mechanistic models have been created in a more theoretical sense; to model tree and stand growth based on known biological phenomena and their responses to environmental processes rather than relying on time (Landsberg and Waring, 1997; Landsberg and Sands, 2011). Generally composed of many functional submodels that define physiological processes and their interactions at a range of hierarchies, their required inputs include a variety of environmental variables to drive the biological responses that are eventually summarized into measurable forest growth (Makela et al., 2000). These models are heavily parameterized and require a plethora of data and assumptions to function.

To bridge the gap between the traditional empirical equations and the complex processed based models, hybrid models have been invoked to provide more biologically sound equations that are simpler in nature as compared to the processed based models but do have biological interpretability in contrast with some empirically based models. A strong, well-accepted relationship potentially defining forest productivity has been shown to be the correlation between annual volume increase and the peak leaf area index observed (Albaugh et al., 1998). Leaf area index, the one-sided summation of leaf surface area over a fixed ground area, represents the surface in which energy is exchanged in the environment. Peak leaf area index, the highest attained leaf area index on an annual basis, represents the maximum amount of leaf surface area available to absorb incoming photosynthetically active radiation (PAR). Consistently, variability in forest productivity has been attributed to variation in radiation interception, which is a function of leaf area, stand structure, foliage orientation and dispersion (Russell et al., 1989; Vose and Swank, 1990; Tang et al., 2019).

The objective of this research was to determine the effect of intercepted radiation as a function of leaf area index on the growth and development of dominant height in loblolly pine plantations across the southeastern United States. The two main sub-objectives of this research were to derive a dominant height model that depends on intercepted radiation for height growth and simultaneously fit a leaf area index and dominant height model.

4.2 Methods

4.2.1 Data

Data from the Plantation Management Research Cooperative's (PMRC) Coastal Plain Culture Density (CPCD) and South Atlantic Gulf Slopes (SAGS) studies were selected for this analysis. The CPCD study was installed in 1995 and 1996 as a split-plot design to study the effects of two cultural regimes (main plot) across a range of planting densities (subplot) in seventeen installations throughout the lower coastal plain of Florida, Georgia, and South

Carolina (Zhao et al., 2011). The cultural regimes for the CPCD study included: 1) operational, consisting of bedding and banded chemical site preparation, herbaceous weed control after the first growing season, and fertilization at planting, before the 8^{th} growing season, and before the 12^{th} growing season and 2) intensive, consisting of bedding and broadcast chemical site preparation, tip moth control, complete vegetation control, and multiple fertilization treatments throughout the first twelve growing seasons (Zhao et al., 2011). The planting densities ranged from 741 to 4448 trees per hectare on study plots sized from 0.12 to 0.23 hectares to accommodate the minimum number of measurement trees. Further details for the CPCD study can be seen in (Harrison and Kane, 2008). Complementary to the CPCD, the SAGS study was installed in 1997 and 1998 at 23 installations throughout the Piedmont/Upper Coastal Plain physiographic region across five southeastern states (Zhao et al., 2010). Experimental design for the SAGS study was similar to the design of the CPCD study, testing the effects of cultural treatments (plot) across planting densities (subplot). Biennial field measurements for dominant height on both studies were collected between ages 2 and 12, with some additional measurements occurring annually between ages 13 and 18, followed by a three year re-measurement period. The age range available for dominant height measurements was 2-21, with dominant height being defined as the average height of the measured study trees with diameters greater than the plot average diameter Zhao et al. (2008). A total of 23 sites, 10 from the CPCD study and 13 from the SAGS study were selected for analysis (Figure 3.1). Only plots under the intensive treatment were selected to remove potential sources of variability associated from differing treatment responses.

Due to the range of subplot size in the CPCD and SAGS studies, subplots from the 741 trees per hectare planting density (3.66 m x 3.66 m spacing) receiving intensive treatments were selected to maximize the subplot size (0.23 hectares) to increase the geographic footprint for remotely sensed data acquisition. All available surface reflectance scenes from the Landsat 5 TM and Landsat 7 ETM+ sensors were queried in the Google Earth Engine platform from site establishment to end of year 2019, masked for clouds, and exported as individual band

values averaged for the subplot (Gorelick et al., 2017). Data were further filtered based upon cloud quality attributes, pixel quality attributes, and radiometric saturation. Using near-infrared (band 4) and short-wave infrared (band 5) bands, LAI was estimated using the model proposed in the previous chapter. The number of available scene observations per study plot ranged from 334 to 1005 from plot establishment to December 31, 2019. Additionally, erroneous and outlying plots within the first five years of development were removed to reduce the amount of noise observed in the time series. The time series of individual plots were shortened to remove any observations that included post-thinning or increased mortality to insects. To provide increased temporal consistency and determine the underlying LAI development trend and seasonality, a thin plate spline (TPS) using the fields package was used to smooth the observed data (Savoy and Mackay, 2015; Nychka et al., 2017). LAI time series for each plot were fitted using the TPS function with the cost function set to 1.2. Using the fitted TPS model, a new time series for each plot was created on a 0.05 year time step for the age range observed at each plot.

Environmental variables for the individual study sites were sourced from the University of East Anglia's Climate Research Unit (CRU) (Harris et al., 2020). Using CRU's TS v. 4.04 NetCDF product, monthly values at the individual study plots were acquired from establishment to December 2019. Water storage capacity for the first 1.5 meters of soil was estimated at the individual site using data from SSURGO spatial data base. Additional indices, including water deficit, water deficit index, excess water, and excess water index were calculated from the environmental variables in association with water storage capacity at the individual study plots. Water deficit for the site is calculated by subtracting the monthly precipitation from the monthly estimated potential evapotranspiration. Conversion of water deficit to water deficit index incorporates subtracting out the water storage capacity and summing the water deficits. Excess water and its index are calculated in similar fashion by subtracting the monthly potential evapotranspiration from the monthly precipitation. High

resolution daily solar radiation (W/m^2) data was sourced from Daymet and summarized to monthly and annual amounts for each plot (Thornton et al., 2016).

4.2.2 Model Formulation

To reflect the amount of energy captured by the stand, a LAI prediction model using a delayed differential equation framework and environmental modifiers was used to provide the trajectory of LAI growth and development (Equation 4.1).

$$\frac{dL}{dt} = \left[r * L_t * \left(1 - \frac{L_{t-\tau}}{K + A * \sin\left(\omega * t + \mu\right)} \right) - h * L_t \right] * f(ENV_1) * f(ENV_2) \tag{4.1}$$

Where r, τ, A, ω, μ, and h are global parameters to be estimated, K is a local parameter to be estimated at the individual plot level, $f(ENV_n)$ are environmental modifiers, and L is LAI. The environmental modifiers chosen were monthly maximum temperature (°C) and monthly excess water (mm) included as a beta and double logistic function, respectively. Using the predicted LAIs across the range of observations, peak LAI, the highest attained LAI value within the year, was identified and selected for all years under observation at the plot level. Using Beer's law, the estimated amount of intercepted radiation (IPAR) was calculated with the given peak LAI on an annual basis. The Chapman-Richards generalization of the von Bertalanffy model was chosen for dominant height growth (Equation 4.2).

$$\frac{dH}{dt} = \eta * H^m - \kappa * H \tag{4.2}$$

Where η, m, and κ are estimated parameters. To test the effect of light interception on dominant height growth, radiation interception, both total intercepted photosynthetically active radiation (IPAR) and the fraction intercepted photosynthetically active radiation (FPAR) were evaluated on how they improved model fit by incorporating them as an additional component in either the growth rate parameter ($m + \beta_1 * PAR$) or asymptote parameter ($\eta + \beta_1 * PAR$). Converting the component to an interactive term was also evaluated.

4.2.3 Model Parameterization

The LAI and dominant height models were parameterized simultaneously using maximum likelihood estimation. To account for differences between the individual plots, parameterization included global variables for the base models and local variables for initial starting values for LAI and dominant height at the individual plot level. The total number of global parameters for the two models was 21. A total of 69 local variables, 23 for local LAI carrying capacity K, 23 for LAI starting values, and 23 for dominant height starting values, were estimated. The dede function in the deSolve package was used to solve the DDE model and a combination of the BFGS, Nelder-Mead, and CG methods were used for optimization methods to maximize the log likelihood of the functions as implemented in the optimx package (Soetaert et al., 2010; Nash and Varadhan, 2011). Starting values for the local parameters were defined by the first observations of LAI at the individual plot level. For the models using local K values, starting values were the global K parameterized in the model fitting procedure. Global parameters starting values were selected through a trial and error approach to define the model form and starting values for the environmental modifiers were selected based on the reported literature response between the environment and loblolly pine (Nedlo et al., 2009; Albaugh et al., 2004; Teskey et al., 1987). Evaluations for parameter uncertainty estimates required the fixing of local variables to allow for the final standard error calculations.

