Determining South Mississippi forest susceptibility to windthrow and shear damage in a hurricane environment through data mining of meteorological, physiographical, pedological, and tree level data

Jared Seth Allen

DETERMINING SOUTH MISSISSIPPI FOREST SUSCEPTIBILITY TO WIND THROW AND SHEAR DAMAGE IN A HURRICANE ENVIRONMENT THROUGH DATA MINING OF METEOROLOGICAL, PHYSIOGRAPHICAL, PEDOLOGICAL, AND TREE LEVEL DATA

By

Jared Seth Allen

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in Geosciences
in the Department of Geosciences

Mississippi State, Mississippi

December 2009

DETERMINING SOUTH MISSISSIPPI FOREST SUSCEPTIBILITY TO WIND THROW AND SHEAR DAMAGE IN A HURRICANE ENVIRONMENT THROUGH DATA MINING OF METEOROLOGICAL, PHYSIOGRAPHICAL, PEDOLOGICAL, AND TREE LEVEL DATA

By

Jared Seth Allen

Approved:

William Cooke III
Professor of Geosciences
(Director of Thesis)

David L. Evans
Professor of Forestry
(Committee Member)

Zhaofei Fan
Assistant Professor of Forestry
(Committee Member)

Jamie Dyer
Associate Professor of Geosciences
(Committee Member)

Chris Dewey
Graduate Student Coordinator
Department of Geosciences

Gary L. Myers
Professor and Dean
College of Arts and Science

Name: Jared Seth Allen

Date of Degree: December 11, 2009

Institution: Mississippi State University

Major Field: Geosciences

Major Professor: Dr. William H. Cooke III

Title of Study: DETERMINING SOUTH MISSISSIPPI FOREST SUSCEPTIBILITY TO WIND THROW AND SHEAR DAMAGE IN A HURRICANE ENVIRONMENT THROUGH DATA MINING OF METEOROLOGICAL, PHYSIOGRAPHICAL, PEDOLOGICAL, AND TREE LEVEL DATA

Pages in Study: 178

Candidate for Degree of Master of Science

An estimated 39 million m^3 of timber was damaged across the Southeast Forest District of Mississippi due to Hurricane Katrina. Aggregated forest plot-level biometrics was coupled with wind, topographical, and soil attributes using a GIS. Data mining through Regression Tree Analysis (RTA) was used to determine factors contributing to shear damage of pines and wind-throw damage of hardwoods. Results depict Lorey's Mean Height (LMH) and Quadratic Mean Diameter (QMD) are important variables in determining the percentage of trees and basal area damaged for both forest classes with sustained wind speed important for wind-throw and peak wind gusts for shear. Logistic regression based on stand damage classification compared to RTA revealed LMH, stand height to diameter ratio, and sustained wind variable concurrence. Reclassification of MIFI plot damage calls based on percentage of trees damaged increased predictability of wind-throw and shear classification

DEDICATION

This Thesis is dedicated to my parents, James and Sandy Allen who have always firmly believed and encouraged me from the beginning. Without their moral and financial support for my education, I would not be the person nor the student of life I am today. I also dedicate this writing to my sister, Christine and my brother, Erek for without them, my life and college years would be less rich with fond memories and great moments. Also part of those great memories is my best-friend, my wife to be, Heather Eschete. I am but dust in the wind without you in my life and I am grateful for everything you are that I am not. I humbly thank and praise God for his grace and all that he has provided for me.

ACKNOWLEDGEMENTS

. I would like to express my sincere gratitude to my major thesis advisor, Dr. William H. Cooke for his impartation of wisdom and guidance for not only my thesis project but to life's idiosyncrasies. I have learned through Dr. Cooke, the more answers are searched for, the more questions arise along the path and sometimes the best place to find answers is being contemplative. Furthermore, I thank Dr. Cooke for my Research Assistantship (RA) through the GeoResource Institute (GRI) and the Department of Geosciences through my graduate career. I also wish to acknowledge Dr. Joseph (Zhaofei) Fan of the College of Forest Resources for his assistance and availability for implementation of R-code for CART Modeling. I thank Dr. David Evans, also of the College of Forest Resources, for his input and suggestions of tree-level variables to be tested to develop multi-faceted results that could be used in both forest management as well as emergency management. I also thank Dr. Jamie Dyer for his expertise and contributions relating to stochastic boundary layer dynamics in a hurricane environment and the DeGrib program

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Hurricane Katrina caused unprecedented damage to the natural and built environment along Mississippi's Gulf Coast. In addition to the oppressive human toll, the strength and size of Katrina devastated the timber industry of Mississippi. Timber losses alone tallied $1.3 billion, approximately worth two years annual harvest for the state of Mississippi (MFC 2005). Severe forest damage was concentrated within 60 miles of the coast while moderate to light damage extended over 150 miles inland. Forests hit the hardest were located in portions of southern Mississippi that comprise the Southeast Forest District.

The Southeast Forest District was inventoried in 2006 by the Mississippi Institute of Forest Inventory (MIFI). From the inventoried data, Hurricane Katrina was responsible for damaging over 500,000 acres (82%) of hardwood, 1.3 million acres (49%) of pine and just under 100,000 acres (49%) of mixed forest. MIFI field crews collected and measured several tree metrics and noted damage types incurred to each tree located on established plots across the Southeast Forest District. Tree metrics and categories inventoried include: species group, product, diameter at breast height (DBH), total height, type of tree damage, and base to live crown height. From these individual tree metrics, basal area per plot, Lorey's mean height, trees per acre, and quadratic mean diameter (QMD) were calculated and aggregated to stand level. These tree and stand metrics are knows as biotic factors. These biotic measurements were coupled with the storm intensity, topography, and soil data of the region and known as abiotic factors. The interrelationship between the biotic and abiotic factors creates a heterogeneous pattern of forest damage across the landscape. Understanding the relationships among biotic and abiotic factors to forest damage will help both forest managers and emergency managers to better mitigate potential future forest damage from similar events.

Damaged and downed timber increases the risk of wildfire occurrence and insect infestation (Everham and Brokaw 1996, Cooke et al. 2006). The relative importance of biotic and abiotic factor interrelationships as contributors to wind-throw of hardwoods and shear of pines is investigated for the purpose of modeling the spatial distribution of vegetative debris. Emergency response and recovery decisions regarding effective vegetative debris removal could be enhanced via these spatial models. Emergency managers in charge of debris removal could assess damage via potential damage spatial distribution and optimize allocation of personnel and equipment to dispose of vegetative debris in a more efficient manner.

CHAPTER 2

LITERATURE REVIEW

The focus of this literature review is to investigate the existing state of knowledge of biotic and abiotic factors that contribute to the mechanics of wind-throw occurrence for hardwoods and shear occurrence for pines during hurricane conditions. First, literature of the physical mechanics of wind-throw and shear are examined at the individual tree-level and stand-level. Secondly, literature on differing biomechanical and physical properties of pines and hardwoods that influence unique damage occurence are identified and reviewed. Next, literature of biotic and abiotic factors that predispose hardwoods to wind-throw occurrence and pines to shear occurrence are reviewed at the landscape scale. Numerous studies have been published across wide-ranging geographical locations, forest types, and tree species. Methods and results of these studies that relate to forest damage in southeastern Mississippi generated by Katrina are reviewed.

Physical Forces Applied to Trees during Hurricanes:

Though it might initially appear that the process by which wind blows a tree over or stem breakage occurs is simple, there are applied and resistive forces at work that determine each respective damage type's probability of occurrence (Figure 2.1). Wind-throw occurs when the applied lateral forces on a tree are transmitted down the trunk to create a torque force that exceeds the resistance to turning of the root/soil plate (Stathers *et al*. 1994; Moore 2000). Shear occurs when a tree is subject to lateral forces that exceed the stem strength but that are not strong enough to dislodge or break the roots and roll the root ball (Putz *et al.* 1983). Two applied lateral forces that cause wind-throw or shear are the force of the wind on the crown and stem and the force of gravity acting downward on the crown and stem during dynamic swaying (Trousdell *et al.* 1965, Fredericksen *et al.* 1993, Stathers *et al*. 1994, Everham and Brokaw 1996, Peltola 2006). These two applied forces act in the same manner on all trees (Peltola 2006). However, it is the different biotic and abiotic factors present that create differing damage types. For simplistic reasons, these physical relationships are best described on the individual tree-level that can then be applied to the stand-level (Stathers *et al*. 1994).

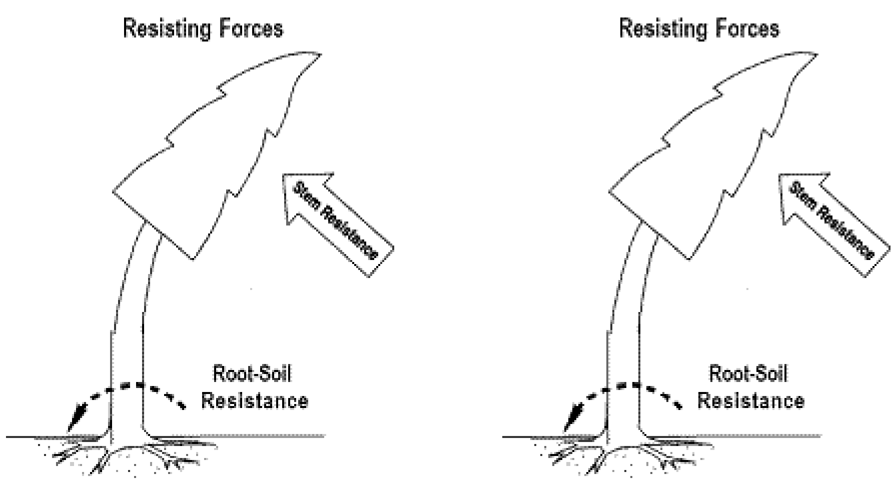

Figure 2.1 Applied and Resistive Forces on Trees from Wind Loading

Individual Tree-Level Physical Forces and Dynamics:

Peltola (2006) and Stathers et al. (1994) describe the following physical equations that subsequently determine the inherent susceptibility of trees to wind-throw or shear. The first lateral force that contributes to the overall torque is a function of wind on the crown and stem at i-th increments of height (Eq. 2.1):

$$F_{iw} = (\rho A_i C_{Di} u_i^2)/2 \qquad \text{Eq. 2.1}$$

where ρ is the density of the air, A_i is the streamlined projected area of the crown and stem perpendicular to wind flow (in square meters) at the i-th increment of height, C_{Di} is the drag coefficient of the crown, and u_i is the wind speed at height (i) above the ground. The second lateral force of gravity increases once substantial bending and swaying of the tree occurs (Peltola 2006). The horizontal displacement of the crown and stem from its vertical axis translates into a gravitational force acting to pull the tree downward (Eq. 2.2).

$$F_{ig} = m_i x_i g \qquad \text{Eq. 2.2}$$

where m_i is the mass of the i-th height increment, x_i is the horizontal displacement from the vertical, and g is the gravitational acceleration (9.8 m/s) (Stathers *et al*. 1994). Based on Eq. 2.1 and Eq. 2.2, several biotic and abiotic factors that affect wind-throw potential can be deduced. Biotic factors that influence the mechanics of wind-throw include crown density, crown size, stem elasticity, wood density, and stem mass, while abiotic factors are wind speed, wind duration, and precipitation intensity. Summing the wind and gravitational forces for each i-th height increment of the tree equals the torque, or bending moment of the tree stem (Eq. 2.3).

$$BM = \Sigma(F_i h_i) \quad \text{Eq. 2.3}$$

where h_i is the height of the i-th increment and F_i is the summed horizontal force (from wind and gravity) on that increment (Stathers *et al*. 1994). As a tree becomes taller, the bending moment along the main stem increases and the tree becomes increasingly prone to wind-throw or shear. Thus, the biotic factors of height, diameter, and the height to diameter ratio of a tree have a role in wind-throw occurrence (Curtis 1943, Everham and Brokaw 1996, Merry *et al.* 2009). These biotic factors have been found to be significant in many studies (Jacobs 1936, Coutts 1983, Petty and Swain 1985, Moore 2000, Meunier et al. 2002, Fredericksen 1993). Opposite of bending moment, an applied force, there are resistive forces that attempt to keep the tree up-right.

Resistive forces counteract the applied wind and gravitational forces acting on the crown and stem (Peltola 2006). A tree must rely on its root anchorage to resist these forces (Frederickson 1993, Stathers *et al*. 1994, Ray and Nicoll 1998, Dupey 2005, Peltola 2006). Resistance to up-rooting or stem breakage is estimated based on tree pulling experiments by determining regressions between resistive bending moments and various tree characteristics (Frederickson 1993, Ray and Nicoll 1998). These biotic factors include root depth, root mass, weight of the root-soil plate, and stem mass (Stathers *et al*. 1994, Peltola 2006). Although, the root-soil strength and root-soil mass relationships to wind-throw and shear occurrence have been measured through static winching/pulling tests, relationships between root-soil properties and wind-throw and shear can not be tested at the landscape scale. In static winching/pulling tests, only a select number of trees are tested at a specific site and few studies have correlated differing soil types to differing critical bending moments (Ray and Nicoll 1998, Dupuy *et* *al*. 2005, Cucchi *et al.* 2004). However, proxy measurements that affect root-soil strength and root-soil mass are more readily studied across the landscape.

Proxy measurements of the root-soil interface include soil depth, soil texture, soil bulk density, and soil infiltration rate (Petty and Swain 1985, Ruel 1992, Frederickson 1993, Cucchi *et al.* 2004, Peltola 2006). Trees growing in deep, well-drained soils produce much larger root systems than those in soils where anaerobic conditions, high bulk density, stoniness, hard pans, or near-surface bed-rock restrict root development (Touliatos and Roth 1970, Stathers *et al*. 1994). Soil texture varies in the amount of sand, silt and clay present which can also influence the root development (Trousdell 1965). Trees growing in sandy well-drained soils are generally deeply rooted than trees growing in clay soils, soils with inhibiting clay layers, or soils with high water tables. When soils become saturated, the shearing strength decreases dramatically with the loss of adhesion and cohesion of the soil granules with the roots. Ray and Nicoll (1998) performed winch/pull tests on Sitka Spruce (*Picea sitchebsis*) and found on soils with higher water tables, the critical bending moment was two to three times less than those winched/pulled on well-drained, dry soils.

Sandy, loamy, and silty soil textures have greater pore space allowing for less restriction of root growth than clay soils. Even though sandy and sandy loam soils are well drained (i.e. larger pore spaces), this allows for faster saturation of the soils in high precipitation events (i.e. hurricanes). Trousdell (1965) observed 30% of trees on sandy profiles where wind-thrown, opposed to only 5% of trees located on silt and clay profiles. Other abiotic factors that increase wind-throw or shear include the presence of pathogens which induce weakness of the stem or root rot (Thompson 1983, Webb 1988, Putz and

Sharitz 1991, Everham and Brokaw 1996). These factors at the individual-tree scale all influence the type and amount of damage sustained.

At the individual tree scale, there are many biotic and abiotic variables that can be identified that modify wind-throw potential based on the physical metrics of a tree set in motion by wind (Cremer 1982, Putz *et al.* 1983, Petty and Swain 1985, Frederickson 1993, Lidemann and Baker 2002, Kupfer 2008). Available literature suggests that important biotic factors include: crown density, crown size, stem elasticity, stem density, stem mass, height, diameter, height to diameter ratio, root depth and root mass. Abiotic factors identified include: wind speed, wind duration, total precipitation, precipitation intensity, soil texture, soil infiltration rate, and soil bulk density. Interaction among all these variables predisposes an individual tree to the type of damage incurred (Everham and Brokaw 1996, Merry *et al.* 2009). Yet, often trees are concentrated in forests and plantations, causing additional biotic and abiotic dynamics to affect the pattern of tree damage (Foster and Boose 1992).

Stand-Level Physical Forces and Dynamics:

A forest stands' ability to withstand strong winds such as hurricanes has been found to be associated with a number of topographical, forest, and silvicultural factors (Wang and Xu 2008). Topographical properties including elevation, slope, aspect, and surface roughness determine the forest stand's relative exposure to wind (Lugo 1983, Foster 1988, Foster and Boose 1992, Boose and Foster 1992, Everham and Brokaw 1996, Lindemann and Baker 2002, McNab *et al.* 2004). Forest factors include species composition, average stand height, average stand density, and natural canopy gaps (Smith *et al.* 1987, Stathers *et al*. 1994, Everham and Brokaw 1995, Xi 2008). Silvicultural management of trees can increase or decrease the susceptibility of individual tree damage through thinning, pruning, or edge feathering (Alexander 1964, Fredericksen 1993, Stathers *et al*. 1994, Stanturf 2007).

The topographical position of trees relative to terrain features such as slopes, ridges, and valleys can either protect or expose them to strong winds. Governing factors of topography include the elevation, slope, aspect, surface roughness and their relative location to topographical features. Observed forest damage has typically been greater on windward sides of slopes than on leeward sides (Bellingham 1991, Reilly 1991, Walker 1991). However, forests on the leeside are still prone to turbulent eddies that develop (O'Cinneide 1975). Anderson (1954), Gloyne (1968), and Foster and Boose (1992) found lee-slopes greater than 10% were protected (Everham and Brokaw 1996). Higher elevations are generally susceptible to higher wind speeds, thus greater damage (Boose *et al*. 1994). However, trees located in river valleys have also suffered high amounts of damage due to wind funneling (Alexander 1967, Walker 1991). Steep slopes correlate with more wind-throw, but not necessarily more overall damage (Putz *et al.* 1983). This could be a result of shallow soils resulting in shallow rooting which enhances the chances of wind-throw. Exposure is a complex relationship of aspect, slope, surface roughness, and topographic position, while damage itself can reflect topographic correlates with species characteristics (Everham and Brokaw 1996).

There are several aspects of forest stand conditions and composition that influence the severity of wind damage across a landscape. These aspects include even-aged vs. mixed-aged stands, single-species vs. mixed-species stands, and maturity of various trees within the stand (Everham and Brokaw 1996). Even-aged stands produce a more uniform canopy height which does not readily allow for wind gusts to penetrate deep into the canopy. With all trees being relatively the same height, no one tree located within the interior of the stand is more exposed than any other given elevation, slope, and aspect stay constant across the stand. Whereas in mixed-age forest stands, the taller trees have greater exposure to winds and are more likely to experience damage. The forest composition of mixed-aged stands are made up of dominate and co-dominate species.

Dominants are early successional species that establish themselves more readily after a disturbance while co-dominants exist in the understory (Everham, and Brokaw 1996, Xi 2008). Dominants are subjected to greater exposure by their height, thus creating greater bending moments along their stems and increasing their chances of wind-throw or shear. Co-dominants are typically less mature, with less exposure to the wind (Everham and Brokaw 1996). Within the Southeast U.S., dominate species within mixed-aged stands typically includes Pine species (*Pinus*), while late-succesional species present in the understory are hardwood species (Edeburn 2009, Xi 2008) (Figure 2.2). Pines species are more shade intolerant and grow at a faster rate than many other hardwood species (Martin and Gower 1996). Pines are typically dominates of a given forest stand and are exposed to higher winds while hardwoods are generally intermediates or under-story in managed pine stands. Mortality tends to be higher in early sucessional species, thus differences in susceptibility among broadleaf species may be due to their successional class (Webb 1986, Zimmerman et al. 1994, Everham and Brokaw 1995). Hardwoods are dominants in lower river basins with wetter soils where pines grow better in sandier, well-drained soils. Without clearing and prescribed burning of managed forests, hardwoods would have a much larger aerial extent in southern Mississippi. The stand density of dominants and co-dominants also contributes to the relative windfirmness of each tree and the stand overall.

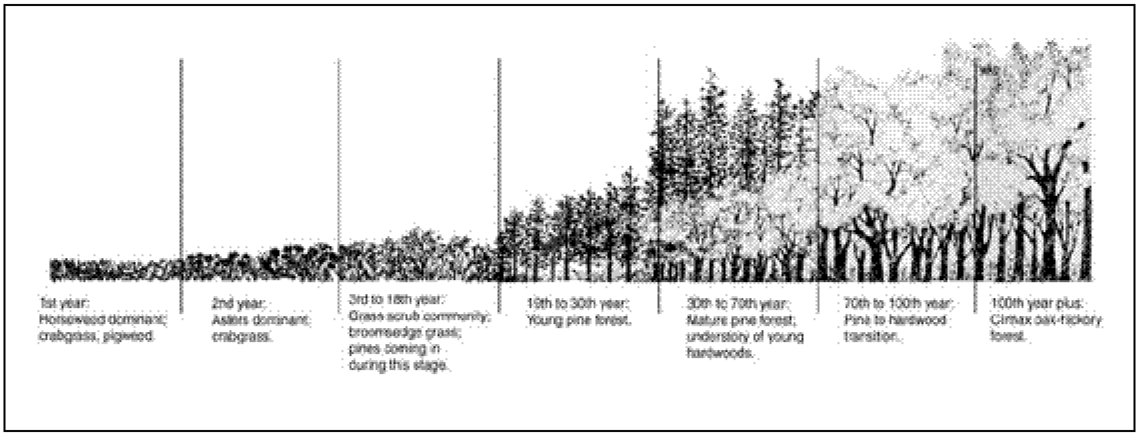

Figure 2.2 Succession of the Gulf Coastal Plain.

Greater density stands tend to have less damage due to interlocking root-systems, inter-crown damping during swaying, and the affect of dense crowns reducing wind penetration into the stand (Stathers *et al*. 1994). However, the individual tree that is part of a dense stand is not wind-firm in isolation because of restricted rooting, and higher height to diameter ratios due to competition for natural resources (Everham and Brokaw 1996). A dense stand is wind-firm as a whole but if a few trees are wind-thrown or sheared and the canopy is opened up, subsequent tree failure is more likely to happen (Merry et al. 2009). One location where tree failure is most likely to occur is along the edge of forest stands. Stanturf *et al.* (2007) determined the Modulus of Rupture (MOR) threshold for stem breakage of longleaf and loblolly pine occurs at lower wind speeds in open stands (7.5m spacing) than in more dense stands (2.5m spacing) (Figure 2.3). Conversely, tree height (i.e. exposure), was the primary factor in determining stem failure of interior longleaf and loblolly pines (Figure 2.4). Figures 2.3 and 2.4 also depict longleaf pine having a greater resistive bending moment than loblolly pine, thus agreeing with Touliatos and Roth (1970), Hughes (2006), and Oswalt and Oswalt (2008).

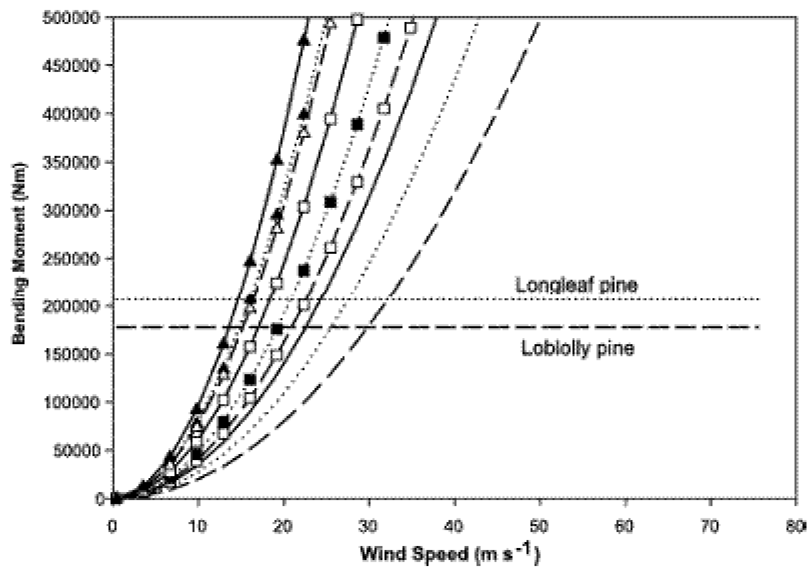

Figure 2.3 Bending Moment expressed as Newton Meters (Nm) of trees at the Stand Edge as a function of Wind Speed.

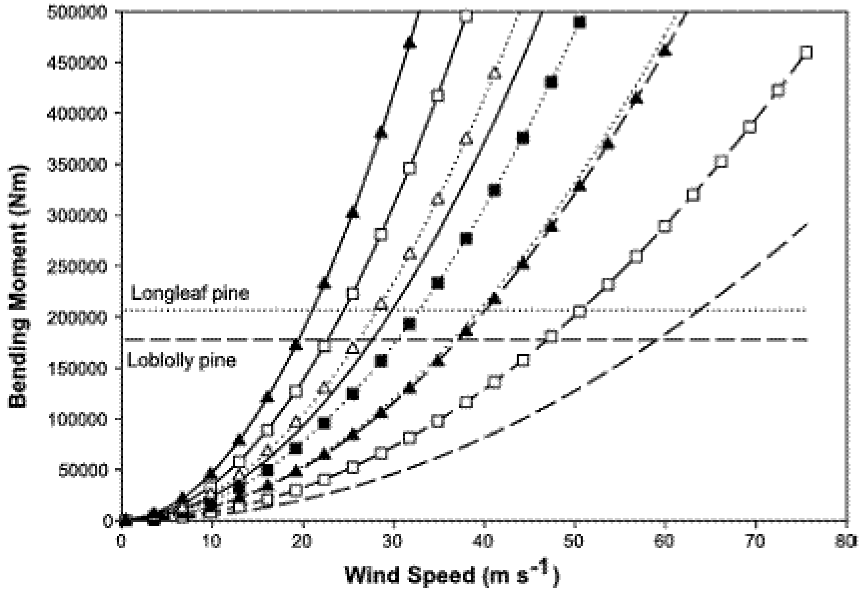

Figure 2.4 Bending Moment expressed as Newton Meters (Nm) of trees at the Stand Interior as a function of Wind Speed.

Trees along the edge of stands have greater exposure to high winds during hurricanes than trees located deeper in the stand. The adaptive growth response of trees to mechanical forces like wind is known as thigmomorphogenesis and causes pre-conditioning of trees (Jaffe 1973, Meng *et al.* 2008). The greater exposure to natural wind regime has been known to pre-condition trees along edge boundaries (Coutts 1986, Cucchi 2004). Pre-conditioning naturally builds up the stem and root strength over time in the direction of the predominate wind (Stokes 2002). Cucchi (2004) investigated the root anchorage differences of maritime pine trees that grew along the edge of forest stands and those which grew towards the center through static winching. Results indicated the soil-root plate of edge trees was three times more asymmetric, with a soil-root plate 30% wider on the windward side and nearly two times larger that those in inner trees. Concurring, Stokes (1995a,b) subjected Sitka spruce to wind 12 hours a day in a wind tunnel experiment for one growing season and found root growth on the windward and leeward side were double compared to growth perpendicular to the wind direction. Leeward roots were thicker than roots found elsewhere in the rooting structure, whereas windward roots were more branched with increased surface area. Thicker and longer roots add to the over all root-soil mass which increases the resistance force to counteract the bending moment placed on a tree during high winds. This, however, places more force along the stem and favors shear damage to occur. Stanturf *et al.* (2007) states the threshold for damage due to stem breakage is much lower along stand edges rather than in the interior of the stand and conversely tree height (i.e. exposure), was the primary factor in determining stem failure of interior longleaf and loblolly pines (Figure 2.3 and Figure 2.4 respectively). If the strong wind differs by 90 degrees from the predominant wind regime, the tree will be more prone to wind-throw (lack of root support). Wind-throw associated with hurricane Katrina most frequently occurred along roads, powerline corridors, open fields and other landscape corridors (Stanturf *et al.* 2007). Wind-throw was also found frequently in forested areas with edges created by adjacent stands with varying tree heights and densities (Marion *et al.* 2005). To assuage the abrupt surface roughness experienced along forest stand edges, proactive silvicultural measures have been taken to mitigate wind-throw and shear through silivicultural treatments.

Silviculture refers to the art and science of controlling the establishment, growth, composition, health, and quality of forests to meet diverse needs and values of many landowners, societies, and cultures (Boone 2005). Specific silvicultural procedures can either exacerbate or mitigate losses of timber in a catastrophic wind event. Such procedures are thinning, pruning, and edge feathering (Stathers *et al*. 1994). The temporal proximity of thinning to the catastrophic wind event greatly impacts the amount of damage incurred.

Thinning is a necessary step in the management of either natural or planted forestland. The three goals of thinning are to promote growth of the residual trees, to promote health of the residual stand and provide an economic return to the landowner (DeLoach 2007). Initially a site may be able to support 600-700 trees per acre, however over time as the trees become larger and use more nutritional resources, the health and growth rates of the trees are affected. To protect both the investment and the forest ecology, landowners and forest managers thin the forest stand. The first thinning of pines in Mississippi typically occurs between 12-18 years of age when they reach an average height of 40 feet (Traugott 2002). In order to maintain growth rates, about 40 percent of the trees are removed (Traugott and Dicke 2006). Since dense stands compete for resources, individual trees have high height-to-diameter ratios. When the canopy is opened up through thinning, the individual trees are highly susceptible to damage during the next high wind event. Trees become more wind-firm after a few years of exposure as they develop reaction wood in response to swaying and take up more nutrients (Stathers *et al*. 1994, Meng *et al.* 2008).

Recently thinned hardwood and pine stands suffer the greatest amount of damage during a hurricane (Alexander 1964, Trousdell *et al.* 1965, Touliatos and Roth 1970, Cremer *et al.* 1977, Stathers *et al*. 1994, Lindemann and Baker 2002). Cremer *et al.* (1977) studied the impact of catastrophic wind on plantations in Australia with stands thinned within 5 years before the disturbance having 22% wind-throw damage opposed to 0.2% for stands greater than 5 years. In addition, un-thinned stands downwind of clearcuts had 38% windthrow, but if these downwinds stands where thinned less than 6 months prior, wind-thrown damage increased to 88%. Cremer *et al.* (1977, 1982) suggest that five years is required to recover thinning, which is supported by Weidman's (1920) results that two-thirds or more of damage occurred in the first five or six years after thinning. Alexander (1964) reviewed previous work that has shown windfall losses to be increased by any kind of partial cutting. Toursdell (1965) noted the presence of an impervious clay layer overtop sandy soils at the surface lead to severe wind-throw of loblolly pines in recently thinned stands. When Camille struck the Mississippi coast in 1969, Touliatos and Roth (1970) found that recently thinned loblolly pine stands with little taper were severely sheared, while open pine stands fared better. Thinning of forest stands can have severe consequences if they coincide with a hurricane. Cremer *et al.* (1982) suggest margins of stands should be thinned and pruned to make them more wind-permeable and therefore prevent extreme turbulence at the edges of stands. In this manner, edge thinning, known as edge feathering can moderate the abrupt forest edge interface to reduce turbulent eddies.

Stabilization treatment of the forest edge has been shown to reduce the amount of wind-throw and shear damage (Stathers *et al*. 1994, BCMF 2003). The frequency of strong winds and damage occurrence of forests in British Columbia, Canada have lead to the development of wind-throw handbooks by Stathers *et al.* (1994) and BCMF (2003). To enhance the stability of forest edges, Stathers *et al.* (1994) and BCMF (2003) developed recommended procedures for edge feathering and crown pruning. The procedures can be different depending if the edge is composed of mixed-age species of varying height or for even-aged stands of uniform height. In multi-storied stands, where smaller, windfirm trees can be left, removal of taller trees at the stand edge may help crate a profile that gradually lifts the wind up over the edge. For even-aged stands, the edge can be thinned by removing dominants having large crowns, large height-to-diameter ratios, and physical deformities or visible disease. In either case, feathering is best if confined to the forest tree length into the stand from the edge and not exceed 20-30% of the total basal area.

Another approached described by Stathers *et al.* (1994) and BCFM (2003) is crown pruning. Crown pruning occurs when climbers prune back the larger branches and reduce the surface area of the crown. As with edge feathering, the trees selected for crown pruning depends on the relative windfirmness of each tree to begin with. Typically, dominants with large crowns and trees with larger height-to-diameter ratios are pruned to reduce the force of the wind on the crown. Removal of 20-30% of the crown is suggested by Stathers *et al.* (1994). Primary and secondary branches that do not come in contact with other trees are removed while branches in close proximity to other trees are kept. Branches proximal to other trees are kept because of the inter-crown dampening effects from the dynamic swaying of the tree.

The physical forces of a hurricane place great stress on the individual tree and the forest stand. The type of damage depends on the complex interaction of biotic and abiotic variables identified through the physical relationships of individual trees and stand dynamics. Yet, damage observed through post-storm inventories, static winching/pulling, and wind tunnel experiments reveal hardwoods are more likely to wind-throw while pines are more likely to shear (Curtis 1943, Petty and Swain 1985, Putz *et al.* 1983, Foster 1988a, Everham and Brokaw 1996, Hook *et al.* 1996, Peltola *et al.*

2000, Rodgers *et al.* 2006, Merry *et al.* 2009). However there have been cases were pines where wind-thrown more than hardwoods (Trousdell 1965, Xi 2008). Differing species are more or less susceptible to a particular type of damage based on their inherent biomechanical and physical properties and the specific properties of the site.

Differing Biomechanical and Physical Properties of Trees

The following section reviews the differences between species groups (pine and hardwoods) that predispose each class to suffer unique damage patterns. Different biomechanical properties are inherent to each species, much less differing taxonomy classes of trees (Everham and Brokaw 1996). These biomechanical and physical properties include: tree height, weight, wood density, height-to-diameter ratios, drag coefficients, crown size, crown density, and those described above in sections 2.1.1 and 2.1.2 (Petty and Swain 1985, Everham and Brokaw 1996, Peltola *et al*. 2000, Merry *et al.* 2009). Since pines species comprise the dominant cover type across the study area, their biomechanical and physical properties that induce shear or snapping of stems are examined. The biomechanical and physical properties that induce increased susceptibility to wind-throw of hardwoods are also reviewed.

Biomechanical and Physical Properties of Pines:

Temperate, sub-tropical, and tropical coniferous forests have some definitive attributes that predispose pine species to shear rather than wind-throw (Everham and Brokaw 1996, Petty and Swain 1985, Merry *et al.* 2009). Pine species typically have greater height-to-diameter ratios as they grow compared to hardwoods which increases the susceptibility of snapping (Petty and Swain 1985). The shape of the bole and thus the taper function of pines are generally are less than hardwood species, leading to a cylindrical shape rather than a more conical shape. The cylindrical shape of most pine species places much greater strain on the stem during high wind events (Peltola *et al*. 2000, Merry *et al.* 2009). Many hardwood species have characteristically denser wood and tend to grow more slowly while pine species which grow faster resulting in softer, less dense wood (USFS 1994). Due to their high growth rate, the tree rings of pines are spaced wider, and result in decreased density of wood fibers (Rodgers *et al.* 2006). One particular study focused on the varying amount of damage between differing pine species in southern Mississippi as a result of Hurricane Katrina (Hughes 2006).

Hughes (2006) found that not all pine species are affected equally, noting loblolly pine (*Pinus taeda*) suffered more damage than slash pine (*Pinus elliottii*) and longleaf pine (*Pinus palustris*). Two plantations in Forest County, Mississippi were planted in the same year (1985) at the same tree farm, three miles apart. The plantations consisted of three different species of pine and both had been thinned in 2002. At both sites, 84% of loblolly pine (*Pinus taeda*) was damaged compared to 47.6% slash pine and 36% longleaf pine. Of the 84% damaged loblolly pines, 75.9% had snapped stems, whereas slash pine had 38.1% and longleaf pine at only 8.9%. Worth noting is that slash and longleaf pines are found more along the coast than loblolly pine suggesting slash and longleaf have adapted to a windier coastal climate and are thus less susceptible to damage. This agrees with Touliatous and Roth (1970) after they surveyed damage of during Hurricane Camille and observed loblolly had the least resistance to stem breakage followed by slash pine and longleaf pine. In Hughes (2006), longleaf pine actually had a greater percentage of wind-throw than shear (10.2% vs. 8.9%). Longleaf pine has a higher wood specific gravity (0.59) than loblolly (0.51) but is equal to slash (0.59) (USFS 1994). As noted with hardwoods, trees with higher specific gravities have greater density wood and are more likely to fail at the root-soil plate than in the main stem. The higher shade-intolerance of longleaf pine allows for better self-pruning of bottom branches than in slash pine (USFS 1994). More branches, coupled with a higher, denser crown, increases the surface area of mature slash pine compared to mature longleaf pine. This greater surface area contributes to a greater lateral force placed on the stem by the wind. Hughes (2006) suggests longleaf pine's survival success pertained to its smaller diameter and shorter height compared to loblolly and slash pine. The smaller diameter and height enabled it to be more flexible in high winds. Secondly, Hughes (2006) notes that the thinning of longleaf pine was not a traditional fifth-row thinning as in the loblolly and slash pine due to smaller size and smaller basal area. Despite longleaf pine's resiliency, the smaller size and smaller basal area make longleaf pine less desired for forest investors. Since timber value is based on weight, an investor would want to grow a large tree as fast as possible to get a quicker and greater economic return (Beckwith 1997). Therefore, in lieu of its slow growth rate compared to loblolly and slash pine, longleaf pine is not favored by forest landowners despite its hurricane resiliency.

Summarizing, conifers tend to have small taper ratios and thus larger height-to-diameter ratios that focuses more stress on the main stem rather than the soil-root plate (Putz *et al.* 1983, Petty and Swain 1985, Foster 1988a, Everham and Brokaw 1996, Merry *et al.* 2009). Taproots of conifers can provide strong anchorage when growing deep, well drained soils with no restrictive layers (Trousdell 1965). Since the root-soil anchorage strength is large compared the stem strength, pines are more predisposed to shear damage than wind-throw although it can occur. Hardwoods, however, have higher density wood and tend to be more dense when compared to wood from pine trees (Cutis 1943, Petty and Swain 1985, Everham and Brokaw 1996, Merry *et al.* 2009). The stronger stem allows for more kinetic energy from the dynamic swaying to be transferred to the ground. This in turn, places stress on the root-soil plate.

Biomechanical and Physical Properties of Hardwoods:

Catastrophic wind storms affect temperate, sub-tropical, and tropical broadleaf deciduous forests through out the world (Everham and Brokaw 1996). Yet, some conclusive biotic factors have been identified across a broad spectrum of hardwood species that render it more susceptible to wind-throw as opposed to shear (Curtis 1943, Putz *et al.* 1983, Foster 1988a, Everham and Brokaw 1996, Hook *et al.* 1996, MIFI 2006, Oswalt and Oswalt 2008, Merry *et al.* 2009). In relation to wood density, Curtis (1943) found the bending force required for hardwoods is greater than twice that required for white pine (*Pinus stobus* L.). Weaver (1989) and Walker (1992) found that higher specific gravity (greater density wood) is associated with less stem breakage. This concurs with Putz *et al.* (1983) which found trees with shorter, thicker stems and denser wood tend to uproot, whereas species with low-density wood tend to snap and have higher mortality rates. Foster (1988a) reported that species with full crowns and shallow roots are more susceptible to wind-throw than species with a vertical distribution of canopy, flexible branches, and a less-tapered shape. Several post-storm inventories in Mississippi following hurricane land-falls are consistent with the above literature with Touliatos and Roth (1970) observing densely-crowned trees to be prone to wind-throw opposed to open-foliage crowns from felled trees due to Hurricane Camille.

A more recent inventory of 1,349 one-fifth acre sample plots done by Glass and Oswalt (2006) in the 6 southern-most counties of Mississippi after Katrina showed oak-gum-cypress forests had the most basal area damaged at 42%. This was followed by oak-hickory (30%), oak-pine mix (28%), and loblolly-shortleaf pine (25%). Oswalt and Oswalt (2008) inventoried 1581 hardwoods and found more wind-related damage across the entire state than softwoods ($p<0.0001$). Oswalt and Oswalt (2008) also determined that hardwoods suffered 1.9 times more wind-throw occurrence than softwoods. Touliatos and Roth (1970) noted hardwood trees most susceptible to wind-throw included pecan (*Carya illinoensis)*, hickory (*Carya tumentosa*, *Carya glabra*), dogwood (*Cornus florida*), red maple (*Acer rubrum*), and water oak (*Quercus nigra*) across southern Mississippi. Oswalt and Oswalt (2008) surveyed southern Mississippi following Katrina and found water oak, and red maple hardwoods susceptible along with yellow poplar (*Liriodendron tulipifera)* and sweetgum (*Liquidamba styraciflua*) agreeing with Touliatos and Roth (1970). Species typical of upland sites (yellow poplar and *loblolly pine*) were among the most damaged while select bottomland species (*pondcypress*, *swamp tupelo*, and *blackgum*) were among the least damaged (Oswalt and Oswalt 2008). Pondcypress and swamp tupelo are resilient in hurricanes due to their buttressed trunks, extensive root network, and deciduous habit which greatly reduces the surface area exposed to high winds (Touliatos and Roth 1970, Putz *et al.* 1983, Peterson 2000, Gresham 1991). Pecan, yellow poplar, red maple, and water oak are more susceptible due to their large crowns and location in hydric soils along river valleys in Mississippi (Hook et al. 1996). The hydric soils of river valleys and bottomlands restrict the root depth allowing for less anchorage to occur (Everham and Brokaw 1996, Wang and Xu 2008). Similar results by

Hook et al. (1996) were found in bottomland hardwoods in South Carolina with the passage of Hurricane Hugo.