4.3 Results

4.3.1 Simultaneous Model Parameterization

Incorporating radiation interception as an additive term rather than interactive proved to be the most flexible option for both IPAR and FPAR at both locations. Evaluation of the location to incorporate the radiation interception component found that incorporating it in the m term, the growth rate of the Chapman-Richards model, to best increase model

performance as compared to the other location. Simultaneous parameterization of the LAI and dominant height models resulted in an overall RMSE for LAI of 0.36 m^2/m^2 and 1.10 m for dominant height with IPAR included as an additive term (Table 4.1). To improve overall dominant height model performance, an additional parameter (β_2) was added to provide a power transformation to the entire dominant height model. Results from the power transformation improved predictions of dominant height across the range of observed heights, with a slight under-prediction at heights greater than 25 m (Figures 4.3, 4.4-B). Residuals for both dominant height and LAI showed little evidence of heteroscedasticity (Figure 4.4). RMSEs for the dominant height model at the individual plot level ranged from 0.52 to 2.69 m while LAI ranged from 0.24 to 0.62 m^2/m^2. The site with the lowest RMSE for LAI coincided with the site that had the lowest RMSE for dominant height, but that relationship did not carry over to the sites with the highest LAI and dominant height RMSEs. Comparisons between predictions and observations for both variables showed a strong 1:1 relationship (Figures 4.2, 4.3).

Parameter estimates in the LAI model showed slight deviations in values as compared to stand alone parameterizations. The r parameter showed an increase to 1.47 while the τ declined to 0.03, an approximately 11-day delay (Table 4.1). The additive term to the m parameter in the dominant height model showed an increase to the growth rate in conjunction with the intercepted radiation, indicating that the growth rate will increase with increased radiation interception, a relationship defined by the peak LAI (Table 4.1). The power transformation of the mode (β_2) approached a $\frac{2}{3}$ transformation, helping reduce overpredictions at lower heights and underpredictions at greater heights.

4.3.2 Theoretical Stands

To test the sensitivity of the parameterized model as a conceptual tool, two theoretical stands with different long term LAI carrying capacity values (K parameter in the LAI model), 3.5 for the low stand and 5.0 for the high stand, were evaluated with identical starting values to see

the effect of the intercepted radiation component on dominant height growth. With starting values, environmental variables, and global variables all held constant with the exception of the K parameter, LAI and dominant height projections were made to determine the effect of differences in radiation interception capabilities between the low and high LAI stands. LAI trends for the low stand reached an upper asymptote that oscillated around an LAI of approximately 2 m^2/m^2 while the high stand oscillated around approximately 4 m^2/m^2 (Figure 4.5). The resulting difference in radiation intercepted in the parameterized dominant height model was approximately 3 m at age 23 (24.02 m for low, 27.02 m for high).

Additionally, to compare the effect of changing the long-term LAI carrying capacity value on estimated base age 25 site index (m), stands with a range of K values from 3.7 to 5.0 were simulated under identical environmental conditions. The simulations were carried out to age 25 to determine the resulting site index of the stands with varying Ks. The relationship between estimated site index and K show a positive, curvilinear relationship (Figure 4.6). The slope of the line decreases as K increases, indicating a saturation effect at higher LAI carrying capacities, which is expected due to canopies with higher LAI becoming less efficient from self-shading (Monsi and Saeki, 1953; Terashima et al., 2005). Simulations to age 25 extrapolated outside the data set's age range; older data must be utilized in the model fitting procedure to verify results.

4.4 Discussion

While forest growth is a result of the interaction in resource availability, we were able to show that light interception is a critical component of the equation in explaining variation in forest productivity. Similar to Binkley et al. (2010); Teskey et al. (1987); Vose and Allen (1988); Vose and Swank (1990), we found that stands with higher LAIs had greater growth. While other factors, such as light use efficiency or other stand structural components, weren't considered in the modeling scheme, LAI was shown to be an adequate variable for relating

stand level conditions to observed growth characteristics in dominant height (Stenberg et al., 1994).

The relationship between the biophysical variable LAI and the dominant height growth trajectory provided an increased understanding of the relationship between forest productivity and a stand's ability to intercept light. As LAI increased and more light was estimated to be intercepted by a stand's canopy, an increase in dominant height growth was shown to occur. Although the LAI model was not able to capture the magnitude of changes between peak LAI and minimum LAI throughout the year, further research should focus on improving the sensitivity of the LAI model. Further investigation to modeling non-dominant trees may provide the necessary conditions in which intercepted radiation affects height growth, rather than just focusing on dominant height of a stand, where it's been shown that a trade-off between light use efficiency and light interception efficiency exists between tall and short trees (Onoda et al., 2014). Improvements in the availability of high resolution satellite data can provide greater accuracy and precision of growth to occur at a pixel level rather than summarized at a stand or tract level that is common in loblolly pine plantation systems. Continued development of the system will investigate the effect of radiation interception on other stand level variables, including basal area growth, diameter growth, volume growth, and mortality to understand the relationship, if any, with radiation interception.

While LAI and the estimated intercepted radiation are relatively straightforward variables to acquire and estimate, other factors in determining productivity of trees remain fundamental to truly understanding the relationship between forest productivity and radiation interception, such as light use efficiency and partitioning (Russell et al., 1989). Limitations of the data set used in the study include the low trees per acre initial planting density (300 trees per acre) that may have introduced variation in the satellite imagery due to competing vegetation, soil reflectance, etc.

The potential applications of this research for forest growth and yield modeling revolve around the fundamental exchanges of energy between earth and the individual stand. Under-

standing the effect of silvicultural treatments, such as fertilization, thinning, or competition control on a stand's LAI can provide a deeper understanding to how a stand's growth responds via light interception. Further development of this system to incorporate mortality, basal area, and volume may gain additional insight into to the mechanisms of forest productivity and their relations to light interception (Courbaud, 2000). As future climate scenarios remain uncertain in the potential magnitude of changes and the associated variability, the modeling scheme proposed here allows for the testing of effects on loblolly pine plantation productivity in the southeastern United States.

4.5 Conclusions

A simultaneous model for leaf area and dominant height was estimated to account for changes in dominant height growth as a function of light interception. The model shows good agreement between predicted values and field observations for both leaf area and dominant height with no bias in the prediction. Environmental variables were incorporated into the leaf area model to account for changes in leaf area growth over time, showing that monthly maximum temperature and monthly excess water are able to explain some of the observed variation. It was shown that accounting for annual intercepted radiation through a function of leaf area index was important in determining dominant height growth in loblolly pine plantations, opening a new avenue to develop simple models that include this ecological indicator.

4.6 References

Albaugh, T. J., Allen, H. L., Dougherty, P. M., Kress, L. W., and King, J. S. (1998). Leaf area and above- and belowground growth responses of loblolly pine to nutrient and water additions. *Forest Science*, 44(2):317–328.

Albaugh, T. J., Lee Allen, H., Dougherty, P. M., and Johnsen, K. H. (2004). Long term growth responses of loblolly pine to optimal nutrient and water resource availability. *Forest Ecology and Management*, 192(1):3–19.

Binkley, D., Stape, J. L., Bauerle, W. L., and Ryan, M. G. (2010). Explaining growth of individual trees: Light interception and efficiency of light use by Eucalyptus at four sites in Brazil. *Forest Ecology and Management*, 259(9):1704–1713.

Courbaud, B. (2000). Comparing light interception with stand basal area for predicting tree growth. *Tree Physiology*, 20(5-6):407–414.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202:18–27.

Harris, I., Osborn, T. J., Jones, P., and Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, 7(1):109.

Harrison, M. and Kane, M. (2008). PMRC Coastal plain culture/density study: age 12 analysis. *PMRC Technical Report*, 1:74.

Landsberg, J. and Waring, R. (1997). A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. *Forest Ecology and Management*, 95(3):209–228.

Landsberg, J. J. and Sands, P. J. (2011). *Physiological ecology of forest production : principles, processes and models.* Terrestrial ecology series: v. 4. Elsevier/Academic Press.

Makela, A., Landsberg, J., Thomas, A. R., Burk, E., Goran, M. T.-M., Agren, I., Oliver, C. D., and Puttonen, P. (2000). Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation. *Tree Physiology*, 20(September):289–298.

Monsi, M. and Saeki, T. (1953). On the Factor Light in Plant Communities and its Importance for Matter Production. *Japanese Journal of Botany*, 14:22–52.

Nash, J. C. and Varadhan, R. (2011). Unifying optimization algorithms to aid software system users: optimx for R. *Journal of Statistical Software*, 43(9):1–14.

Nedlo, J. E., Martin, T. A., Vose, J. M., and Teskey, R. O. (2009). Growing season temperatures limit growth of loblolly pine (Pinus taeda L.) seedlings across a wide geographic transect. *Trees - Structure and Function*, 23(4):751–759.

Nychka, D., Furrer, R., Paige, J., and Sain, S. (2017). fields: Tools for spatial data. R package version 10.3.

Onoda, Y., Saluñga, J. B., Akutsu, K., Aiba, S.-i., Yahara, T., and Anten, N. P. R. (2014). Trade-off between light interception efficiency and light use efficiency: implications for species coexistence in one-sided light competition. *Journal of Ecology*, 102(1):167–175.

Russell, G., Jarvis, P., and Monteith, J. (1989). *Absorption of radiation by canopies and stand growth*, page 21–40. Society for Experimental Biology Seminar Series. Cambridge University Press.

Savoy, P. and Mackay, D. S. (2015). Modeling the seasonal dynamics of leaf area index based on environmental constraints to canopy development. *Agricultural and Forest Meteorology*, 200:46–56.

Soetaert, K., Petzoldt, T., and Setzer, R. W. (2010). Solving differential equations in R: Package deSolve. *Journal of Statistical Software*, 33(9):1–25.

Stenberg, P., Kuuluvainen, T., Kellomäki, S., Grace, J. C., Jokela, E. J., and Gholz, H. L. (1994). Crown Structure, Light Interception and Productivity of Pine Trees and Stands. *Ecological Bulletins*, (43):pp. 20–34.

Tang, L., Yin, D., Chen, C., Yu, D., and Han, W. (2019). Optimal design of plant canopy based on light interception: A case study with loquat. *Frontiers in Plant Science*, 10(March):1–11.

Terashima, I., Araya, T., Miyazawa, S. I., Sone, K., and Yano, S. (2005). Construction and maintenance of the optimal photosynthetic systems of the leaf, herbaceous plant and tree: An eco-developmental treatise. *Annals of Botany*, 95(3):507–519.

Teskey, R. O., Bongarten, B. C., Cregg, B. M., Dougherty, P. M., and Hennessey, T. C. (1987). Physiology and genetics of tree growth response to moisture and temperature stress: an examination of the characteristics of loblolly pine (Pinus taeda L.). *Tree Physiology*, 3(1):41–61.

Thornton, P., Thornton, M., Mayer, B., Wei, Y., Devarakonda, R., Vose, R., and Cook, R. (2016). Daymet: Daily surface weather data on a 1-km grid for north america, version 3.

Vose, J. M. and Allen, H. L. (1988). Leaf area, stemwood growth, and nutrition relationships in loblolly pine. *Forest Science*, 34(3):547–563.

Vose, J. M. and Swank, W. T. (1990). A conceptual model of forest growth emphasizing stand leaf area. In Dixon, R. K., Meldahl, R. S., Ruark, G. A., and Warren, W. G., editors, *Process Modeling of Forest Growth Responses to Environmental Stress*, chapter 24, pages 278–287. Timber Press, Inc., Portland, OR.

Zhao, D., Kane, M., and Borders, B. E. (2010). Development and applications of the relative spacing model for loblolly pine plantations. *Forest Ecology and Management*, 259(10):1922–1929.

Zhao, D., Kane, M., and Borders, B. E. (2011). Growth responses to planting density and management intensity in loblolly pine plantations in the southeastern USA Lower Coastal Plain. *Annals of Forest Science*, 68(3):625–635.

Zhao, D., Kane, M., and Harrison, W. M. (2008). SAGS culture/density study: results through age 10. *PMRC Technical Report*, 3:33.

Table 4.1: Global parameter estimates for the leaf area index (Equation 4.1) and dominant height (Equation 4.2) models fitted to 23 plots across the southeastern United States with the environmental variables included as a beta and double logistic function, and a local carrying capacity parameter (K). Root mean square error (RMSE) was calculated across all observations.

Parameter	Estimate	Model Form	Environmental Variable
r	1.4747	LAI Base	N/A
τ	0.0344		
A	2.9007		
ω	6.3140		
μ	−2.0876		
h	0.1561		
μ_2	−5.8073	Beta	Monthly Maximum Temperature
T_b	1.4637		
α	1.4743		
T_c	37.9576		
β	0.3705		
α	2.0187	Double Logistic	Monthly Excess Water
β	0.0040		
δ	0.0079		
T_b	−0.0005		
T_c	329.9668		
η	1.8660	Dominant Height	Intercepted Radiation
m	0.1960		
κ	0.2013		
β_1	0.1293		
β_2	0.6212		
RMSE-LAI	0.36		
RMSE-HDOM	1.10		

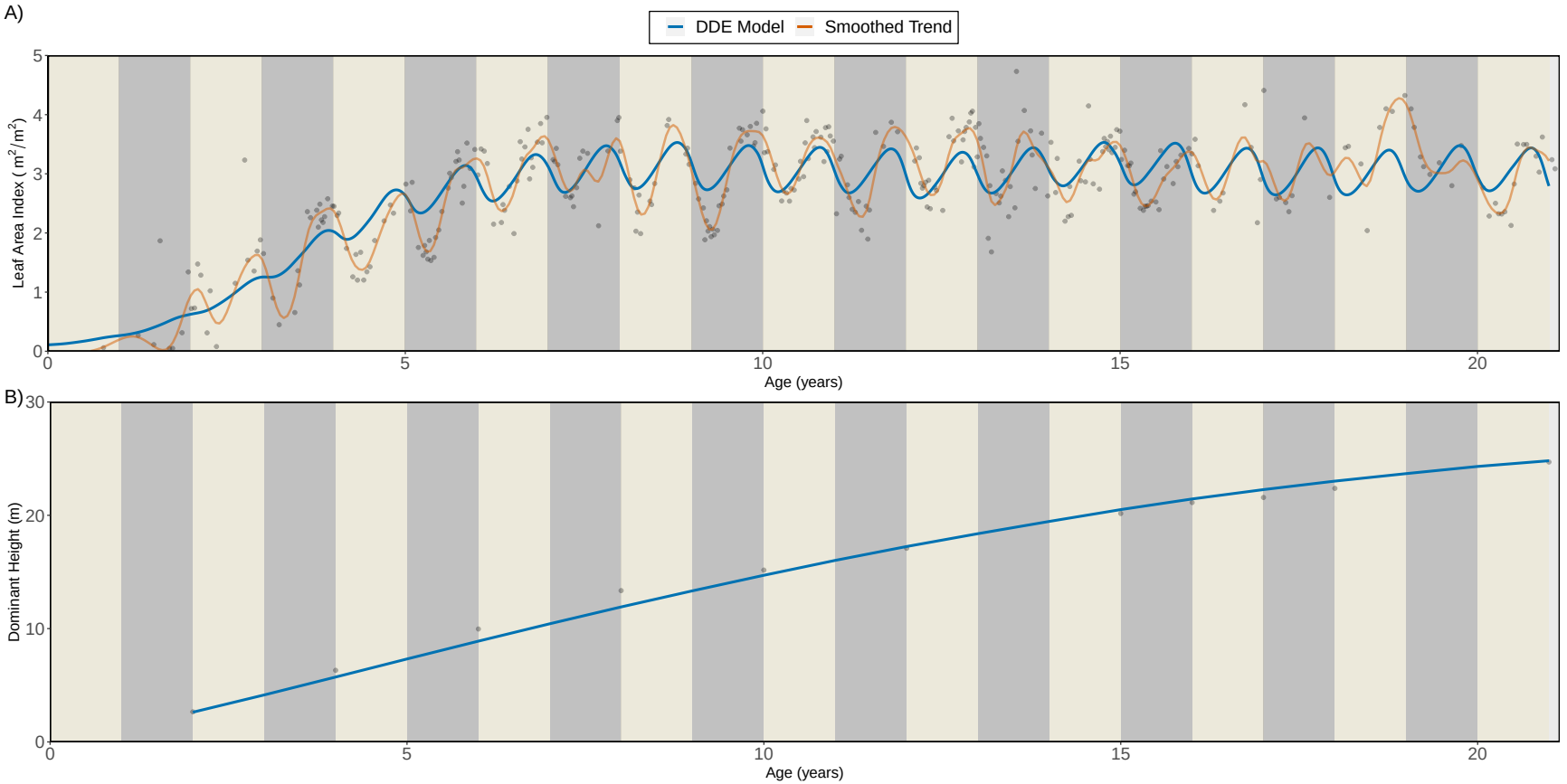


Figure 4.1: Model results for one observed plot from the simultaneous fitting of A) leaf area index and B) dominant height model. In addition to the model results, observed LAI and dominant heights are displayed. Root mean squared error (RMSE) for dominant height was 0.67 m and LAI was 0.35 m^2/m^2.