To summarize, the dense broadleaf crown of a larger and thus older deciduous hardwood given same site characteristics and higher density wood with a lower center of gravity compared to coniferous trees results in greater probability of exceeding the root-soil strength than the stem strength. The susceptibility among hardwood species to wind-throw differs based on the characteristics of that species (Touliatos and Roth 1970, Putz *et al.* 1983, Peterson 2000, Gresham 1991, Everham and Brokaw 1996, Hook *et al.* 1996, MIFI 2006, Oswalt and Oswalt 2008, Wang and Xu 2008). The tree characteristics of pines and hardwoods have been identified through small-scale post-storm inventories, static winching/pulling testes, and wind tunnel experiments. Only recently have studies used remote sensing and GIS to study biotic and abiotic variables on the landscape scale as they relate to the pattern of damage observed (Foster and Boose 1992, Jacobs and Eggen-McIntosh 1993, Boose 1994, Nix 1996, Ramsey et al. 1997, Ramsey *et al.* 2001, Ayala-Silva and Twumasi 2004, McMaster 2005, Murrah 2007, Kupfer 2008, Xi 2008, Wang and Xu 2008).

Review of Statistical Methodologies to Depict Threats to Forests

Forests are subject to multiple threats which can jeopardize their health, ecology, biodiversity, and resources (SRS 2006). Such threats can be natural or anthropogenic. Natural disturbances include wildfire, catastrophic wind events, drought, ice storms, insect infestation, fungal/pathogen outbreaks, and invasive plants. Anthropogenic disturbances include pollution, forest fragmentation, and urbanization (SRS 2006). Sections 2.1 and 2.2 of this literature review identified many biotic and abiotic variables that affect the amount and type of damage incurred to hardwood and pines during a hurricane.   A large database is required in order for hidden patterns of these variables to be extracted and understood. A method that extracts hidden patterns in large amounts of data is known as Knowledge Discovery in Databases (KDD) or simply data mining (Imberman 2000). The concept of data mining is based off well-known mathematical algorithms and techniques but its application is relatively new due to higher computational power, data storage logistics, and cost (Alexander 1997). Existing data mining technologies used to identify causal variables of each respective threat to forests using predictive analytics can be categorized into two types: (a) Artificial Neural Networks (ANNs) and (b) Classification and Regression Trees (CART) (Lindemann and Baker 2002, Hanewinkel *et al.* 2004, Miller et al. 2005, Kupfer 2008). These technologies have been combined with standard linear regression, binary logistic regression, and forward stepwise logistic regression to further depict the subsequent roles of biotic and abiotic variables and their predictive ability for forest damage (Hanewinkel *et al.* 2004, Wang and Xu 2008). Each data mining type has its advantages and disadvantages.

Artificial Neural Networks (ANNs)

An Artificial Neural Network (ANN) is composed of interconnected processing elements known as nodes in parallel structure working in unison to solve specific and complex problems (Stergiou and Siganos 1996). A typical ANN consists of three layers including: input, hidden, and output layers (Stergiou and Siganos 1996, Rounds 2002) (Figure 2.3.1). Stergiou and Siganos (1996) further state the input layer contains multiple input units and their respective raw information and unit of measure. The raw information originates from a large database compiled by the user. Each input layer unit then connects to each hidden layer. The activity of each hidden unit is determined by the activities of the input units and the weights on the connections between the input and hidden weights. The weights are determined by the amount and quality of training data that goes into the input layers. Since the system is artificial, the interaction of the input units and the weights allows the hidden units to freely interpret their own representations of the input. Once the hidden units have been determined, they combine to form the output layer. The behavior of the output units also depends on the activity of the hidden units and the weights between the hidden and output units. The user only defines the raw information of the input units with the rest of the process being automated. This leads to the advantages and disadvantages of ANNs (Silva 2003) (Table 2.1).

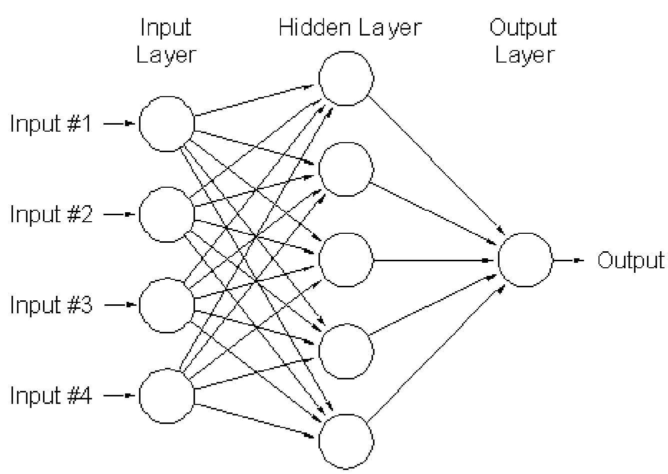

Figure 2.5 Schematic Diagram of a Neural Network

Table 2.1 Advantages and Disadvantages of ANN

ANN Advantages	ANN Disadvantages
Capable to learn non-linear and very complex relations	Long training time requirement and possible over-fitting of data
Ability to handle noisy data	Limiting Analytical Abilities: Can not identify significance level of different input variables.
Easy to use, implement, and integrate results in a GIS	Inconsistent results due to the initial weights and learning parameters
Good predictive capabilities	Difficult to understand internal behavior

Artificial Neural Networks have been used to model natural and anthropogenic forest threats. Primarily, natural threats such as wildfires, insect outbreaks, and climate change have been studied using ANNs (Gardner and Dorling 1998, Ozesmi and Ozesmi 1999, McCormick 2000, Schmildt 2001, Miller *et al.* 2005). Rarely, have ANNs been utilized to model forest damage from a hurricane (Hanewinkel *et al.* 2004). Hanewinkel *et al.* (2004) compared ANN to binary logistical regression using several biotic and abiotic factors to classify forests susceptible to wind damage. Biotic and abiotic factors included stand age, tree species, dominant height, and aspect. Results of Hanewinkel *et al.* (2004) did not reflect the actual interactions of the biotic and abiotic variables but focused on the accuracy of the differing methodologies using the mean squared sensitivity error (MSSE). ANNs had lower MSSEs in four of the five datasets tested. Hanewinkel *et al.* (2004) concluded however, that due to the black-box nature of ANNs, identification of the hidden units and their weights could not be determined. Thus, the significance level of the different input variables (biotic and abiotic) could not be identified. If one is to depict the variables responsible for occurrence of wind-throw in hardwoods or shear of pines, and their interaction, then ANN methodology cannot be used. This is due to the lack of depicted variable weights necessary for GIS model implementation. Silva (2003) only uses the predicted ANN output as a GIS layer without knowing the underlying contributions of each variable. There are, however, other data mining techniques that generate decisions based on the classification or regression of input variables of a large dataset. These data mining techniques use predictive analytics that examine current and historical data to make predictions about future events (SAS 2009).

Predictive Analytics and Data Mining

Predictive analytic modeling can be categorized in three types: fixed models, parametric models, and non-parametric models (Lynn 2006). Fixed models are used when the exact input-output relationships are known and thus cannot be used to derive relationships between biotic and abiotic variables to forest damage. Parametric models are used when a parametric mathematical relationship can be obtained. Many small scale inventories and static/pulling studies used parametric statistics (linear and logistic regression) to determine the relationships of measured or derived tree characteristics and observed damage (Everham and Borkaw 1996). Non-parametric models are used when the relationships between the input variables and the associated output are not well understood (Lynn 2006). Different non-parametric algorithms include Multivariate Adaptive Regression Splines (MARS) and Classification and Regression Tree (CART) analysis. MARS and CART analysis are regression procedures that make no assumption about the underlying functional relationship between the dependent and independent variables (Statsoft 2008). Yet, they do have their respective advantages and disadvantages.

Multiple Adaptive Regression Splines (MARS)

MARS is a flexible data-mining tool that automates the building of predictive models for continuous and binary dependent variables (Salford 2009). Given a driving variable (independent variable), and a set of candidate predictor variables, MARS can determine the interactions between predictor variables (Salford 2009). An important concept associated with MARS is that it contains hinge functions or knots where one local regression model gives way to another (Statsoft 2008). These hinges allow MARS to operate as a multiple piecewise linear regression where each breakpoint, or hinge, defines the region of application for a particular linear regression equation between the dependent and independent variables (Statsoft 2008). At the end of the model-building process, the straight lines at each node are replaced with a smooth function called a spline. MARS is a computationally intensive procedure with CART being much faster (Lynn 2006). The main disadvantages of MARS compared to CART are the interpretability of the interacting predictor variables and the cross-validation procedure. There is no graphical output of MARS showing the interaction of all the variables. The splines are only models for each predictor variable and the interactions are displayed using mathematical equations. MARS software also requires manual cross-validation the individual predictor's over-fit splines of the training data whereas CART automatically performs a 10-fold cross-validation of the decision tree. Since CART cross-validates the decision tree automatically, classification accuracy R^2 and cross-validation R^2 values are generated for the interaction of the independent variables. Additional advantages of CART over MARS are that CART is not affected by outliers, collinearities, or distributional error structures that affect parametric procedures.

Classification and Regression Tree (CART) Analysis

Data-mining statistics utilizing Classification and Regression Tree (CART) analysis have been more readily used than ANNs to identify important variables and their interaction with each other (Lindemann and Baker 2002, Kupfer 2008). Additionally, CART analysis has been coupled with binary logistic regression and stepwise logistic regression to predict the probability of forest damage (Lindemann and Baker 2002, Oswalt and Oswalt 2008, Xi 2008, Wang and Xu 2008). CARTs are analytic procedures for predicting the values of a categorical or continuous response variable given a categorical or continuous response variable (Statsoft 2009). There are however, several analogous procedures that emulate CART analysis, each having advantages and disadvantages. Such procedures include: TreeNet, RandomForests, and C-5/Cubist applications. All three applications utilize the same concept behind CART analysis by using decision trees. First, TreeNet, generates thousands of small decision trees built in a sequential error-correcting process to converge to an accurate model (Salford 2009). TreeNet models are very complex and thus the software generates a number of reports that must be meticulously analyzed. The random forests procedure is similar to TreeNet procedure in that many decision trees are grown. However, the procedures diverge because random forests grow classification trees and not regression trees (Breiman and Cutler 2009). Only categorical data can be used for the random forests procedure. In addition, each classification tree is fully grown and not pruned, tending to cause over-fitting of the data. Lastly, C5/Cubist most closely approximated CART methodology by only growing single classification and regression trees. The decision tree program known as C5 has been used readily by the United States Geological Survey (USGS) for inventorying agricultural land use and large area forest mapping including the National Land Cover Database (NLCD) (McNairn *et al.* 2004, Huang *et al.* 2001). Advantages of C5, like CART include: being non-parametric, can handle both continuous and nominal data, generates interpretable classification rules, and is fast to train (USGS 2003). Both the software used in this study, [R 2.8.0], and C5 would have likely produced similar results because of similar methodology. C5/Cubist was not used simply because it was not readily available.

CART was first introduced by Breiman (1984) allowing for an easy to interpret method of data mining using both categorical and continuous variables. When the response variable of interest is categorical in nature, the technique is referred as classification tree analysis. Conversely, if the response variable is continuous, then the method is referred as regression tree analysis (Statsoft 2009). CART analysis is a form of binary recursive partitioning that forms a decision tree (Lewis 2000). A decision tree is comprised of a root from which all decisions are then grown into subsequent child and terminal nodes (Figure 2.6).

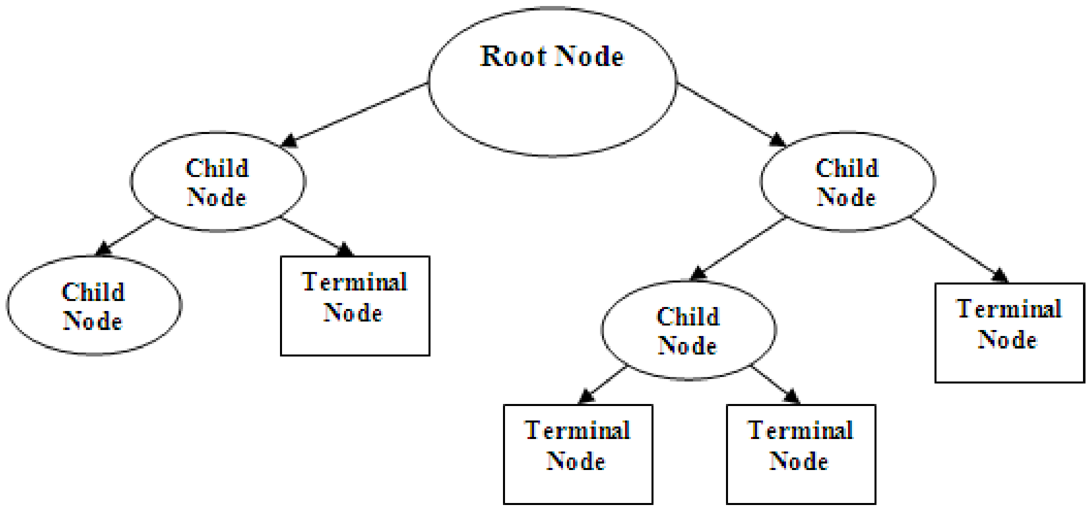

Figure 2.6 Example of Classification or Regression Tree comprised of a Root Node, Child Nodes, and Terminal Nodes.

The root node contains the entire dataset to be tested according to an independent variable. The subsequent splits that best predict the relationship between the dependent variable and independent variable are based of logical "if-then" statements. Each "if-then" statement separates the data that best depicts the relationship into child nodes. After many logical "if-then" statements, a node becomes heterogeneous and cannot be split further, resulting in a terminal node. A decision tree is constructed based on several logical "if-then" statements constructed of one root node and several child and terminal nodes. The term "binary" implies that the root and each child node can only be split into two groups, or child nodes, in which each case the original node is called the parent node (Lewis 2000). The term recursive refers to the binary partitioning, or splitting process, that can be applied over and over again. CART analysis consists of four basic steps. The first step consists of tree building, as described above, using recursive splitting of nodes. Each resulting node is assigned a predicted class or value based on the distribution of classes or values in the learning dataset. A splitting criterion, which is used to decide which variable gives the best split, is determined by identifying which variables and their respective values has the greatest between-groups sum-of-squares in a simple analysis of variance between the left and right child nodes (Therneau *et al.* 1997). In a sense, the decision tree seeks to maximize the average "purity" of the two child nodes (Lewis 2000). Once a decision tree is built, there comes a point where building must stop. Building stops when: (1) there is only one observation in each of the child nodes; (2) all observations within each child node have the identical distribution of predictor variables, making splitting impossible; or (3) an external limit placed by the user on the maximal depth of the tree. Every classification or regression tree built over-fits the dataset (Lewis 2000, Therneau *et al.* 1997). This means that the maximal tree follows every idiosyncrasy in the dataset with later splits in the tree more likely to represent to over-fitting than earlier splits. This leads to the third step of tree pruning. In order to generate simpler trees that provide a better fit to the dataset, a method of cost-complexity pruning is used (Therneau *et al.* 1997). This method relies on a cost-complexity parameter (CP), usually denoted ά, which gradually increases during the pruning process. Terminal and child nodes at the last levels of the decision tree are pruned away if the resulting change in misclassification cost is less than ά times the change in tree complexity. Thus, ά is a measure of how much additional accuracy a split must add to the entire tree to warrant the additional complexity. As the cost-complexity parameter (ά) value increases, a greater number of nodes (decreasing to increasing importance) are pruned away, resulting in a simpler decision tree. Now the question becomes, what is the optimal tree that fits the dataset the best? Since the maximal tree over-fits the dataset provided, the performance of the tree on an independent dataset is likely to be poor (Lewis 2000). Therefore, the goal is selecting the optimal tree, defined by its performance on an independent dataset, is to find the correct cost-complexity parameter (ά) so that that the information in the learning dataset is fit but not over-fit. A method known as cross-validation is used to determine the number of nodes the optimal tree should be.

Cross-validation is a computationally-intensive method for validating a procedure for model building which avoids the requirement for a new or independent validation dataset (Lewis 2000). In cross-validation, the learning dataset is randomly split into *N* sections, stratified by the driving variable of interest. One of the *N* subsets is reserved for use as an independent test dataset, while the other *N*-1 subsets are combined for use as the leaning dataset in the model-building procedure (Figure 2.7). Then the entire modeling procedure is repeated *N* times, with a different subset of the dataset reserved for use as the test dataset each time (Lewis 2000). The average performance of these *N* models produces an average decision cost versus the complexity or number of nodes the optimal tree should have (Figure 2.8). The minimum average decision cost within one standard-error is then used to prune the over-fit tree to the optimal tree that best fits subsequent untested independent datasets. By using the 1-SE rule, it becomes a balance of higher decision cost for each split and the amount of information and interpretability of the results. Ripley (2009) suggests a good choice for CP pruning is often the left-most value for which the mean lies below the horizontal line. Furthermore, the corresponding CP value to the minimum cross-validation error (xerror) can be used to prune the regression tree. The minimum cross-validation pruning method is known as minimum

CP pruning and generates a tree with the highest R^2 correlation between the depending response variable tested against many independent variables.

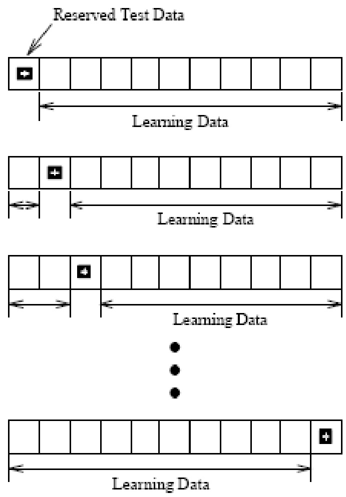

Figure 2.7 Cross-Validation Procedure to Determine Average Decision Cost.

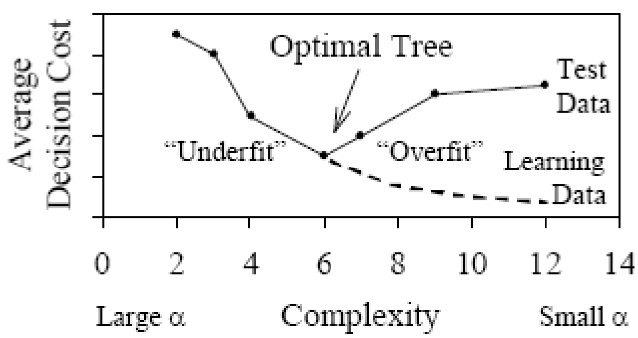

Figure 2.8 Selection of Optimal Tree based on Minimum Average Decision Cost and Complexity

Relatively few studies have used predictive analytics to determine the significant causal biotic and abiotic variables responsible for forest damage from historical or current data (Lindemann and Baker 2002, Kupfer *et al.* 2008). Only CART analysis has been used due to the interpretable results, quick processing, and accurate classifications

(Lindemann and Baker 2002, Kupfer *et al.* 2008). CART has been coupled with univariate analysis and parametric techniques such as frequency distributions, binary logistical regression, and forward stepwise logistical regression to determine the associated coefficients and trends of each biotic and abiotic factor on the measured and recorded damage pattern (Lindermann and Baker 2002, Kupfer *et al.* 2008). These Geographical Information Studies (GIS) studies also use classified remotely sensed (RS) imagery that also aids in change detection of the landscape before and after the catastrophic wind event.

Forest Damage Pattern across the Landscape Using Remote Sensing and GIS

Literature herein will focus on remote sensing and GIS techniques and applications investigating relative roles of biotic and abiotic factors for mapping landscape-scale forest damage patterns. Remote sensing platforms (including aircraft and satellite) provide a valuable perspective on disaster situations that allows for rapid assessment of damage and losses on the natural and built environment (Womble *et al.* 2006). However, remote sensing studies that measure the health and subsequent damage to forest vegetation following a hurricane and relate it to causal biotic and abiotic variables are minimal and scant (Ayala-Silva and Twumasi 2004, McMaster 2005, Wang and Xu 2008). Moreover, there are limited remote sensing studies that utilize low to moderate spatial resolution necessary to view change in forest health and damage at the landscape scale following a hurricane (Jacobs and Eggen-McIntosh 1993, Nix 1996, Ramsey *et al.* 1997, Ramsey *et al.* 2001, Murrah 2007, Collins *et al.* 2008).

Similar to remote sensing applications, there are a select few studies that incorporate GIS techniques that depict biotic and abiotic variables at the landscape scale

(Lindemann and Baker 2002, Kupfer 2008, Wang and Xu 2008). These few studies utilize inventoried forest biotic metrics and measured or derived abiotic variables across the landscape. Unfortunately, biotic and abiotic causal variables and their complex interaction have not been studied across the southeastern landscape of Mississippi. The Southeast Forest District inventory of 2006 and the resulting damage from Katrina provide a unique research opportunity to identify the biotic and abiotic relationships for hardwoods and pines across the southeastern Mississippi landscape.

Remote Sensing Studies

The use of remote sensing platforms allows for rapid assessment of forest damage following a hurricane. Furthermore, the size and scope of forest damage can be determined much quicker than ground surveys and inventories. However, to relate the derived damage from remote sensing applications, ground truth data must be compiled. The time, personnel, and resource strain to conduct a tree-level inventory following a hurricane however is much more than the time and money expenditure for landscape scale remote sensing applications. Thus very few landscape scale forest inventories are compiled following a hurricane except by select state agencies and the USFS through FIA plots (MIFI 2006, Oswalt and Oswalt 2008). Thus, remote sensing studies used to classify forest damage across the landscape as they relate to forest coverage type and other abiotic variables are reviewed. Such studies have related the derived forest health and impact from a hurricane to the inherent biotic metrics and varying abiotic factors across the landscape (Ramsey III et al. 1998, Ayala-Silva and Twumasi 2004, Wang and Xu 2008). Ayala-Silva and Twumasi (2004) used Advanced Very High Resolution Radiometer (AVHRR) to detect changes in the Normalized Difference Vegetation Index

(NDVI) and how NDVI values correlated with Hurricane Georges' path across Puerto Rico. Ramsey III et al. (1998) combined Landsat TM and NOAA AVHRR data to assess the vegetation biomass impact and recovery from hurricane Andrew in coastal Louisiana. Wang and Xu (2008) performed a change detection analysis, along with derived NVDI and Tasseled Cap Wetness (TCW) values for the Pearl River Basin following Hurricane Katrina. The assessed factors of forest and site conditions consisted of landform characteristics of elevation, slope, and aspect, soil great groups, buffer zones along river channels, forest types, and forest attributes derived from Landsat 5 TM vegetation indicies (Wang and Xu 2008).

AVHRR satellite data have been readily used for derived vegetation indices and landcover classifications to determine the relative vegetation health and coverage over a particular area (Roy *et al.* 1997, Ramsey III *et al.* 1998, Ayala-Silva and Twumasi 2004). Ayala-Silva and Twumasi (2004) assessed the impact in different forests in Puerto Rico following hurricane Georges using AVHRR NDVI data. NDVI is the most widely used vegetative index which is a function of the red and near-infrared spectral bands with higher index values correlated to healthier vegetation (Ayala-Silva and Twumasi 2004) (Eq. 2.4).

$$\mathrm{NDVI} = \frac{NIR_{BAND} - RED_{BAND}}{NIR_{BAND} + RED_{BAND}} \qquad \text{Eq 2.4}$$

Data expressed as NDVI provides information on the vegetation health as well as the spatial-temporal changes that coincide with a forest disturbance such as a hurricane (Sader et. al. 2001). Some of the most common change-detection approaches include (a) post-classification (supervised or unsupervised) change detections, (b) spectral-temporal change classification, and (c) the NDVI (Michener and Houhoulis 1997, Ayala-Silva and Twumasi 2004).

The objective of Ayala-Silva and Twumasi (2004) was to determine whether AVHRR data could be used to assess the effects of hurricanes with respect to vegetation across the landscape scale of Puerto Rico. Visualizing damage over such a broad scale has the potential for forest damage estimates and volumes to be determined and resources for recovery allocated. The study areas include two protected forested areas, the Luquillo

Experimental Forest (LEF) located on the northeast side and Guanica Forest located on the southwestern side of the island and five cities spread through out the island. AVHRR data from September 1997 and September 1999 were used for vegetative monitoring, temperature and percent reflectance (albedo) to evaluate the NDVI across the study areas, before and after hurricane Georges which make landfall in September 1998. Results show that Georges affected the Guanica forest more than LEF (Table 2.2).

Table 2.2 Change in mean NDVI values for the LEF and Guanica Forests before and after Hurricane Georges.

Year	Study Area	Mean NDVI	Standard Deviation	Δ NDVI
September 1997	LEF	0.0094	0.1659	-
“	Guanica	0.3152	0.0629	-
September 1999	LEF	0.0467	0.0733	+0.0373
“	Guanica	0.1017	0.0409	-0.2135

Forests on the south side of the island that were exposed to higher wind speeds for longer durations were more adversely affected than the LEF on the northeast side of the island. Cities studies showed similar results with those located on the south side of the island experiencing the greatest negative change in NDVI values. Based on these values, a regression equation was fit to the NDVI changes for each city and their distance from the track of hurricane Georges. Despite the second AVHRR image taken a year after the hurricane, there were detectible and significant relationships ($p<0.025$) between the distance from the track and ΔNDVI as measured by low-resolution data. Ayala-Silva and Twumasi (2004) shows AVHRR data can reveal significant impacts from the passage of hurricanes at the landscape scale. The distance from the center of the hurricane track does show a correlation to the amount of forest damage sustained; suggesting areas closer to the center of the track are prone to greater amount of damage. Understanding the dynamic relationship of differing types of forest and differing types of damage to the distance away from the center has the potential to identify forested areas more or less susceptible to damage base of their location relative to the land-fall and hurricane track location.

Ramsey III et al. (1998) also combined AVHRR images that were transformed into NDVI values and coupled with Landsat TM data to map the association between forest types and hurricane effects in the Atchafalaya River Basin, Louisiana, USA. Hurricane Andrew provided a practical application for a combined use of AVHRR and Landsat TM data. Temporal curves of mean NDVIs for three forest sites for the dates before and shortly after the hurricanes' passage and aggregate curves of the impacted forest to the average NVDIs of the study region were compared. The results identified changes in NDVI values reflected differences in damage severity and type across the studied landscape and spatial covariation between increased impact magnitudes and river corridors dominated by open hardwood forest. Open hardwood forests in bottomland river locations are more susceptible to damage and infrastructures in these areas are placed at greater risk by wind-thrown trees. Thus, appropriate mitigation and recovery procedures should be put into practice in these areas to minimize damage caused by falling trees.

Overall, four AVHRR images were acquired for the time periods before and after landfall to determine the change in NDVI values shortly after landfall and during a recovery period. Then unsupervised classification with an overall accuracy of 85.9% and Kappa Statistic of 0.81 was performed on the Landsat imagery with the aid of color-infrared photography as reference data (Figure 2.9). For comparison to the Landsat TM, the AVHRR NDVI images were subset to an area coincident with the Atchafalaya River Basin extent and resampled with a nearest-neighbor selection criteria to a 25 meter spatial resolution. Therefore, the cross tabulation and spatial extent of the impact and recovery magnitudes were created for open, cypress-tupelo, and hardwood forest classes.

A vast majority of the damage occurred on the east side of the Andrew's Path with very little impact to hardwood forests to the west of the path (Figure 2.10). The distribution of higher impact magnitudes was concentrated along the southwestern edge of the basin closely following the mixed hardwood/cypress-tupelo corridor. In addition, there was moderate to high impact to cypress-tupelo in the extreme southeastern portion of the river basin. Conversely, there was a large area of minimally impacted but highly dense cypress tupelo centered in the lower river basin. This finding agrees with Oswalt and Oswalt's (2008) inventory of cypress-tupelo's resiliency in the Pearl River basin following Katrina and other studies (Touliatos and Roth 1970, Putz *et al.* 1983, Gresham 1991, Doyle *et al.* 1995, Peterson 2000).

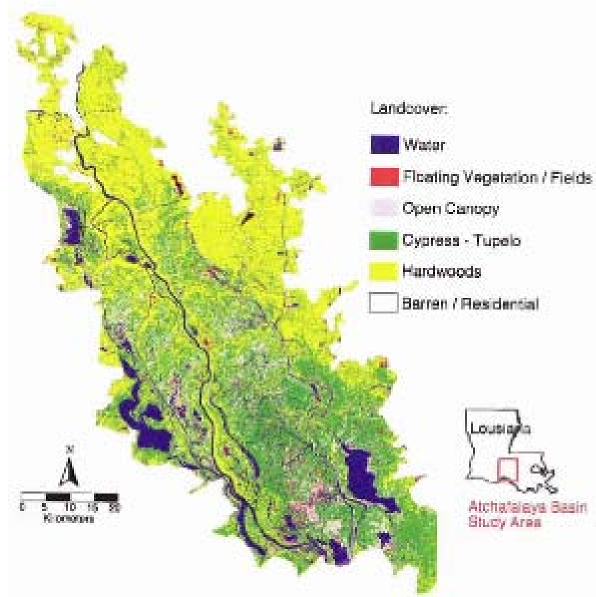

Figure 2.9 Unsupervised Classification of the Atchafalaya River Basin

The spatial distribution of recovery (increase in NDVI) generally follows the spatial distribution of the impacts. Hardwood and cypress-tupelo areas moderately impacted in the central portions of the Achafalaya River Basin showed subsequent moderate to high recovery two months later. Basin wide statistical measures implied nearly equal average impact to all three forest classes, while recovery distribution measures implied higher recovery among cypress-tupelo and open classes and lower recovery in hardwood classes. Based on reconnaissance of the area, some of the highest impacted areas were identified to be solely cypress-tupelo. Ramsey III et al. (1998) then states that if cypress-tupelo were more resistant to wind damage than hardwoods, higher impacts could be a result of longer duration of wind speeds. However, without a wind field distribution availible, the change in NDVI spatial distribution can not be directly correlated to one of the most important abiotic factors, the wind intensity.

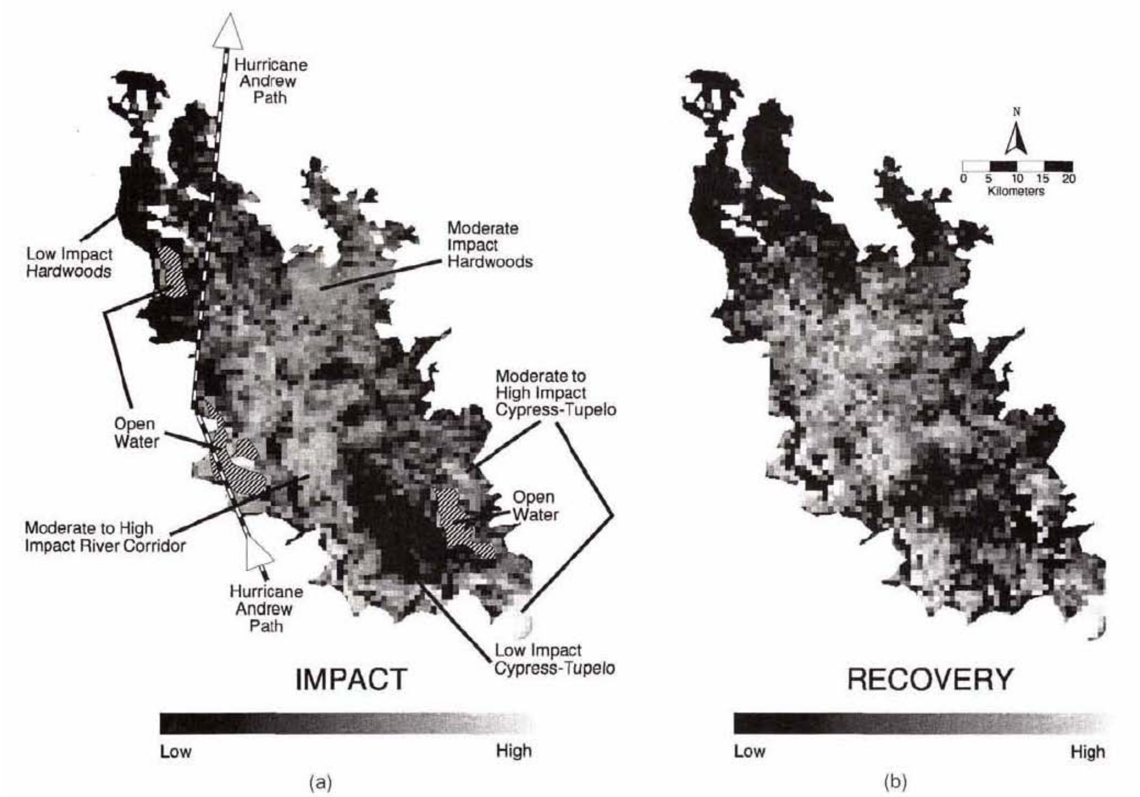

Figure 2.10 NDVI Impact and Recovery Maps of the Atchafalaya River Basin following Hurricane Andrew.

This caveat is negated in Wang and Xu (2008) which used Landsat 5 TM data coupled with several biotic and abiotic factors through GIS of a known and strong wind field of Hurricane Katrina along the Lower Pearl River Basin, USA.

In addition to derived NDVI values that Ayala-Silva and Twumasi (2004) and Ramsey III et al. (1998) used, Wang and Xu (2008) also used Tasseled Cap Wetness (TCW) to supplement other ecological factors to be correlated to forest damage across the landscape scale. Wang and Xu (2008) use logit regression on forest type, forest coverage, stand density, soil great group, elevation, slope, aspect, and stream buffer zones to identify causal factors best predicting forest damage. The factors with continuous attributes were classified into discrete categories for analysis together with other categorical factors. The relative effects of the categories of each factor pertaining to hurricane disturbance were assessed by comparing percentages of the disturbed forest areas, percentage of the disturbed forest areas at three severity levels (light, moderate, and high), and odds ratio. The relative effects of each factor on the disturbance were determined by comparing the full and reduced logit models.

Wang and Xu (2008) noted a heterogeneous forest damage pattern with bottomland hardwood forests on the floodplains most severely affected. Soil groups and stand factors including forest types, forest coverage, and stand density contributed to 85% accuracy in modeling the probability of the hurricane disturbance to forests in this region. The study area included the counties of Hancock and Pearl River in Mississippi and Washington and St. Tammany parishes of Louisiana that were centered around the Lower Pearl River Basin. Forests in this area occupy 370,000 ha, which is roughly 53% of the total area investigated in this study. The main forest cover types are wetland forests, upland forests, and urban forests. In wetland forests, over 50% of the stands were comprised of water tupelo, swamp tupelo, sweet gum, oaks, and bald cypress (Rosson 1995, USDA Forest Service 2007). Upland forests are predominantly mixed groups of loblolly-shortleaf pines, longleaf pine, slash pine, hickories, and oak–pine mix forests (Rosson 1995, USDA Forest Service 2007). To classify these forest coverage types and the other continuous variables across the landscape, Wang and Xu (2008) used national land-use land-cover data (NLCD), state soil geographic (STATSGO) data, digital elevation (DEM) data (30m x 30m), national hydrography data (NHD), NDVI, and TCW derived values. The STATSGO data consisted of soil great groups across the study region. Five of the six land cover types were studied including urban forests, evergreen forests, mixed forests, shrub/scrub, and wetland forests to explore the differing damage amounts occurring to each forest type (Figure 2.11).

Wang and Xu (2008) performed univariate analysis along with full and reduced logit models to analyze the effects of forest characteristics and site conditions on hurricane disturbance and to model probabilities that forests would be disturbed by hurricanes. Forest disturbance (disturbed forests were coded as 1 and undisturbed forests were coded as 0) was entered as the dependent binominal variable and the coded factors of forest type, soil, elevation, aspect, buffer zone, NDVI, and TCW were entered as independent variables. The authors never define the word “disturbance” and how it directly relates to the amount of damage incurred. Each variable was categorized according to equal intervals of its respective value range. Thus the differing categories could be compared to each other through logistical regression for the determination of their relative resiliencies to forest disturbance. Variable comparison is then made between categories coded as lower numbers and the variable coded as the highest number. These comparisons result in odds ratios that determine the relationship between the categories and thus the probability of forest disturbance based on that variable’s values. Odd ratios less than one indicate forested areas with attributes of the lower rank are experience less disturbance when compared to the higher category. Evaluation of the fit models included percentages of correctly classified undisturbed and disturbed forests, percentage of concordance of predicted probabilities, and Akaike’s information criterion using SAS 9.1.

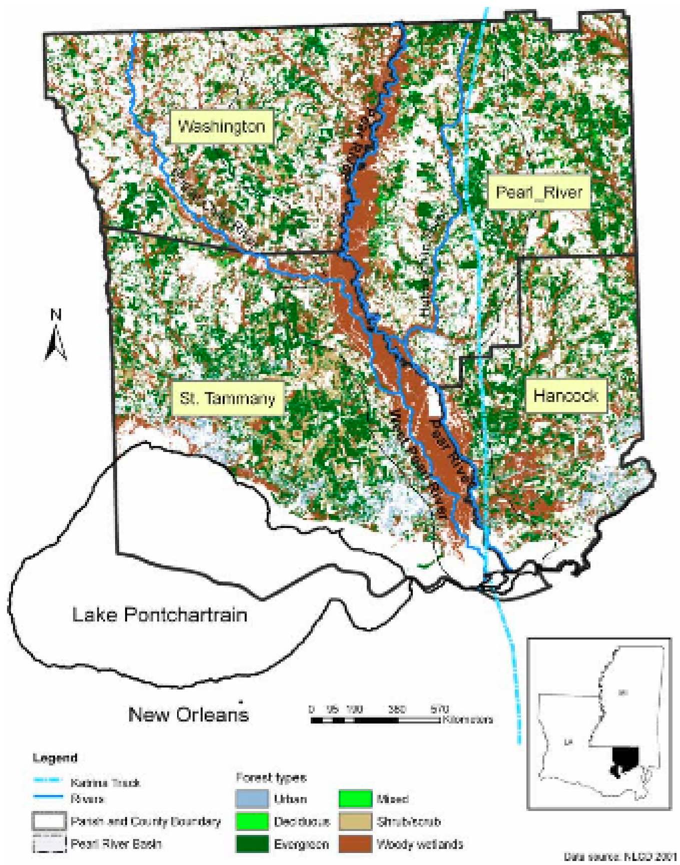

Figure 2.11 Forest Coverage Types across the Lower Pearl River Basin

Through univariate analysis, Wang and Xu (2008) determined that approximately 60% of forested land, 18% with highly, 35% with moderate, and 7% lightly disturbed forest land occurred in the Lower Pearl River Basin (Figure 2.12). Similar to forest disturbance, the definitions of highly, moderate, or lightly disturbed forests is not specified within the paper. Wetland forests located along the Pearl River northward from Hancock County to northern Pearl River County and Washington Parish were among the most heavily damaged. Bottomland hardwoods of the Lower Pearl River Basin were especially prone to greater damage opposed to pine and mixed forests (Figure 2.12). The damage of bottomland hardwoods corroborates the surveys by MIFI (2006) and Oswalt and Oswalt (2008). The largest amounts of highly disturbed forested areas were located with 100 meters of river channels. With increasing distance from the river channels, the percentage of highly disturbed forests declines, while percentages of moderately and lightly disturbed forests escalate slightly. Although a majority of moderate disturbance occurred along Pearl River County and St. Tammany Parish, sporadic pockets also exist across the landscape. This applies with light disturbance to forests as well agreeing with most all previous research stating many biotic and abiotic variables create a heterogeneous forest damage pattern.