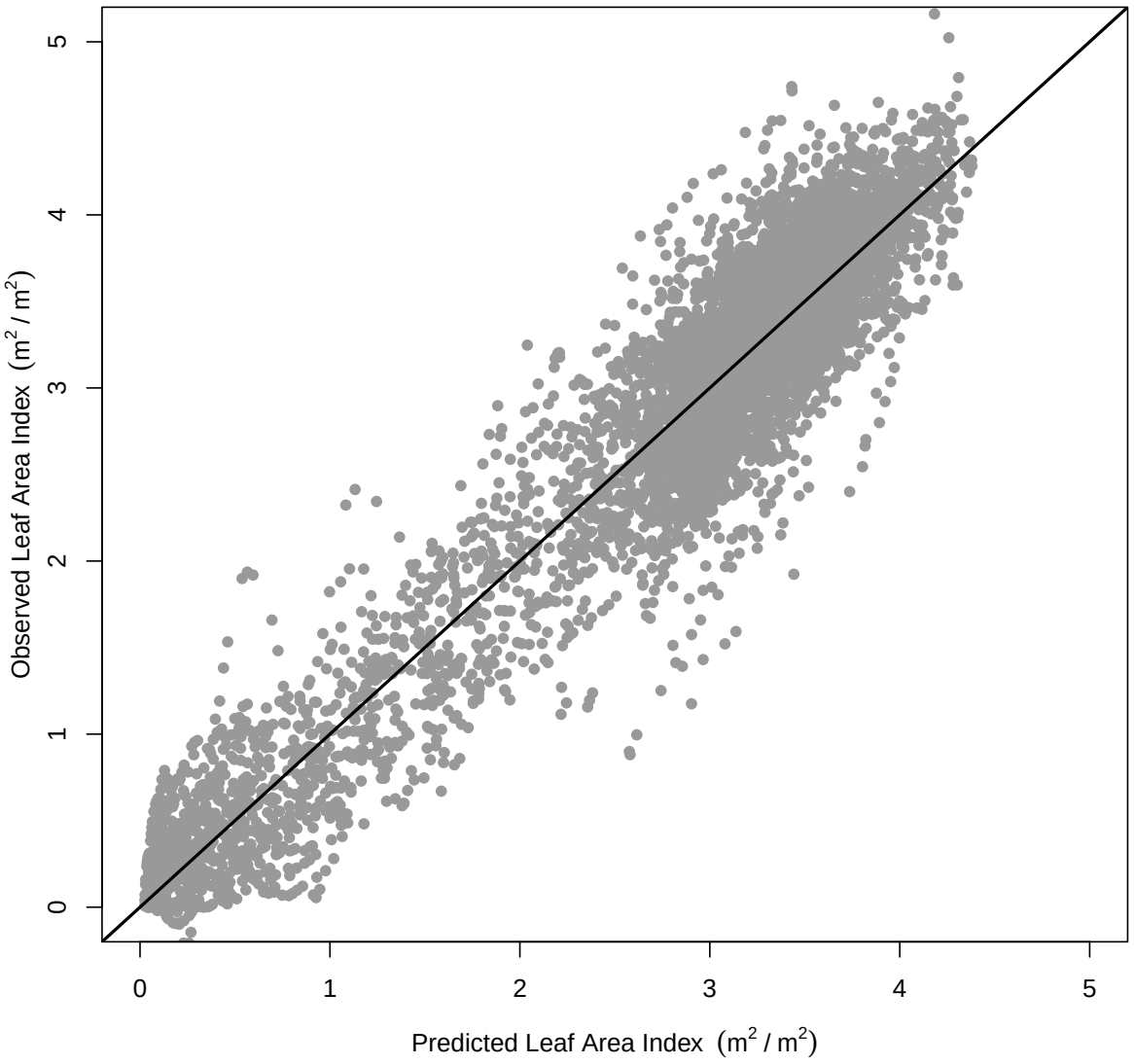


Figure 4.2: Observations vs. predictions for leaf area index values. A solid black line indicates a 1:1 relationship.

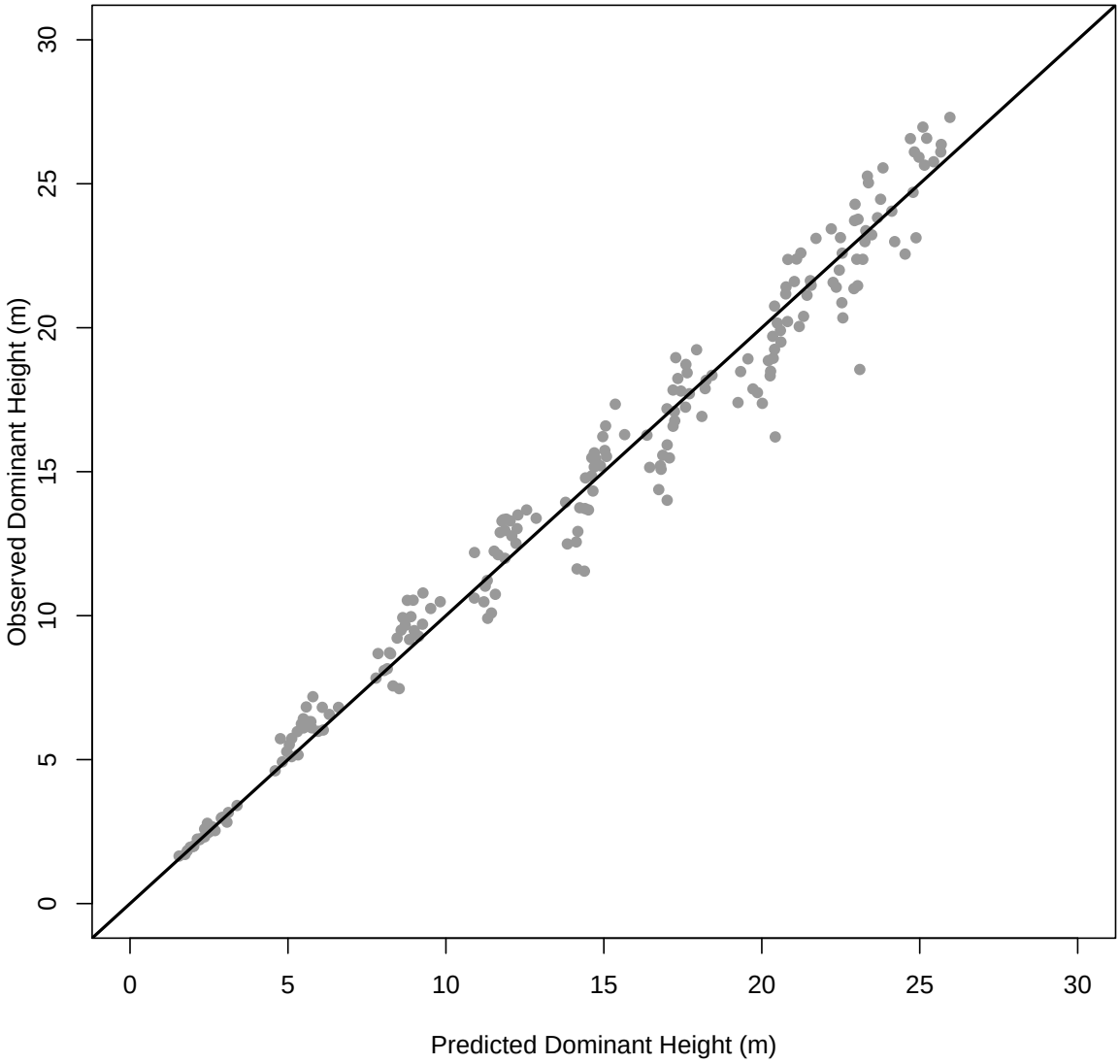


Figure 4.3: Observations vs. predictions for dominant height values. A solid black line indicates a 1:1 relationship.

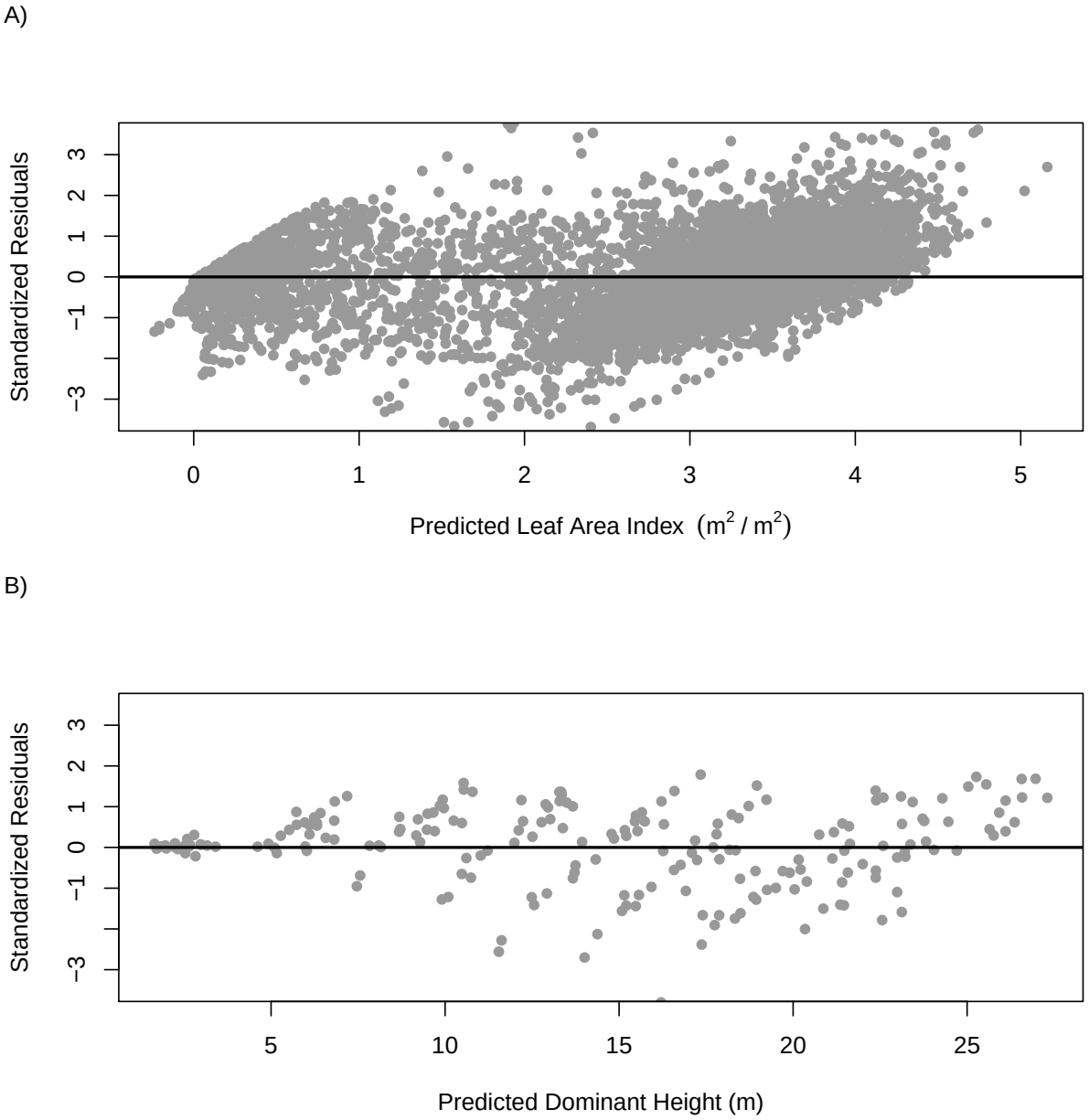


Figure 4.4: Studentized residuals for A) leaf area index (m^2/m^2) predictions and B) dominant height predictions (m).

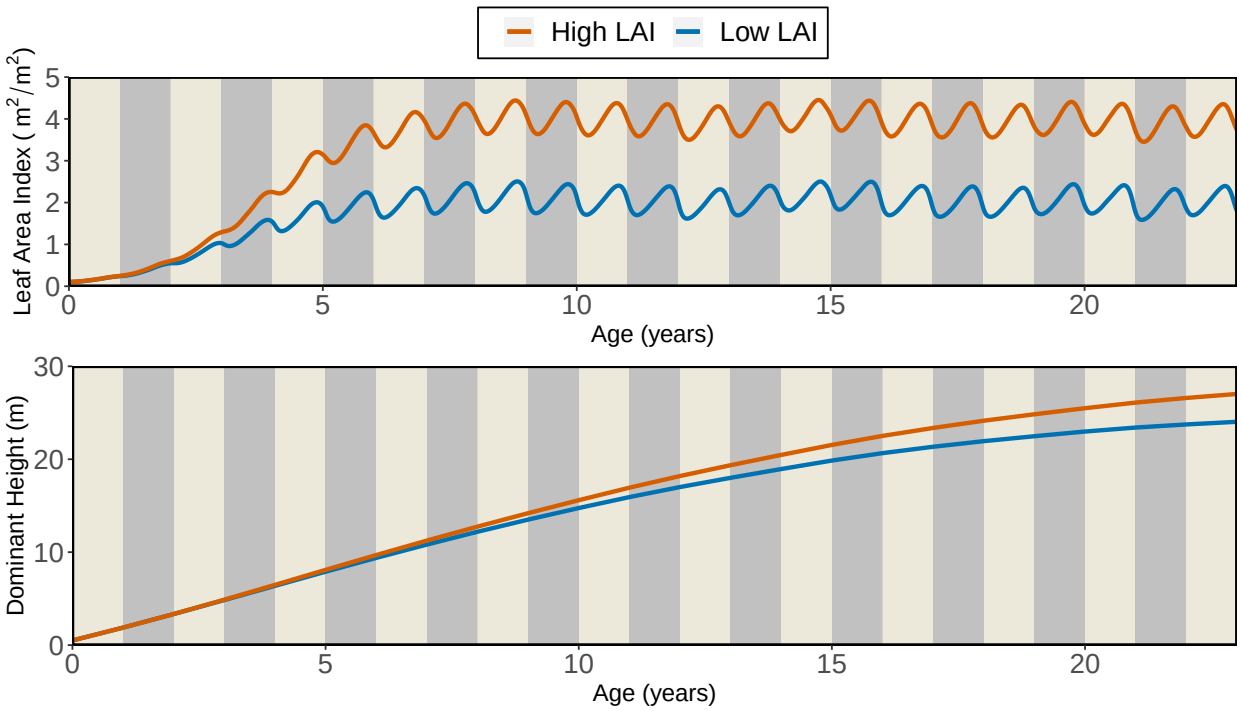


Figure 4.5: Using the model estimated parameters, a sensitivity analysis was performed to compare the dominant height growth trajectories for a site with low LAI (3.5 m^2/m^2) and high LAI (5.0 m^2/m^2), with identical starting values.

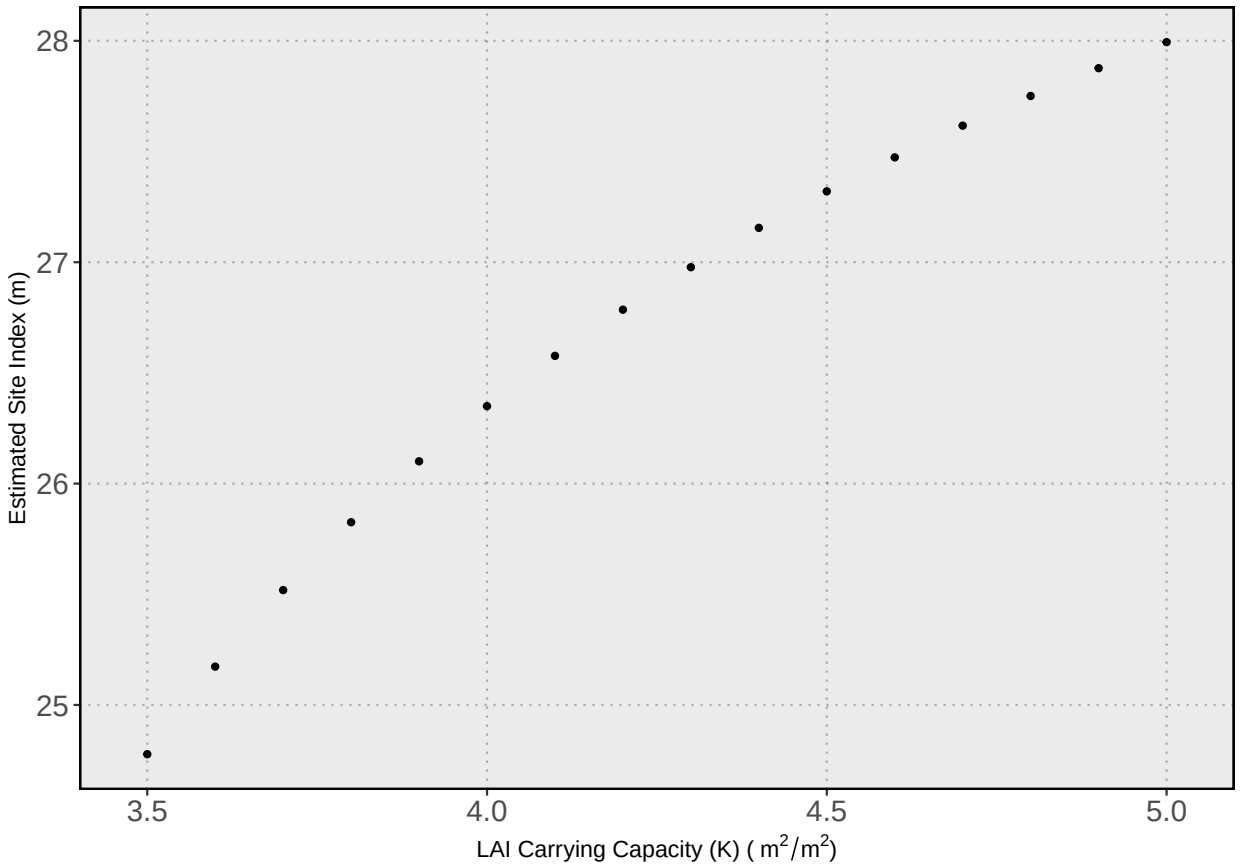


Figure 4.6: Estimation of site indices (m) base age 25 using a range of leaf area index carrying capacity values (m^2/m^2) at a site in south Alabama.

Chapter 5

Overall Conclusions

Understanding variability in forest productivity through its relationship with light interception is essential for efficiently managing forest resources for current and future obligations. As we gain increased understanding about the factors that influence light interception, management criteria can adapt to the factors that can be controlled while better predicting the influence of exogenous factors . Before being able to study those influences, Chapter 2 provided a LAI estimation model that provided a way to correct for observation errors in remotely sensed data to retrieve an unbiased model. The LAI estimation model provided improvements over the current industry standard for Landsat 5 + 7 imagery, which now allows for access to almost 40 years of imagery with unbiased estimates of loblolly pine LAI. Novel to that chapter is the use of the Kalman filter to quantify observation errors and the use of error-in-variable methods to reduce the influence of those observation errors on parameter estimates (Kalman, 1960; Cook and Stefanski, 1994). Use of the normalized difference moisture index provided improvements in model fitting and prediction as compared to more popular vegetation indices, such as the normalized difference moisture index or simple ratio.

Chapter 3 derived a LAI growth model with modifiers to account for influence of the environment on LAI trends. Predictions from this model can provide a measure of productivity in areas with scant field data while relying on only remotely sensed imagery. Following the framework outlined by Powers et al. (2003), various environmental variables and model forms were evaluated on their effects of LAI growth and model improvement. The inclusion of environmental variables, monthly maximum temperature and monthly excess water, provided an increased understand of the factors affecting light interception, while providing a

way to model changes in future climate scenarios and how they might affect forest growth and development.

Chapter 4 used the increased understanding of LAI growth and development to drive a dominant height model by calculating the amount of radiation intercepted for a given year to govern the anabolic term of the dominant height model. By accounting for the annual intercepted light and the dominant height growth that occurred, we were able to provide an updated measure of forest productivity that accounts for some of the biological processes that determine forest growth. Quantifying the exchange of energy between the environment and a forest provided biological significance to growth, allowing for the comparison of forests with low and high light interception capabilities and determining the difference in dominant height growth. While using LAI as a management tool has been a well-defined concept (Vose and Swank, 1990; Sampson et al., 1997), we can now show the effects that changing LAI have on forest productivity, defined by dominant height (Teskey et al., 1987; Vose and Allen, 1988; Albaugh et al., 1998).