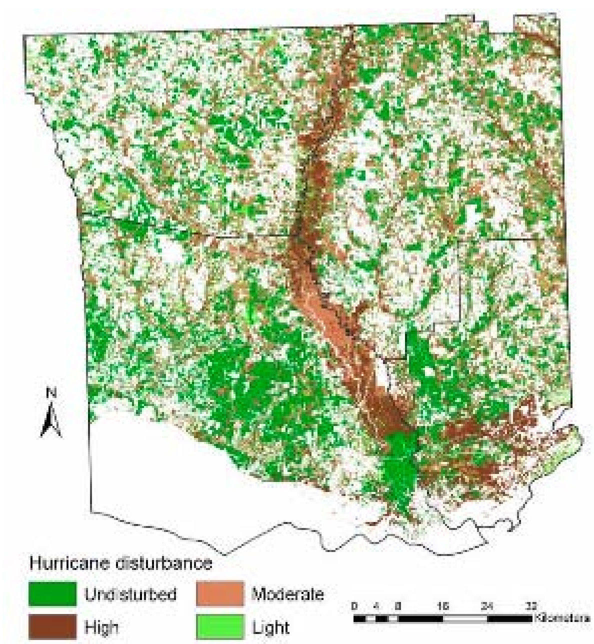

Figure 2.12 Spatial Distribution of Hurricane-Induced Forest disturbances by Severity Levels.

The heterogeneity of disturbance for Katrina-impacted forests and degree of severity varied by forest type. The largest proportion of wetlands forests were disturbed

(78%), followed by urban (73%), mixed (69%), evergreen (47%), and shrub/scrub (43%). Among the three severity levels, the greatest proportion of wetland and urban forests were highly disturbed, 39% and 30% respectively. Additionally, forests having greater NDVI and TCW derived values were less likely to be disturbed whereas areas with lower values were more prone to moderate or high damage. NDVI is not an intrinsic physical measurement of vegetation characteristics, but is indeed correlated with vegetation properties such as percent vegetation cover, Leaf Area Index (LAI), and stand density (Ormsby et al. 1987, Purevdorj et al. 1998). Wang and Xu (2008) conclude the relationships of higher NDVI/TCW derived values pertaining to higher percentage of forest cover and stand density are correlated to lower forest susceptibility to hurricane damage.

Soil properties in Wang and Xu (2008) were found to correlate with forest disturbance susceptibility, agreeing with Trousdell (1965) and Everham and Brokaw (1996). Glossaqualt soils typically do not have a fragipan, duripan, or nitric horizons, thus allowing for greater root growth and anchorage (USDA 1795). Sulfaquants are almost permanently saturated while endoaquuepts are primarily wet soils located in flood plains or have very high water tables (USDA 1975). Sulfaquants are almost permanently saturated while endoaquuepts are primarily wet soils located in flood plains or have very high water tables (USDA 1975). Lower forest susceptibility occurred on glossaqualt soils with less than half of forests disturbed at any severity level. Forests on endoaquept and sulfaquept soils were found to be most susceptible to hurricane damage with over 80% forests disturbed. Since higher forest disturbance was found on wet soils in the bottomlands, higher disturbed forests to be in the lower elevations. However, forests experiencing light and moderate disturbance increased as elevation increased. Relationships between forest disturbance and other topographical variables such as slope and aspect as relating to amount of forest disturbance were inconclusive. Forest disturbance did not correlate to aspects although it was noted that there were greater amounts of highly disturbed and moderately disturbed forests on aspects of 315-360°. Percentage of disturbed forested area did not correlate well with varying of slopes. Full logit regression results concurred with the univariate analysis with wetland forests having the greatest probability of being disturbed compared to all other forest types. Odds ratio of the forest types with codes less (1-5) than wetland (6), indicated that woody wetlands had the highest probability of being disturbed compared to all other types of forest coverage. Percent forest coverage as related to NDVI and TCW values statistically showed that higher NDVI and TCW derived values from Landsat imagery were correlated to less forest disturbance by the hurricane which agrees with Ramsey *et al.* (1998). High odds ratio for endoaquepts and the subsequent low odds ratio for glossaqualfs agreed with the univariate analysis. The topographic location of the endoaquepts also lead to low odds ratios for low elevations between 0-48 meters having a greater likelihood of forest disturbance than those at higher elevations. Odds ratio of slope and aspect were close to one, indicating forest disturbance was not well correlated to any single category.

The lack of topographical factors associated with percentage of disturbed forests is not surprising in Wang and Xu (2008). The study area is located along the coast and southern Gulf Coast Plain with characteristically little relief. If the study area was expanded northward, more terrain variability would be included that may alter the relationships of forest disturbance to various topographical factors. Another limitation of Wang and Xu (2008) is the lack of damage definition. The authors never define the word "disturbance" and how it directly relates to the amount of damage incurred. Moreover, the type of damage incurred to various types of forests was not explored as it related to the biotic and abiotic factors. Wang and Xu (2008), however, begins to shed light on biotic and abiotic variables' relative role on different forest types using remote sensing and GIS techniques. Two other studies, Lindemann and Baker (2002) and Kupfer (2008) also delve into the physical and biological factors contributing to the heterogeneity of forest damage. These two studies use an extensive GIS database of biotic and abiotic factors and data mining software to extract the significant causal biotic and abiotic factors and their respective unit of measure.

Geographical Information Systems (GIS) Studies

Within the past 20 years, there has been little research utilizing GIS to analyze forest damage and the associated biotic and abiotic variables (Foster and Boose 1992, Boose 1994, Lindemann and Baker 2002, Kupfer 2008, Xi 2008, Wang and Xu 2008). Only two studies have used CART methodology to isolate significant variables across the landscape (Lindemann and Baker 2002, Kupfer 2008). One study focused on a severe windstorm in the Rocky Mountains that affected 10,000 hectacres (ha) of forest while the later study focuses on observed tree damage in the Desoto National Forest as a result of Hurricane Katrina.

In October 1997, easterly winds at 120-155 mph (200-250 km/hr) from a strong baroclinic low caused the largest forest blowdown known to occur in the southern Rockies (Wesley *et al.*1998). The disturbance, known as the Routte-Divide blowdown is located in the Routte National Forest in north-central Colorado's Park Range. Lindemann and Baker (2002) built an extensive GIS database that included many abiotic variables. However, only one biotic variable, vegetation cover type was used as a driving variable. The authors' primary objective was to determine which potential predictors can be used to predict the severity blowdown and where the blowdown occurred using univariate and CART analysis. A list of all variables tested can be seen in Table 2.3.

Table 2.3 List of Predictor Variables for a Severe Forest Disturbance.

Biotic Variables
Vegetation Cover Type
Engelmann Spruce/Subalpine fir
Grasslands/forblands
Lakes
Lodgepole Pine
Rockland, talus, scree
Rush Species, Sedge Species
Sagebush
Shrublands
Streams
Wetlands
Willows
Abiotic Variables
Topographic Factors
Elevation (m)
Slope (°)
Aspect (°)
Roughness
Exposure 5°
Exposure 10°
Exposure 15°
Exposure 30°
Terrain Complexity (300m)
Terrain Complexity (600m)
Terrain Complexity (1km)
Terrain Complexity (2m)
Terrain Complexity (33m)
Soil and Geological Characteristics
Soil Texture
Rooting Depth
Soil Permeability
Water Holding Capacity
Geology
Human Influences
Distance to Roads
Past Disturbances
Past Beetle Outbreaks
Other Factors
Distance to Natural Edges
Distances to Continental Divide

The univariate analysis showed the most prominent predictors related to blowdown were in the middle elevation ranges, on east-facing aspects, shallow slopes, the middle bands of the distance from the Continental Divide, the engelmann spruce landcover type, exposed areas and on the west side of the Continental Divide (Lindermann and Baker 2002). Blowdown was not increased by high terrain complexity or moist and thin soils. Areas with low terrain complexity and greater rooting depths actually showed a slightly greater tendency of being blowndown. Soil permeability, water holding capacity, distance to nearest ridge and distance to natural edge do not have a significant relationship with blowdown.

A 22-node classification tree with a dependent variable of 5 categories of percent blowdown was produced (Figure 2.13). The classification tree bases the first split on the distance from the Continental Divide. The second and third splits narrow the elevation bands at which the most blowdown occurred. The greatest amount of blowdown actually occurred in the middle elevations and not near ridges. The first terminal node with the highest number of sample points (347 of 401) predicted to be blown down includes points 4km-9km from the Continental Divide at elevations greater than 2737.5m and in areas of wind exposure at a 15° inflection angle. Conversely, the terminal node with the highest number of sample points regarding non-blowdown were less than 4km or grater than 9km from the Continental Divide and at elevations greater than 3288m. The remaining sample points are spread out through many splits of the predictors, indicating that the explanations for areas not blowndown are quite complex (Lindermann and Baker 2002). As for vegetation cover, aspen trees showed resistance to blowdown while Engelmann Spruce was more susceptible.

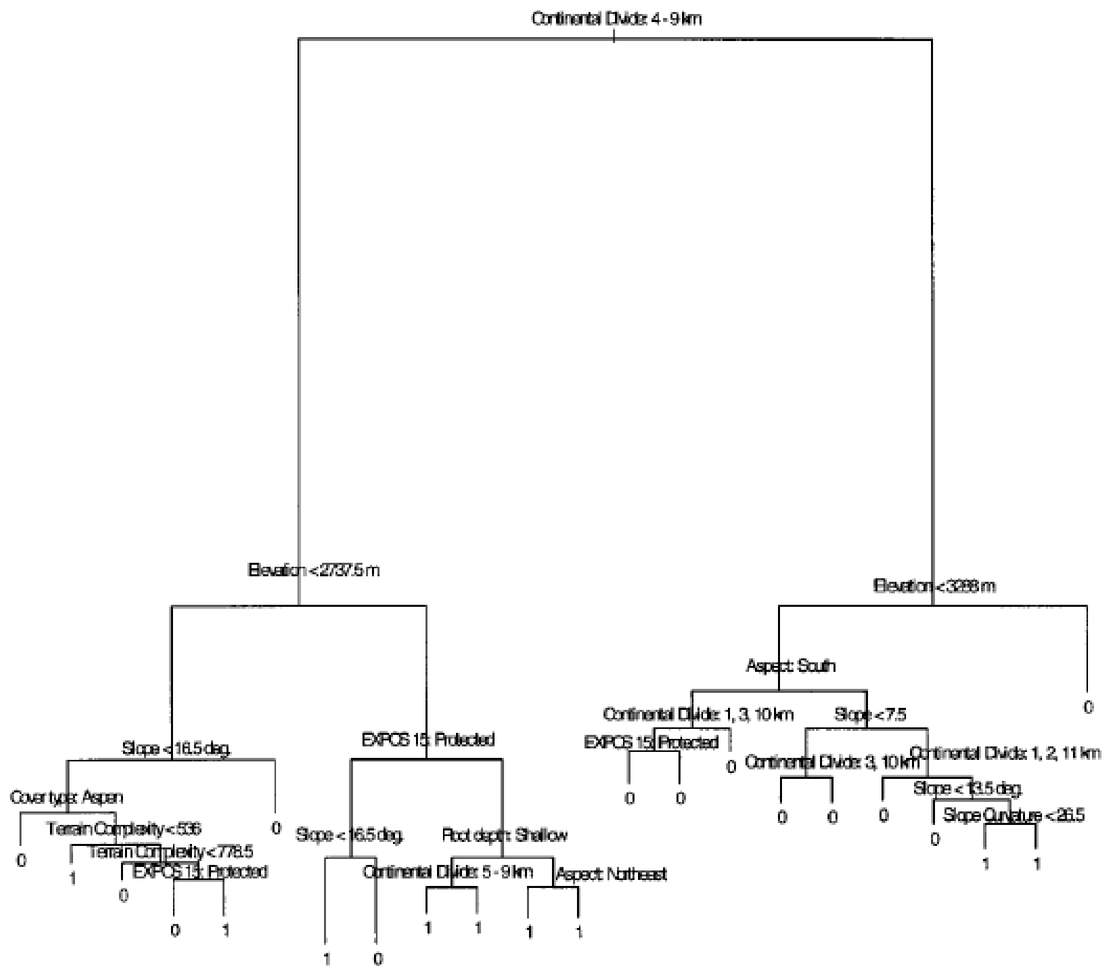

Figure 2.13 22-Node Classification Tree of Predictor Variables for Blowdown and Non-blowdown areas in the Routte National Forest.

According to Lindermann and Baker (2002), relatively few variables contribute to the overall pattern of blowdown with the primary variables defining blowdown locations were topographical features, rather than vegetation and geological or soil features. Vegetation-cover type played a small role in the blowdown location. The authors attribute this to most of the area being the Engelmann Spruce/Subalpine fir cover type (58%). However, in the forested areas, Englelmann Spruce/Subalpine fir did show a tendency to being blown down, whereas Aspen was more resistant and lodgepole pine was moderate. Several studies have shown that broadleaf species are more resistant to blowdown than pines (Curtis 1943, Foster 1988a, Everham and Brokaw 1996, Ramsey III et al. 1997, Peterson 2000). Lindermann and Baker (2002) did not explore the biotic aspect as related to forest damage in detail since only one biotic variable was examined. Lindermann and Baker (2002) is one of the first to use CART analysis to depict the interaction of measured and derived biotic and abiotic variables across a large landscape. In doing so, they showed an accurate assessment across a10,000 ha (38 sq. mile) area by having an overall classification accuracy of 81% for blowdown areas. Therefore the CART methodology has shown to be an effective method in determining forest susceptibility using governing biotic and abiotic variables. Similar to Ayala-Silva and Twumasi (2004), the authors explicitly state two variables that could have improved CART analysis would have been an accurate coverage showing stand height and fine resolution model showing wind speeds over the landscape. This caveat is negating in Kupfer *et al.* (2008) using H*Wind analysis and CART methodology to analyze biotic and abiotic factors that affected the forests of the Desoto National Forest in Mississippi as a result of Hurricane Katrina.

Hurricanes alter landscape-scale patterns of forest structure and composition, habitat availability and distribution, and susceptibility to subsequent disturbances (Kupfer *et al.* 2008). Nonetheless, they are normal, integral parts of long-term forest dynamics in the Caribbean, Atlantic, and Gulf Coast Regions which means that forest management plans need to recognize their effects and include the potential for such events to occur (Dale *et al.* 1998). Understanding the biotic and abiotic factors that govern tree and thus forest susceptibility to hurricane winds must be known before mitigation practices can be implemented. Both forest managers and emergency managers can institute mitigation efforts on two fronts, proactively before the storm, and retroactively following the storm. An interesting question is how large can an area become before the CART procedure cannot resolve the biotic and abiotic interactions and the respective classification and cross-validation values show weak correlation? Kupfer *et al.* (2008) begins to explore the biotic and abiotic dynamics for the Desoto National Forest covering an area of 153,000 ha (590 sq. miles). This area is much larger than the Routte-Divide area from Lindermann and Baker (2002). Kupfer's *et al.* (2008) objective was to develop and test a CART model of forest damage resulting from Katrina for a 153,000 ha forest management unit in southern Mississippi. Kupfer *et al.* (2008) used classification tree analysis to develop a model of damage severity on the basis of storm meteorology, stand conditions, and site characteristics for more than 400 locations. By including variables addressing wind speed, rainfall, species composition, forest structure, age, topography, and flood plain conditions, Kupfer *et al.* (2008) attempted to contrast the importance of different types of factors associated with landscape-scale patterns of forest damage.

The study was conducted in the DeSoto Ranger District of the DeSoto National Forest (NF) that was located in right-front quadrant of Katrina as is move north over Mississippi. Kupfer et al. (2008) modeled the wind swath of Katrina as it made landfall by using the H*Wind product from the Hurricane Research Divison (HRD) of the National Hurricane Center (NHC). Pioneered by Powell *et al.* (1996) and furthered by Powell et al. (2004), the H*Wind product is a downloadable GIS uniform gridded 4km point shapefile.   Attributes contained for each grid point include the maximum sustained wind speed, duration of hurricane force winds, direction of maximum wind, and the steadiness of the wind. The attributes of each point are compiled through remotely sensed satellite, in situ airborne, and surface data collected through an array of platforms and sensors. The wind swath of hurricane according to Kupfer *et al.* (2008) is displayed in Figure 2.14. The details of how the gridded point shapefile was interpolated to produce a continuous wind-field are not detailed in Kupfer et al. (2008). There seems to be disconnect between the stronger winds over Harrison and Stone Counties and the actual path of the hurricane over Hancock and Pearl River Counties. According to the official write-up on Katrina from the NHC, the higher sustained winds were realized farther west then over the DeSoto National Forest as portrayed in Figure 2.14 (NHC 2009).

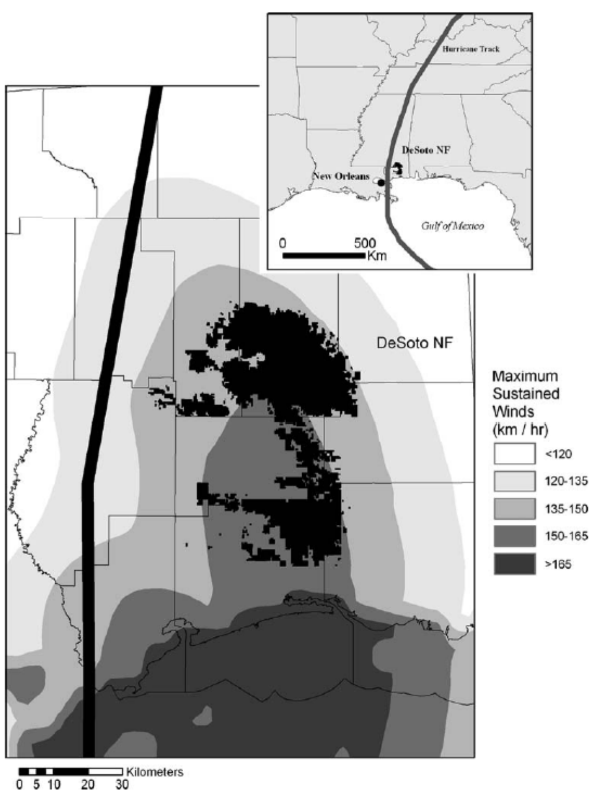

Figure 2.14 Location of DeSoto NF relative to Katrina’s storm track and Wind Speeds.

Kupfer *et al.* (2008) described the topography as characterized by broad and gently sloping uplands dissected by numerous streams and rivers, the largest of which have mature flood plains (Pearl River and Pascagoula River Flood plains). As depicted by Oswalt and Oswalt (2008) and Glass and Oswalt (2006), the uplands are dominated by pines, especially loblolly pine, slash pine, longleaf pine, and shortleaf pine. Bottomlands are dominated by hardwoods, including various species of oak, sweetgum, and many others.

Assessments of damage severity to forests were conducted using ADS40 aerial photographs taken by the U.S. Army Corps of Engineers (USACE) in September of 2005 and were retrieved from the Mississippi Automated Resource Information System (MARIS). For pre-storm imagery, Kupfer *et al.* (2008) used the USDA National Agriculture Imagery Program (NAIP) imagery taken in 2004 of the same area. After image acquisition and rectification, 415 random points were selected for forest damage analysis using Hawths tools extension for ArcGIS. A 17.8m buffer around each point was created to be consistent with field assessments. Field assessments were taken on 40 0.1 hectacre (~0.25 acre) circular plots in February 2006. All standing and fallen trees greater than 10 cm DBH were measured and identified to species. Kupfer *et al.* (2008) used the 40 field assessment plots to train CART and supplement information for the other 415 random points that were chosen through Hawths Tools. The damage was classified using a four-point scale based on the estimated percentage of downed overstory trees: (1) no discernable downed trees, (2) light damage (<33% blowdown), (3) moderate damage (33-67% blowdown), and (4) heavy damage (>67% blowdown). Damage was classified by two independent researchers with 88% agreement.

Predictor variables used for input of CART can be viewed in Table 2.4. Four measures of storm meteorology were used, along with estimated cumulative rainfall, topographic settings (elevation, slope angle, and aspect), distance to nearest streams, distance to hurricane track, forest type, stand age, stand condition, hardwood basal area, pine basal area, and total basal area. Wind speeds were derived from the H*wind product pioneered by Powell et al. (1996) and furthered by Powell et al. (2004). Pine Forest types are categorized as ≥ 70% of basal area with dominant and co-dominant crowns were softwoods; Pine-Hardwoods: 51-69% softwood basal area; Hardwood-Pine: 51-69% hardwood basal area, Hardwoods ≥ 70% hardwood basal area. Stand Conditions were extracted from the DeSoto NF Continuous Inventory of Stand Conditions (CISC) database.

Table 2.4 Variables used to predict Forest Damage from Hurricane Katrina at Desoto National Forest.

Variables	Description/Source
Storm Meteorology	
Maximum Sustained Winds (km/h)	Extracted from NOAA H*Wind Model
Duration of Hurricane Force Winds (h)	Extracted from NOAA H*Wind Model
Directional Steadiness	Extracted from NOAA H*Wind Model
Cumulative Precipitation (in)	NOAA Climate Prediction Center
Site Topography and Location	
Elevation (m above sea-level)	Extracted from 10m DEM
Slope Angle (°)	Extracted from 10m DEM
Aspect (Classified to Direction)	Extracted from 10m DEM
Distance to Nearest Perennial Stream (m)	Based on U.S. Geological Hydrography data
Distance to nearest Stream (m)	Based on U.S. Geological Hydrography data
Distance to Hurricane Track (km)	Extracted Euclidian Distance based on NOAA map of hurricane track
Stand Characteristics	
Age	From CISC Database, defined as years since stand origination
Forest types	From CISC Database, grouped into: 1. Pine: P. palustris dominant 2. Pine: P. taeda dominant; 3. Pine: P. elliottii dominant; 4. Pine: mixed yellow pines; 5. Pine-hardwood: non-P. taeda dominant; 6. Pine-hardwood: P. taeda dominant; 7. Hardwood-pine:
Stand Condition	From CISC Database, grouped into 1. sparse pole- and sawtimber; 2. mature sawtimber and poletimber; 3: immature poletimber; 5: regeneration; 6: immature sawtimber
Hardwood Basal Area (ft^2/acre)	From CISC Database
Pine Basal Area (ft^2/acre)	From CISC Database
Total Basal Area (ft^2/acre)	From CISC Database

To relate each biotic and abiotic variable from Table 2.4., Kupfer *et al.* (2008) used classification tree analysis (CTA), a non-parametric, probabilistic machine learning method that recursively partitions observations with a categorical response variable based on binary splitting criteria. A single classification tree was grown and subsequently pruned such that the cross-validated error cost of the smaller tree was no more than one standard error from the minimal cross-validated error. The one standard error (1-SE) pruning method and minimum cross-validation error pruning method were used for grown regression trees across the Southeast Forest District as stated in the methods section. Additionally, Kupfer *et al.* (2008) predicted forest damage severity using a form of CTA termed stochastic gradient boosting (SGB). SGB starts by fitting an initial tree but the residuals for the first tree are fed into a second tree which attempts to reduce the error, a process that is repeated through a series of successive trees (Lawrence *et al.* 2004). The final predicted value is formed by adding the weighted contribution of each tree. Both CTA and SGB models in DTREG software provide a variable importance score to clarify the relationships between forest damage and predictor variables. This score is calculated based on the improvement in classification gained by each split that used the predictor.

Overall classification accuracy of the optimal single tree model was 71.5% with a producer's accuracy ranging from 65-82% and user's accuracy varying from 58 to 82%. The use of SBG increased the overall accuracy of predictions to 81% and improved the producer's and user's accuracy over values of the single tree model. Age was the best predictor of forest damage in both the single tree and SGB models. Forest type, stand condition, aspect, and distance to the nearest perennial stream had moderate to high importance scores as evident in the first successive splits (Figure 2.15). The storm meteorology variables were generally of limited importance in the single-tree model but moderate in SBG model.

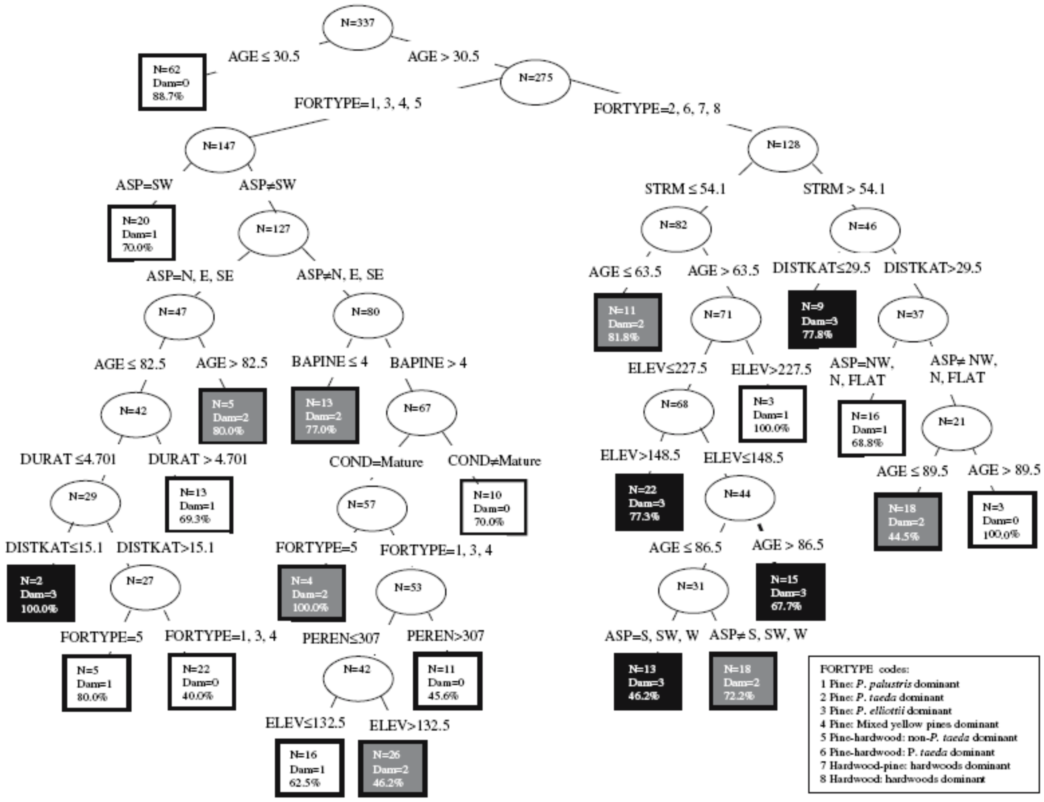

Figure 2.15 Classification tree for forest damage caused by Hurricane Katrina at DeSoto National Forest.

The first split states forests stands younger than 30.5 years were least likely to suffer damage. The greatest number of plots that experienced severe damage were older than 63.5 years, comprised of predominantly loblolly pine, at elevations between 148.5 feet and 227.5 feet with streams less than 54.1 meters away. Other terminal nodes showing severe forest damage are trees greater than 86.5 years of age and elevations less than 148.5 meters and stands on south, southwest, and west facing slopes. However, other stands comprised of longleaf, slash, mixed yellow poplar, and loblolly dominant mixed pine forests on north, east, and southeast slopes younger than 82.5 years, experiencing less than 4.7 hours of hurricane force winds, but were less than 15.1 km away from the storm track were also severely damaged. Mature longleaf, slash, mixed yellow pine , and loblolly dominant mixed pine forests greater than 5 inches DBH with greater 4 ft^2/plot basal area were more likely to be damaged than immature longleaf, slash, mixed yellow pine, and loblolly dominant mixed pine forests (less than 5 inches DBH). Based of these classifications, Kupfer et al. (2008) depicted the spatial pattern of forest damage across the DeSoto National Forest (Figure 2.16). Not explained, however, is the variable of basal area being incorporated into the spatial model since it is not a continuous variable as are all other variables used.

Kupfer *et al.* (2008) found that damage patterns following Hurricane Katrina over a large, heterogeneous landscape were mostly strongly associated with stand conditions and site characteristics such as age, forest type, aspect, and distance to nearest perennial stream. Increasing severity of damage was positively correlated to increasing age, increasing hardwood component of each stand, and increasing damage on aspects facing the dominant wind flow (east, southeast, south, and southwest), agreeing with Sheffield and Thompson (1992), Everham and Brokaw (1996), Oswalt and Oswalt (2008). Increasing distance from a perennial stream had a negative relationship on severity of damage with greatest severity coming from hardwoods located in bottomlands and along river channels. Plots classified as pine had increased damage with increased pine basal area.

Figure 2.16 Pattern of forest damage predicted for Desoto National Forest by the Single Tree Classification Model. Inset shows damage mapped by USDA Forest Service.

Concurring with Trousdell *et al*. (1965), Touliatos and Roth (1971), and Xi (2008), Kupfer *et al*. (2008) found cumulative precipitation and maximum sustained winds had a positive relationship with amount of damage incurred. Upland pine and pine-hardwood stands only suffered heavy damage within 15 km of the hurricane's track whereas heavy damage on less-resistant bottomland hardwood and hardwood-pine stands extend 30 km from the track. This agrees with multiple studies summarized in Everham and Brokaw (1996) and post-storm inventories from Sheffield and Thompson (1992) taken in South Carolina following Hurricane Hugo. Even though the bottomland hardwood forest stands were further away from the stronger winds, other biotic and abiotic factors still contributed to their susceptibility of damage. Despite the many findings of Kupfer *et al.* (2008), there are some limitations and potential sources of error. One obvious potential source of human error is potential misclassification of damage from aerial photographs even though two independent researchers classified the damage. Another potential source of human error is image rectification and registration errors of the aerial imagery causing plot locations to be inaccurate. Also, damage was classified on 0.1 ha plots but corresponding predictor variables extracted from the CISC database were aggregated to entire forest stands. This may lead to some errors in assigning stand-level properties to individual plots. Kupfer *et al.* (2008) does not specify the type of damage recorded for each measured tree on each 0.1 ha plot. The authors state their long term objective is to project the susceptibility of forest stands to future events as an aid to forest management activities. Yet, without specifically documenting the type of damage incurred across the landscape, the implications for management are less clear. Differing damage types and amounts to differing species require different salvage and management practices (DeLoach and Dicke 2005, Long *et al.* 2005).

DeLoach and Dicke (2005) of the Mississippi State Extension Service detail questions facing landowners with salvage and management decisions of damaged forest stands due to Katrina (Table 2.5). One main controlling factor in determining if a stand remains manageable is the amount of basal area damaged within the forest stand. Basal area is the cross sectional area of a tree at DBH. The basal area of all trees on a plot is then summed to the plot or acre level to determine the amount, and thus the value of wood growing in a given area.

Table 2.5 Questions facing Landowners with Salvage and Management Decisions.

Question	Information needed to Answer
Do I have a manageable timber stand left undamaged?	The Basal Area (Density) of undamaged standing timber.
Will I be able to make a timber sale in the future when prices are better?	The volume or tonnage of undamaged standing timber.
Can I salvage the damaged timber?	The amount of damage – volume or tonnage of damaged timber.

The amount of basal area damaged will direct the order of operations of salvage efforts. DeLoach and Dicke (2005) present a timber stand salvage decision model to aid forest landowners operations following a catastrophic wind event (Figure 2.4.2.5). The decision tree is based off the basal area of undamaged timber and splits are then made based on the estimated weight of undamaged and damaged sawtimber and pulpwood leading to the management decision. Figure 2.17 depicts forest stands with 50 ft^2/acre undamaged timber with 15 tons/acre of sawtimber or 25 tons/acre of pulpwood undamaged but has less than 15 tons/acre sawtimber or 25 tons/acre leads to salvaging the downed timber yet having a remaining manageable stand. Conversely, if there is only 30 ft^2/acre undamaged timber, meaning greater amounts of damaged basal area, with more than 15 tons/acre of sawtimber or 25 tons/acre of pulpwood undamaged yet the majority of timber is damaged then the landowner is to salvage all timber, site prep, and then replant the stand.

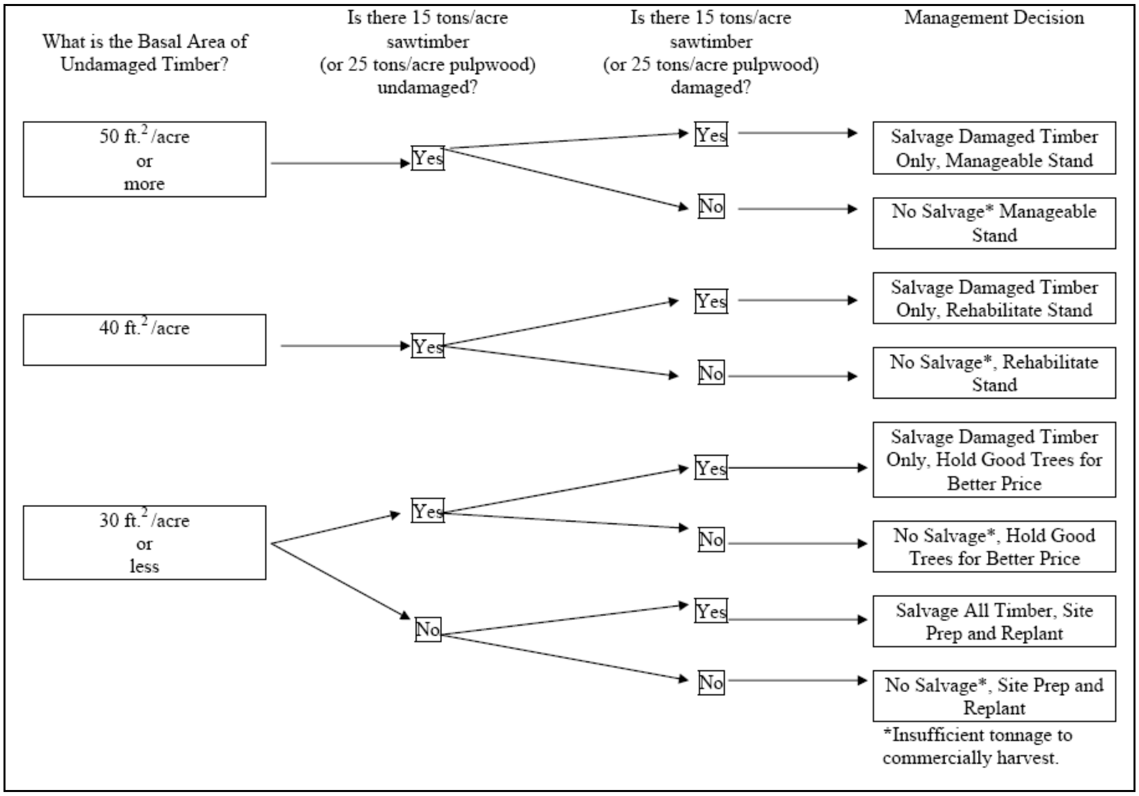

Figure 2.17 Timber Stand Salvage Decision Model.

For a land-owner, downed timber must be salvaged as soon as possible to provide maximum return from investment (MFC 2009). Timber left down, particularly snapped trees which are typically pines, degrade in quality and lose considerable value in the first 60-90 days following a hurricane (Hughes 2005). Trees still attached to the root ball or that are uprooted retain their value for more than 90 days before losing value (Hughes 2005). Understanding the causal biotic and abiotic factors that govern the type of damage for pines and hardwoods at the landscape scale, aid both emergency managers and forest managers decision process for recovery efforts. Lessons learned from previous land-falling major hurricanes illustrate that coordination and communication are critical to successfully mitigating immediate effects (Stanturf 2007). Understanding the relative importance of each biotic and abiotic factor and its contribution to wind-throw in hardwoods and shear of pines may enable spatial modeling distribution of vegetative debris. Rapid assessment of damage via spatial models can guide recovery efforts and mobilize the political and financial support necessary to meet both short-term and long-term needs (Compass 2008). Unfortunately, there are few rapid assessment predictive models that map forest damage extent and severity for use by emergency managers and forest managers.

Mississippi Institute of Forest Inventory (MIFI)

Legislation was passed in July of 2002 for the creation of the Mississippi Institute for Forest Inventory (MIFI) by Ronnie Musgrove (Sewell 2002). Four years later, MIFI was incorporated in the Mississippi Forest Commission (MFC). The responsibilities of MIFI include developing statewide forest resource inventory protocols, policies to maintain inventory information and reports, and analyzing data on forest resources in support of new and existing forest industries. Mississippi was split into five different forest districts based on geography, physiography, economic, and political characteristics (Figure 2.18) (MIFI 2006).

Figure 2.18 Forest Districts as established by MIFI for the State of Mississippi

Each district is inventoried every five years with one district inventoried each year to keep an updated statewide forest inventory (MIFI 2006). Years of Mississippi forest district inventories can be viewed in Table 2.6.

Table 2.6 Forest District Inventory Dates

Forest District	**Year First Inventoried**
Southwest Forest District	2004-2005
Southeast Forest District	2005-2006
Central Forest District	2006-2007
North Forest District	2007-2008
Delta Forest District	2008-2009

Between the time of inception of MIFI in the summer of 2002, to the first inventory in 2004, the inventory methodology was established primarily by MSU researchers aiding state officials (Parker et al. 2005). MIFI directs sampling of forest resources in a two stage process, (1) remote sensing and (2) ground surveys. The remote sensing effort utilizes the spectral reflectance of vegetation captured by spectral bands of the Landsat 7 ETM+ satellite. Through a combination of band analysis, and image classifications techniques, water, non-forest, and forest classes are obtained (Figure 2.19) (MIFI 2008). Imagery from previous years is used in conjunction with classification products to map the dynamic change of land cover

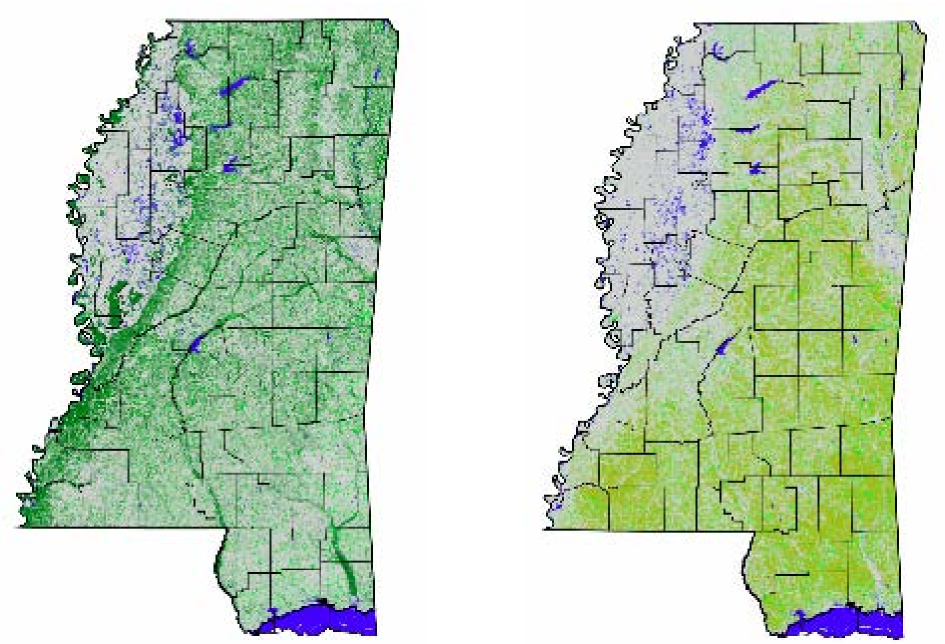

Figure 2.19 Classified Pine Forest (Left) and Hardwood Forest (Right)

Ground-based measurements are implemented on one-fifth acre fixed radius plots located randomly within the forest cover strata obtained from the remotely sensed data (Figure 2.20) (MIFI 2006). Saw-timber, pole, and veneer volume product classes are sampled and characteristics associated with stand dynamics are measured. A one-tenth acre plot is incorporated to measure the volume of products classes that produce fiber for the pulp industry. Finally, a one-twentieth acre plot is inventoried to measure non-merchantable stems that range from 1.0 to 4.5 inches in diameter at breast height (DBH). In the event there are no merchantable trees located on a plot, such as following a harvest, a one-hundredth acre plot is established to measure the reproduction that will develop into a future timber stand (MIFI 2006). Individual tree attributes measured include species, product, observable damage, DBH, total height, height to base of live crown, five and ten year radial growth, and age (MIFI 2006). Stand level attributes recorded include slope, size class, apparent stand level damages, over-story composition, and physiographic region (MIFI 2006). These individual and stand-level attributes were to be inventoried for the Southeast Forest District starting September 1, 2005. However, Mother Nature intervened with Hurricane Katrina making landfall on the Mississippi/Louisiana border on August 29th, 2005.