The results from this research allows for researchers, forest managers, investors, and other interested stakeholders to take advantage of the ever growing catalog of freely available, high resolution remotely sensed satellite imagery to improve and advance the decision making process. Access to historical data will provide better understanding of long term trends observed in loblolly pine productivity and how it relates to the environment. By using remotely sensed data to estimate LAI and ecophysiological concepts to predict long term trends, the productivity of future stands under different climate scenarios can be modeled. The methods proposed in this research provide a framework to expand understanding of light interception and forest productivity to other species and stand types.

5.1 References

Albaugh, T. J., Allen, H. L., Dougherty, P. M., Kress, L. W., and King, J. S. (1998). Leaf area and above- and belowground growth responses of loblolly pine to nutrient and water additions. *Forest Science*, 44(2):317–328.

Cook, J. R. and Stefanski, L. A. (1994). Simulation-extrapolation estimation in parametric measurement error models. *Journal of the American Statistical Association*, 89(428):1314–1328.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45.

Powers, S. J., Brain, P., and Barlow, P. W. (2003). First-order differential equation models with estimable parameters as functions of environmental variables and their application to a study of vascular development in young hybrid aspen stems. *Journal of Theoretical Biology*, 222(2):219–232.

Sampson, D. A., Vose, J. M., and Allen, H. L. (1997). A conceptual approach to stand management using leaf area index as the integral of site structure, physiological function, and resource supply. *Proceedings of the ninth biennial southern silvicultureal research conference*, 2:25–27.

Teskey, R. O., Bongarten, B. C., Cregg, B. M., Dougherty, P. M., and Hennessey, T. C. (1987). Physiology and genetics of tree growth response to moisture and temperature stress: an examination of the characteristics of loblolly pine (Pinus taeda L.). *Tree Physiology*, 3(1):41–61.

Vose, J. M. and Allen, H. L. (1988). Leaf area, stemwood growth, and nutrition relationships in loblolly pine. *Forest Science*, 34(3):547–563.

Vose, J. M. and Swank, W. T. (1990). A conceptual model of forest growth emphasizing stand leaf area. In Dixon, R. K., Meldahl, R. S., Ruark, G. A., and Warren, W. G., editors, *Process Modeling of Forest Growth Responses to Environmental Stress*, chapter 24, pages 278–287. Timber Press, Inc., Portland, OR.

Appendix A

Chapter 2 Code

```
1
2  ## kalman filter on peak lai data
3
4  nlkflik = function(param,  LAI,  irrigate, fert, PLOTID,age)
5
6  {
7  kfs <- foreach(dfs= levels(max_all_data$PLOTID), .combine = rbind, .
       export = c("Kfilter")) %dopar%{
8     #nplots      = as.numeric(dfs)
9     laiss        = LAI[PLOTID == dfs]
10    agess          = age[PLOTID == dfs]
11    pred         = Kfilter(param = param,
12                           LAI   = laiss,
13                           age   = agess,
14                           PLOTID= dfs
15    )}
16
17   kfmean <<- unlist(kfs)[1:ndata]
18   kfvars <<- unlist(kfs)[(ndata+1):(2*ndata)]
19
20   resu   <- sum(dnorm(x           = LAI,
21                       mean         = kfmean,
22                       sd           = sqrt(kfvars),
23                       log          = TRUE))
24   return(-resu)
25
26 }
27
28 Kfilter<- function(param, LAI, age,PLOTID)
29 {
30   #param holds the list of parameters
31   #lai holds the LAI
32   #age holds age at which lai was measured
33   ndata1     <- length(LAI)
34   lai.nobs   <- numeric(ndata1)
35   var.nobs   <- numeric(ndata1)
36   proc.var   <- exp(param["procvar"])
37   obs.var    <- exp(param["obsvar"])
38   lai.n      <- exp(param["M.lai.start"])
39   var.n      <- exp(param["var.lai.start"])
40   xmid       <- param["xmid"]
41   scale      <- param["scale"]
42   lai.nobs[1]<- lai.n
43   var.nobs[1]<- var.n + obs.var
44
45
```

```
  for (i in 2:ndata1) {
    #Specify growth function.
    lai.ni      <- (lai.n )*
      (1+exp((xmid - age[i-1])/scale))/
      (1+exp((xmid - age[i])/scale))
    #Here the first derivative with respect
    #of the variable to be increased.
    slope       <- (1+exp((xmid - age[i])/scale))/
      (1+exp((xmid - age[i-1])/scale))
    var.ni      <- slope^2*var.n + proc.var
    lai.nobs[i] <- lai.ni
    var.nobs[i] <- var.ni + obs.var
    lai.n       <- lai.ni + var.ni /
      var.nobs[i] * (LAI[i] - as.numeric(lai.nobs[i]))
    var.n       <- var.ni * (1- var.ni/var.nobs[i])
  }
  return(list(lai.est = lai.nobs, var.est = var.nobs))
}

## starting values
param<- c(procvar       =log(0.006),
          obsvar        =log(0.005),
          M.lai.start   =log(0.89),
          var.lai.start =log(0.02),
          xmid          = 0.2,
          scale         = 8.0)

## run filter in parallel environment
no_cores <- detectCores()
cl <- makeCluster(no_cores, type="SOCK")
registerDoSNOW(cl)

reg.fit<-with(max_all_data,
              optim(param,nlkflik,
                    LAI      = NDMI_shift,
                    age      = AGE,
                    PLOTID   = PLOTID,
                    method   = "BFGS",
                    control  = list(maxit=10000)
              ))
stopImplicitCluster()



# simex for final model - using error estimate from Kalman output
loglik7.simex<-function(parameters,LAI,NDMI,NDMI_error,Fert,noise.
   level)
{
  ndata <- (length(LAI))
  B0   <-parameters[1]
  B1   <-parameters[2]
  B2   <-parameters[3]
  A1   <-parameters[4]
  A2   <-parameters[5]
  sig1 <-parameters[6]
  sig2 <-parameters[7]

  total.noise <- sqrt(noise.level*NDMI_error)
  means <- rep(0,ndata)
```

```
  NDMI2 <- rnorm(n=means,mean=NDMI,sd=total.noise) # create
     pseudodata

  exp.value <- ((B0+B1*NDMI2+Fert*(A1*NDMI2**A2))/(1+B2*NDMI2))
  sigma <- (sig1*exp.value^sig2)


  result<-sum(dnorm(LAI,mean=exp.value,sd=sigma,log=T))
  return(-result)
}

inits_noise1<-c(B0= -32.2802480,
                B1= 37.7618371 ,
                B2= 5.8551406 ,
                A1 = 1.7856614,
                A2 = 5.3384160,
                sig1= 0.3739386 ,
                sig2= 0.1408014)

## run simex for one level of noise
## NDMI_var_Total includes combination of spline error and Kalman
   Filter error - error

run_simex1 <- function(noise.amount){
  maxlik.noise<-with(subset(vi_data,!is.na(LAI)),
                     optim(inits_noise1,
                           loglik7.simex,
                           LAI=LAI,
                           NDMI=NDMI_shift,
                           Fert=fert,
                           NDMI_error=NDMI_var_Total,
                           noise.level = noise.amount,
                           method="SANN"))
  return(c(noise.amount,maxlik.noise$par))
}

noise.levels <- rep(0:20/10,each=1000)

# look at one level - need to repeat for all levels
noise.levels2 <- subset(noise.levels,noise.levels==0.1)

system.time(for (i in 1:length(noise.levels2)){
  run_simex1(((noise.levels2)[i]))

  if(i==1) simex_results2<-(run_simex1((noise.levels2[i])))
  if(i>1) simex_results2<-rbind(simex_results2,(run_simex1((noise.
     levels2[i]))))
})


simex_results2.df <- na.omit(as.data.frame(simex_results2))
names(simex_results2.df)<-c("V1","B0","B1","B2","A1","A2","sig1","
   sig2")

# get updated estimates for one level of noise - repeat for all
   levels
run2_summary <- ddply(simex_results2.df, .(V1), summarize,
                      B0 = mean(B0),
                      B1 = mean(B1),
```

```
                          B2 = mean(B2),
                          A1 = mean(A1),
                          A2 = mean(A2),
                          sig1=mean(sig1),
                          sig2=mean(sig2))


## combine all runs in simex_results.df
simex_results.df$lambda <- simex_results.df$V1

simex_results.summary$lambda <- simex_results.summary$V1

# function to evaluate all runs - use quadratic model to
   extrapolation to lambda=-1
simex_plotter <- function(df,betas){
  model <- lm(substitute(i ~ lambda + I(lambda^2), list(i = as.name(
     betas))), data = df)

  new.df <- data.frame(lambda=c(-1))
  unbiased <- predict(model, newdata=new.df)
  xmin <- -1.01
  xmax <- 2.05
  plot(get(paste0(betas))~lambda,data=df,
       type="p",pch=1,cex=.8,
       xlim=c(xmin,xmax),
       xlab="", ylab="")
  title(ylab = expression(paste("Estimated Parameter ", beta[0])),
        xlab = expression(lambda),
        cex=2)
  newdata <- seq (from=0,to=3, by=0.01000)
  newdata2 <- seq (from=-1,to=0, by=0.01000)
  lines(newdata, predict(model, data.frame(lambda=newdata)), col=2,
     lwd=2)
  lines(newdata2, predict(model, data.frame(lambda=newdata2)), col
     =2,lwd=2,lty=3)
  points(-1,as.numeric(unbiased),pch="+",cex=2)
  #points(results.summary$V1,get(paste("results.summary",sep=""))[[
     as.name(betas)]],cex=2,pch="x",col="blue")
  print(summary(model))
  print(format(round(unbiased,20),nsmall=20))
}

# run for B0, repeat for other parameters, output returns unbiased
   value
simex_plotter(simex_results.df,"B0")
```

Appendix B

Chapter 3 Code

```
library(deSolve)
library(numDeriv)
library(optimx)
library(doParallel)
library(doSNOW)
library(snow)
## fit lai model to all plots
## with local

##-----------------------------
## the derivative function
##-----------------------------
logistic <- function(t, y, parms,env,env2) {
  r = (parms["r"])
  K = (parms["K"]) # this is local
  tau = (parms["tau"])
  h = (parms["h"])
  A = (parms["A"])
  omega = (parms["omega"])
  mu = (parms["mu"])
  mu2 = parms["mu2"]
  temp.b = parms["temp.b"]
  alpha  = parms["alpha"]
  temp.m = parms["temp.m"]*100
  beta   = parms["beta"]

  temp.b2 <- parms["temp.b2"]
  temp.c2 <- parms["temp.c2"]*100
  alpha2  <- parms["alpha2"]
  beta2   <- parms["beta2"]
  delta2  <- parms["delta2"]


  tlag = tau

  lag =ifelse(tlag <= 0,0.0,
              ifelse(t-tlag <=0, 0.0,lagvalue(t-tlag)))

  # environmental modifier - beta function
  H.temp <- exp(mu2) * (((env(t)-temp.b)**alpha)*((temp.m-env(t))**
     beta))

  # environmental modifier - double logistic
  H.temp2 <- alpha2*(1/((1+exp(-beta2*(env2(t)-temp.b2)))*(1+exp(
     delta2*(env2(t)-temp.c2)))))

```

```
    dy <- (r*y[1]*(1-lag[1]/((K+A*sin(omega*t+mu)))) - h*y[1])*H.
       temp*H.temp2

  list(dy=dy,lag=lag)
}



y.dinamic<-function(local.start,times,IDs,m,interpol.env,interpol.
   env2){
  y.d<- dede(y = local.start,
             times = times,
             func = logistic,
             parms = m,
             env = interpol.env,
             env2= interpol.env2,
             control=list(mxhist = 10000000))
  return(y.d)}



global <- c(         tau      = 0.1142544229 ,
                     h        = 0.0646616036 ,
                     A        = 2.6254156196 ,
                     omega    = 6.2876493039 ,
                     mu       = -1.7420027704 ,
                     r        = 1.1968129114 ,
                     sds = (0.3800999736 ))

local <-c(           ystart = 0.3425833) # Local LAI starting value


local.K <-c(         K =  4.065111 ) # LAI K starting value

# function imports parameters, LAI data, number of plots,
   environmental data
nlkflik.mon.tmx = function(parameters, data.LAI,nplots,data.clim,
   local.L,local.K){
  if(parameters["tau"]<0||parameters["r"]<0||parameters["h"]<0||
     parameters["A"]<0||parameters["omega"]<0){ # enforce positive
     values
    value <- (-10e100)}
  else{

    ndata <- length(data.LAI$fit)
    lower <- 17+1
    upper <- 17+nplots
    local <- parameters[lower:upper] # Local lai starting value
    lower.K <- 17+nplots+1
    upper.K <- 17 +nplots+nplots
    local.K <- parameters[lower.K:upper.K]
    m<-parameters[1:17] # global values for lai + environmental
       modifiers
    LAI <- data.LAI$fit
    times<-data.LAI$new_x
    IDPlots<-unique(data.LAI$Name)
    IDs<-data.LAI$Name
    IDclim = data.clim$Name
    AGE = data.clim$AGE
```

```
    TEMP = data.clim$mon.tmx
    TEMP2 = data.clim$mon.EW

    kfs <- foreach(dfs= 1:length(IDPlots), .combine = rbind,.
       packages = c("deSolve"),.export = c("y.dinamic","logistic"))
       %dopar%{
      plotid      = IDPlots[dfs]
      times.ss    = times[IDs%in%plotid]
      local.ss    = local[dfs]
      local.K.ss  = local.K[dfs]
      IDs.ss      = IDs[IDs%in%plotid]

      # pass environmental data as spline function - interpolates at
         timestep
      interpol.env <- approxfun(AGE[IDclim%in%plotid],TEMP[IDclim%in
         %plotid],rule=2,method="linear")
      interpol.env2 <- approxfun(AGE[IDclim%in%plotid],TEMP2[IDclim%
         in%plotid],rule=2,method="linear")

      pred        = y.dinamic(local.start = local.ss, # this is the
         localized starting value for LAI series
                              times = times.ss,
                              IDs = IDs.ss,
                              m = c(m,local.K.ss),
                              interpol.env=interpol.env,
                              interpol.env2=interpol.env2)
    }

    kfmean <<- kfs[1:ndata,2]

    value  <- sum(dnorm(x          =  LAI,
                        mean        = kfmean,
                        sd          = (parameters["sds"]),
                        log         = TRUE))
  }
  return(-value)

}

inits.mon.tmx = c(global,
                  alpha=1.5926632564   ,
                  beta = 0.3418526183   ,
                  temp.b =1.4592300055  ,
                  temp.m = (0.4468866257 )  ,
                  mu2    =  -6.0358585068   ,

                  alpha2=1.5984363773   ,
                  beta2 = 0.0063069909        ,
                  delta2=0.0070890486    ,
                  temp.b2 =-0.0003954878  ,
                  temp.c2 = (2.7636195940 ) )

# combine global model, environmental model, and local starting
   values
all.inits <- c(inits.mon.tmx, local, local.K)

no_cores <- detectCores()
cl <- makeCluster(no_cores, type="SOCK")
registerDoSNOW(cl)
```

```
149
150 result.interaction <- optimx(par=all.inits,
151                                fn=nlkflik.mon.tmx,
152                                data.LAI=smooth_combine, # lai data
153                                nplots=length(unique(smooth_combine$
                                   Name)),
154                                control=list(follow.on=T,trace=T,REPORT
                                   =10,maxit=500,kkt=T,
155                                             parscale=round(sqrt(inits.
                                                mon.tmx**2),5)),
156                                data.clim = climate.dat, # east anglia
                                   cru data
157                                method=c("CG","Nelder-Mead","BFGS"))
158 print(result.interaction)
159
160 stopImplicitCluster()
```