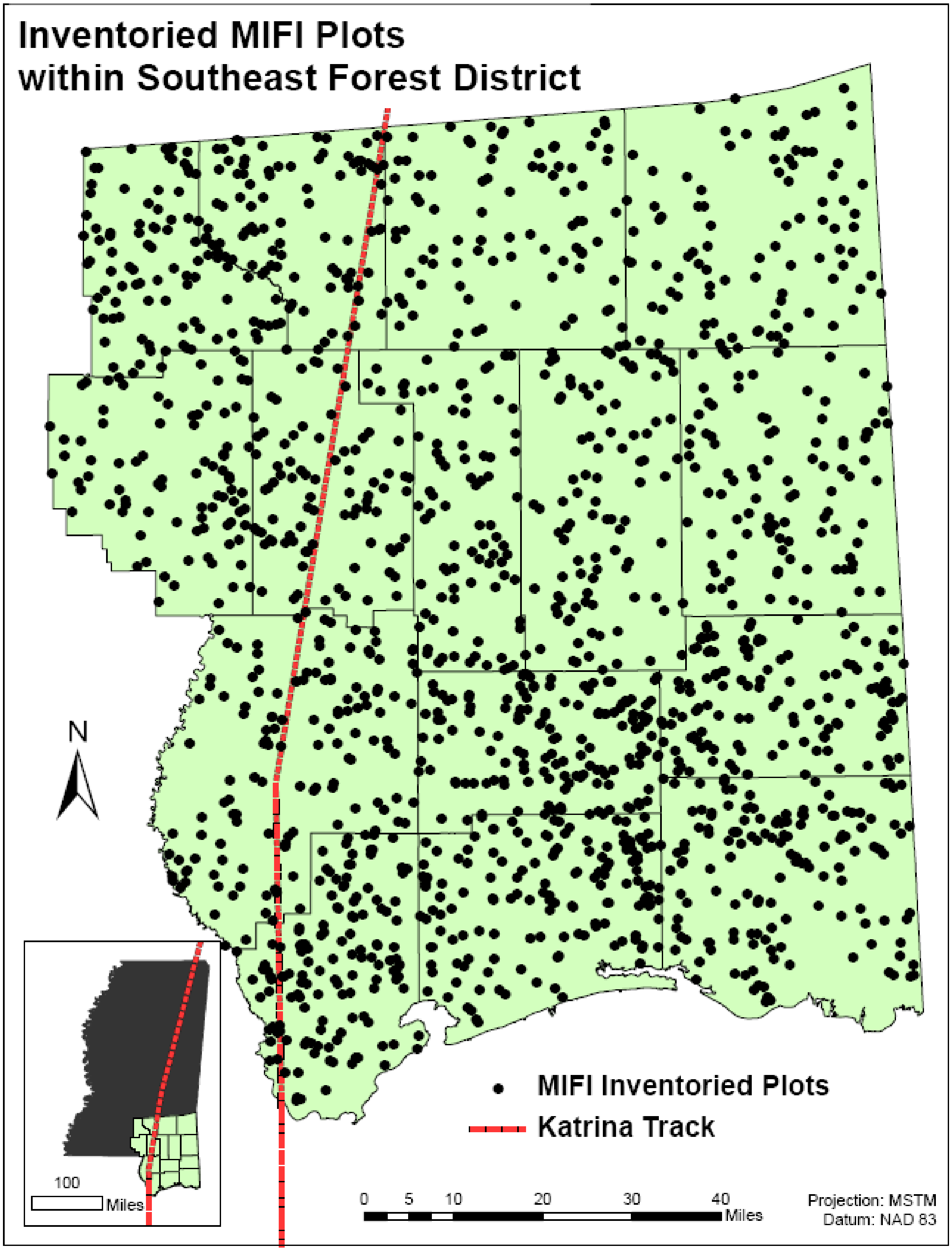

Figure 2.20 Inventoried MIFI Plots of the Southeast Forest District 2005-2006 with Hurricane Katrina track overlaid.

Hurricane Katrina and Related Forest Damage

The record setting year of 2005 in terms of named storms (28), number of hurricanes (15), and Category 5 storms (4), caused great destruction of property, natural resources and loss of human life (NOAA 2006). Hurricane Katrina was particularly devastating becoming the costliest natural disaster in U.S history (NCDC 2009). However, Katrina is not the first major hurricane to strike the Mississippi/Louisiana border. Emergency management and forest management lessons were also partially learned in 1969 with the land-fall of Hurricane Camille (Touliatos and Roth 1971). There have only been three recorded Category 5 hurricanes to make landfall in the U.S. since 1851. One of them was Hurricane Camille, which slammed into the Mississippi coast on August 17th, 1969. Although small in diameter, Camille came ashore with 1-minute sustained wind speeds of 190 mph (305 km/h) coupled with 10 inches of rain falling in Hancock County (NOAA 2009). This resulted in a loss of 8.2 million m^3 of felled timber in Mississippi. In Touliatos and Roth's (1971) survey of damage from Camille concentrated on stem breakage versus uprooting, they found shallow-rooted species along the coast were most susceptible to uprooting (dogwood, water oak, pecan, bay and red maple). They also concluded that recently-thinned pine stands with little taper were severely broken with loblolly pine being the most susceptible. Hurricane Camille became the "benchmark" storm for Gulf coast residents and for forest damage, until August 29th, 2005.

Hurricane Camille's hurricane force wind radius was only 60 miles whereas Hurricane Katrina became a very large and powerful category five hurricane with maximum sustained winds at 175 mph (282 km/h) and hurricane force winds extending

125 miles from the center. In addition to its large size, the presence of a double eyewall feature in Katrina lead to higher sustained winds realized over a greater portion of southern Mississippi for a longer duration (Blackwell *et al.* 2007). Concentric double eyewalls develop most often in more powerful hurricanes and broaden the wind-field (Willoughby 1990). Interestingly, the highest unofficial peak wind gust for Katrina was recorded in Poplarville, Mississippi, with winds of 135 mph (217km/h) (Blackwell *et al.* 2007). Poplarville, located in heavily forested Pearl River County, is 40 miles north of the coast. The outer eyewall of Katrina contracted as it moved north into Pearl River and Lamar Counties causing locally higher sustained and peak gusts to occur. The initial expansion of the wind field associated with the double eyewall structure coupled with the contraction of the outer eyewall inland resulted in tremendous timber damage across Mississippi.

Hancock and Harrison Counties suffered between 51-60% county level timber damage (Wayne 2006). Jackson County suffered 41-50% while Pearl River, Stone, Lamar, Forrest, and Perry Counties suffered 31-40% county level timber damage (Wayne 2006). Compared to Camille's 8.2 million m^3 of timber damage, Katrina felled or damaged an estimated 49 million m^3 (MFC 2005). This volume of downed timber was estimated to be enough to build 800,000 single family homes or produce 25 million tons of paper and pulp products (MFC 2005). Roughly 39 million m^3 (~83%) of timber damage occurred in the 10 southernmost counties of the Southeast Forest District (MFC 2005, Jacobs 2007).

Study Area

Hurricane Katrina made landfall at the mouth of the Pearl River Basin along the Mississippi and Louisiana border. This landfall location brought the heaviest rain and highest winds as the right-front quadrant of the hurricane traversed southeast Mississippi. The study area is the Southeast Forest District which contains 15 counties in the hardest hit region of Mississippi (Figure 2.21).

Figure 2.21 Southeast Forest District with Depiction of Hurricane Katrina Track

The Southeast Forest District's climate is marked by mild, short winters and hot, humid summers (NCDC 2009). Normal annual temperatures range between 64-67° Fahrenheit (F) with normal January temperatures ranging 48-51°F and normal July temperatures ranging 80-84°F (NDCD 2009). However, average maximum temperatures in July average routinely average over 90°F across the entire region (NCDC 2009). Average precipitation across the southern tier of Mississippi ranges 58-64 inches and evenly distributed through out the year (NCDC 2009).

The Southeast Forest District is located in the Gulf Coastal Plain which is characterized by broad and gently sloping uplands dissected by numerous streams and rivers, the largest of which have mature floodplains and noticeable topographic changes in areas to upland systems (Chapman 2004). Elevation of the area ranges from sea-level along the coast with little elevation rise northward along mature floodplains of the Pearl and Pascagoula River Basins to around 500 feet in Jefferson and Covington Counties. As seen in Figure 2.22, the lowlands of the Pearl River Basin and uplands in Lamar, Jones and Covington Counties were directly in the path of Katrina.

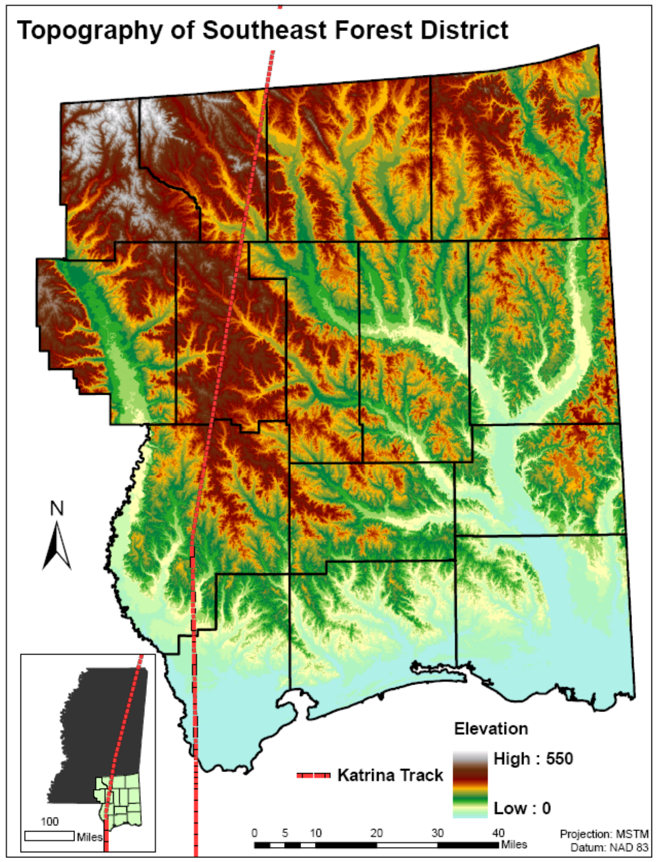

Figure 2.22 Topography of the Southeast Forest District

The varying climate and physiography of the Southeast Forest District supports several different ecoregions (Chapman 2004). The study area is comprised of five ecoregions: Coastal Marshes, Gulf Coast Flatwoods, Southeastern Floodplains and Terraces, Southern Pine Plains and Hills, and Southern Hilly Gulf Coastal Plain (Figure 2.23) (Chapman 2004).

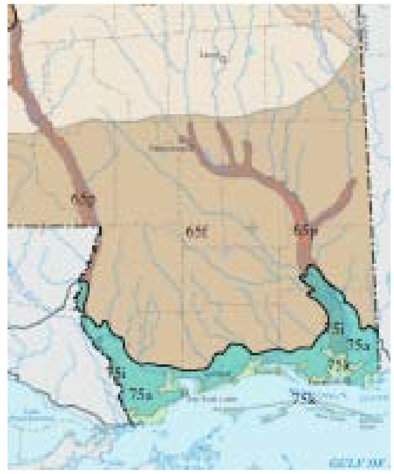

Figure 2.23 Level-IV Ecoregions encompassing the Southeast Forest District in Mississippi. Nomenclature; 75k (Gulf Coastal Marshes), 75a (Gulf Coast Flatwoods), 75i (Floodplains and Low Terraces), 65p, (Southeastern Floodplains and Low Terraces), 65f (Southern Pine Plains and Hills), and 65d (Southern Hilly Gulf Coastal Plain).

The Costal Marshes region contains salt and brackish marshes, dominated with cordgrass (*Spartina patens*) and saltgrass (*Distichlis spicata*) with live oak and laurel oak hardwoods (Chapman 2004). Slightly further north are the Gulf Coast Flatwoods which are characterized by level terraces and delta deposits lined with hardwoods such as swamp tupelo, bald cypress, oaks, hickories, and in some higher drained sites, longleaf and slash pine (Chapman 2004). Continuing along the river basins, the Southeastern Floodplains and Low Terraces is a riverine ecoregion of large rivers and backwaters with ponds, and oxbow lakes (Chatman 2004). Similar to the coastal flatwoods, water tupelo, bald cypress, and oak-dominate the bottomland hardwood forests. The Southern Pine Plains and Hills cover most of the study area consisting almost entirely of loblolly, slash, and longleaf pine plantations. The northern tier of the study area is known as the Southern Hilly Gulf Coastal Plain. This region is characterized by rolling topography, higher elevations, and more relief than regions mentioned above. Vegetative cover consists primarily of longleaf and loblolly pine with some mixed forests of oak, hickory, and co-dominant pine. Summarizing, the vegetative cover of the uplands is primarily dominated by pines, while bottomlands along low terraces and river floodplains are flanked with hardwoods (Figure 2.24).

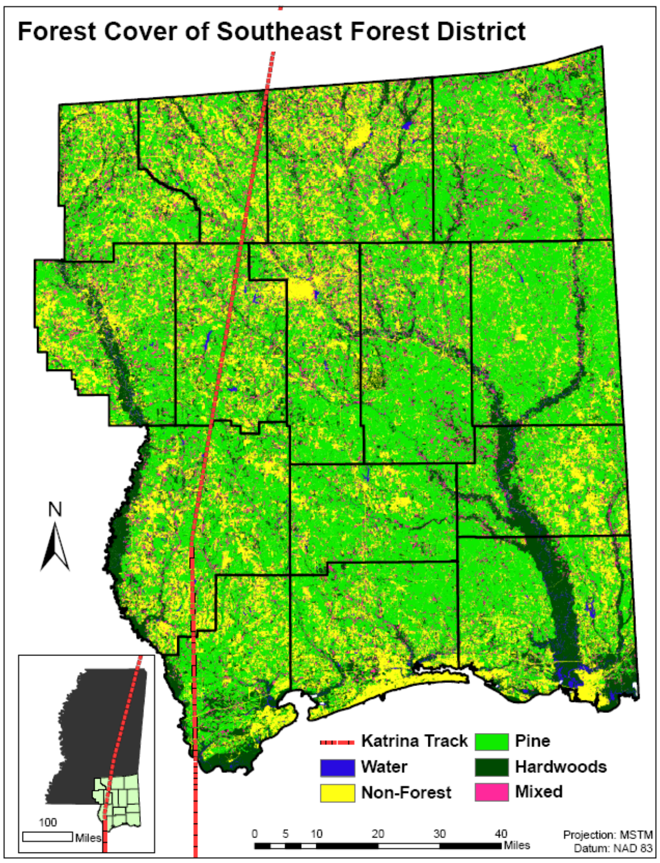

Figure 2.24 Forest Cover of Southeast Forest District in Mississippi.

CHAPTER 3

RESEARCH OBJECTIVES

The primary objective of this research was to employ data mining techniques to ascertain the interactive affects of the biotic and abiotic factors that were significantly related to the damage patterns in the Southeastern Forest District of Mississippi as a result of Hurricane Katrina. The primary objective investigated two separate dependent response variables against independent biotic and abiotic variables which maximize wind-throw of hardwoods and shear in pines:

- Percentage of Stems Damaged per Plot
- Percentage Basal Area Damaged per Plot

A second objective was to compare CART analysis with stepwise forward logistic regression in the identification of important and significant variables for wind-throw of hardwoods and shear of pines.

A tertiary objective was to determine the accuracy of MIFI's original classification of damaged plots. Through personal communication with Patrick Glass, MIFI Director, plot-level classification of wind-throw or shear is subjectively based on aerial images and ground truth opinions of various surveyors. The original MIFI plot damage classifications were reclassified to wind-throw or shear based on greatest percentage of damage type recorded per tree per plot.

CHAPTER 4

DATA AND METHODS

The following section details the materials and methods used to develop a GIS database of biotic and abiotic variables across the Southeast Forest District following Hurricane Katrina. First, the raster and vector GIS layers needed to develop the GIS database using ArcMAP 9.2 are described. Next, the operation and coding of the statistical data mining software [R 2.8.1] is given. Then, forward stepwise logistic regression was used to compare significant variables to variables identified through CART analysis. Then, the statistical methodology investigating the original MIFI plot damage classifications to the reclassified damage plots using SPSS 15.0 is reviewed. Lastly, simple binary logistical regression was used to predict the probability of each individual variable and its respective data values.

Raster and Vector Data

The GIS raster layers used to develop the GIS database where obtained from the Mississippi Automated Resource Information System (MARIS) and the College of Forest Resources (CFR) at Mississippi State University. The geographic coordinate system of each raster dataset was projected in North American Datum (NAD) 1983 with the projected coordinate system of Mississippi Transverse Mercator (MSTM). The raster description, resolution, and respective source can be viewed in Table 4.1.

Table 4.1 Raster Datasets, Resolution, and Sources

Raster Dataset	Resolution	Source
Mississippi Topography		
DEM of Mississippi	10 Meters	MARIS/Dr. Cooke
Elevation (m)→(ft)	10 Meters	MARIS/Dr. Cooke
Slope (°)	10 Meters	MARIS/Dr. Cooke
Surface Roughness	10 Meters	MARIS/Dr. Cooke
Aspect (°)→(direction)	10 Meters	MARIS/Dr. Cooke
Forest Land Cover		
Classified Landcover	30 Meters	CFR

The 10 meter Digital Elevation Model (DEM) of Mississippi was mosaiced together by Dr. Cooke using a 9x9 focalmean window to fill in data gaps in between county borders. The DEM was subset to the Southeast Forest District study area. The units of the original DEM (meters) was converted to feet and GIS functions were used to derive slope, aspect, and surface roughness.

The GIS vector layers used to develop the GIS database where obtained through the Mississippi Automated Resource Information System (MARIS), Untied States Geological Survey (USGS), Hurricane Research Division (HRD), and Mississippi Institute for Forest Inventory (MIFI). The vector shapefiles and their respective sources can be viewed in Table 4.2.

Table 4.2 Vector Data and Respective Sources

Vector Shapefile Name	Vector Geometry	Source
Geographical Shapefiles		
State of Mississippi	Polygon	MARIS
Mississippi Counties	Polygon	MARIS
Mississippi Coast	Line	MARIS
Perennial Streams	Line	MARIS
Meteorological Shapefiles		
Katrina Path	Line	USGS
Sustained Wind (MPH)	Point	HRD/H*Wind
Cumulative Wind (MPH)	Point	Derived
Peak Wind Gusts (MPH)	Point	USGS
Max. Surface Direction (°)	Point	HRD/H*Wind
Duration (hr)	Point	HRD/H*Wind
Precipitation Totals (in)	Point	NCAR/EOL
Precipitation Intensity (in/hr)	Point	Derived
Soil Shapefiles		
Texture (Categorical)	Polygon	Eduardo (2008)/SWAP/NCRS
Infiltration Rate (Categorical)	Polygon	Eduardo (2008)/SWAP/NCRS
Bulk Density (g/cc)	Polygon	Eduardo (2008)/SWAP/NCRS
MIFI Shapefiles		
Forest Inventory Plots	Point	MIFI/Patrick Glass

The H*Wind and precipitation point shapefiles were converted to continuous raster fields across the study area. An Inverse Distance Weighting (IDW) interpolation technique was implemented with the 4km x 4km point shapefile containing the attributes of sustained wind speeds, peak wind gusts, cumulative wind speeds, duration, and maximum surface direction to create continuous wind-flow layers (Figures 4.1-4.5). The IDW technique uses the inverse distance as a weight multiplied by the sample value for the calculation of unknown values between the known values (DeMers 2005). The IDW technique results in the output of a continuous grid that represents general trends in the values across the landscape. Kriging, another commonly used interpolation technique, incorporates the spatial correlation of the sample data opposed to a simple weighted approach (Largueche 2006). This spatial correlation is measured through a semivariogram graph. The vertical graph axis contains the distance between points and the horizontal axis contains the square of the standard deviation between each sample value and its neighbors (Gilreath 2007). A best-fit curve is applied to the points to estimate the spatial correlation along a certain direction or gradient. Points located along the isotropic gradient are weighted higher than along the anisontropic gradient. The differing weights across the sample points can alter the interpolation compared to an IDW of the same sample data. However, the uniform distribution of the 4km x 4km HRD H*Wind and hourly MPE grid points provides an excellent sample grid for applying IDW interpolations. If one was to only have access to surface wind observations at fewer locations opposed to an equally distributed grid, then kriging would be a better method.

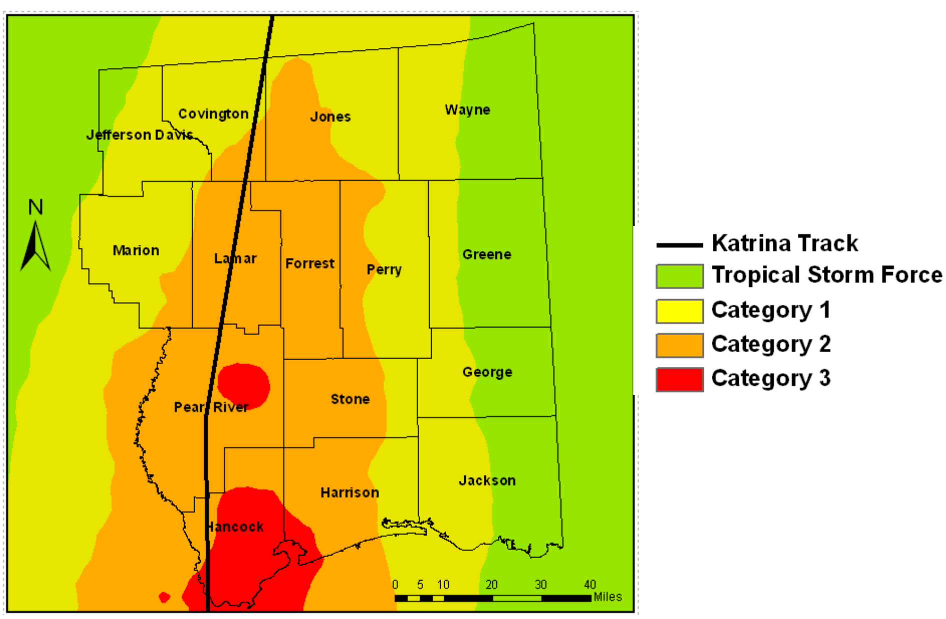

Figure 4.1 Sustained Wind Speeds derived from H*Wind Product

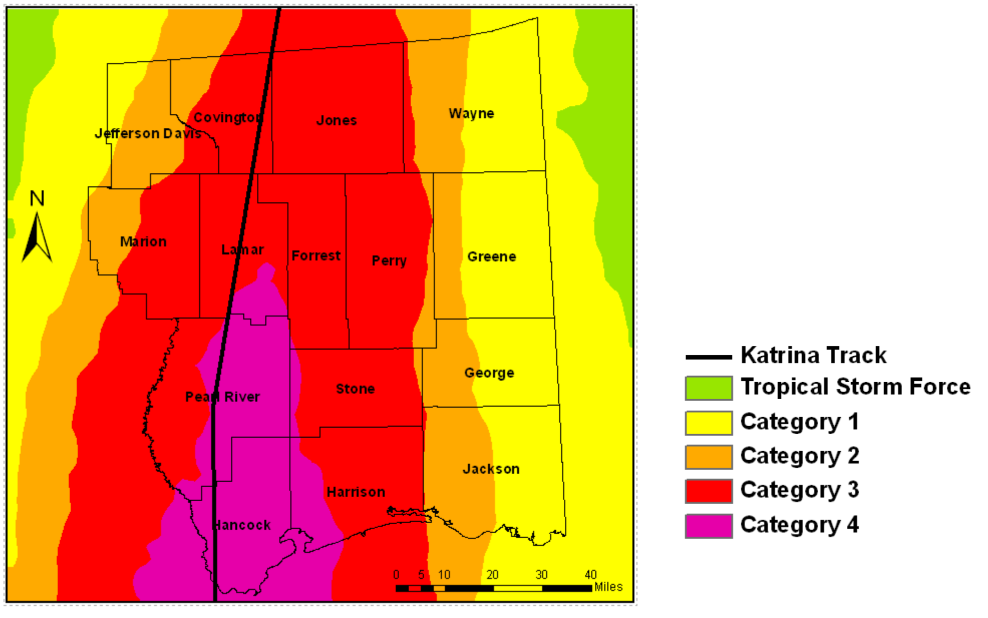

Figure 4.2 Peak Wind Gusts derived from USGS Point data.

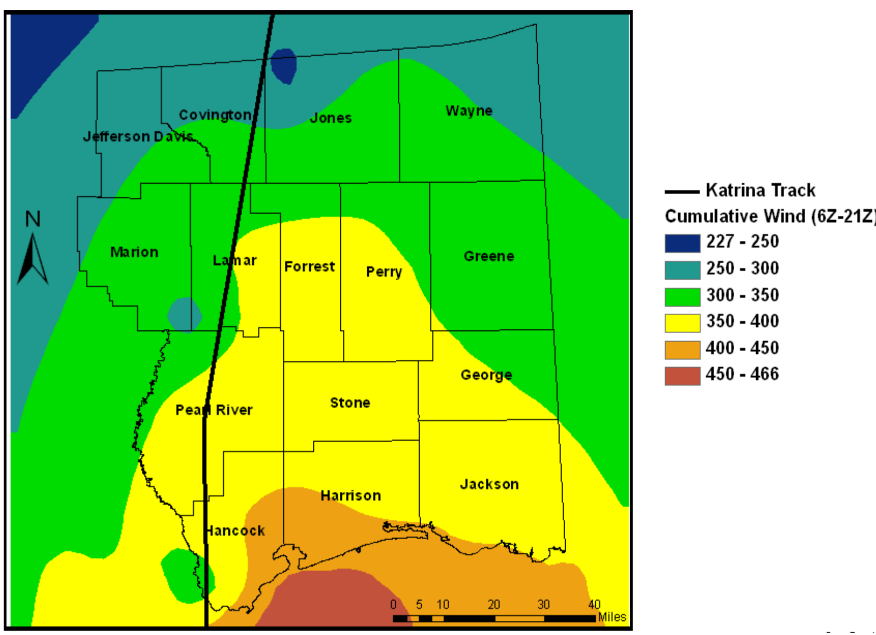

Figure 4.3 Cumulative Wind Experienced from 06Z-21Z on August 29th, 2005.

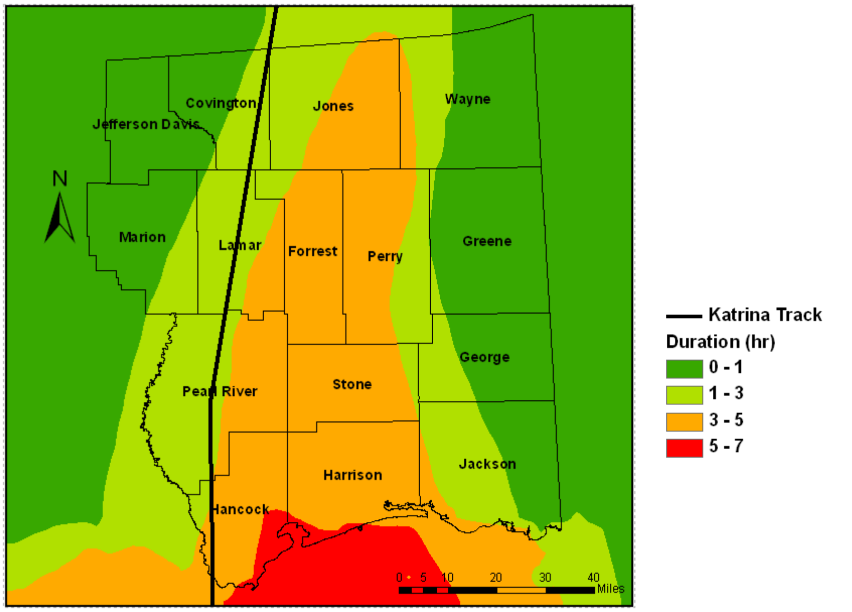

Figure 4.4 Duration of Hurricane Force Winds (hr) derived from H*Wind

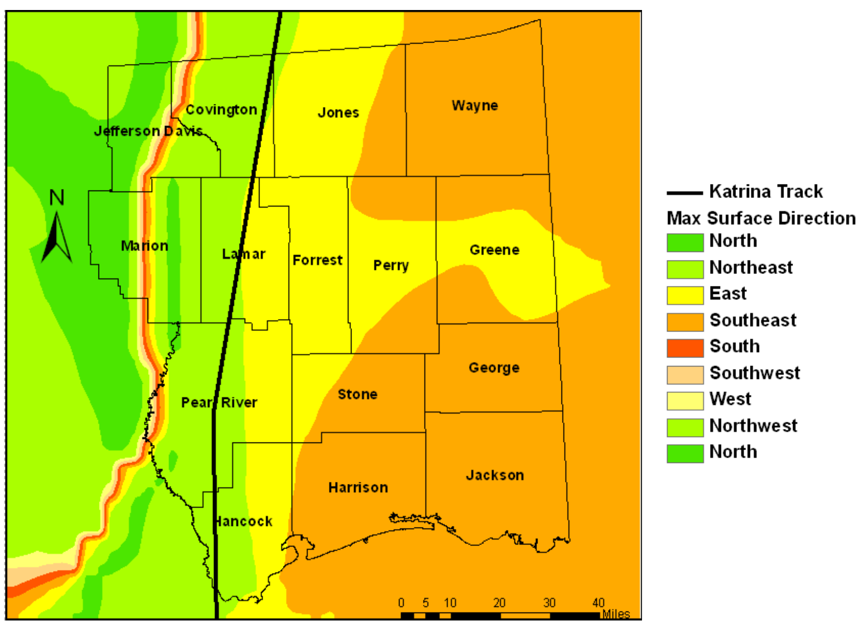

Figure 4.5 Maximum Surface Direction derived from H*Wind Product.

Soil polygon shapefiles were created by importing soil texture, infiltration rates, and mean bulk density values compiled by Eduardo (2008) into a downloaded soils associations database file. Polygons of each soil variable with similar attributes were aggregated using the dissolve function in ArcMap 9.2.

The MIFI Southeast Forest District (SFD) inventoried plot point shapefile was used for model fitting and validation. Original attributes include: plot number, forest coverage type, stand condition, physiographic location, type of damage estimated for the stand-level, stand origin, and x,y coordinates for each plot. Tables 4.3-4.5 detail the forest coverage types, stand conditions, physiographic location, and stand-level damage classification and their respective descriptions. Initial damage classifications were investigated to determine if pine plots suffered more shear damage and if hardwoods suffered more wind-throw damage (Figure 4.6).

Table 4.3 Forest Type and Descriptions.

Coverage Type	**Coverage Description**
Hardwood	Canopy composition is less than 20% Coniferous
Mixed	Canopy Composition is between 40%-60% coniferous
Pine	Canopy composition is greater than 80% Coniferous

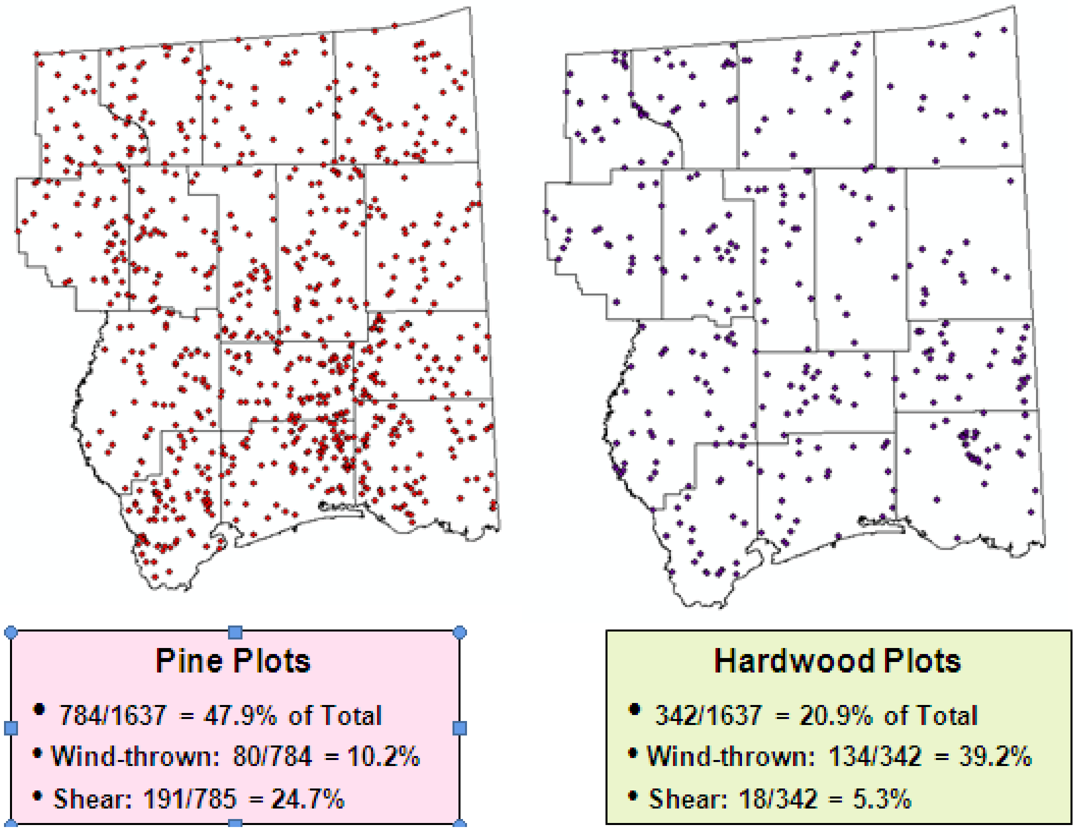

Figure 4.6 Pine and Hardwood Plots across the Study Area and their respective wind-thrown and Sheared Percentages.

Table 4.4 Stand Condition/Size Class and Descriptions.

Stand Condition/Size Class:	Class Description
Reproduction	No Commercial tree species greater than 1 inch in DBH are encountered with the radius of 1/5th acre plot.
Sub-Merchantable	No commercial tree species greater than 4.5 inches in DBH are encountered with the radius of 1/5th acre plot.
Pulpwood	The majority of commercial tree species occupying the 1/5th acre are 4.6 to 10.6 inches in DBH.
Pallet	The majority of commercial hardwood stems occupying the 1/5th acre plot at 7.6 to 10.5 inches in DBH.
Chip'n Saw	The majority of commercial pine stems occupying the 1/5th acre plot are 7.6 to 10.5 inches in DBH.
Saw Timber	The majority of commercial tree species occupying the 1/5th acre plot are greater than 10.6 inches in DBH.
Non-Timber	The site has been converted to non-forestry application.
Non-Stocked	There is less than 20 square feet of basal area on a 1/5th acre plot.

Table 4.5 Topographic Position and Description.

Topographic Position	Topographic Description
Upland	Drier, xeric sites found on top of ridges and side slopes.
Bottomland	Wet, Hydric sites found along rivers and streams.
Terrace	Mesic sites that by default not Upland or Bottomland.

Table 4.6 Stand-Level Damage Classifications and Descriptions

Stand-Level Damage Categories	Damage Descriptions
Shear	Stems have been sheared of twisted off with no crown material present above the break.
Blowdown	Tree stem intact but laying on the ground with the root system attached.

Table 4.3 reveals a potential error in the classification of forest types with forests greater than 20% but less than 40% coniferous trees not clearly defined by the coverage definition MIFI provides in their technical manual. The issue arises again when the percentage of coniferous trees is greater than 60% but less than 80%. Table 4.4 depicts the stand condition and classes that are primarily associated with DBH criteria. Table 4.5 identifies the location of each plot according to physiographical location. Table 4.6 depicts the only two damage types investigated for this study and their respective descriptions according to MIFI (2007).

Southeast Forest District Inventory Tree-Level Data

In addition to the stand-level attributes present in the SFD inventoried plot shapefile, tree-level data for the corresponding plots were also obtained from MIFI through the Director of Operations, Patrick Glass. The tree-level data were originally sent as a large spreadsheet containing tree characteristics of each individual tree measured on plots according the methodology explained above. A total of 46,848 trees were individually sampled by: species group, DBH, total height, type of tree damage, diameter at break in stem, height to base of live crown, and diameter of main stem one foot above the ground being. The fields of height to base of live crown, diameter at break in stem, and diameter of main stem at one foot above the ground were incomplete for all tree records and a decision was made to exclude these fields from the analysis. Therefore, only the DBH, tree height, and damage type statistics were utilized. Basal Area (BA) per plot, trees per acre, wind-thrown basal area, sheared basal area, height to diameter ratio, Quadratic Mean Diameter (QMD), and Lorey's mean height were derived from the utilized fields. Basal area is the cross-sectional area of a tree measured at breast height (DBH) (Emanuel 2009) and is calculated using Eq. 4.1 for English units (in).

$$BA = 0.005454 * DBH^2 \qquad \text{Eq. 4.1}$$

Quadratic mean diameter was computed since it has been shown to be more strongly correlated to stand volume than the arithmetic mean (Brack 1999)

$$QMD = \sqrt{\sum d_i^2 / n} \qquad \text{Eq. 4.2}$$

where d_i is the DBH of an individual tree and n is the total number of trees in the stand. Lorey's mean height was calculated because it weights the contribution of trees to the stand height by their basal area (Brack 1999). Lorey's mean height is determined by multiplying the tree height (h) by the basal area (g), and then dividing the sum of this calculation by the total stand basal area (Eq 4.3) (Brack 1999, Dr. David Evans, personal communication, 2009).

$$H_L = \frac{\sum g * h}{\sum g} \qquad \text{Eq. 4.3}$$

GIS Database Construction using ArcGIS

ArcMap 9.2 and Excel were used in conjunction to build a database containing biotic and abiotic variables to be subsequently analyzed through data-mining techniques, forward stepwise logistic regression, and binary logistic regression to depict causal variables for wind-throw of hardwoods and shear of pines (Figure 4.7). Tree-level data of DBH, height, and type of damage were aggregated to the plot-level. Seven biotic variables were calculated from the tree-level data using Eq. 4.1-Eq. 4.3 in Excel while the sheared and wind-thrown basal areas were determined using SAS 9.1 (Fan 2009). Ninety-one of the plots did not have latitude or longitude coordinates and were excluded from the dataset. An attribute join of the tree-level database file (.dbf) to the MIFI inventory plot (.dbf) was performed using the common plot number field. The attribute join function links two separate tables (.dbf format) based on a common field in their respective attribute tables. With biotic variables now entered into the MIFI inventory point shapefile, abiotic variables were then addressed.

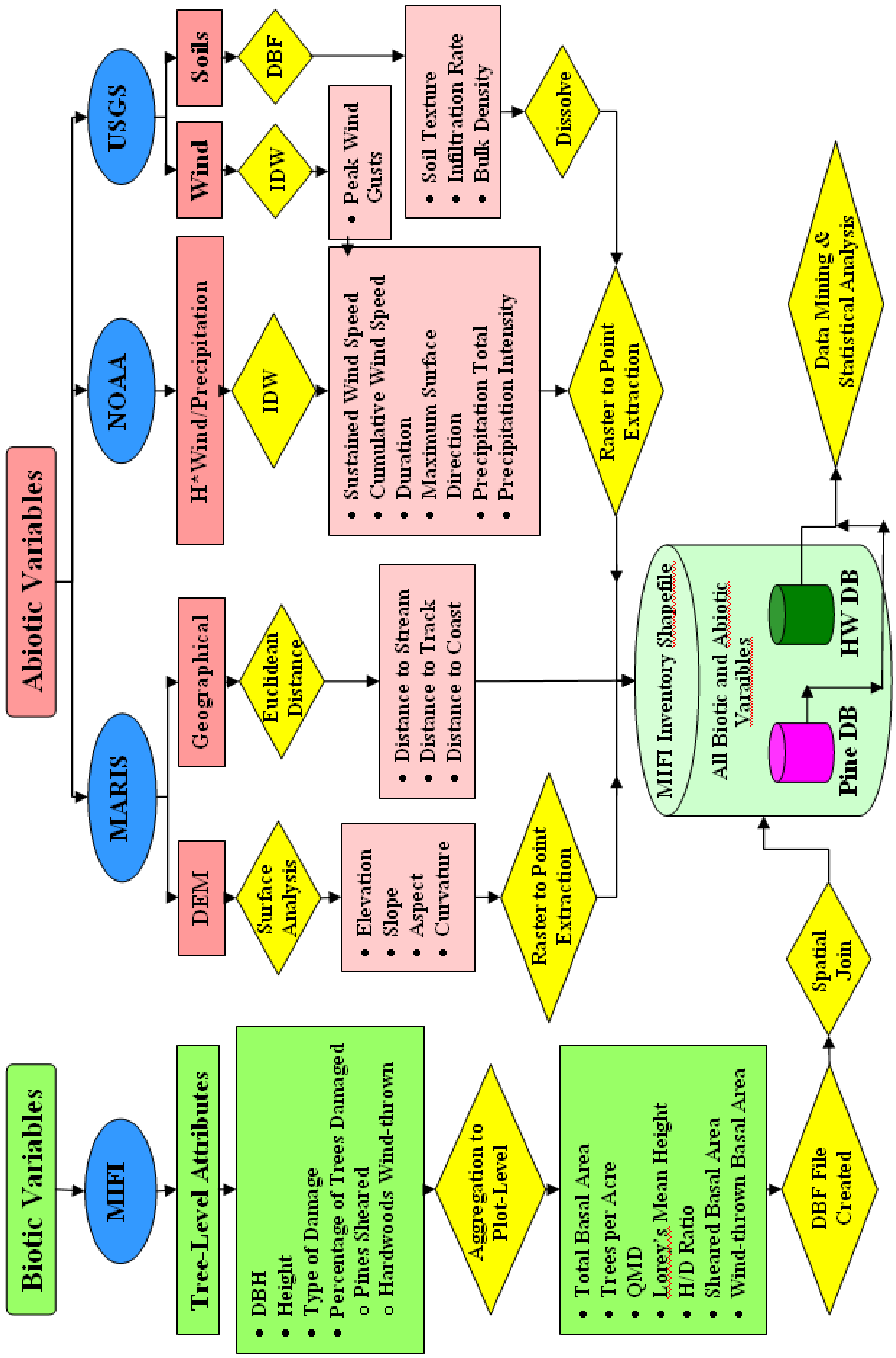

Figure 4.7 Flowchart describing the Database Building Process.

The abiotic factors consist of multiple variables, including both raster and vector datasets. The H*Wind product was accessed from the Hurricane Research Division (HRD) and is a compilation of satellite, aerial, and surface observing platforms (Powell *et al.*1998). The HRD produces 3-hourly intervals of wind-surface analysis available in image and shapefile format. A shapefile stores non-topological geometry and attribute information for the spatial features in a dataset and consists of a main file, an index file, and dBase file (ESRI 2009). Each H*Wind point shapefile contains the attributes of latitude and longitude, maximum sustained wind, and maximum surface direction. The HRD also produces a separate point file which includes the swath of maximum sustained winds, duration, and maximum surface direction. Inverse Distance Weighting (IDW) interpolation was implemented for each H*Wind point file resulting in a continuous raster grid with a resolution of 4km. Multiple time steps of Katrina H*Wind analysis were downloaded to develop the cumulative wind speed layer over from 06Z (1am) the morning of Katrina's landfall (~12Z or 7am) to 00Z (7pm) that evening. The IDW procedure was also applied to these point shapefiles in the same manner. The continuous raster wind-fields were summed within the GIS. Peak wind gusts were not included in the HRD H*Wind Analysis product, however, the USGS provided a point shapefile based on surface observations stations. IDW was again implemented for peak wind gust field at 4km x 4km resolution to produce a similar grid to the H*Wind grids. In addition to continuous wind fields, continuous precipitation fields were also generated across the study area.

Precipitation data were accessed from the National Center for Atmospheric Research (NCAR) Earth Observing Laboratory (EOL). The Stage IV data are based on the multi-sensor hourly/6-hourly State III analyses produced at twelve River Forecast Centers (RFCs) around the U.S (NCAR/EOL 2009). The original downloaded files are projected to a 4km polar-sterographic Grided Binary (GRIB) grid that was subsequently re-projected to Mississippi Transverse Mercator (MSTM) and decoded. Time series grids spanning from 13Z (7am) August 28th through 00z (7pm) August 29th were compiled and decoded using "Degrib." Degrib is a software program maintained by the NWS to decode the National Digital Forecast Database (NDFD) output which uses grib2 files. Degrib allows the user to create large or small polygon, or point shapefiles from the input grib2 file and specify a file path and location. A total of 37 hours were decoded and converted to point shapefiles. IDW interpolation was again utilized to generate continuous precipitation fields across the study area at 4km x 4km resolution. Since the precipitation grids units were in inches and the time step of one hour was known, the precipitation intensity could then be calculated (in/hr). Precipitation intensity was determined using a maximum function in the GIS. The maximum function determines the greatest value of each pixel present in a series of overlaid raster datasets. When this function is executed, the pixel values are reassigned to the maximum value present at each cell location over the time period. A GIS cannot process all 37 grids at one time so the maximum function was used on 6-hour blocks of hourly data. The final step was finding the maximum pixel value of each of the 6 hour blocks that generated a continuous precipitation intensity field across the study area. H*Wind products, peak wind gusts, and precipitation variables were assembled into continuous fields and their respective values at MIFI field plot locations were extracted for analysis. This extraction method was also utilized for the topographical variables.

The topographical variables of elevation, slope, aspect, and curvature were derived from a 10 meter mosaiced DEM of the study area. The original DEMs used were 10 meter county DEMs acquired from MARIS. They were mosaiced in the GIS. Once mosaiced, data gaps existed at some county boundaries. This was remedied by constructing and moving a 9x9 averaging window across the extent of raster data using the a GIS focalmean function. For each cell location in the input grid, focalmean uses the mean values within a specified neighborhood (9x9) and places that value in the corresponding cell location in the output grid (Easson and Robinson 2008). The mosacied grid of the original DEMs and the focalmean grid using the 9x9 window were then merged, only filling the pixels that originally had no data values. Once all data gaps had been filled, the FILL function was used to fill in sinks and level peaks that represented probable errors in the DEM data. Using the focalmean and FILL functions creates and averages out erroneous data present in the DEM. Surface Analysis under spatial analyst in ArcMap 9.2 was then used to determine the slope (degrees), aspect (degrees), and the surface roughness or curvature. These continuous values, like the wind and precipitation values, were extracted at MIFI inventory plot locations. In addition to the DEMs, MARIS was a source for several vector GIS files.

Vector files acquired from MARIS included perennial streams, Hurricane Katrina's track, and a polyline file depicting the coast of Mississippi. Previous research of Everham and Brokaw (1996), Wang and Xu (2008), and Kupfer et al. (2008) suggested forests closer to streams, to the hurricane track, and the coast were more likely to be damaged. Three new attribute fields were created in the MIFI inventory plot file and the distance from the coast, from streams, and from the hurricane track for each plot was determined from these polyline files for the study region. Everham and Brokaw (1996) and Wang and Xu (2008) also have found soil properties such as texture and infiltration rate to be significant in determining susceptibility to wind-throw and shear.

The soil properties of texture, infiltration rate, and bulk density were obtained through previous work of Arias (2008) which used Soil Water Atmosphere Plant (SWAP) software that simulates transport of water, solutes, and heat in the vadose zone in interaction with vegetation development. Texture, infiltration rates, and bulk densities were determined for all soil associations in Mississippi by Arias (2008). The data from Arias (2008) was in Excel spreadsheet format and organized according to soil associations. Therefore, a soil association polygon GIS file was accessed from MARIS and sorted in alphabetical order to match that Excel database of Arias (2008). The information from the Excel spreadsheet was then copied and pasted into the attribute file of the newly downloaded soil association polygon GIS file. Three copies were made of the soil association file, each one displaying one of the 'distance from' attributes. A dissolve function was used to aggregate the smaller soil association polygons to larger polygons based on the attribute field. The categorical values of soil texture and infiltration rate and numerical bulk density values were extracted to the MIFI plot locations.

Through modification, the MIFI inventory plot point file contains all biotic and abiotic variables and their association with type of damage recorded for 1637 plots across the 15 counties of the Southeast Forest District. The plots were then queried by hardwood and pine forest type and exported as new GIS layers with hardwood and pine forest types separated into their own databases (Figure 4.2.1.1). Table 4.7 identifies all biotic and abiotic variables now present for all 1637 MIFI Inventory plots that were analyzed using [R 2.8.0] for CART analysis and by SPSS 15.0 for logistic regression.

Table 4.7 Biotic and Abiotic Variables, Units, and Resolution aggregated to the MIFI Inventory Plot point shapefile.

Variable	Units	Resolution(if applicable)
Biotic		
Percentage of Trees Wind-Thrown	%	NA
Percentage of Trees Sheared	%	NA
Sheared Basal Area	ft^2/plot	NA
Wind-Thrown Basal Area	ft^2/plot	NA
Total Basal Area	ft^2/plot	NA
Trees per Acre	-	NA
Quadratic Mean Diameter (QMD)	In	NA
Lorey's Mean Height (LMH)	Ft	NA
Height to Diameter Ratio (H/D)	-	NA
Forest Coverage Classification	Categorical	NA
Stand Condition Classification	Categorical	NA
Stand-Level Forest Damage	Categorical	NA
Abiotic		
Meteorological		
Sustained Wind Speed	MPH	4km x 4km
Peak Wind Gusts	MPH	4km x 4km
Cumulative Wind Speed	MPH	4km x 4km
Duration	Hr	4km x 4km
Maximum Surface Direction	°	4km x 4km
Precipitation Storm Total	In	4km x 4km
Precipitation Intensity	In/hr	4km x 4km
Topographical		
Elevation	Ft	10 Meters
Slope	°	10 Meters
Aspect	Categorical	10 Meters
Surface Roughness/Curvature	-	10 Meters
Topographic Position	Categorical	NA
Geographical		
Distance to Perennial Streams	Miles	NA
Distance to Hurricane Track	Miles	NA
Distance to Coast	Miles	NA
Pedological		
Soil Texture	Categorical	NA
Infiltration Rate	Categorical	NA
Bulk Density	g/cc	NA

Data Mining Operation and Code using [R 2.8.1]

Prior to implementing analysis procedures, the MIFI inventory point database was queried first for only hardwood plots and then for pine plots in order to separate both spatial and tabular data by forest type (Figure 2.19). The respective database (attribute table) files of both hardwood and pine GIS files were imported into Excel and saved as both a comma separated value (.csv) format and workbook (.xls) format in separate directories. The (.csv) files where then brought into [R 2.8.1], a free software environment for statistical computing and graphics developed at Bell Laboratories (formerly AT&T) by John Chambers and colleagues (Leisch 2003). [R 2.8.1] was downloaded from the Comprehensive R Archive Network (CRAN) which hosts the R-code as it is updated. For this project, the CRAN host selected was the University of California, Berkley (http://cran.cnr.Berkeley.edu). Concurrent with downloading [R 2.8.1], three extensions were also downloaded. These three extensions included rpart, boot, and bootstrap. Rpart, short for Recursive Partitioning, is the primary package used for recursive portioning and regression trees (i.e. CART analysis) (Ripley 2009). Boot and bootstrapping allow for summary statistics and Confidence Intervals (CIs) of each terminal node value to be determined.

Growing Regression Trees

To complete the first objective, rpart, boot, and bootstrap packages were loaded into the active R-Console window within [R 2.8.1]. The methodology of entering R-code to yield a regression tree will be described below using the dependent response variable of percentage of sheared pines. The same methodology was used for the dependent response variables of: percentage of hardwoods wind-thrown on hardwood plots, basal area of pines sheared on pine plots, and basal area of hardwoods wind-thrown on hardwood plots. Example R-code from previous research by Dr. Fan of the College of Forest Resources at Mississippi State University was utilized for this study. The R-Code was tailored to fit the dependent response and independent driving variables amassed in the GIS Database along with R-code determining the confidence intervals for each successive split. This allowed for R-code to be copied and pasted from Notepad into the R-Console. The R-code containing the pine (.csv) file was copied and pasted into the R-Console which then reads all data contained within the (.csv) file. Next, the name for the over-fit regression tree is set as PN_shear_P.tree and the rpart package is called to action with the dependent response variable stated first, followed by all other biotic and abiotic variables (Figure 4.8). The method of ANOVA was specified since the recursive partitioning was based on a continuous dependent variable. If the dependent variable had been categorical in nature, the method would have been specified as CLASS for classification tree. This option was not used in this study. After the rpart package was called to action, [R] subsequently data mines all variables to determine which biotic and abiotic variables and their respective values were important and produced an over-fit regression tree. To view the regression tree, a post script file was created and a title was specified, “Fully-Grown Regression Tree for Pine Shear Percentage.” The post script file was then opened using Adobe Acrobat 8.0.

```
> local({pkg <- select.list(sort(.packages(all.available = TRUE)))
+ if(nchar(pkg)) library(pkg, character.only=TRUE)})
> local({pkg <- select.list(sort(.packages(all.available = TRUE)))
+ if(nchar(pkg)) library(pkg, character.only=TRUE)})
> local({pkg <- select.list(sort(.packages(all.available = TRUE)))
+ if(nchar(pkg)) library(pkg, character.only=TRUE)})
> wind=read.table(file="G:/Jared_Thesis_Work/Forest_Data/Forest_Correlations/Tr
> PN_shear_P.tree=rpart(SHEAR_P~CU_TOTAL+DURATION+HRD_GUST+HRD_SUS+ELEVATION+SL
> post(PN_shear_P.tree,title="Fully-Grown Regression Tree for Pine Shear Percen
>
```

Figure 4.8 R-Code depicting Package Loading, (.CSV) File Loading, Defining Regression Tree Name and Input and Creation of Post Script File.

The over-fit regression tree was printed out and each node was numbered starting with root node. Nodes were numbered in ascending order as nodes split to the left and then to the right. The terminal nodes' respective position (i.e. number) was later used in R-coding to determine the confidence interval for each successive split. The next task was to prune the over-fit regression tree using two different methods.

Pruning Regression Trees

Initial pruning for the over-fit regression tree used the highest cross-validation error less than one standard error above the minimum cross-validation error (Thereau *et al.* 1997). This was determined by analyzing the cost-complexity parameter (CP) table in [R]. This table was displayed by entering "printcp(PN_shear_P.tree) (Figure 4.9)." The rpart regression tree formula that constructed the tree, the variables actually used in the tree, the root node error, sample size, and summary statistics for different sized regression trees are shown in Figure 4.9. The minimum "xerror" or cross-validation error was added to the "xstd" (standard deviation) creating the one standard error (1-SE) bar. The resulting value was then used to determine the proper number of splits the optimal tree should have before over-fitting occurred.

```
> plotcp(PN_shear_P.tree)
> printcp(PN_shear_P.tree)

Regression tree:
rpart(formula = SHEAR_P ~ CU_TOTAL + DURATION + HRD_GUST + HRD_SUS +
    ELEVATION + SLOPE + ASPECT + SCF_RF + PRECIP_INS + PRECIP_T +
    K_TRACKD + STREAM_D + COAST_D + TEXTURE + STAND_COND + PHYS_POS +
    BULK_DEN + INFIL_RATE + M_SFC_DIR + QMD + TREES_ACRE + H_D_RATIO +
    L_M_H, data = wind, method = "anova")

Variables actually used in tree construction:
[1] ELEVATION H_D_RATIO HRD_GUST  L_M_H     M_SFC_DIR QMD

Root node error: 89059/784 = 113.60

n= 784

        CP nsplit rel error  xerror    xstd
1 0.112644      0   1.00000 1.00059 0.14158
2 0.064274      2   0.77471 0.89240 0.11044
3 0.050965      3   0.71044 0.87129 0.11007
4 0.040547      4   0.65947 0.86951 0.10849
5 0.038230      5   0.61893 0.89889 0.11152
6 0.036371      6   0.58070 0.90580 0.11123
7 0.032002      7   0.54433 0.90567 0.11370
8 0.010633      8   0.51232 0.89873 0.11497
9 0.010000      9   0.50169 0.89680 0.11057
>
```

Figure 4.9 PrintCP Table showing [R] Output of Important Variables used and the relative CP values for varying sized Regression Trees for the Percentage of Pines Sheared.

In addition to this value, the number of splits was also determined by plotting the cross-validation relative error against the Cost-complexity Parameter (CP) value (Figure 4.10). The CP value is a measure of how much additional accuracy a split must add to the entire tree to warrant the additional complexity. As the cost-complexity parameter value increases, a greater number of nodes (decreasing to increasing importance) are pruned away, resulting in a simpler decision tree. A third axis on the graph depicts the number of splits that an optimal tree will have depending on the CP value used to prune the tree. The relative error is equal to 1-R^2, with R^2 being the correlation coefficient that measures how well the data fit the dependent response variable. Therefore, the lower the relative error, the better the fit the data are to the dependent response variable. However, as the number of splits increase, the R^2 will be higher due to over-fitting of many parameters to the dataset. By using the 1-SE rule, it becomes a balance of higher decision cost for each split and the amount of information and interpretability of the results. Ripley (2009) suggests a good choice for CP pruning is often the left-most value for which the mean lies below the horizontal line. Based on Figure 4.10, the 1-SE rule suggests a tree supporting 6 terminal nodes or 5 splits can be used as the optimal tree. However, quantitatively this tree was not the best-fit model according to the xerror values. The minimum xerror values showed the optimal tree is one that only has 4 terminal nodes with three splits (Figure 4.10). This led to the second method of pruning by using the minimum xerror and its corresponding CP value. This method is known as minimum CP pruning. Minimum CP pruning was carried out on all regression trees for each dependent response variable. R-Code created for each dependent variable was imported in the R-Console and the CP value corresponding to the lowest cross-validation error was used to automatically prune the regression trees.

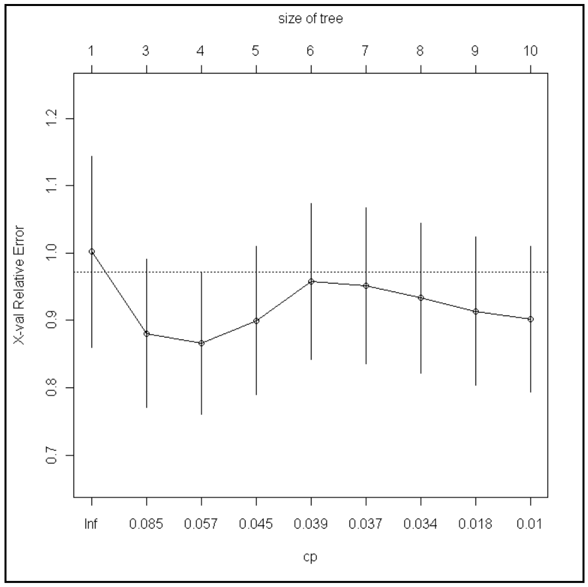

Figure 4.10 Example of PLOTCP in [R 2.8.1] to Determine the Optimal Size of the Regression Tree.

In addition to the lowest cross-validation error being used to prune the over-fit regression tree, the R^2 values of each split was also used to determine when successive splits stopped increasing the correlation to percentage of damaged trees or basal area on hardwood and pine plots. The apparent R^2 value is derived by subtracting the relative error by one and the X-Relative R^2 is determined by subtracting one from the cross-validation error (Difford 2008). The apparent R^2 value represents the classification accuracy of the data being utilized through that respective split. The X-Relative R^2 depicts the optimal tree when this value is at its maximum. Recall that when the minimum CP method of pruning is used, this is actually using the minimum cross-validation error (xerror). Since the X-Relative error is calculated by subtracting one from the cross-validation error, the maximum value of the X Relative graph in Figure 4.11 depicts the optimal tree size. This command was entered into the active R-Console as rsq.rpart(PN_shear_P.tree) and produces a R^2 and CP graphs in a new [R] window (Figure 4.11).

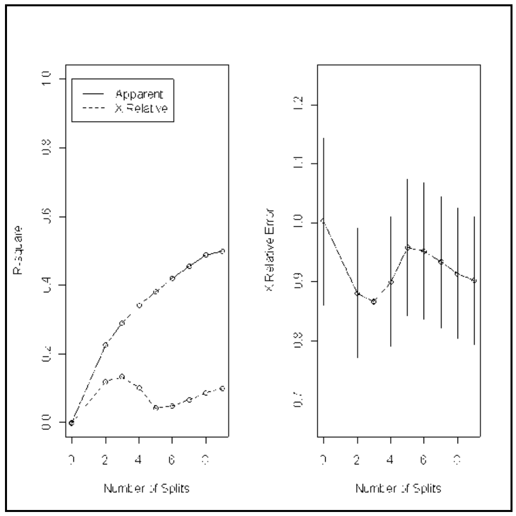

Figure 4.11 R^2 and Cost-complexity Parameter (CP) graphs depicting Split Correlation and Decision Cost.

Determination of 95% Confidence Intervals

As mentioned above, three extensions were downloaded and installed along with [R 2.8.1]. Two of those extensions, boot, and bootstrap were used to determine the 95% confidence intervals of the means of each node. Both extensions compile a large number of datasets and a distribution of the mean value for each node can be determined. Boot and Bootstrap packages allow for resampling of the dataset to estimate the standard error of the mean by repeatedly drawing “bootstraping samples” from the original data, re-evaluating each median value for each bootstrap sample, and estimating the standard error of the original median by the observed variability in the bootstrap means (Efron 2009). This re-sampling method allows for the estimation of the distribution of the sample's statistics. Potential statistics derived using the bootstrapping method include the mean, median, standard deviation, and quartiles of a given dataset. A schematic of how bootstrapping works is shown in Figure 4.12. The biggest advantage of bootstrapping over other summary statistic analytics is the straight-forward application to derived estimates of standard errors and confidence intervals for complex estimators of complex parameters within the distribution (Burns 2008).

As mentioned above, three extensions were downloaded and installed along with [R 2.8.1]. Two of those extensions, boot, and bootstrap were used to determine the 95% confidence intervals of the means of each node. Both extensions compile a large number of datasets and a distribution of the mean value for each node can be determined. Boot and Bootstrap packages allow for resampling of the dataset to estimate the standard error of the mean by repeatedly drawing "bootstraping samples" from the original data, re-evaluating each median value for each bootstrap sample, and estimating the standard error of the original median by the observed variability in the bootstrap means (Efron 2009). This re-sampling method allows for the estimation of the distribution of the sample's statistics. Potential statistics derived using the bootstrapping method include the mean, median, standard deviation, and quartiles of a given dataset. A schematic of how bootstrapping works is shown in Figure 4.12. The biggest advantage of bootstrapping over other summary statistic analytics is the straight-forward application to derived estimates of standard errors and confidence intervals for complex estimators of complex parameters within the distribution (Burns 2008).

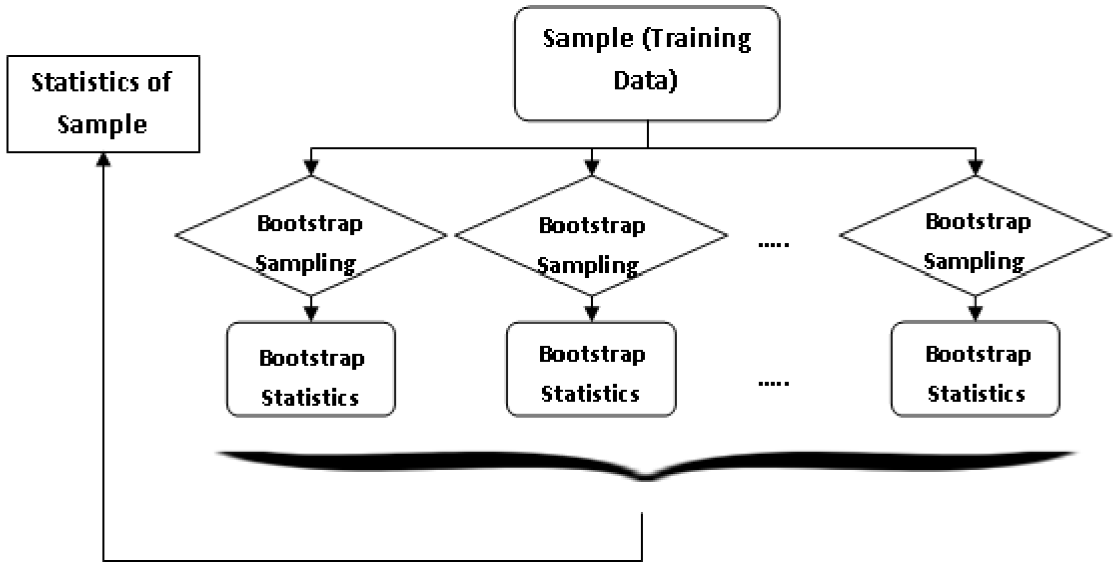

Figure 4.12 Schematic Diagram illustrating the Concept of Bootstrapping.

Previously compiled R-code was obtained from Dr. Fan of the CFR at Mississippi State University for boot and bootstrapping packages. The R-code was tailored to each dependent response variable tested. The nodes numbering from the printouts were used from each full, 1-SE, and min CP regression tree to determine the 95% Confidence Interval (CI) for the mean values of all terminal nodes. Using the bootstrap package, the sample data set of pine (N = 784 plots) was resampled 1000 times to determine the upper and lower 95% confidence intervals. The same number of resamplings was carried out on each dependent response variable. The mean values, along with their upper and lower bounds, of the full, 1-SE, and minimum CP values were input into an Excel spreadsheet and graphed.

Forward Stepwise Logistic Regression using SPSS 15.0

Logistic regression, a multivariate technique that uses a logit function to predict the outcome of a dichotomous or polytomous response, was utilized to complete the second and third objectives of this study. The second objective was to compare the CART analysis with forward stepwise logistic regression and their respective identification of significant variables for wind-throw of hardwoods and shear of pines. In the amassed MIFI Inventory plot point shapefile, pine plots were queried for plots that were field-classified as sheared plots and were given the code as one. Pine plots not classified as sheared plots were coded as zero. The same coding was applied to hardwoods with respect to wind-throw plot classification. Hardwood plots deemed by MIFI to have suffered wind-throw were coded with a value of one, while hardwood plots that were not wind-thrown were coded as zero. The attribute tables were saved as workbooks (.xls) in Excel and opened in SPSS 15.0 for statistical analysis on the bivariate plot damage classifications as they related to the biotic and abiotic variables.

Unlike CART methodology, logistic regression has been utilized more frequently for bivariate analysis on stand-level damage classifications (Peterson, 2007, Oswalt and Oswalt 2008, Xi et al. 2008). Logistic regression is the best technique for relating binary response variables (wind-throw occurrence or not, shear occurrence or not) to categorical or continuous biotic and abiotic independent variables (Peterson 2007). Once files were imported into SPSS 15.0, a forward stepwise logistic regression analysis was initiated. When forward logistic regression is employed, the computer begins with a model that includes only a constant value and then adds single predictor variables into the model based on the significance of the Wald statistic. The Wald statistic, otherwise known as

Wald χ^2 statistic is used to test the significance of individual coefficients in the model and is calculated by Eq. 4.4

$$Wald = \left(\frac{B}{SE}\right)^2 \qquad \text{Eq. 4.4}$$

where B is the coefficient of the independent variable and SE is the standard error of the variable. In addition to the Wald statistic, the log-likelihood ratio was used to compare each logistic regression model. Lower log-liklihood ratio values result in better model performance (Wuensch 2008). Models were built using forward selection of variables, with a threshold of $p<0.05$ to enter a variable into the equation and $p>0.10$ to remove a variable from the equation. The overall significance of the logistic regression models was tested using the -2 log likelihood ratio, while the Wald statistic was used to evaluate the significance of individual variables within the models for sheared pine and wind-thrown hardwoods. Pseudo R^2 values known as Cox and Snell R^2 and Nagelkerke R^2 values are also used as a measure of strength of association. Cox and Snell's R^2 is an attempt to imitate the interpretation of of multiple R-Square based on the log likelihood of the final model versus log likelihood for the base line model, but its maximum can not be one (Garson 2009). Nagelkerke's R^2 is a modification of the Cox and Snell coefficient to assure that it can vary from zero to one (Garson 2009). This approach divided Cox and Snell R^2 by its maximum in order to achieve a measure that ranges from zero to one. Thus, Nagelkerke's R^2 will normally be higher than the Cox and Snell measure and will tend to run lower than the corresponding ordinary least squares R^2. The results of the forward stepwise logistic regression models were compared to the biotic and abiotic variables that were important in the CART analysis for pine and hardwood shear and wind-throw percentages. MIFI personnel used aerial photographs and conditions observed in-route to plots to classify plot damage types and did not base estimates on aggregation of tree-level damage incurred to plot-level statistics. This fact leads to the third objective which was an assessment of the accuracy of MIFI's original classification of plot-level damage classes.

To determine the accuracy of MIFI's original classification, the binary plot-level information damage was analyzed using the same forward stepwise logistic regression procedure applied for the second objective. However, to determine the accuracy of predicting whether a plot should be categorized as sheared or wind-thrown, the binary plot damage classifications were recoded according to a set threshold value. This threshold value was chosen on the basis percentage of wind-throw and shear per plot from the tree-level data. Histograms of all plots and the respective wind-throw and shear percentages were generated and identified a majority of plots suffered less than 10% for each respective damage type (Figures 4.13 and 4.14).

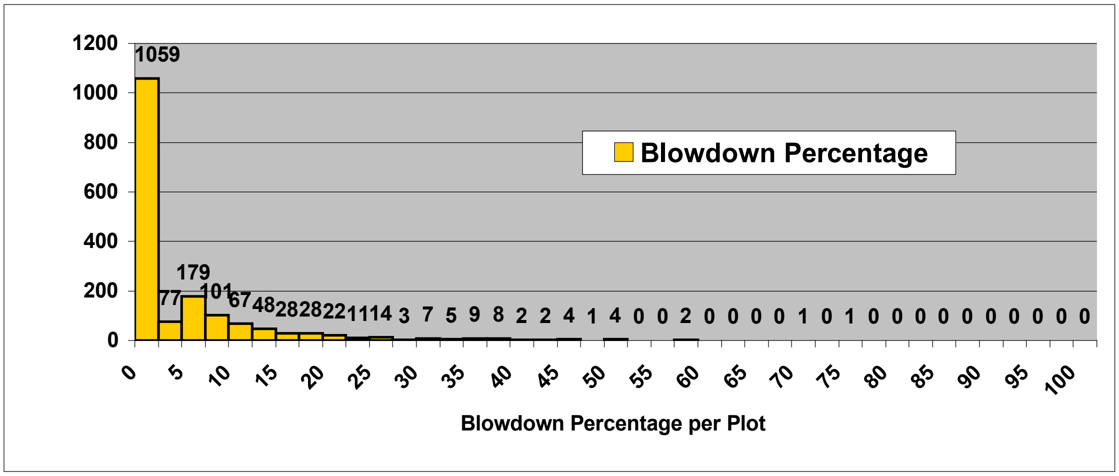

Figure 4.13 Plot Distribution of Blowdown Percentage for All Plots from Tree Level Data.

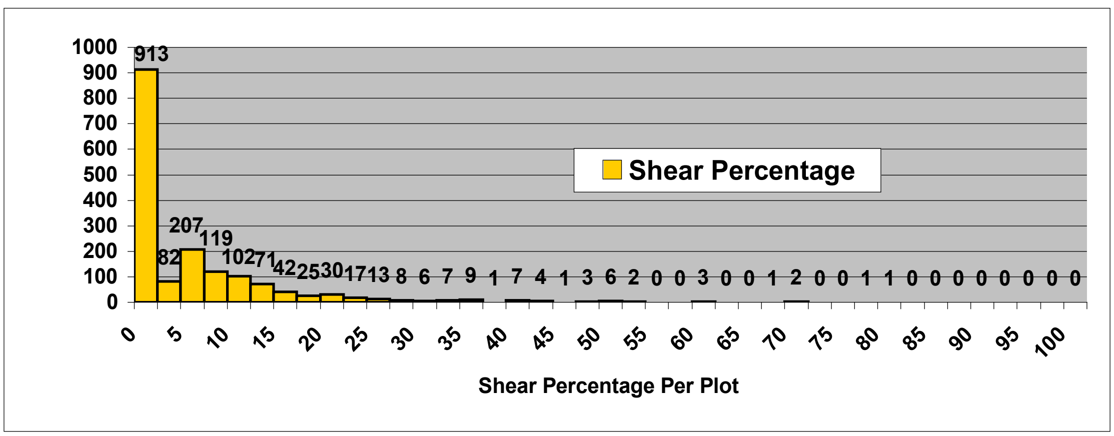

Figure 4.14 Plot Distribution of Sheared Plot Percentage for All Plots from Tree Level Data.

Since a majority of plots suffered less than 10% of wind-throw or shear damage, plots that experienced more then 10% damage were reclassified as wind-throw damage if the plot was hardwood and shear damage if the plot was pine. A new attribute field was created for the recoded plot classifications to the MIFI Inventory plot point shapefile. Again using SPSS 15.0 forward stepwise logistic regression was performed on the recoded classifications and compared to MIFI's original damage plot classifications. Forward stepwise logistic regression was used to determine which variables and their interactions were significant to predicting the damage classification of a given plot. In addition, each individual biotic and abiotic variable was assessed in determining its role in plot-level classification accuracy through binary logistic regression. Each variable was analyzed using the logistic regression equation through SPSS 15.0. Since the coefficients in the logistic regression equation are in log-odds units the use of Equation 4.5 allows for a back-transformation to a predicted probability where P is the probability, *e* is the base of the natural logarithm and *a* is the constant and *b* adjusts how quickly the probability changes with changing X a single unit.

$$P = \frac{e^{a+bX}}{1+e^{a+bX}} \qquad \text{Eq. 4.5}$$

This allowed for each individual biotic and abiotic variable value to be correlated to a predicted probability of wind-throw or shear at the stand-level.

CHAPTER 5

RESULTS

This chapter focuses on the biotic and abiotic variable interaction and the relationship to wind-throw damage of hardwood plots and shear damage of pine plots across southeastern Mississippi following Hurricane Katrina. Results of each fully-grown regression tree, 1-SE pruned tree, minimum CP pruned tree and R^2 split-correlation and decision-cost graphs are presented. The 95% confidence intervals of the terminal nodes mean values are also displayed. Next, forward stepwise logistic regression identification of significant biotic and abiotic variables is compared to CART analysis findings. Original MIFI damage plot classifications are compared to the quantitatively reclassified damage plots.

CART Results

The CART results below are primarily displayed and explained based on the regression tree diagrams. An example below can be used as an aid to understand the regression tree layout and numerical output shown in each node (Figure 5.1).

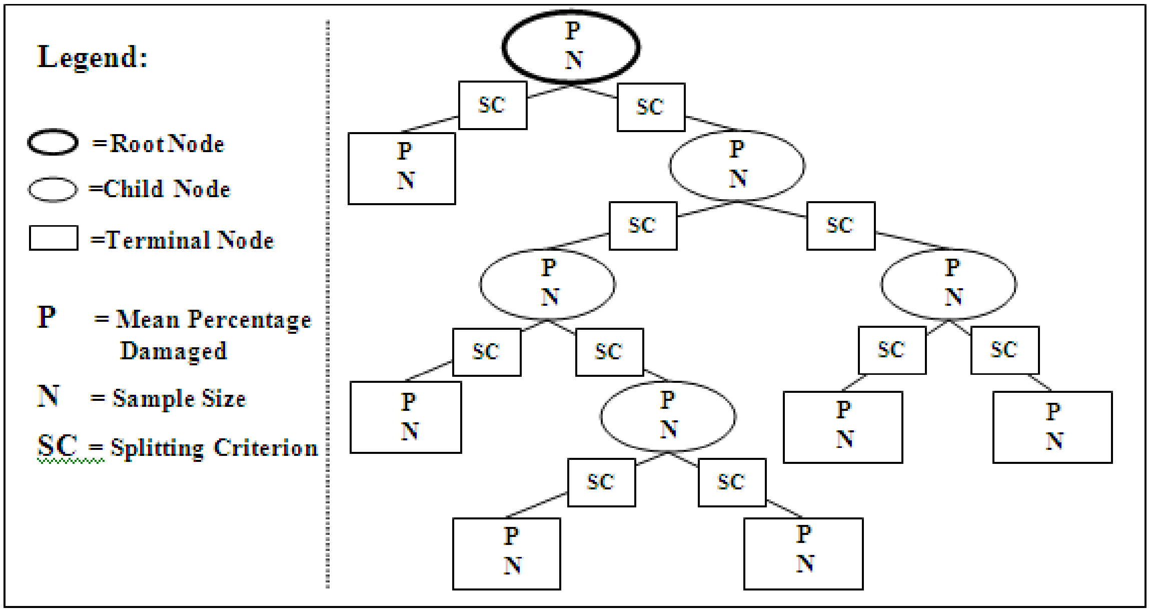

Figure 5.1 Example of Regression Tree Output

The oval at the top of the regression tree is the root node which displays the mean percentage of the dependent response variable tested (P) and the sample size (N) contained within that specific node. Therefore, the sample size (N) value of the root node contains the entire dataset for either pine or hardwood plots, depending on which is being tested. Each splitting criterion (SC) showing the biotic or abiotic variable responsible and its corresponding value is displayed between the parent node and subsequent child or terminal node. As a general rule, as nodes split to the right, the mean percentage of trees or basal area damaged increases, while splits to the left show decreases of mean percentage of trees or basal area damaged.

Percentage of Pines per Plot Sheared

The fully-grown regression tree for percentage of sheared pines per plot reveal pine plots having greater than 8 inches QMD had 3.5 times more damage than plots that averaged less than 8 inches QMD (Figure 5.1). The next split shows pine plots with a Lorey's Mean Height (LMH) less than 43.5 feet experienced 3.9 times the shear damage than those with LMH vales greater than 43.5 feet. The regression tree then separates into two main branches but interestingly uses the same variable in the following split, peak wind gusts. For the left branch, pine plots greater than 43.5 feet LMH that experienced over 130 mph had 3 times the amount of damage opposed to plots that experienced less than 130 mph. QMD is again used in the next split further reinforcing that pines with greater diameters suffered greater shear damage. Further splits in the left branch are made but are most likely the result of over-fitting of the data. Focusing on the right branch were sheared damaged to pine plots is maximized shows pine plots less than 43.5 feet LMH and prone to peak wind gusts greater than 138 MPH had the highest percentage of shear damage at 47%. Pine plots experiencing wind gusts less than 138 mph but having higher height to diameter (H/D) ratios of 54 had 15 times the amount of shear damage than on plots where H/D ratio was less than 54. The last split on the right branch, even through it over-fit to the dataset, depicts the higher the elevation of the stand, the greater chance of shear damage was to occur. The values stated above are means of a vast data mining array not accessible or viewable by the user. However, bootstrapping statistics can be performed on this array to determine the upper and lower bounds (i.e. confidence intervals). Knowledge of the upper and lower 95% confidence intervals (CI) reveal the relative robustness of each mean node value.

The fully-grown regression tree CIs for each terminal node of percentage of pines per plot experiencing shear is illustrated in Figure 5.2. The first four terminal nodes (Nodes 2, 5, 7, 10) had relatively small intervals leading to greater robustness and confidence in the mean values. Nodes further down the tree and those that split right tend to have greater intervals. The largest interval occurred for Node 18. Referencing back to Figure 5.1, the bottom right terminal node splits right and is based on an elevation splitting criteria. A good majority of values from elevation terminal nodes equal or surpass the mean value of node the peak wind gust terminal node which was also split right and based on peak wind gust speed. However, node 18 is most likely over-fit to the data and will subsequently be pruned away. The presence of elevation however, hints that is this variable may be important for other dependent response variables or other statistical procedures. The statistics of the cptable were then examined to determine the classification accuracy of the full regression tree.