Appendix C

Chapter 4 Code

```
1
2  library(deSolve)
3  library(numDeriv)
4  library(optimx)
5  library(doParallel)
6  library(doSNOW)
7  library(snow)
8  library(plyr)
9
10 # lai model
11 dL<- function(t, y, parms,env,env2) {
12   r     = (parms["r"])
13   K     = (parms["K"])*10
14   tau   = (parms["tau"])
15   h     = (parms["h"])
16   A     = (parms["A"])
17   omega = (parms["omega"])
18   mu    = (parms["mu"])
19
20   mu2    = parms["mu2"]
21   temp.b = parms["temp.b"]
22   alpha  = parms["alpha"]
23   temp.m = (parms["temp.m"])*100
24   beta   = parms["beta"]
25
26   temp.b2 <- parms["temp.b2"]
27   temp.c2 <- (parms["temp.c2"])*100
28   alpha2  <- parms["alpha2"]
29   beta2   <- parms["beta2"]
30   delta2  <- parms["delta2"]
31
32   tlag = tau
33
34   lag =ifelse(tlag <= 0,0.0,
35               ifelse(t-tlag <=0, 0.0,lagvalue(t-tlag)))
36
37   # environmental modifiers
38   H.temp <- exp(mu2) * (((env(t)-temp.b)**alpha)*((temp.m-env(t))**
       beta))
39   H.temp2 <- alpha2*(1/((1+exp(-beta2*(env2(t)-temp.b2)))*(1+exp(
       delta2*(env2(t)-temp.c2)))))
40
41   dL <- (r*y[1]*(1-lag[1]/((K+A*sin(omega*t+mu)))) - h*y[1])*H.temp*
       H.temp2
42
43   list(c(dL=dL))
```

```
}

# dominant height model
dH<- function(t, y, parms,env,lai) {
  eta    = parms["eta"]
  m      = parms["m"]
  kappa  = parms["kappa"]
  b1     = parms["b1"]
  b2     = parms["b2"]

  ipar <- env(t)-env(t)*exp(-0.5*(lai(t))) # calc intercepted
     radiation

   #fpar <- env(t) - ipar

  dH <- ((eta)*y[1]**(b1*(ipar/100000)+m)-kappa*y[1])**b2  # dom
     height model with IPAR scaled

  list(c(dH=dH))
}

# dynamic LAI formulation for plot level
y.dinamic.L<-function(local.start,times,IDs,m,model,interpol.env,
    interpol.env2){
  y.d <- dede(y = local.start,
              times = times,
              func = model,
              parms = m,
              env = interpol.env,
              env2= interpol.env2,
              control=list(mxhist = 1000000))
    return(y.d)}

# dynamic HDOM formulation for plot level
y.dinamic.H<-function(local.start,times,IDs,m,interpol.env,lai,model
    ){
  y.d <- rk4(y = local.start,
             times = times,
             func = model,
             parms = m,
             env = interpol.env,
             lai = lai)
  return(y.d)}

nlkflik = function(parameters, data.LAI,nplots,data.clim,data.clim.
    lai,data.Hdom,y.dinamic.L,y.dinamic.H,model1,model2,local.vars,
    fixed.parms){

    ndata <- length(data.LAI$fit)
    ndataH <- length(data.Hdom$HDomM)
    lower1 <- 23+1
    upper1 <- 23+nplots
    lower2 <- 23+nplots+1
    upper2 <- 23+nplots+nplots
    lower3 <- 23+nplots+nplots+1
    upper3 <- 23+nplots+nplots+nplots
    local.L<-parameters[lower1:upper1] # Local parameters - LAI
    local.H<-parameters[lower2:upper2] # Local parameters - Height
    local.K<-parameters[lower3:upper3]
```

```
97      m<-parameters[1:23] # global parameters
98      LAI <- data.LAI$fit # get all LAI data
99      Height <- data.Hdom$HDomM # get all HDom data (meters)
100     times<-data.LAI$new_x # get all the times for LAI observations
101     timesH <- data.Hdom$AGE # get times for HDom observations
102     IDPlots<-unique(data.LAI$Name) # list of all 23 plot names for "
           foreach" cycle
103     IDs<-data.LAI$Name # list of all names associated to LAI data
104     IDsH<-data.Hdom$Name # list of all names associated to LAI data
105     IDclim = data.clim$Name # list of all names associated with
           solar radiation data
106     AGE = data.clim$floorage # age associated with solar radiation
           data
107     TEMP = data.clim$ann.Solrad # get the solar radiation data
108
109     IDclim.lai = data.clim.lai$Name ## lai env data
110     AGE.lai = data.clim.lai$AGE ## lai env data
111     TEMP.lai = data.clim.lai$mon.tmx## lai env data
112     TEMP2.lai = data.clim.lai$mon.EW## lai env data
113
114     kfs <- foreach(dfs= 1:length(IDPlots), .combine = rbind,.
           packages = c("deSolve","plyr"),
115                    .export = c("dL","dH")) %dopar%{
116
117     PLOT.ID <- IDPlots[dfs]
118
119     analyze <- function(dfs){
120       plotid       = IDPlots[dfs]  # get plot id
121       times.ss     = times[IDs%in%plotid] # subset times for plot id
122       local.ss     = c(local.L[dfs]) # get the local starting values
             for LAI and Height
123       local.K.ss     = c(local.K[dfs]) # get the local starting
             values for LAI and Height
124       IDs.ss       = IDs[IDs%in%plotid] # get IDs but don't use
125       # spline the env data to pass as function
126       interpol.env <- approxfun(AGE.lai[IDclim.lai%in%plotid],TEMP.
             lai[IDclim.lai%in%plotid],rule=2,method="linear")
127       interpol.env2 <- approxfun(AGE.lai[IDclim.lai%in%plotid],TEMP2
             .lai[IDclim.lai%in%plotid],rule=2,method="linear")
128
129       pred         = y.dinamic.L(local.start = local.ss,
130                                  times = times.ss,
131                                  IDs = IDs.ss,
132                                  m = c(m,local.K.ss),
133                                  model = model1,
134                                  interpol.env=interpol.env,
135                                  interpol.env2=interpol.env2) #
                                        interpol.env is the solar radiation
                                         data in function form
136       return(pred)
137     }
138
139
140     kfsL <- analyze(dfs)
141
142     kfmeanL <- kfsL[,2]  # get all the predicted LAI
143     LAI.plot <- LAI[IDs%in%PLOT.ID]
144     value.L  <- ifelse(is.na(sum(dnorm(x          =  LAI.plot,
145                                        mean        = kfmeanL,
```

```
                                        sd          = (parameters["
                                           sdsL"]),
                                        log         = TRUE)))==T,-10
                                           e200,
                                        sum(dnorm(x          =  LAI.
                                           plot,
                                        mean        = kfmeanL,
                                        sd          = (parameters["
                                           sdsL"]),
                                        log         = TRUE)))

    LAI.E <- kfsL[,2]
    LAI.E.time <- kfsL[,1]

    LAI.estimated <- data.frame(kfsL)
    LAI.estimated$age <- floor(LAI.estimated$time) # get rounded AGE
        to summarize LAI
    peak.lai <- ddply(LAI.estimated,.(age),summarize,
                      plai = max(ystart)) # identify peak LAI for
                         each AGE

    analyze.H <- function(dfs){

      plotid      = IDPlots[dfs]  # get plot id
      local.ss    = c(local.H[dfs]) # get the local starting values
         for LAI and Height
      IDs.ss      = IDsH[IDsH%in%plotid] # get IDs but don't use
      LAI.E.ss    = peak.lai$plai # get the peak LAI for each age
      LAI.E.time.ss= peak.lai$age # get the ages
      timesH.ss    = timesH[IDsH%in%plotid]

      interpol.env <- approxfun(AGE[IDclim%in%plotid],TEMP[IDclim%in
         %plotid],rule=2,method="linear",ties = mean) #
         interpolation function to pass to dH for solar radiation
         data at any timestep
      interpol.lai <- splinefun(LAI.E.time.ss,LAI.E.ss,method="
         natural") # interpolation function to pass to dH for peak
         LAI data at any timestep

      pred2        = y.dinamic.H(local.start = local.ss,
                                 times = timesH.ss,
                                 IDs = IDs.ss,
                                 m = m,
                                 model = model2,
                                 lai = interpol.lai,
                                 interpol.env=interpol.env) # interpol.
                                    env is the solar radiation data in
                                    function form

      return(pred2)
    }

    kfsH <- analyze.H(dfs)

    kfmeanH <- kfsH[,2]  # get all the predicted HDom
    Height.plot <- Height[IDsH%in%PLOT.ID]

    value.H  <- ifelse(is.na(sum(dnorm(x          =  Height.plot,
                                       mean        = kfmeanH,
```

```
                                                sd          = (parameters["
                                                   sdsH"]),
                                                log         = TRUE)))==T,-10
                                                   e200,
                                                sum(dnorm(x          =
                                                   Height.plot,
                                                  mean        = kfmeanH,
                                                  sd          = (
                                                     parameters["sdsH"]),
                                                  log         = TRUE)))


    return(-(value.L+value.H))
    }
    big_data = sum(kfs)

    return(big_data)
}




### global starting values ###
global <- c(r =1.4746959081 ,
            tau   =0.0344250168  ,
            h     =0.1560551105  ,
            A     =2.9006717349  ,
            omega =6.3139684502 ,
            mu    =  -2.0875889025 ,
            sdsL = 0.3615596429,
            sdsH =1.1002179618  ,
            eta =1.8659879902  ,
            m =0.1960352862 ,
            kappa =0.2012837343  ,
            b1    =0.1292565632  ,
            b2    =0.6212351008 ,
            mu2    =-5.8073016794  ,
            temp.b =1.4637299680 ,
            alpha  =1.4742685513  ,
            temp.m =0.3791976448  ,
            beta   =0.3705493181 ,
            temp.b2 =-0.0005003456  ,
            temp.c2 =3.2927438890 ,
            alpha2  =2.0186663774  ,
            beta2   =0.0039990726  ,
            delta2  =0.0079399434 )


### local - dominant height starting values ###
### these come from fitting the dominant height model independently
h.start.vals <- c(hstart = 3.1127218559 )

### local - LAI starting values ###
### these come from fitting the LAI model independently
l.start.vals <- c(ystart =0.2527804271)

## local - K starting values
local.K <-c( K = 0.4503376466)
```

```
246
247
248
249 ## merge the local starting values
250 local<- c(l.start.vals ,h.start.vals,local.K)
251
252 ## combine the global and local starting values
253 inits <- c(global,local)
254
255 scales <- round(sqrt(inits**2),4)
256
257 no_cores <- detectCores()
258 cl <- makeCluster(no_cores-1, type="SOCK")
259 registerDoSNOW(cl)
260
261 result.2 <- optimx(par=inits, ## pass the inits
262                          fn=nlkflik, ## overall function to minimize
263                          data.LAI=smooth_combine, ## lai data to
                                pass
264                          data.Hdom=study_data, ## hdom data to pass
265                          nplots= length(unique(smooth_combine$Name))
                                , ## total number of plots - 23
266                          y.dinamic.L = y.dinamic.L, ## pass function
267                          y.dinamic.H = y.dinamic.H,
268                          model1 = dL, ## pass function
269                          model2 = dH, ## pass function
270                          data.clim = env.data.ann, ## solar
                                radiation data to pass
271                          data.clim.lai = climate.dat,
272                          control=list(trace=6,REPORT=10,maxit=1000,
                                follow.on=T,kkt=F,
273                                       parscale=scales),
274                          hessian=F,
275                          method=c("Nelder-Mead","CG","BFGS"))
276
277 stopImplicitCluster()
```