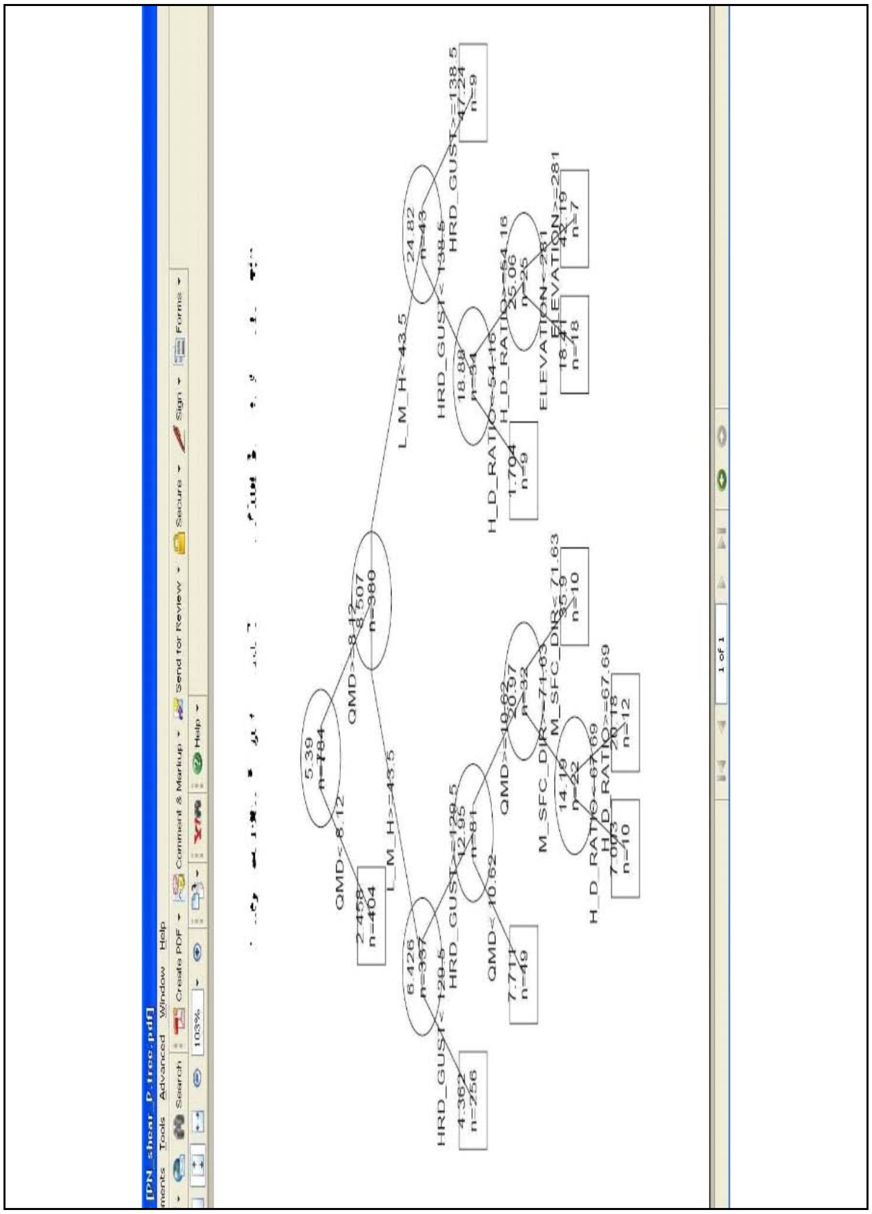

Figure 5.2 Fully-Grown Regression Tree for Pine Shear Percentage.

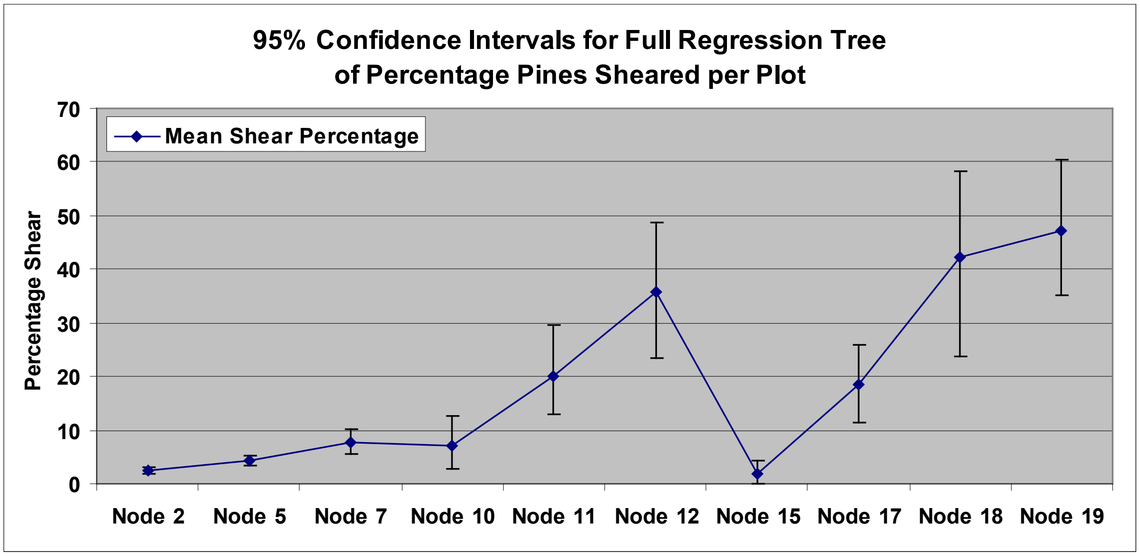

Figure 5.3 95% Confidence Intervals for Full Regression Tree of Percentage Pines Sheared per Plot.

The overall classification accuracy for the full regression model for biotic and abiotic variables to predict the percentage of sheared pines per plot is 50% based off the cost-complexity parameter table (cptable) (Table 5.1). Further investigation of the cptable shows the greatest decrease in the first and second splits with another small drop for the third split for both relative error and xerror columns. The minimum xerror is reached after three splits, suggesting the optimal tree should be pruned to three splits. However, if additional information following the third split was sought out, the CP plot indicates up to 6 splits could be used using the 1-SE rule of pruning (Figure 5.1.1.3). In concurrence with the cptable, the R-Square graph shows that only the first three splits help explain the correlation to percentage of pines sheared per plot (Figure 5.1.1.4). The X Relative R^2 is only 0.13 at its greatest value, thus alluding that percentage of pines sheared per plot are not well correlated to the measured and derived independent biotic and abiotic variables.

Table 5.1 Cost-Complexity Parameter (CP) Table for Full Regression Tree for Percentage of Pines per Plot Sheared.

CP	NSplit	Relative Error	Xerror	Xstd
0.112644	0	1.000	1.000	0.14188
0.064274	2	0.77471	0.89224	0.12492
0.050965	3	0.71044	0.86520	0.12589
0.040547	4	0.65947	0.96200	0.12822
0.038230	5	0.61893	0.94419	0.12560
0.036371	6	0.58070	0.94419	0.12560
0.032002	7	0.54433	0.92225	0.12441
0.010633	8	0.51232	0.91347	0.11963
0.010000	9	0.50169	0.92496	0.11470

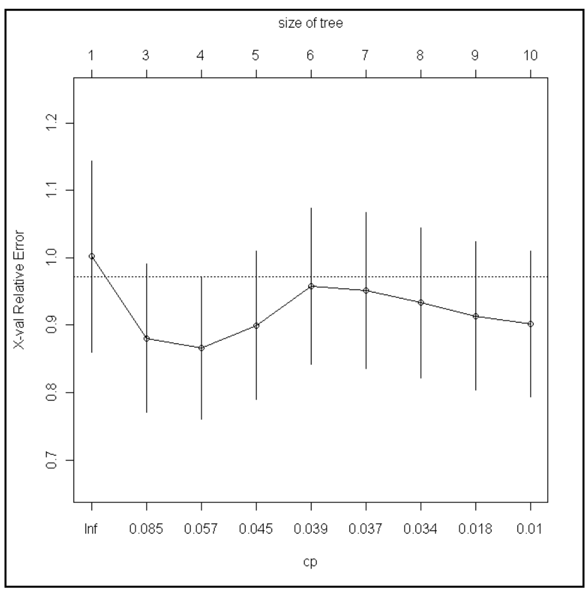

Figure 5.4 Cross-Validation Relative Error Graph with Standard Error Lines for Percentage of Pines per Plot Sheared.

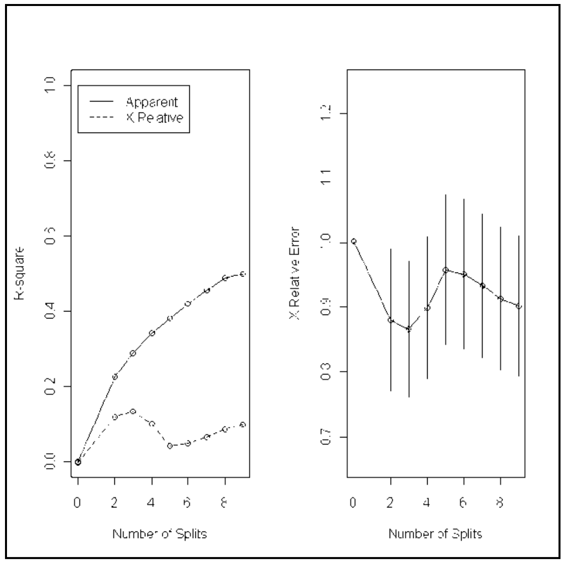

Figure 5.5 Apparent, X Relative R-Square and Cross Validation Relative Error Graphs of 1-SE Pruned Regression Tree for Shear of Pine Plots Percentage.

The fully-grown regression tree was then pruned using Figure 5.4 and the 1-SE rule. The pruned tree now consists of 6 splits or 7 terminal nodes opposed to the 10 splits or 11 terminal nodes of the fully-grown and over-fit regression tree (Figure 5.6). The mean values of remaining nodes do not change as other nodes are pruned away. However, as the nodes are pruned away, the fewer variables are used in the actual regression tree construction. In Figure 5.6, the only variables used now are QMD, LMH, peak wind gust speed and H/D ratio. By removing the terminal nodes fitting the idiosyncrasies of the dataset, the classification accuracy decreases to 42% (1.0-0.58070) (Table 5.2).

Table 5.2 Cost-Complexity Parameter (CP) Table for 1-SE Pruned Regression Tree for Percentage of Pines per Plot Sheared.

CP	NSplit	Relative Error	Xerror	Xstd
0.112644	0	1.00000	1.00172	0.14165
0.064274	2	0.77471	0.88062	0.11011
0.050965	3	0.71044	0.86662	0.10567
0.040547	4	0.65947	0.89973	0.10999
0.038230	5	0.61893	0.95735	0.11568
0.037000	6	0.58070	0.95116	0.11568

As stated above, the greatest decreases in both relative error and xerror occur within the first three splits with the minimum cross-validation error (0.86662) being attained at the third split. The 95% CIs continued to show early nodes splitting to the left have smaller intervals than nodes splitting right and subsequently going through additional splitting criterion (Figure 5.7). The farthest right terminal node has the largest interval with a mean value of 47.24 but its 95% CI upper bound is 60.7 and lower bound is 34.0.

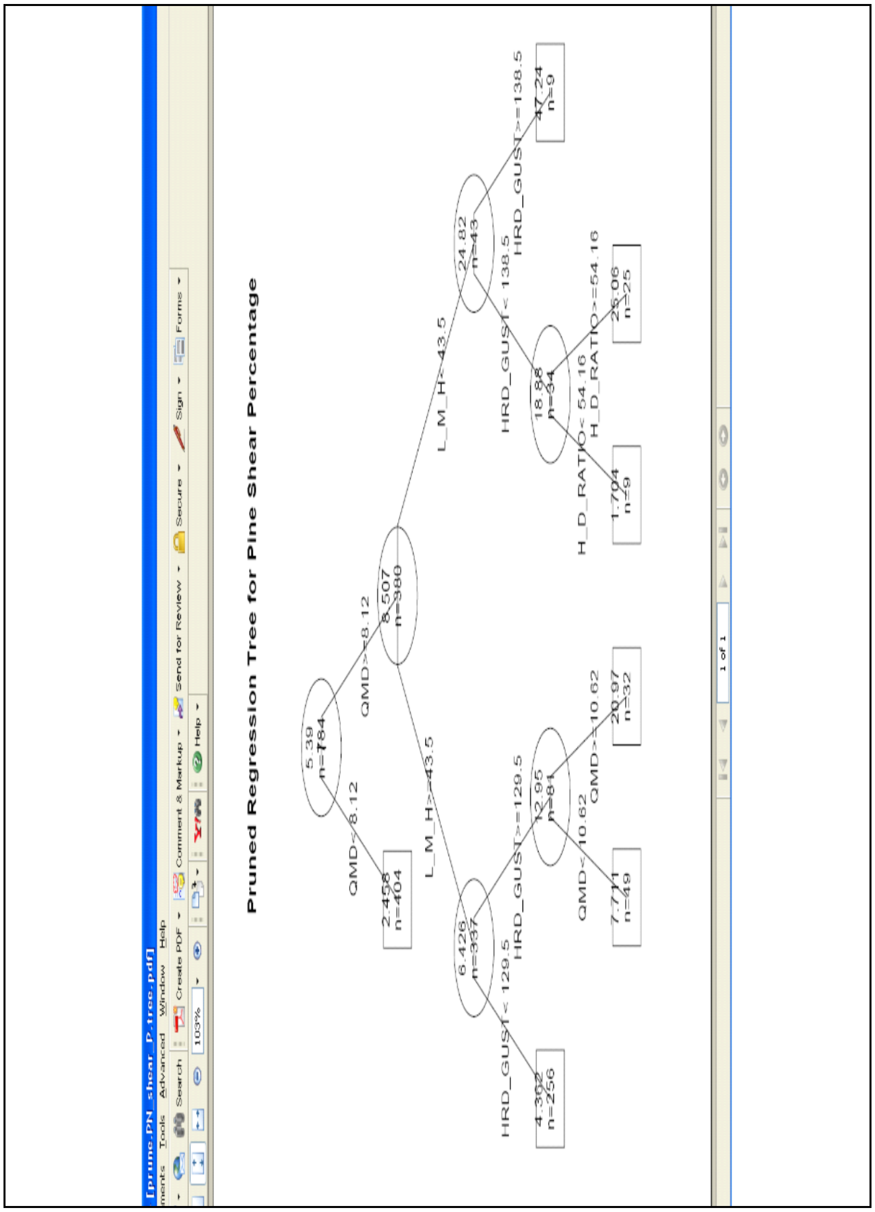

Figure 5.6 1-SE Pruned Regression Tree for Pine Shear Percentage

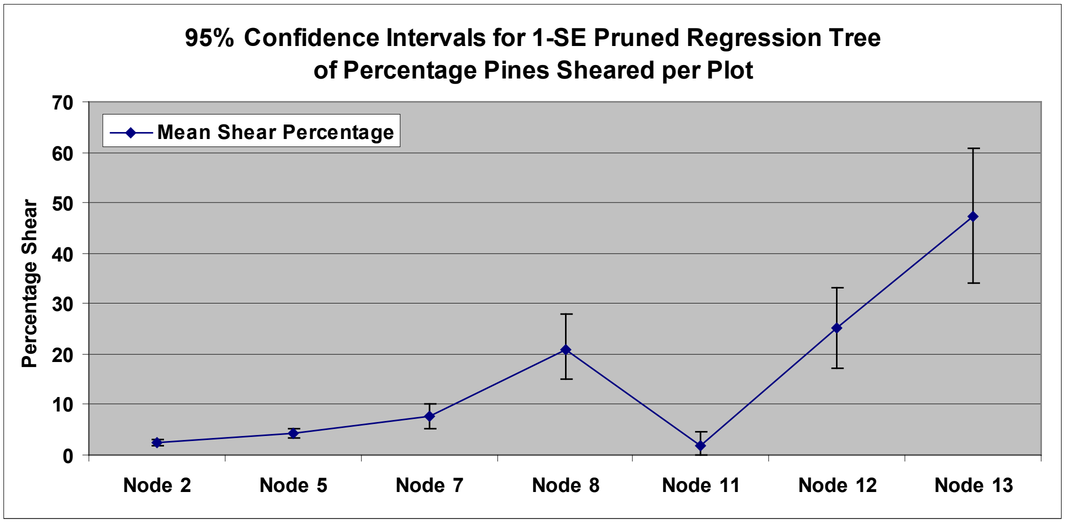

Figure 5.7 95% Confidence Intervals for 1-SE Pruned Regression Tree of Percentage Pines Sheared per Plot.

The 1-SE regression tree was pruned to 3 splits based of the minimum cross-validation error in Table 5.2 and the R-Square graph (Figure 5.8).

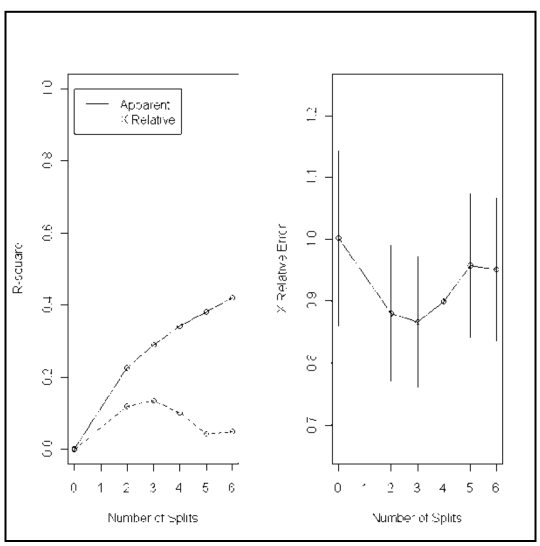

Figure 5.8 Apparent, X Relative R-Square and Cross Validation Relative Error Graphs of 1-SE Pruned Regression Tree for Percentage Pines Sheared per Plot.

The minimum CP pruned regression was then pruned using the CP value of 0.045 (Figure 5.9). This parsimonious regression tree only consists of 3 splits or 4 terminal nodes with only three variables used: QMD, LMH, and peak wind gust speed (Table 5.3). The minimum CP tree had an overall classification accuracy of 39%. Recall the fully-grown regression tree had a classification accuracy of 50% using 6 variables. Only an 11% drop occurred while removing half of the variables used to predict the percentage of pines sheared per plot.

Table 5.3 Cost-Complexity Parameter (CP) Table for Minimum CP Pruned Regression Tree for Percentage of Pines per Plot Sheared.

CP	**NSplit**	**Relative Error**	**Xerror**	**Xstd**
0.112644	0	1.00000	1.00172	0.14165
0.064274	2	0.77471	0.88062	0.11011
0.050965	3	0.71044	0.86662	0.10567

The parsimonious minimum CP regression tree was then smoothed to the 95% CI line with each successive node having a greater interval than the previous node (Figure 5.10). As each regression tree is pruned and subsequent child and terminal nodes are removed, the node number alters according to the number of nodes present.

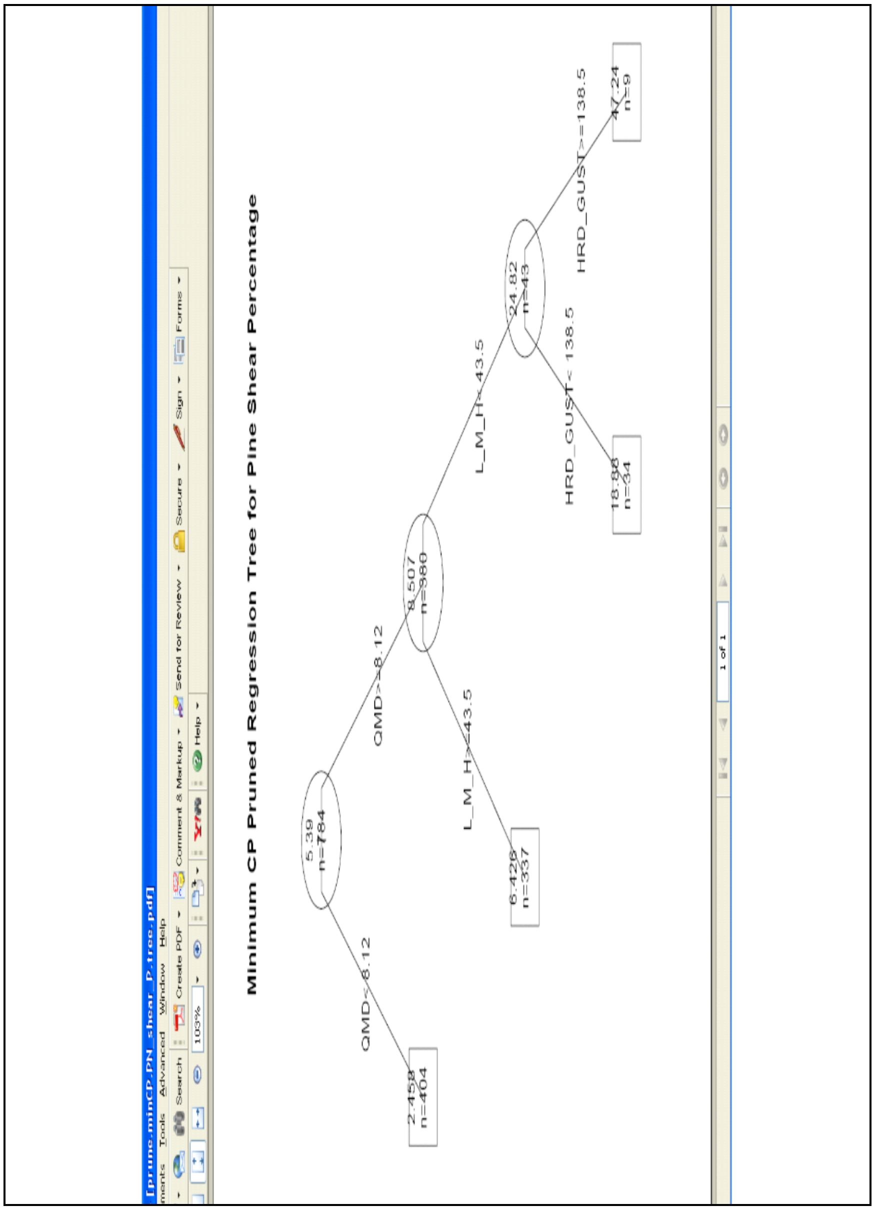

Figure 5.9 Minimum CP Pruned Regression Tree for Pine Shear Percentage

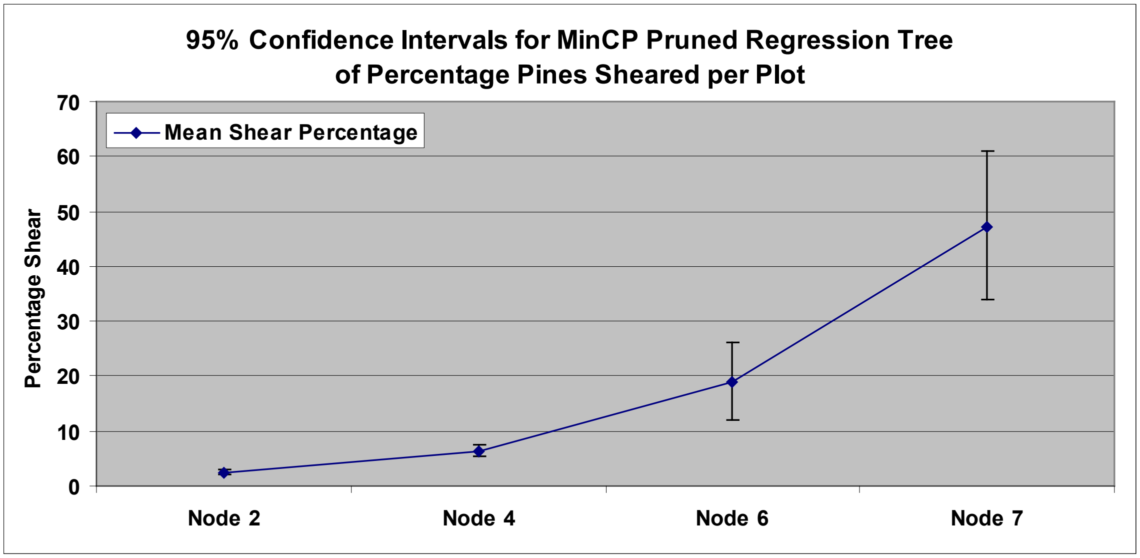

Figure 5.10 95% Confidence Intervals for MinCP Pruned Regression Tree of Percentage Pines Sheared per Plot.

Percentage of Hardwoods per Plot Wind-thrown

The fully-grown regression tree for percentage of wind-thrown hardwoods per plot reveals hardwoods with H/D ratios less than 28 experienced 8.3 times more wind-throw than hardwoods greater than 28 (Figure 5.11). Surprisingly, the plots having less than a 28 H/D ratio only tally 11 plots (3%) of the 342 total hardwood plots incorporated into the regression tree. These plots were identified and removed from the dataset and CART analysis was performed again (Figure 5.12). The new fully-grown regression tree excluded the H/D ratio variable entirely based of the removal of these plots. In addition, the classification accuracy of the first regression tree including the 11 plots is 65% with the second full regression tree having a classification accuracy of 66% (Tables 5.4 and 5.5).

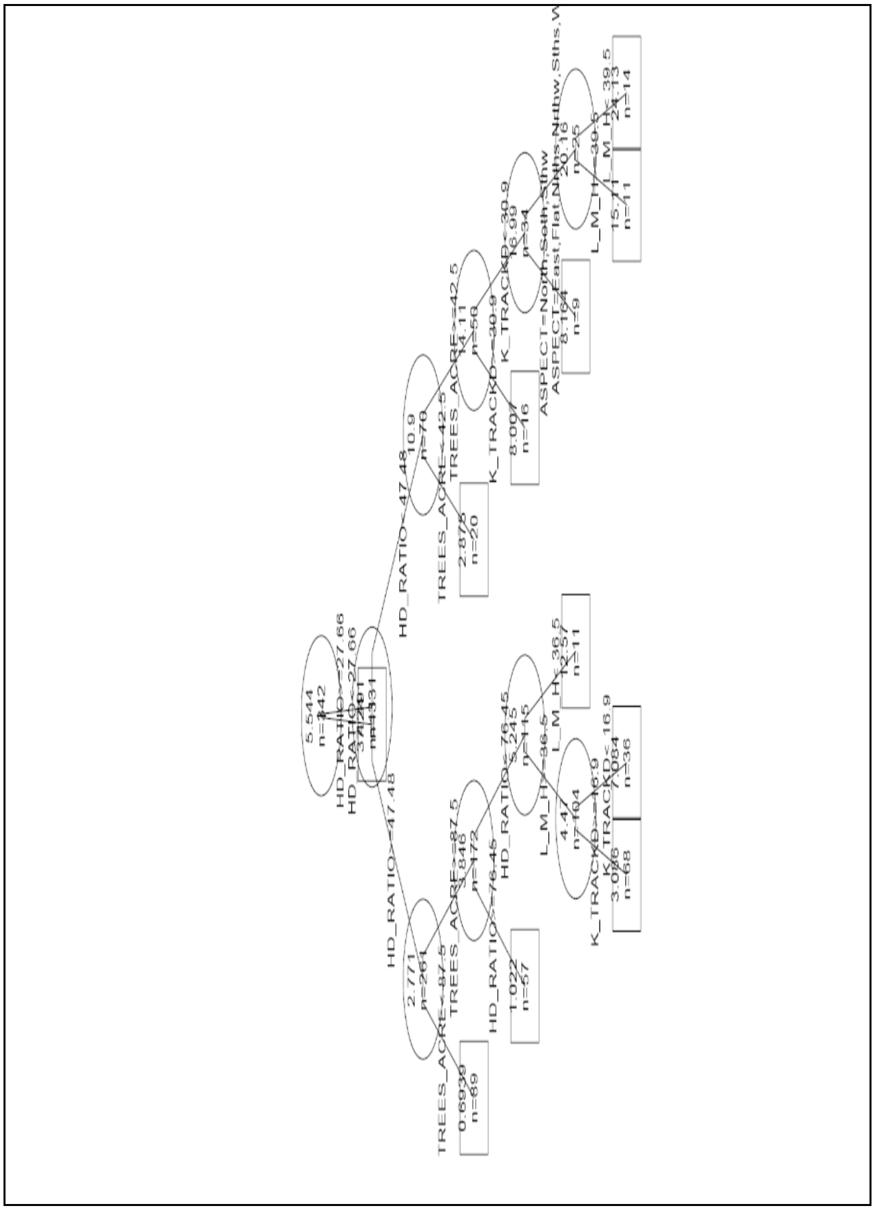

Figure 5.11 Original Fully-Grown Regression Tree for Hardwood Wind-throw Percentage.

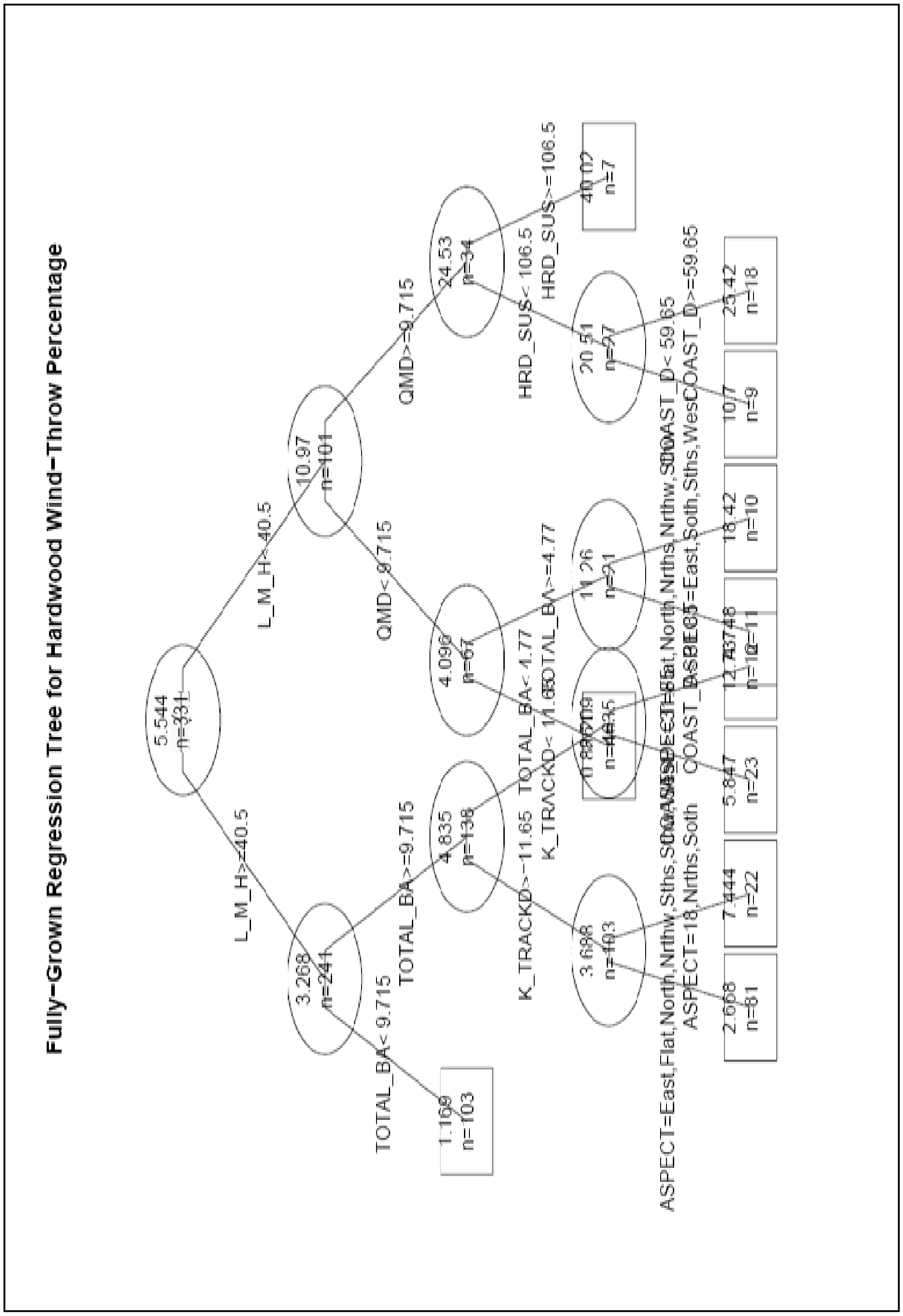

Figure 5.12 Reanalyzed Fully-Grown Regression Tree for Hardwood Wind-throw Percentage with removal of 11 HW Plots.

Table 5.4 Cost-Complexity Parameter (CP) Table for Original Full Regression Tree for Percentage of Hardwoods per Plot Wind-thrown.

CP	NSplit	Relative Error	Xerror	Xstd
0.347901	0	1.00000	1.00114	0.158603
0.111234	1	0.65210	0.72915	0.107995
0.054991	2	0.54086	0.60919	0.092845
0.027891	3	0.48587	0.59865	0.095687
0.019235	5	0.43009	0.66975	0.101877
0.015269	8	0.37172	0.66760	0.100558
0.011461	9	0.35645	0.68798	0.101359
0.010000	10	0.34499	0.68792	0.101295

Table 5.5 Cost-Complexity Parameter (CP) Table for Reanalyzed Full Regression Tree for Percentage of Hardwoods per Plot Wind-thrown.

CP	NSplit	Relative Error	Xerror	Xstd
0.207926	0	1.00000	1.00681	0.158972
0.064454	2	0.58415	0.8142	0.094053
0.047813	3	0.51969	0.90158	0.102799
0.039602	4	0.47188	0.86711	0.106497
0.029830	5	0.43228	0.85813	0.104514
0.024159	6	0.40245	0.86759	0.104267
0.016267	7	0.37829	0.83488	0.103420
0.012024	8	0.36202	0.83307	0.100518
0.011398	9	0.35000	0.84218	0.100266
0.010000	10	0.33860	0.85272	0.100416

The reanalyzed fully-grown regression tree for percentage of wind-thrown pines per plot reveals plots averaging less than 40.5 feet LMH experienced three times more wind-throw damage than plots greater than 40.5 feet. Similar to the fully-grown regression tree for pines, the gully-grown regression tree for hardwoods also splits into two main branches. Exploring the left branch shows hardwood plots having greater than 9.7 ft^2/plot basal area (i.e. larger diameters) had 4 times the amount of wind-throw damage than plots averaging less than 9.7 ft^2/plot. The following split reveals the distance from hurricane track had an influence the percentage of hardwood trees wind-thrown. Hardwood plots greater than 11.6 miles away from Katrina's track had 2.2 times less wind-throw damage than those within 11.6 miles of the hurricane's center path. The next split to the left has the categorical aspect depicting plots on northeast and southern exposures had about two times as much wind-throw then those on other aspects. Splitting right, the plots in the child node less than 11.6 miles from the track were further split by the distance the plots were from the coast. Hardwood plots located within 31 miles of the coast suffered 2 times the amount of wind-throw damage (12.7%) compared to plots greater than 31 miles (5.8%). Aspect and Distance from coast are located near the bottom of the regression tree and are most likely over-fit to the dataset. Focusing on the right branch, the splitting criteria used after LMH, is QMD. Hardwood plots having greater than 9.7 inch QMD experienced 6 times more wind throw damage (24.5%) compared to hardwoods plots averaging less than 9.7 inches QMD (4.1%). Following the left split, the Basal Area (BA) per plot was next used to split the previous child node. Hardwood plots having greater than 4.77 ft^2/plot BA suffered 11 times more wind-throw damage (11.26%) opposed to hardwood plots having less than 4.77 ft^2/plot (0.82%). Aspect was again used as a split criterion variable again depicting plots on east, southeast, and southern exposures had greater instance of wind-throw damage. Following the right split that maximizes the percentage of hardwoods per plot occurrence, the splitting criterion is sustained wind speeds. Hardwood plots subjected to sustained wind speeds greater than 106 MPH had twice the percentage of wind-throw occurrence (40%) than plots experiencing less than 106 MPH (20.5). Interestingly, the next split for plots experiencing less than 106 MPH was the distance to coast variable showing plots greater than 59 miles from the coast had 1.5 times more wind-throw damage (25.4%) then plots less than 59 miles (10.7%) This left split is a discrepancy of a right split that occurred in the left main branch where the closer a plot was to the coast given the previous splitting criterion, the more prone it was to suffering wind-throw damage. However, in both instances where the distance from coast variable is used, the location within the regression tree suggests this variable is over-fitting the dataset.

The 95% CIs of the full regression tree are similar to the full regression tree of pines in that as the number of splits and subsequent nodes increase, so does the 95% CI (Figure 5.13). Nodes 2, 6, 7, 9, and 10, all have small upper and lower bound variations from the mean value suggesting the mean values presented by CART analysis lead to higher confidence of these values. The terminal nodes present in the main right branch have higher 95% CIs with the last node, node 21, having the greatest 95% CI with the upper bound being 49.9 and the lower bound being 30.4.

Through the use of the cptable (Table 5.5) of the newly accepted regression tree over the original and the cpplot graph, the regression tree for wind-throw percentage on hardwood plots was pruned using the 1-SE rule (Figure 5.14 & Figure 5.15). Figure 5.14 depicts the cross-validation graph for the full regression tree and suggests five terminal nodes or four splits should be used before over-fitting begins to occur after the 5th split with a noticeable decreasing trend of the cross-validation error.

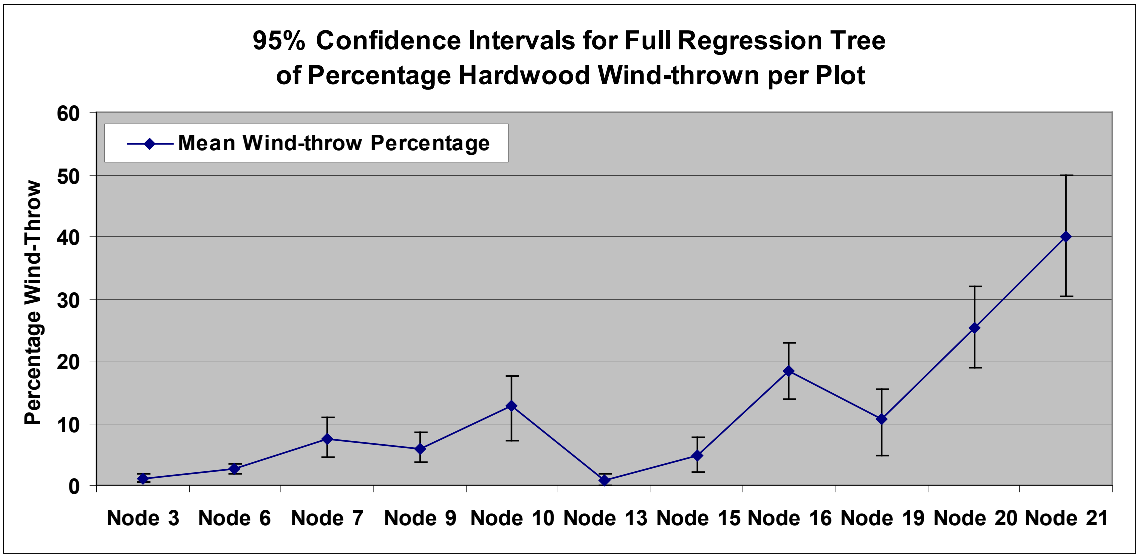

Figure 5.13 95% Confidence Intervals for Full Regression Tree of Percentage Hardwood Wind-thrown per Plot.

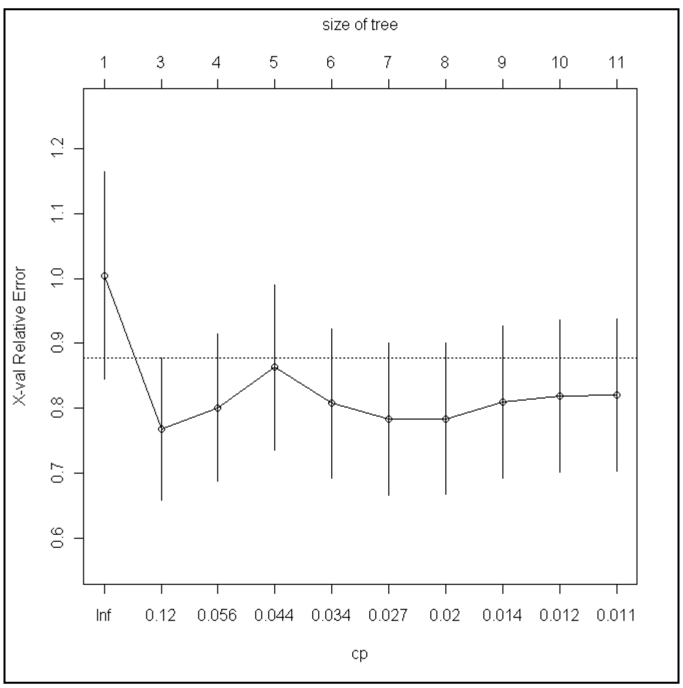

Figure 5.14 Cross-Validation Relative Error Graph with Standard Error Lines for Percentage of Hardwoods per Plot Wind-thrown.

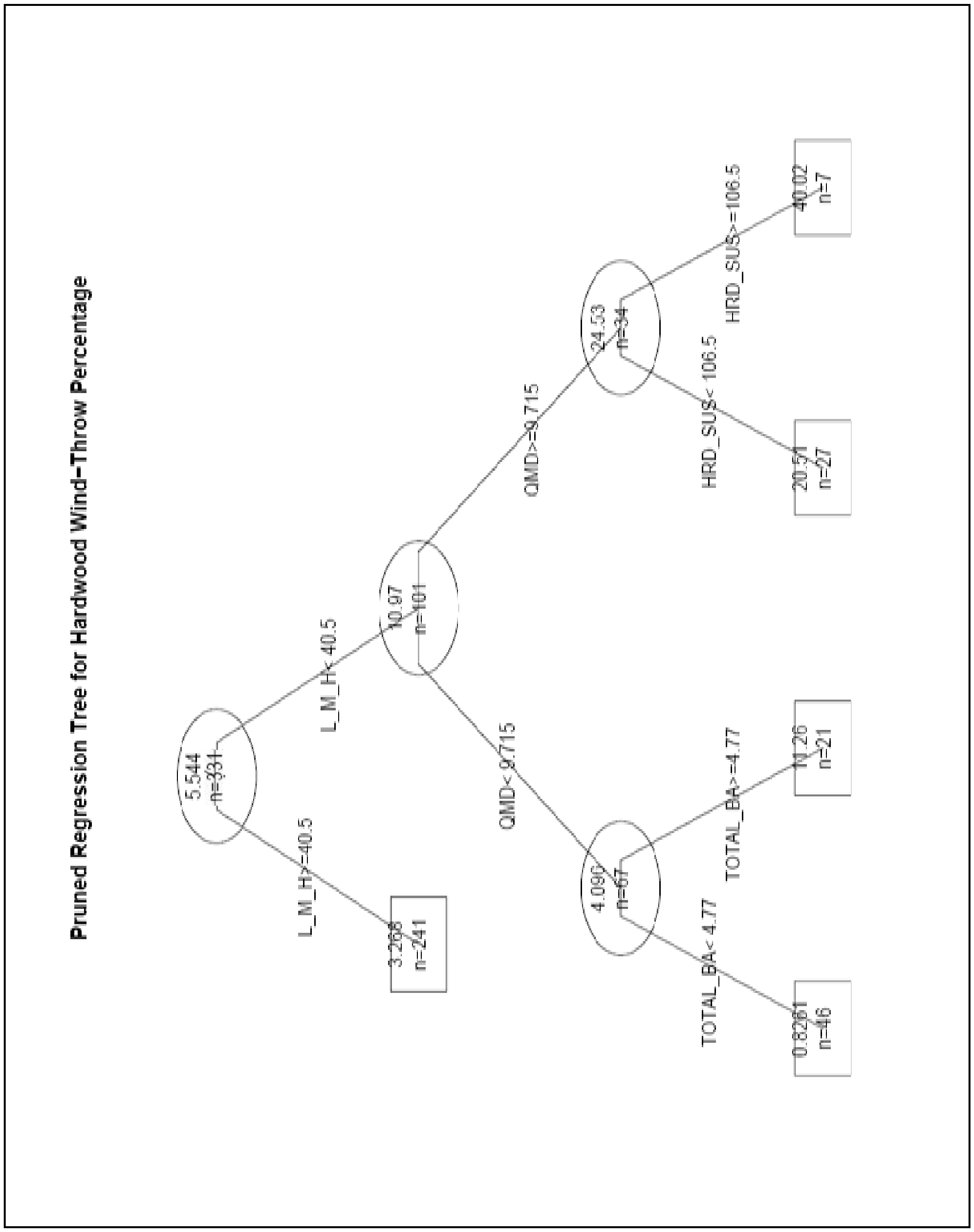

Figure 5.15 1-SE Pruned Regression Tree for Hardwood Wind-throw Percentage

The 1-SE pruned tree in Figure 5.15 shows using a CP value of 0.40 prunes the full regression tree to 5 terminal nodes or 4 splits total. The entire left main branch from the fully-grown regression tree is pruned away, and the splits remaining are the ones contributing to the overall classification accuracy and R^2 values. The 1-SE pruned regression tree now only depicts LMH, QMD, basal area per plot, and sustained wind speeds. The classification accuracy of the 1-SE pruned regression tree is 53% (1.0 – 0.47188) with a cross-validation R^2 value of 0.13 (1.0-0.86295) with all four splits (Table 5.6 and Figure 5.16). The 95% CIs of the 1-SE pruned regression tree again show splits to the right have larger CIs with far most right node, node 9's upper bound is 49.6 and lower bound is 29.4 with a mean value of 40 (Figure 5.17). To increase the cross-validation R^2 value, the tree is further pruned to the minimum cross-validation error, which only contains three terminal nodes (Figure 5.18).

Table 5.6 Cost-Complexity Parameter (CP) Table for Reanalyzed 1-SE Pruned Regression Tree for Percentage of Hardwoods per Plot Wind-thrown.

CP	NSplit	Relative Error	Xerror	Xstd
0.207926	0	1.00000	1.00439	0.158972
0.064454	2	0.58415	0.76786	0.094053
0.047813	3	0.51969	0.80080	0.102799
0.040000	4	0.47188	0.86295	0.106497

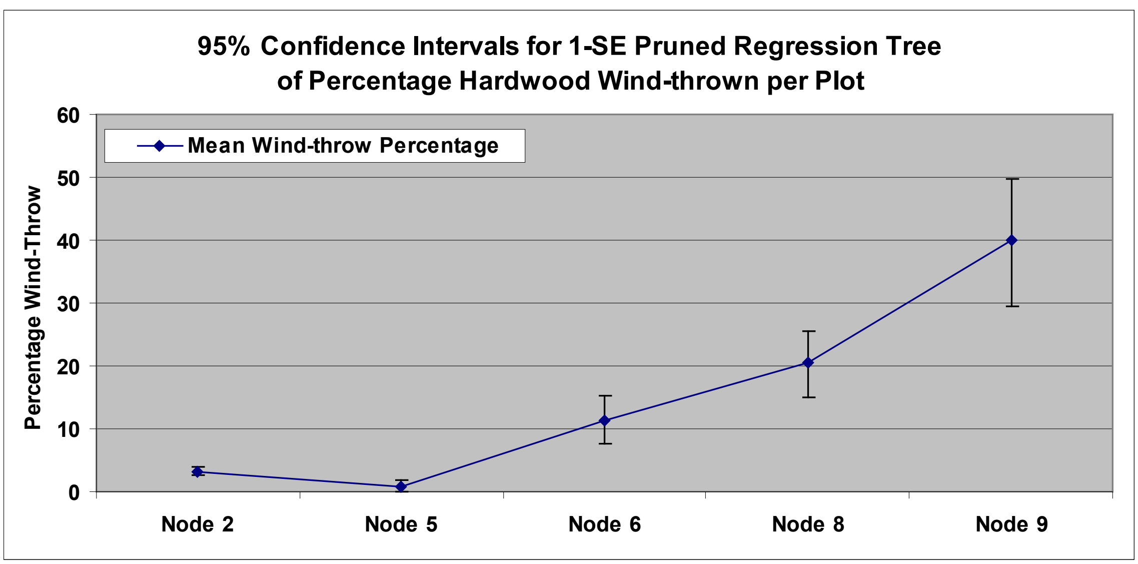

Figure 5.16 95% Confidence Intervals for 1-SE Pruned Regression Tree of Percentage Hardwoods Sheared per Plot.

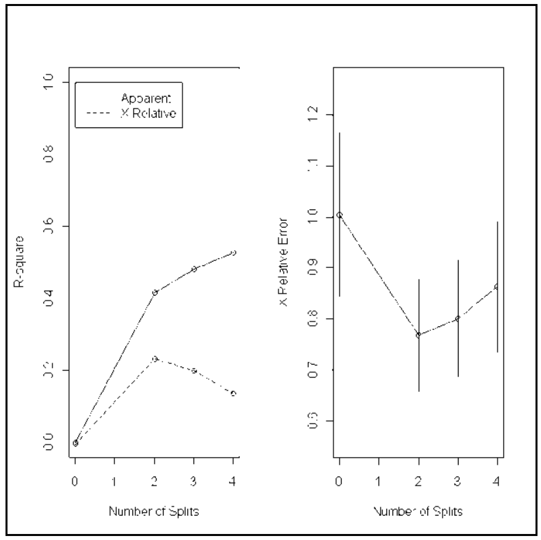

Figure 5.17 Apparent, X Relative R-Square and Cross Validation Relative Error Graphs of 1-SE Pruned Regression Tree for Wind-thrown Hardwood Percentage.

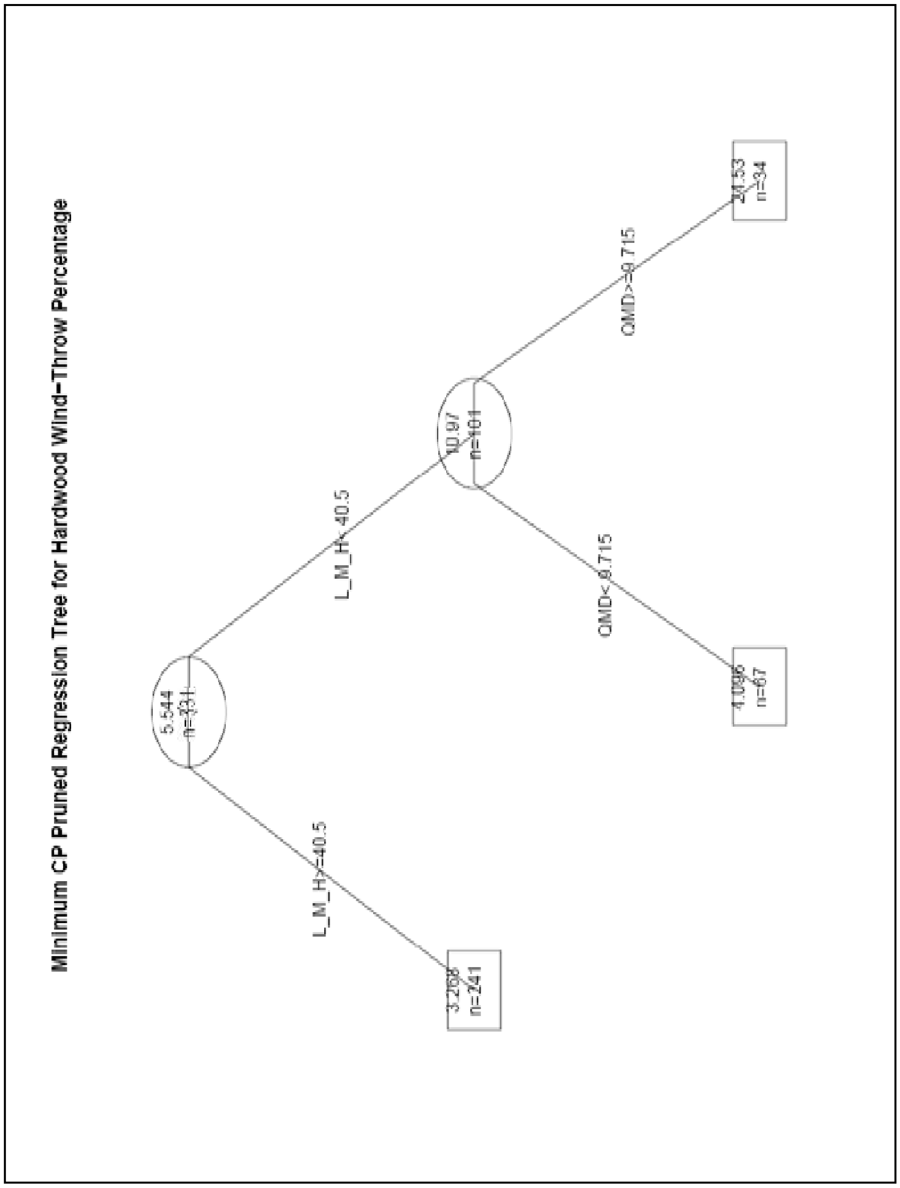

Figure 5.18 Minimum CP Pruned Regression Tree for Hardwood Wind-throw Percentage.

The minimum CP pruned regression tree for percentage of hardwoods per plot experiencing wind-throw only contains two variables: LMH and QMD. Based on the cptable, the classification accuracy dropped to 42% from 53% of the 1-SE pruned tree (Table 5.7). However, the cross-validation R^2 increased to 0.23 from 0.14 by just using the first two splits of the entire regression tree opposed to the first four splits (Figure 5.19).

Table 5.7 Cost-Complexity Parameter (CP) Table for Reanalyzed MinCP Pruned Regression Tree for Percentage of Hardwoods per Plot Wind-thrown.

CP	NSplit	Relative Error	Xerror	Xstd
0.207926	0	1.00000	1.00439	0.158972
0.064454	2	0.58415	0.76786	0.094053

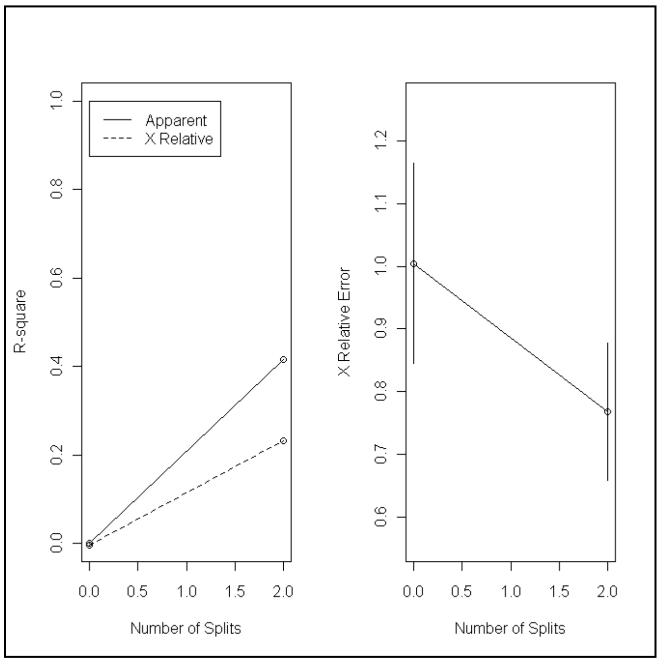

Figure 5.19 Apparent, X Relative R-Square and Cross Validation Relative Error Graphs of MinCP Pruned Regression Tree for Wind-thrown Hardwood Percentage.

The 95% CIs for the minimum CP pruned tree show the three terminal nodes that have not been pruned away (Figure 5.20). There is vary little variance within Node 2 and Node 4, with Node 5 having a larger interval. Node 5's upper bound is 30.1 while its lower bound is 19.1 with a mean value of 24.5.

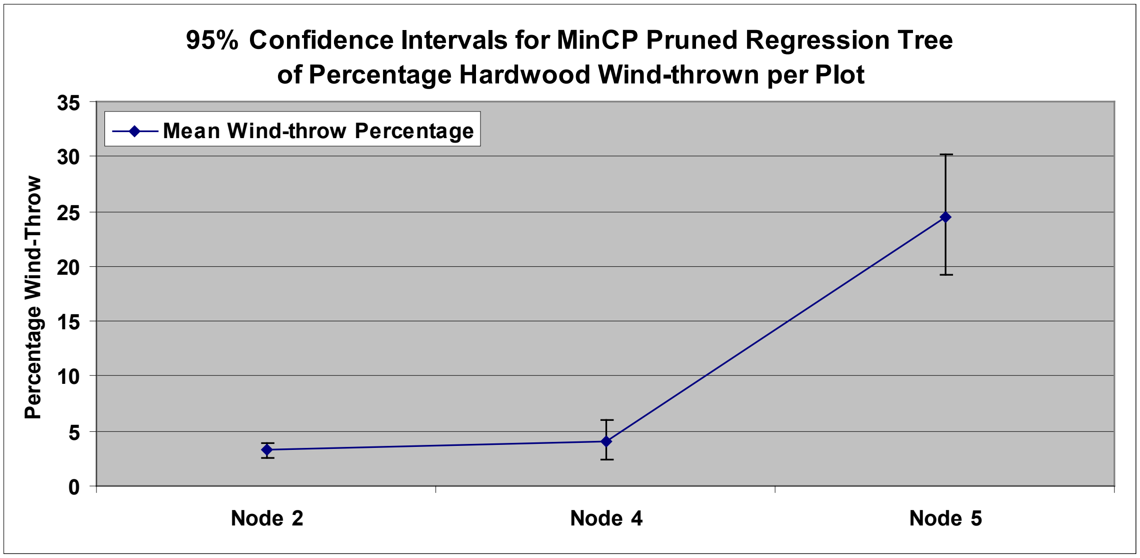

Figure 5.20 95% Confidence Intervals for MinCP Pruned Regression Tree of Percentage Hardwoods Wind-thrown per Plot.

Summarizing both the percentage of pines per plot sheared and hardwoods per plot wind-thrown reveals similarities and differences among the biotic and abiotic variables. For both tree classes, Lorey's Mean Height (LMH) and Quadratic Mean Diameter (QMD) contribute greater to both the classification accuracy of the CART analysis and minimizing the cross-validation error of the dataset. The abiotic variable depicted through CART analysis that differs between pines and hardwoods is the wind regime pattern. Greater shear percentages to pines occur with high peak wind gusts while greater wind-throw percentages occur with higher sustained wind speeds. Sections 5.1 and 5.2 focused on the percentage of both pines and hardwood trees per plot experiencing shear and wind-throw damage respectively. The following sections, 5.3 and 5.4, focus on the percent basal area of pine and hardwood trees per plot that experienced shear and wind-throw respectively. The amount of basal area damaged on established plantations determines the course of action for salvaging efforts of the affect timber and the future manageability of the stand (DeLoach and Dicke 2005, Long *et al.* 2005).

Percentage of Pine Basal Area per Plot Sheared

The fully-grown regression tree of sheared basal area percentage per plot across all pine plots shows the first splitting criteria used is peak wind gusts (Figure 5.21). Pine plots experiencing over 124 mph peak wind gusts had three times the basal area damaged (12.5%) compared to plots experiencing less than 124 mph wind gusts (4%). Peak wind gusts also showed up in the percentage of pine trees per plot damaged minCP regression tree in the earlier section. Following the peak wind gusts split, pine plots averaging greater than 9.8 inches QMD had twice the amount of basal area percentage sheared (21.6%) opposed to pine plots averaging less than 9.8 inches QMD (9.3%). The full regression tree then splits into two main branches. QMD is again used in the next split to the left, reinforcing the fact the larger the diameter, the more prone pine trees and pine plots are to being sheared. The next split uses the number of trees per acre with plots having less trees, thus being more open, had three times the amount of shear (32.8%) compared to greater density pine stands (10.4%). Further splits in the left main branch include the distance from coast and elevation

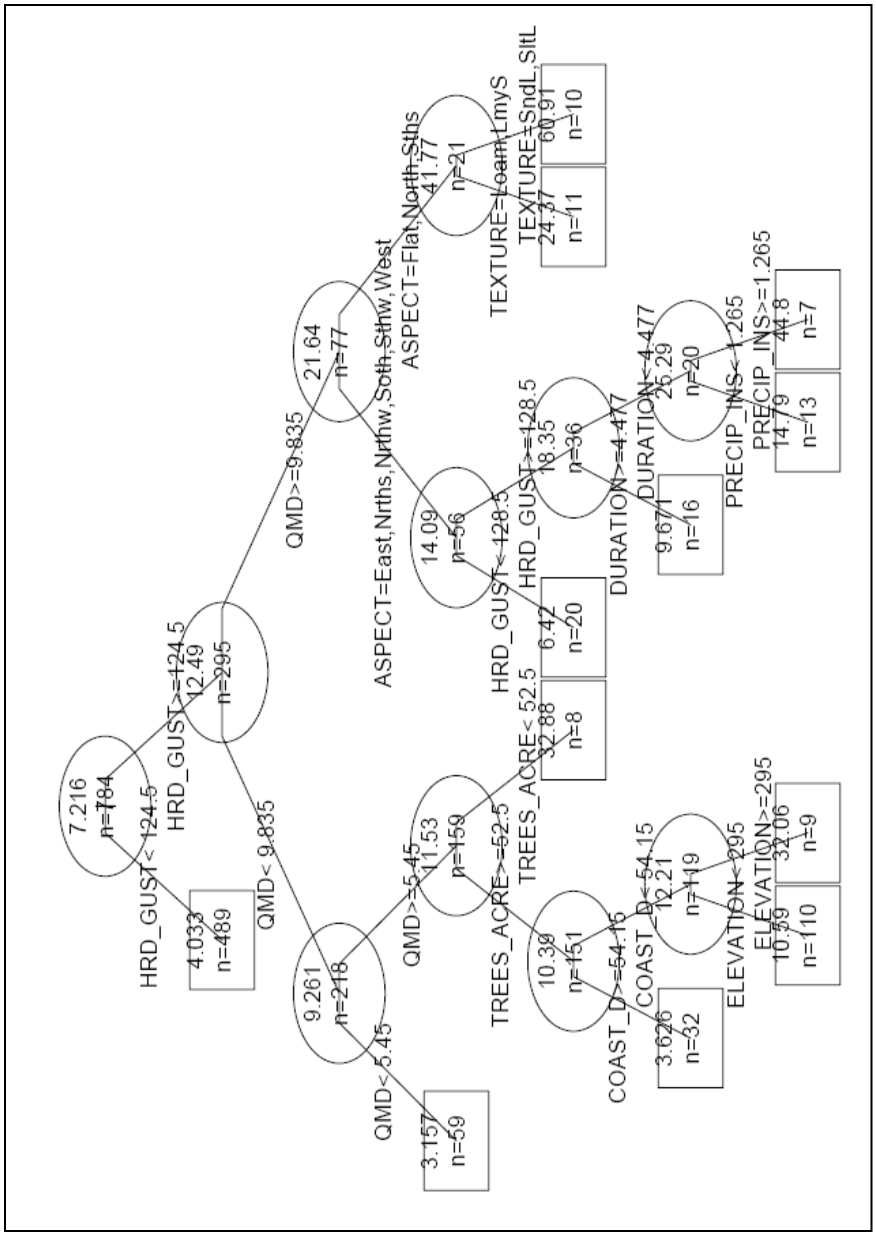

Figure 5.21 Fully-Grown Regression Tree for Sheared Basal Area Percentage per plot across Pine Plots

Pine plots less than 54 miles from the coast averaged almost three times the amount of basal area sheared than plots greater than this distance. Additionally, pine plots higher than 295 feet averaged 2.9 times more basal area sheared than pine trees below 295 feet. Now following the right main branch, aspects that were flat, faced the north and southeast also had 3 times the amount of basal area sheared compared to other aspect directions. Splitting left, peak wind gust is used again and concurs with the first split with higher peak gusts being responsible for higher mean percentages of sheared basal area. The last two splits of the secondary left branch depict two abiotic variables including duration of hurricane force winds and precipitation intensity. Interestingly, durations less than 4.5 hours of hurricane force winds actually caused more sheared basal area damage than in areas with longer durations. Pine plots that experienced over 1.2 inches of rain in one hours time throughout the course of the event had three times more sheared basal area damage (44.8%) then areas not affecting with high precipitation rates (14.8). The last split to the right identifies soil texture as having a large influence on the amount of sheared basal area. Pine plots exposed to greater than 124 peak wind gusts and having QMD grater than 9.8 inches on flat, north, and southern aspects located on sandy-loam and silty-loam soils experienced 61% basal area shear damage per plot. Whereas plots located on loamy and loamy-sand had 24% basal area damage per plot.

The overall classification accuracy of the full regression tree however, in only 39% (Table 5.7). In addition, the maximum cross-validation R^2 is only 0.07 and is reached after three splits, further suggesting that the full regression tree is well over-fit and needs to be pruned (Table 5.8).

Table 5.8 Cost-Complexity Parameter (CP) Table for Full Regression Tree for Percentage of Basal Area of Sheared Pine Trees per Plot.

CP	**NSplit**	**Relative Error**	**Xerror**	**Xstd**
0.084137	0	1.00000	1.00165	0.12997
0.065257	1	0.91586	0.95845	0.11769
0.044708	3	0.78535	0.93602	0.11434
0.021908	4	0.74064	0.96059	0.11426
0.018203	6	0.69683	1.00418	0.11677
0.017251	8	0.66042	1.01758	0.11705
0.010000	11	0.60866	1.02767	0.11694

The 95% confidence intervals suggest the more splitting criteria a node is based off of, the greater the interval (Figure 5.22). Nodes 11, 12, and 15 all had large 95% CIs compared to other nodes and were subsequently pruned away in smaller regression trees. Nodes with 95% CIs less than 24 were typically not pruned away. In order to prune the tree back to the lowest cross-validation error and thus the highest R^2 value, the 1-SE pruned tree is forgone and the MinCP pruned regression tree is plotted (Figure 5.23). The MinCP pruned regression tree was pruned using a cost-complexity parameter value of 0.40 (Table 5.9).

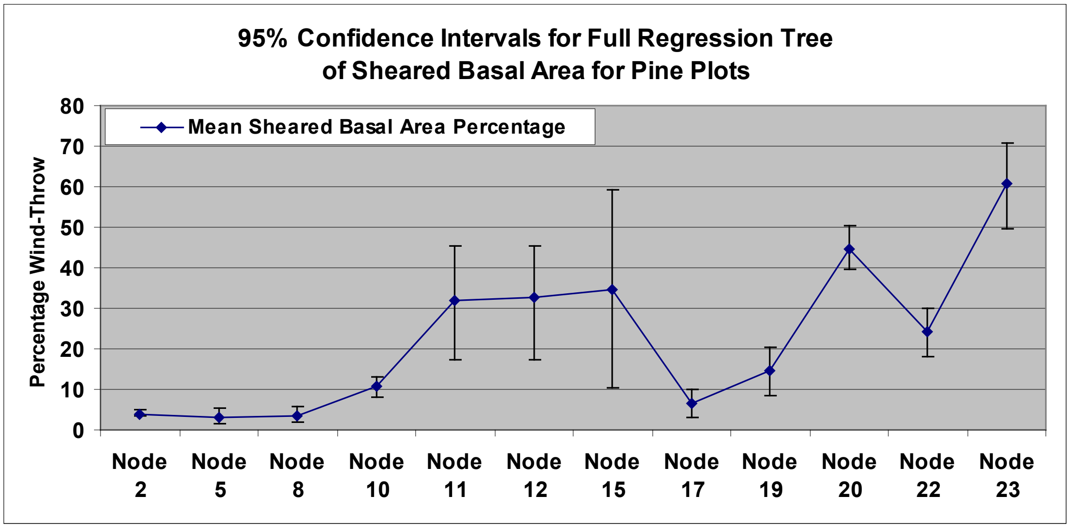

Figure 5.22 95% Confidence Intervals for Full Regression Tree of Sheared Basal Area for Pine Plots.

The parsimonious MinCP regression tree pruned away much of the full regression tree, leaving only 4 terminal nodes (Figure 5.23). The classification accuracy of the MinCP tree dropped from 39% of the full regression tree to 21% and the cross-validation R^2 value is only equal to 0.07 (Table 5.9 and Figure 5.24). Figure 5.25 depicts the 95% CIs for the four remaining nodes of the MinCP pruned regression tree.

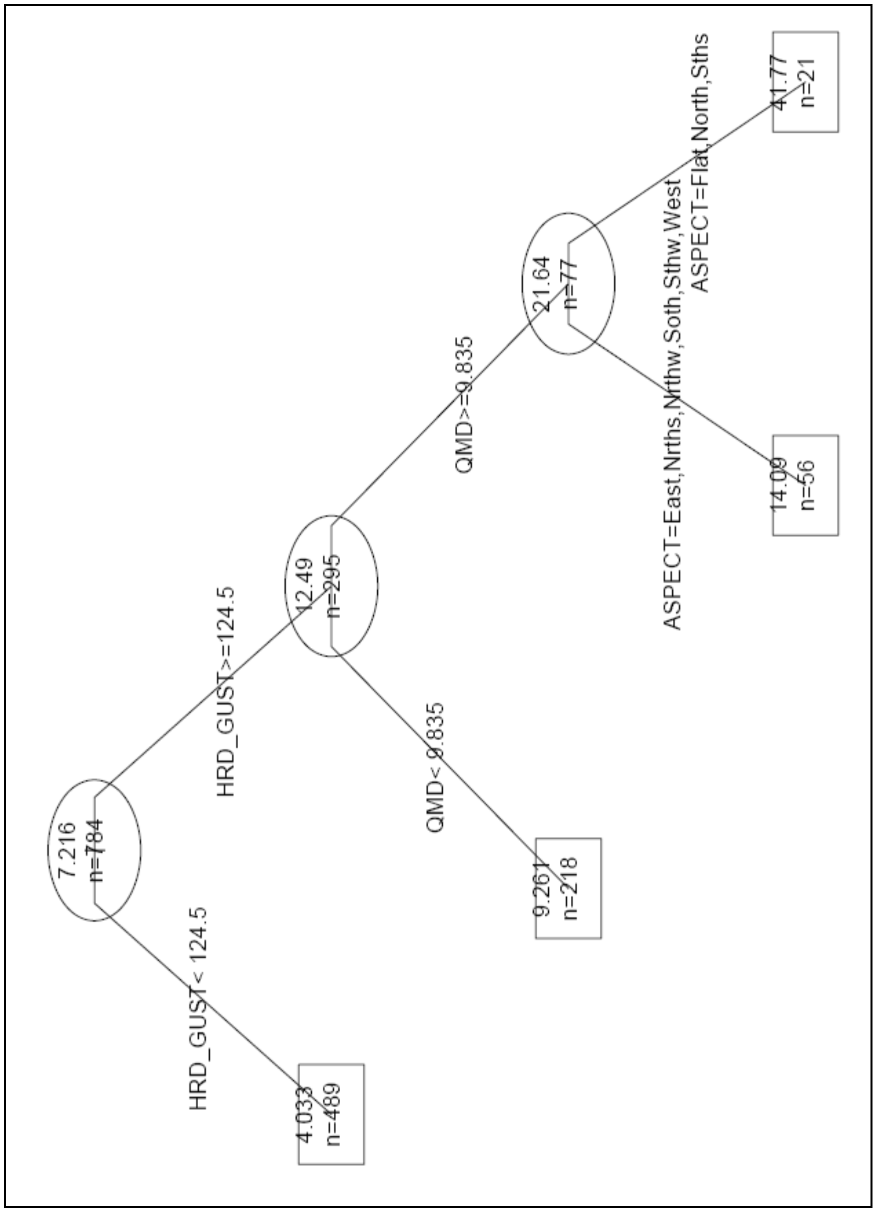

Figure 5.23 MinCP Regression Tree for Sheared Basal Area Percentage per plot across Pine Plots.

Table 5.9 Cost-Complexity Parameter (CP) Table for MinCP Tree for Percentage of Basal Area of Sheared Pine Trees per Plot.

CP	NSplit	Relative Error	Xerror	Xstd
0.084137	0	1.00000	1.00165	0.12997
0.065257	1	0.91586	0.95845	0.11769
0.044708	3	0.78535	0.93602	0.11434

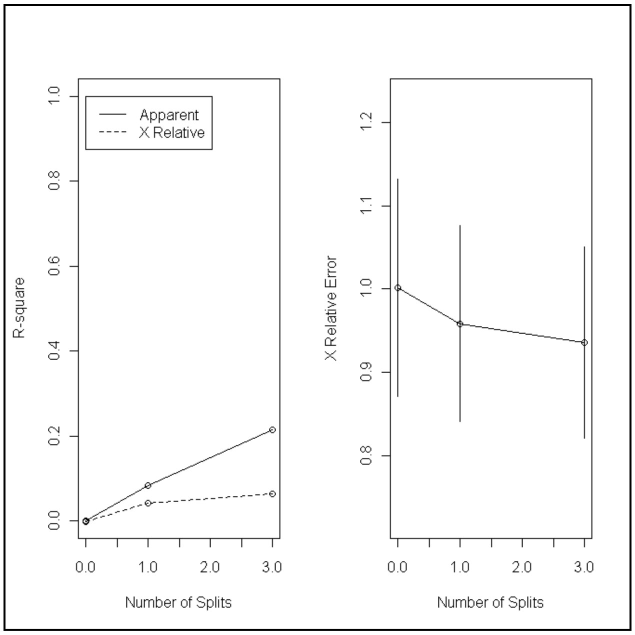

Figure 5.24 Apparent, X Relative R-Square and Cross Validation Relative Error Graphs of MinCP Pruned Regression Tree for Percentage Pines Sheared per Plot.

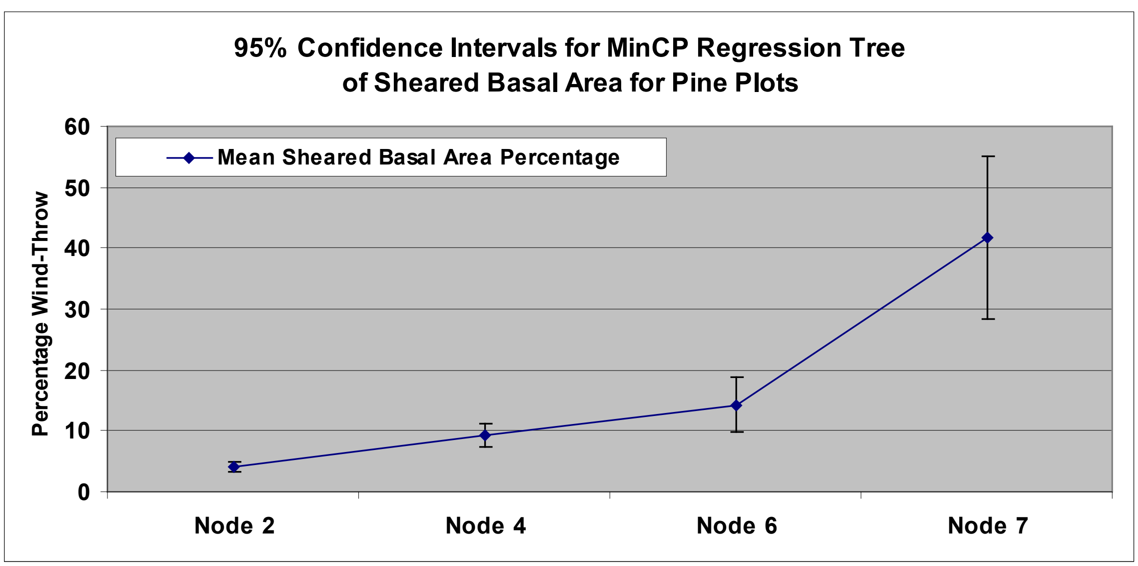

Figure 5.25 95% Confidence Intervals for MinCP Pruned Regression Tree of Sheared Basal Area for Pine Plots.

Percentage of Hardwood Basal Area per Plot Wind-thrown

Similar to the first couple splits in the MinCP regression tree for percentage of hardwoods wind-thrown, are the variables of LMH and QMD reappearing in the fully-grown regression tree of wind-thrown basal area percentage (Figure 5.26). Hardwood plots averaging less than 40.5 feet LMH had a mean basal area wind-thrown of 19.3% compared to hardwood plots greater than 40.5 feet LMH with 5.1% wind-thrown basal area. Splitting left, LMH is again used with hardwood plots less than 48.5 feet LMH are more prone to greater wind-thrown basal areas opposed to hardwood plots averaging higher LMHs. The next split shows QMDs greater than 8.7 inches had 3 times more wind-thrown basal area (13.6%) opposed to hardwoods averaging less than 8.7 inches QMD (3.9%). Sustained wind speeds are used in the next split with hardwood plots expose to winds greater than 74 mph averaged 17.7% wind-thrown basal area opposed to 4% for hardwood stands not experiencing hurricane force winds. Following the main right branch, QMD again appears as an important splitter with larger average diameter hardwood plots has more wind-thrown basal area damage. CART analysis then depict the LMH variable again, showing shorter plots averaged more wind-thrown basal area damage than plots having taller trees. Splitting left, the variables of aspect and stand condition are shown. Northwest, southeast, and western aspects averaged greater wind-thrown basal areas then other aspects. The stand condition split shows hardwood plots classified as sawtimber or sub-merchantable had greater wind-thrown basal area damage (21.5%) compared to pulpwood and pallet classified stands (8%). The splits maximizing the percentage of wind-thrown basal area were aspect and trees per acre. Hardwoods plots averaging less than 33.5 feet LMH and were on southeast or southwest facing slopes averaged 61.5% of wind-thrown basal area per plot. Hardwood plots having greater than 72.5 trees per acre averaged 38.5% wind-thrown basal area wind-thrown opposed to 18.8% with plots averaging less than 72 trees per acre.

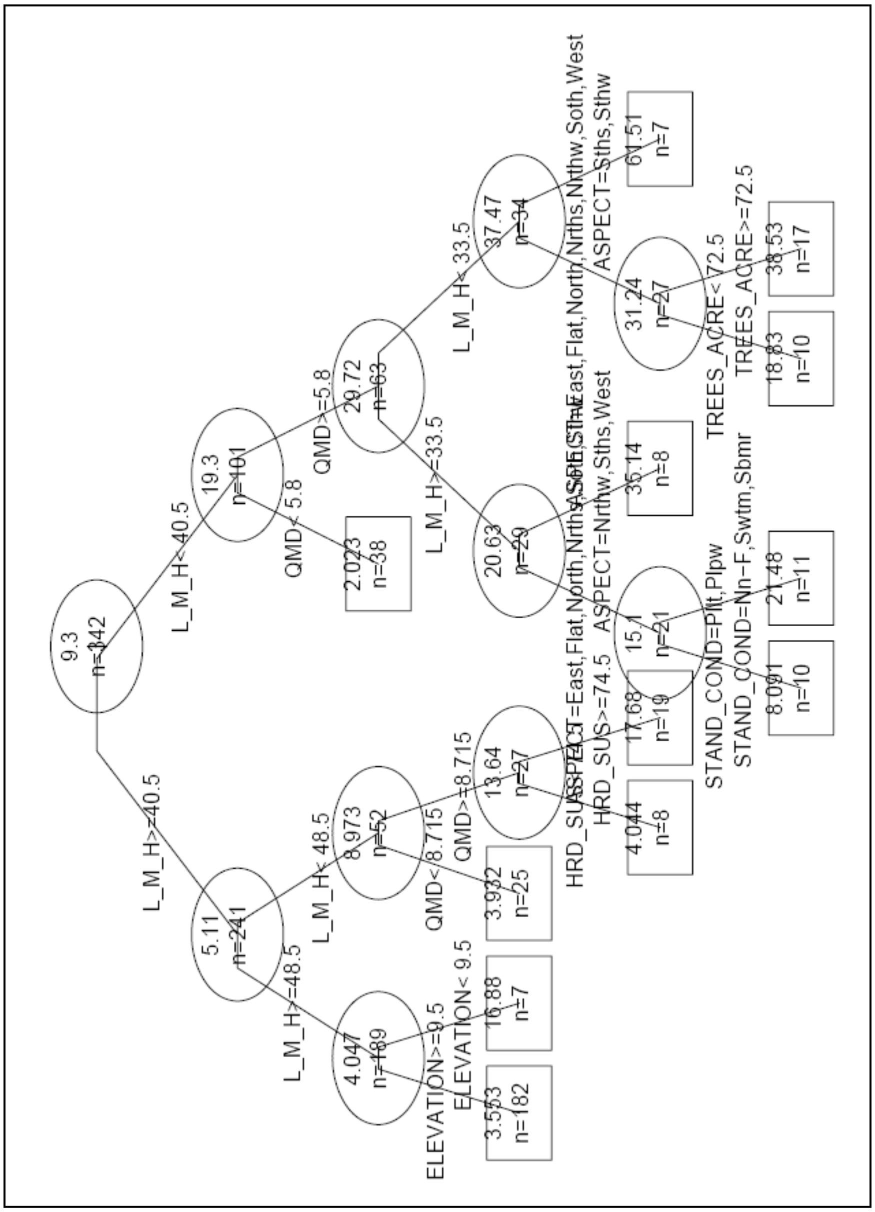

Figure 5.26 Fully-Grown Regression Tree for Wind-thrown Basal Area Percentage per plot across Hardwood Plots.

The classification accuracy of the full regression tree for percentage of basal area wind-thrown per plot for hardwoods is 62% (Table 5.10). However, the full regression tree only has a cross-validation R^2 of 0.07. The highest cross-validation R^2 of 0.24 occurs after two splits and thus the full regression tree is pruned to the minimum cross-validation error and corresponding CP value (Figure 5.27).

Table 5.10 Cost-Complexity Parameter (CP) Table for Full Regression Tree for Percentage of Basal Area of Wind-Thrown Hardwood Trees per Plot.

CP	NSplit	Relative Error	Xerror	Xstd
0.192826	0	1.00000	1.00702	0.12951
0.056549	2	0.61435	0.76633	0.10289
0.028990	4	0.50125	0.82008	0.11814
0.027588	5	0.47226	0.87102	0.12523
0.013487	6	0.44467	0.86702	0.12068
0.012423	9	0.40421	0.91575	0.12262
0.011138	10	0.39179	0.92620	0.12314
0.010000	11	0.38065	0.93445	0.12343

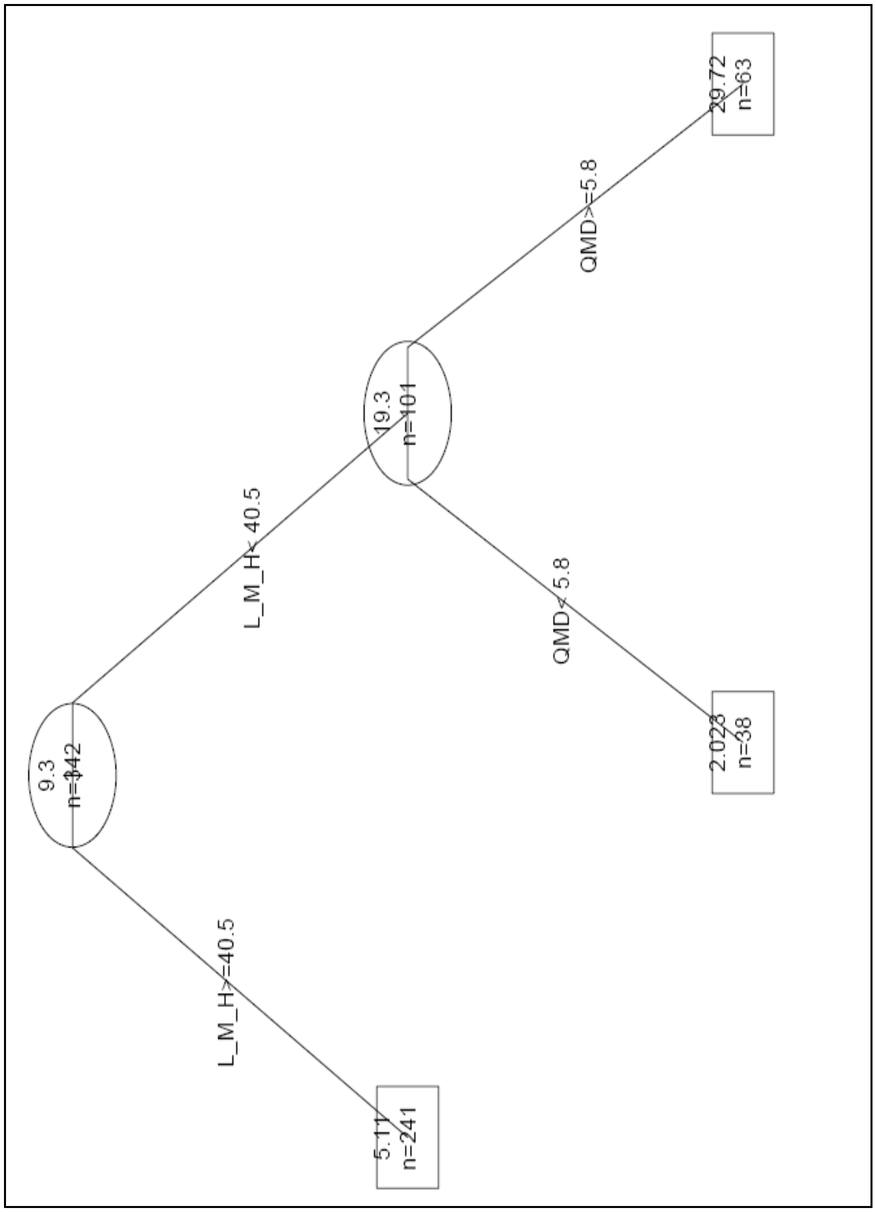

Figure 5.27 MinCP Pruned Regression Tree for Wind-thrown Basal Area Percentage per plot across Hardwood Plots.

The MinCP regression tree only depicts two variables used: LMH and QMD. The classification accuracy drops from 62% to 40% (Table 5.11). The R-square graphs of both the apparent and X-Relative error concur with Table 5.11 (Figure 5.28).

Table 5.11 Cost-Complexity Parameter (CP) Table for MinCP Regression Tree for Percentage of Basal Area of Wind-Thrown Hardwood Trees per Plot.

CP	NSplit	Relative Error	Xerror	Xstd
0.192826	0	1.00000	1.00702	0.12951
0.056549	2	0.61435	0.76633	0.10289

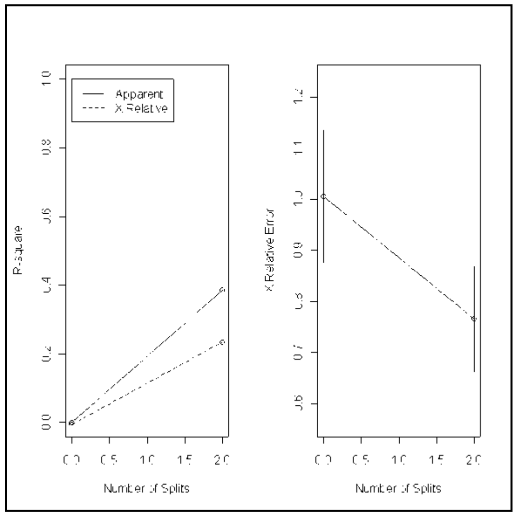

Figure 5.28 Apparent, X Relative R-Square and Cross Validation Relative Error Graphs of MinCP Pruned Regression Tree for Percentage Pines Sheared per Plot.

The 95% CIs for the minCP pruned regression tree depicting the three nodes show relative little variance in the ensemble bootstrapping statistics (Figure 5.29). With such small variances, higher confidence is placed on these mean values and their contribution to the percentage of basal area wind-thrown for hardwood plots.

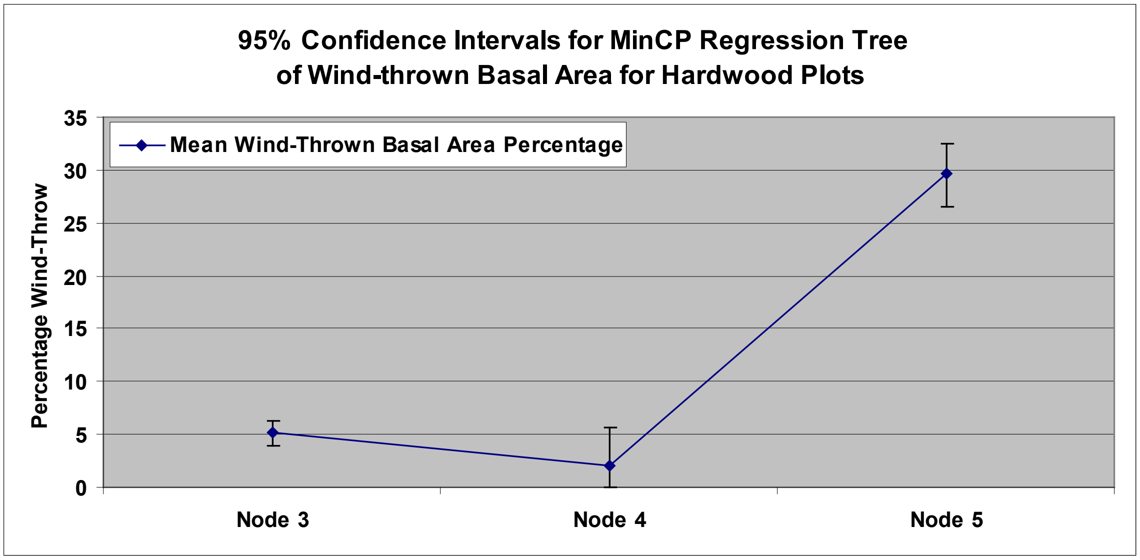

Figure 5.29 95% Confidence Intervals for MinCP Pruned Regression Tree of Wind-thrown Basal Area for Hardwood Plots.

Biotic and abiotic variables contributing the shearing and wind-throw basal area of pine and hardwood plots tend to differ more than the biotic and abiotic variables controlling the percentage of actual trees sheared or wind-thrown. The only biotic factor common to both pine and hardwood plots when percentage of basal area sheared or wind-thrown is tested, is QMD. Peak wind gust speeds and aspect were more contributing factors in pine plots while hardwood plot basal area damaged was governed by the stands' relative height. Classification accuracies through CART analysis were higher for the minCP regression trees of hardwood plots (42% and 35%) for both dependent variables tested versus pine plots (29% and 21%). Greater cross-validation R^2 values were obtained for hardwood plots testing wind-throw (0.24 and 0.24) than shear damage for pines (0.13 and 0.07). Forward Stepwise Logistic Regression (FSLR) was performed in SPSS 15.0 to corroborate the findings of CART analysis.

Forward Stepwise Logistic Regression Results

Forward Stepwise Binary Logistic Regression was performed using the Wald Statistic as the selecting criterion for a variable to be entered or removed from the stepwise regression equation (Peterson 2007). Only one dependent response variable was tested in FSLR opposed to the two tested in CART analysis. The dependent response variable examined here is the binary plot damage classifications tested against the biotic and abiotic independent variables. Binary plot classification based on the percentage of sheared or wind-thrown basal area was not computed.

FSLR Analysis of Sheared Pine Plots

Forward stepwise logistic regression based on the binary classification of sheared and non-sheared plots of pine plots depicts five variables that are significant at $p<0.01$ and one variable was significant at $p<0.05$ (Table 5.12). The coefficient (B) is the value for the logistic regression equation for predicting the dependent variable from the independent variable and expressed in log-odds units. The Standard Error (SE) is associated with the coefficient of the independent variable (B). The Exp(B) is the odds ratio of the independent variable by exponentiating the coefficients (B) using log to base *e*. There are no coefficients, SE, or Exp(B) for the variable physiographic region because it was not entered into the logistic regression equation ($p>0.05$). Only the bottomland physiographic region was entered into the equation for $p<0.05$. However, similar to the logistic regression used in Wang and Xu (2008), the odds ratios of the physiographic region suggest pine plots located in bottomland areas where much more resilient to shear. Where pine plots located on terraces were not as resilient with an odds ratio of 0.829. The positive coefficient of 0.007 of elevation means for every foot higher a pine plot is located, it is 1.007 times more likely to be classified as sheared. The negative coefficient of the distance to coast variable indicate for every mile a pine plot is located inland from the coast it is 0.966 times less likely to be classified as shear. The H/D ratio is also negative, meaning as the average H/D ratio per plot decreases, shear of pines increases. Lorey's Mean Height (LMH) coefficient was negative with an odds ratio of 0.961 indicating that for every foot increase in the average LMH of a 1/5th acre pine plot, it is 0.961 times less likely to be classified as shear than non-shear. This concurs with the results of CART analysis which depicted higher percentage of shear occurrence had lower LMH values. Seventy-six plots had 10% or greater sheared pine trees per plot with LMHs greater than 50 feet whereas 368 pine plots having greater than 50 feet LMH but had less than 10% sheared pines per plot, further reinforcing the results of CART and FSBLR.

Table 5.12 Significant variables and respective statistics for prediction of pine plot damage classification. ** denotes p<0.01 significance and denotes p<0.05 significance.

Variable	Coefficient (B)	SE	Wald	Exp(B)	Sig
Elevation	0.007	0.002	14.509	1.007	0.000**
Trees per Acre	0.008	0.002	30.635	1.008	0.000**
Lorey's Mean Height	-0.040	0.008	25.505	0.961	0.000**
Distance from Coast	-0.035	0.008	20.852	0.966	0.000**
H/D Ratio	-0.019	0.007	7.778	0.981	0.005**
Physiographic Position	-	-	4.963	-	0.084
Bottomlands	-1.421	0.644	4.863	0.241	0.027*
Terraces	-0.187	0.265	0.501	0.829	0.479

The model summary shows 12 steps were taken to determine the significant variables (Table 5.13). However, the lowest -2 log likelihood and highest pseudo R^2 values are shown in the tenth step of the model building process. Peak wind gust speed was entered as a significant variable early on in the modeling process but was removed after the 10th step yet contributed to higher pseudo R^2 values for classification of shear occurrence on pine plots.

Table5.13 Model Summary of Forward Stepwise Logistic Regression of Shear Occurrence on Pine Plots.

Step	**-2 Log likelihood**	**Cox & Snell R Square**	**Nagelkerke R Square**
1	642.555	.023	.040
2	624.509	.045	.079
3	613.013	.059	.104
4	605.979	.067	.118
5	598.063	.077	.135
6	591.960	.084	.147
7	592.016	.084	.147
8	580.695	.097	.170
9	576.241	.102	.179
10	569.382	.110	.193
11	572.050	.107	.188
12	574.369	.104	.183

Percentage of Hardwoods Wind-thrown

Forward stepwise logistic regression based on the binary classification of Wind-thrown and non-wind-thrown plots of hardwood plots depicts five variables that are significant at $p<0.01$ and two are significant at $p<0.05$ (Table 5.14).

Table 5.14 Significant variables and respective statistics for prediction of hardwood plot damage classification. ** denotes p<0.01 significance and * denotes p<0.05 significance.

Variable	Coefficient (B)	SE	Wald	Exp(B)	Sig
Elevation	0.006	0.002	10.481	1.006	0.000**
Trees Per Acre	0.009	0.002	20.250	1.009	0.000**
H/D Ratio	-0.043	0.009	24.971	0.958	0.000**
Cumulative Wind	0.044	0.011	17.423	1.044	0.000**
Sustained Wind	0.029	0.012	6.098	1.029	0.014*
Total Precipitation	-0.394	0.155	6.455	0.674	0.011*
Physiographic Position	-	-	13.561	-	0.001**
Bottomlands	1.529	0.415	13.560	4.613	0.000*
Terraces	1.204	0.523	5.292	3.332	0.21*

The exponentiated coefficient (Exp(B)), shows that for every one foot rise in elevation a hardwood plot is located, it is 1.006 times more likely to be classified as wind-thrown. Related to elevation, the topographical position classification identified hardwood plots classified as bottomland were 4.613 times more likely to be classified wind-thrown opposed to hardwood plots classified as upland. More so, hardwood plots in terrace locations were 3.332 times more likely to be classified as being wind-thrown than hardwood upland plots. The greater the amount of trees per acre, the higher likelihood of a hardwood plot being classified as wind-thrown. Similar to pine plots, the H/D ratio for hardwoods is also negative indicating for each one unit increase, a hardwood plot is 0.958 times less likely to be categorized as wind-thrown. The next variables include cumulative and sustained wind speeds. With an increase in one mile per hour, a hardwood plot is 1.044 and 1.029 times more likely to be classified as wind-thrown opposed to non-wind-thrown. Total precipitation, however is inversely related to the probability of a hardwood plot being classified as wind-thrown. Further investigation of the distribution of wind-thrown classified plots overlaid on the event total precipitation grid does corroborate with the SPSS statistical output. There more hardwood plots classified as non-wind-thrown (0) on precipitation totals greater than 6 inches (82) than plots classified as wind-thrown (1) in areas receiving over 6 inches (69) (Figure 5.30).

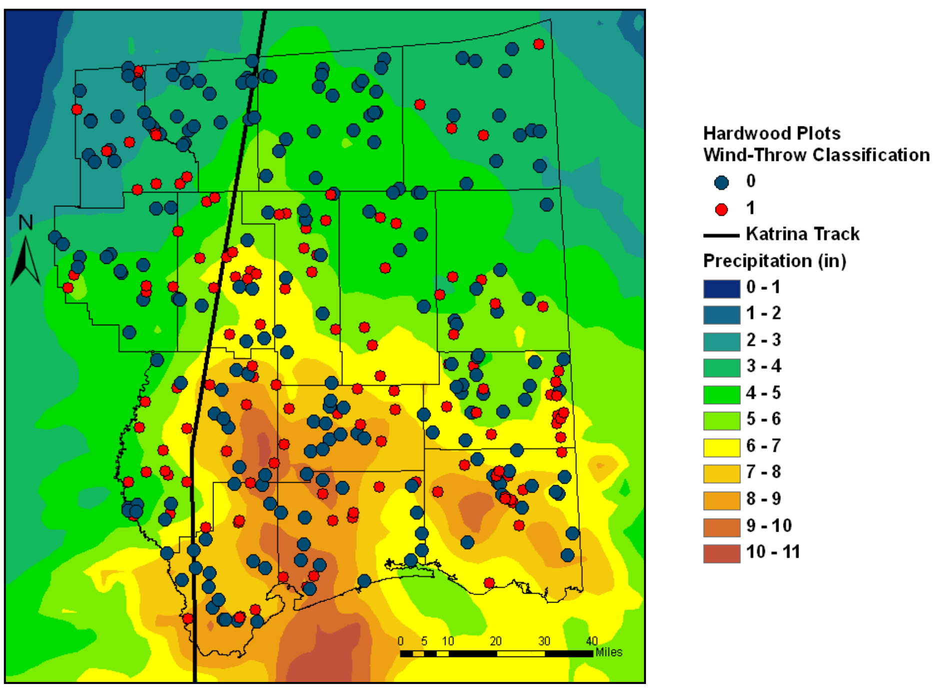

Figure 5.30 Wind-throw and Non-Wind-throw Classified Hardwood Plots overlaid on Total Event Precipitation.

The model summary for significant variables predicting the classification of wind-throw for hardwood plots shows that the logistic regression model built includes 9 steps

(Table 5.15). However, the 9th step does provide the lowest -2 log-likelihood ratio and highest pseudo R^2 values opposed to the model summary for shear classification of pines which occurred in the 10th step and lower pseudo R^2 values.

Table 5.15 Model Summary of Forward Stepwise Logistic Regression of Wind-throw Occurrence on Hardwood Plots.

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	432.333	.072	.098
2	423.290	.096	.131
3	411.078	.128	.174
4	404.264	.145	.197
5	391.703	.176	.239
6	391.743	.176	.239
7	385.728	.190	.258
8	379.855	.204	.277
9	373.454	.219	.297

Comparison of CART analysis and FSBLR Results

A direct comparison of the CART analysis and the FSBLR results can not truly be made since CART investigates a continuous variable while FSBLR examines a categorical binary classification. However, the variables identified as important in CART and the resulting splitting criteria used in the 1-SE and MinCP pruned trees can be compared to the significant exponentiated coefficients through FSBLR.

There is very little concurrence between the important variables of CART analysis and significant variables of FSBLR for predicting sheared percentage to sheared classification (Table 5.16). There are only two variables common to both analyses: H/D Ratio and Lorey's Mean Height. Yet, in the CART methodology, the 1-SE pruned tree has greater H/D ratios splitting right, indicative of increased shear percentage while logistic regression analysis shows a negative coefficient to the likelihood of shear occurrence. Similarly, CART analysis has larger LMH values splitting to the left opposed to FSBLR depicting the greater the LMH height of a pine forest plot, the greater its chance is of being classified as sheared. The cross-validation error R^2 of the 1-SE and MinCP pruned regression trees of percentage pine sheared was only 0.05 and 0.13 respectively. The FSBLR regression equation attained a pseudo R^2 value between 0.12 and 0.2. Hardwoods, however, had similar R^2 values from the CART and FSBLR.

Table 5.16 Comparison of CART Analysis and FSBLR Analysis for Shear of Pine Plots. ** denotes p<0.01 significance and * denotes p<0.05 significance.

Variable	Important in CART	Significant in FSBLR
Elevation		X**
Lorey's Mean Height	X	X**
Distance from Coast		X**
H/D Ratio	X	X**
Trees Per Acre		X**
Physiographic Region		X*
Quadratic Mean Diameter	X	
Lorey's Mean Height	X	
Peak Wind Gusts	X	

The 1-SE and MinCP pruned regression trees for wind-throw percentage cross-validation R^2 values are 0.14 and 0.24 respectively. The FSBLR regression equation attained a pseudo R^2 value between 0.21 and 0.29. Yet, similar to pine plots, CART analysis and FSBLR only share one variable in common for hardwoods wind-thrown (Table 5.17). Both CART and FSBLR identify increasing sustained wind speeds increases the percentage and likelihood that a plot will be wind-thrown.

Table 5.17 Comparison of CART Analysis and FSBLR Analysis for Wind-throw of Hardwood Plots. ** denotes $p<0.01$ significance and * denotes $p<0.05$ significance.

Variable	Important in CART	Significant in FSBLR
Elevation		X**
Trees per Acre		X**
H/D Ratio		X**
Cumulative Wind		X**
Sustained Wind	X	X*
Total Precipitation		
Physiographic Region		X**
Quadratic Mean Diameter	X	
Lorey's Mean Height	X	

MIFI Plot Damage Classification/Reclassification Results

The original binary classifications completed by MIFI through aerial photographs and ground surveys by multiple personnel were recalibrated based on a plot experiencing 10% or greater of shear or wind-throw damage. If shear and wind-throw damage was greater than 10% for a specific plot, the plot was classified by the greater percentage damage percentage. The classification accuracies and pseudo R^2 values were determined through FSBLR of the subjectively based MIFI plot damage classification and the quantification based plot damage classifications for shear damage of pines and wind-throw damage of hardwoods.

FSBLR of the original MIFI plot classifications showed 12 steps (Table 5.13) used to generate the significant variables whereas the reclassified pine plots only used 3 steps (Table 5.18). The average pseudo R^2 values of the reclassified damaged minutely increase compared to the original MIFI damage plot classifications from 0.15 to 0.17.

Table 5.18 Model Summary of Forward Stepwise Logistic Binary Regression of Shear Occurrence on Reclassified Pine Plots.

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	613.511	.058	.103
2	563.924	.116	.204
3	558.82	.122	.214

SPSS 15.0 produces classification accuracy tables detailing the observed binary variable and based on the regression equation built at that step, is used to predict the variable's binary occurrence. The classification accuracies through FSBLR were low when trying to predict whether a pine plot will be classified as shear. Classification accuracies for predicting whether a plot will be classified as sheared was only 5.1% in the 10^{th} step for the original MIFI plot classifications (Table 5.19). Whereas the 2^{nd} step of the reclassified MIFI plot calls is able to correctly classify 10.3% of classified pine plots (Table 5.20). The change in -2 log likelihood from the original binary MIFI damage classifications to the reclassified damaged plots was not statistically significant χ^2 (3, N=784) = 15.543, p>0.05.

Table 5.19 Classification Accuracy of Original MIFI plot damage classifications through FSBLR for Pine Plots.

Observed			Predicted		
			SHEAR		Percentage Correct
			0	1	0
Step 1	SHEAR	0	667	0	100.0
		1	117	0	.0
	Overall Percentage				85.1
Step 2	SHEAR	0	667	0	100.0
		1	117	0	.0
	Overall Percentage				85.1
Step 3	SHEAR	0	665	2	99.7
		1	116	1	.9
	Overall Percentage				84.9
Step 4	SHEAR	0	664	3	99.6
		1	114	3	2.6
	Overall Percentage				85.1
Step 5	SHEAR	0	663	4	99.4
		1	115	2	1.7
	Overall Percentage				84.8
Step 6	SHEAR	0	664	3	99.6
		1	114	3	2.6
	Overall Percentage				85.1
Step 7	SHEAR	0	664	3	99.6
		1	115	2	1.7
	Overall Percentage				84.9
Step 8	SHEAR	0	662	5	99.3
		1	115	2	1.7
	Overall Percentage				84.7
Step 9	SHEAR	0	661	6	99.1
		1	112	5	4.3
	Overall Percentage				84.9
Step 10	SHEAR	0	665	2	99.7
		1	111	6	5.1
	Overall Percentage				85.6
Step 11	SHEAR	0	666	1	99.9
		1	111	6	5.1
	Overall Percentage				85.7
Step 12	SHEAR	0	665	2	99.7
		1	112	5	4.3
	Overall Percentage				85.5

Table 5.20 Classification Accuracy of Reclassified MIFI plot damage classifications through FSBLR for Pine Plots.

Observed			Predicted SHR_0_1 0	Predicted SHR_0_1 1	Percentage Correct 0
Step 1	SHR_0_1	0	664	3	99.6
		1	116	1	.9
	Overall Percentage				84.8
Step 2	SHR_0_1	0	659	8	98.8
		1	105	12	10.3
	Overall Percentage				85.6
Step 3	SHR_0_1	0	655	12	98.2
		1	106	11	9.4
	Overall Percentage				84.9

The same classification procedure was carried out on hardwood plots classified as wind-thrown or non-wind-thrown. The original MIFI plot damage classifications for hardwoods had 9 steps with only 4 steps used for the reclassified MIFI damage classifications (Table 5.21). The average pseudo R^2 values of the reclassified damaged increased compared to the original MIFI damage plot classifications from 0.26 to 0.33.

Table 5.21 Model Summary of Forward Stepwise Logistic Binary Regression of Wind-throw Occurrence on Reclassified Hardwood Plots.

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	250.762	.185	.304
2	237.238	.217	.356
3	226.508	.241	.396
4	219.919	.255	.420

Classification accuracies for predicting whether a plot will be classified as Wind-thrown was 61.2% in the 9th step for the original MIFI plot classifications (Table 5.22), whereas the 2nd step of the reclassified MIFI plot calls decreases to 45.3% (Table 5.23).

However, the overall classification accuracy including both wind-thrown and non-wind-thrown plots increases from the original MIFI plot classifications from 71.1% to 88%. The change in -2 log likelihood from the original binary MIFI damage classifications to the reclassified damaged plots was statistically significant χ^2 (4, N=342) = 153.535, p<0.01.

Table 5.22 Classification Accuracy of Original MIFI plot damage classifications through FSBLR for Hardwood Plots.

Observed			Predicted		
			BLOWDOWN		Percentage Correct
			0	1	0
Step 1	BLOWDOWN	0	172	36	82.7
		1	98	36	26.9
	Overall Percentage				60.8
Step 2	BLOWDOWN	0	171	37	82.2
		1	89	45	33.6
	Overall Percentage				63.2
Step 3	BLOWDOWN	0	168	40	80.8
		1	77	57	42.5
	Overall Percentage				65.8
Step 4	BLOWDOWN	0	166	42	79.8
		1	70	64	47.8
	Overall Percentage				67.3
Step 5	BLOWDOWN	0	168	40	80.8
		1	66	68	50.7
	Overall Percentage				69.0
Step 6	BLOWDOWN	0	168	40	80.8
		1	66	68	50.7
	Overall Percentage				69.0
Step 7	BLOWDOWN	0	165	43	79.3
		1	57	77	57.5
	Overall Percentage				70.8
Step 8	BLOWDOWN	0	164	44	78.8
		1	57	77	57.5
	Overall Percentage				70.5
Step 9	BLOWDOWN	0	161	47	77.4
		1	52	82	61.2
	Overall Percentage				71.1

Table 5.23 Classification Accuracy of Reclassified MIFI plot damage classifications through FSBLR for Hardwood Plots.

Observed			Predicted WTRW_0_1: 0	Predicted WTRW_0_1: 1	Percentage Correct
Step 1	WTRW_0_1	0	273	8	97.2
		1	46	15	24.6
	Overall Percentage				84.2
Step 2	WTRW_0_1	0	273	8	97.2
		1	40	21	34.4
	Overall Percentage				86.0
Step 3	WTRW_0_1	0	273	8	97.2
		1	33	28	45.9
	Overall Percentage				88.0
Step 4	WTRW_0_1	0	272	9	96.8
		1	33	28	45.9
	Overall Percentage				87.7

CHAPTER 6

DISCUSSION AND CONCLUSIONS

The primary focus of this study was to employ data mining and statistical techniques to ascertain the affects of biotic and abiotic factors that are significantly related to damage patterns across the Southeastern Forest District of Mississippi as a result of Katrina. A GIS database was built using tree-level characteristics aggregated to the plot level along with meteorological, topographical, and pedological variables in an attempt to identify the causal factors influencing shear damage of pines and wind-throw damage of hardwoods. Using [R 2.8.1], CART analysis was performed using two dependent response variables, percentage of stems damaged per plot and percentage basal area damaged per plot. Additional statistical analysis was executed on the binary classifications of plot damaged and compared to CART for the identification of causal variables responsible for differing damage types to trees. Reclassification of the binary original plot damage classifications was attempted to increase the classification accuracy through logistic regression. The relative importance of biotic and abiotic factor inter-relationships as contributors to wind-throw of hardwoods and shear of pines was investigated for the purpose of modeling the spatial distribution of vegetative debris. Knowledge of the spatial distribution of vegetative debris could optimize allocation of personnel and equipment to dispose of vegetative debris in a more efficient manner following a hurricane.

CART

CART analysis was used to assess the factors which maximized: the percent of pines per plot sheared, percentage of hardwoods per plot wind-thrown, percentage of pine basal area per plot sheared, and the percentage of hardwood basal area per plot wind-thrown. Each regression tree was fully-grown and pruned twice using the 1-SE method and MinCP method. Subsequent pruning of the grown regression trees revealed that Lorey's Mean Height (LMH) and average Quadratic Mean Diameter (QMD) were important variables for both pine and hardwood plots. Pine plots experiencing greater wind gusts were more prone to snapping while hardwood plots exposed to higher sustained wind speeds maximized the percentage of wind-thrown hardwoods and percent wind-thrown basal area. However the splitting criteria used is confounding when compared to other literature.

The 1-SE pruned regression tree for pine shear percentage (Figure 5.1.1.5) shows the pine plots averaging greater than 8 inches QMD experienced about 2.5 times more sheared pines per plot than pine plots averaging less than 8 inches QMD. This is suggesting the larger diameter, thus taller trees are more prone to snapping then smaller diameter, shorter trees agreeing with Everham and Brokaw (1996) and many others. However, the next split details plots averaging less than 44 feet LMH experienced more shear damage than plots averaging more than 44 feet LMH. Further investigation in ArcMap 9.2 showed only 76 plots having LMHs equal to and greater than 44 feet with over 10% pines sheared, whereas 345 plots existed with LMHs less than 44 feet but still experienced 10% shear, corroborating the findings of CART. Further splits of the 1-SE pruned regression tree seem to justify the prior split however. Higher peak wind gusts values were used as splitting criteria (139 mph) for plots less than 44 feet LMH compared to pine plots greater than 44 feet LMH (130 mph). The shorter the average LMH, the greater the wind speed must be to cause damage. Shorter pine trees have smaller bending moments placed on them due to wind loading compared to larger trees, and for more damage to occur on shorter pine stands, higher winds and peak gusts must be realized (Stanturf 2007). For the taller pine plots, the split following the wind gusts again shows greater QMD plots averaged 20% shear damage per plot. Moreover, the following split for shorter pine plots experiencing less than 139 MPH peak wind gusts depicts plots with H/D ratios greater than 54 had 14.7 times more sheared pines per plot than plots averaging less than 54. According to Pienaar *et al.* (1997) which investigated growth and yield information for loblolly pine in the south, found the average H/D ratio for a loblolly pine located in a 400 tree/acre stand at age 14 is around 74. The average H/D ratio for loblolly pines at age 5 was determined to be 19.9ft/0.35ft equaling 57. Therefore, caution should be used when using the splitting criteria and how it relates to previous splits.

The 1-SE pruned regression tree of percent hardwoods per plot wind-thrown also identified shorter LMHs were prone to more wind-throw but the next splitting criteria depicts hardwoods plots having greater QMD had 6 times the amount of wind-throw. This study incorporates all hardwood species surveyed by MIFI in 2006. Yet it has been shown by Oswalt and Oswalt (2008), Touliatos and Roth (1970), Putz et al. (1983), Peterson (2000), that differing species of hardwoods are more or less resilient in a hurricane environment. Therefore, trees less susceptible to wind-throw damage such as pondcypress and swamp tupelo with greater heights than trees more susceptible like Pecan, Hickory, Dogwood, and Red Maple with lower heights at the time of the inventory, may be leading to this splitting criteria being used. Dogwoods only grow to 20-25 feet, Red Maple grows to 60 feet, Hickories grow to 60-80 feet, with Pecan trees being the taller of the hardwoods growing upwards to 100 feet (Sander 2009). Whereas Pondcypress grows to roughly 80 feet and Swamp Tupelo achieves heights of 120 feet (USDA 1994). The resulting species differences in heights and a respective species susceptibility may contribute the splitting criteria used in CART suggesting hardwood stands lower in LMH are more wind-thrown susceptible.

A larger question at hand for the CART analysis is the low classification accuracies and low cross-validation R^2 compared to other studies which used CART methodology. Lindermann and Baker (2002) were able to correctly classify 82.9% of their dataset to a 22-node decision tree over a 10,000ha (38.6 sq. mi) area. Kupfer *et al.* (2008) correctly classified 71.5% of their single classification tree across 153,000ha (590 sq. mi) area. The Southeast forest District study area is 2,294,607ha (8860 sq. mi) (MIFI 2006). The 15-fold increase in area of the Southeast Forest District compared to the DeSoto National Forest and over 233 times the size of the Routte-Divide area most likely reduced the classification accuracies and cross-validation R^2 values of the regression trees produced (Table 6.1 and Table 6.2).

Table 6.1 Classification Accuracies of each Response Variable for each Regression Tree Generated.

Dependent Response Variable Tested	Full Tree	1-SE Tree	MinCP Tree
Percent Pine Trees Sheared	50.0%	42.2%	29.5%
Percent Basal Area of Pines Sheared	39.7%	30.3%	21.6%
Percent Hardwood Trees Wind-thrown	66.0%	53.0%	41.5%
Percent Basal Area of Hardwoods Wind-thrown	61.9%	55.5%	38.5%

Table 6.2 Cross-Validation R^2 values of each Response Variable for each Regression Tree Generated.

Dependent Response Variable Tested	Full Tree	1-SE Tree	MinCP Tree
Percent Pine Trees Sheared	0.076	0.051	0.133
Percent Basal Area of Pines Sheared	0.039	0.039	0.064
Percent Hardwood Trees Wind-thrown	0.15	0.137	0.232
Percent Basal Area of Hardwoods Wind-thrown	0.065	0.132	0.233

The size of the study area and the generality of tree groupings, i.e. hardwood and pines, may hinder the accurate classification of the dependent response variables to the biotic and abiotic variables and results in low cross-validation R^2 values. The large area used may be deterring the true relationship between the dependent response variables and independent biotic and abiotic variables to be realized through CART analysis. Higher classification accuracies and cross-validation R^2 values may be obtained by partitioning the forest inventory plots spatially or by their attributes. Inventory plots located with a given county can be analyzed through CART and resulting regression trees be compared to the Southeast Forest District to see if classification accuracy and cross-validation R^2 values increase. In addition to spatial segregation, plots could be further separated by their species group attribute. This will allow for single or similar species to be tested against the biotic and abiotic independent factors opposed to the general classes of pine and hardwood.

Further limitations of CART methodology are the hard splitting criterion used to build regression trees. For example in the MinCP pruned regression tree for sheared basal area percentage for pine plots depicts stands exposed to wind gust speeds greater than 124 MPH had more basal area sheared (12.5%) opposed to plots experiencing less than 124 MPH peak wind gust speeds (4%). What if a plot experiences a wind gust speed of 123MPH? Does this mean it will not experience more basal area sheared compared to another pine plot that experienced 125MPH peak wind gust speed? A comparison of CART analysis to MARS analysis would also be of interest since the hinge functions built into MARS allow trends along the range of an independent variable as it compares to the dependent variable.

Forward Stepwise Binary Logistic Regression

There were very few variables that were portrayed in CART analysis and also proved to be significant through binary logistic regression. Variables that did coincide were Loreys' Mean Height and H/D ratio. However, greater H/D ratios in CART split right indicating greater percentage of damage, whereas in FSBLR the coefficient was negative. FSBLR reinforced that topographical classification (bottomland, terrace, or upland) depicted pines plots classified as bottomland were 0.241 times less likely to be classified as shear proving to much more resilient than those classified as terrace or upland plots. The pseudo R^2 values corroborated with the findings of CART analysis with lower values for the prediction of pine shear opposed to hardwood wind-throw.

Forward Stepwise Binary Logistic Regression can only perform statistical analysis on binary coded dependent variables and it cannot identify hard splits rules like CART can. Future utilization of FSBLR, like CART, can be used to analyze various portions of the dataset opposed the entire dataset. As mentioned above, the incorporated data can be separated spatially or by a specific attribute.

Reclassification of MIFI Damage Classifications

FSBLR results of the original and reclassified MIFI plot damage classifications showed a significant ($p<0.01$) increase in classification accuracy for hardwood plots when the reclassified plot damage classifications were used. There was not a significant difference in classification accuracy for pine plots categorized as sheared. However, for both hardwood and pine plots, simpler regression equations were generated using the reclassified MIFI plot damage classifications which in turn generated higher pseudo R^2 values. It is the recommendation of these results that once a ground survey has been completed of the damaged areas, the plots be classified according to the percentage of the majority of damage sustained greater than 10% opposed to subjective analysis.

Future Studies

Much continued work is needed to fill the limitations of this study. The dataset can be partitioned by political or environmental boundaries to analyze and compare smaller sized datasets analyzed through CART to the results from the entire study area. This will test the interaction robustness among the biotic and abiotic factors as they relate to the dependent variables. In order for the important biotic variables identified through CART and significant biotic variables through FSBLR to be implemented into a continuous predictive vegetative debris distribution model, remote sensing studies similar to Ramsey *et al.* (1998) need to be carried out on archived Landsat 5 TM, NAIP imagery, or other aerial imagery. Large area forest inventories using Landsat imagery have shown promise in predicting forest parameters such as basal area, height, health conditions, and biomass across large landscapes (McCombs 2003, Berryman 2004, Schultz *et al.* 2006, Meng *et al.* 2008). Applying continuous forest metrics over the landscape scale with knowledge of their relationship to forest damage will increase the accuracy of potential vegetative debris distribution following a land-falling hurricane. Applying the values used in splits of the full regression trees produced using CART coupled with the logistic regression equations to the continuous raster grids of biotic and abiotic variables of future work will allow for a rapid continuous spatial model depicting vegetative debris distribution to be generated for utilization by emergency managers. This information can potentially improve decisions made in allocation of personnel and equipment for faster recovery following a hurricane.

REFERENCES

Ayala-Silva, T., Twumasi, Y. A. 2004. Hurricane Georges and Vegetation change in Puerto Rico using AVHRR satellite data. *International Journal of Remote Sensing, 25,* 1629–1640.

Berryman, B.N. 2004. Investigation of Relationships Between Landsat ETM+ Data and Ground-Adjusted LiDAR Measurements in Southern Pine Stands. MS Thesis, Mississippi State University, Mississippi State, MS USA, 50 p.

Beckwith, J.R. 1997. Timber Sale volumes, Weights, and Prices. Feb 2009. <http://warnell.forestry.uga.edu/warnell/service/library/index.php3?docID=333>

Bettinger, P., Merry, K., Hepinstall, J. 2009. Average Tropical Cyclone Intensity Along. the Georgia, Alabama, Mississippi, and North Florida Coasts. *Southeastern Geographer* 49.1: 50-66.

Boose, E. R., Foster, D. R., Fluet, M. 1994. Hurricane Impacts to Tropical and Temperate Forest Landscapes. *Ecological Monographs, 64,* 369–400.

Brack, C. 1999. Average Diameter of a Stand. 5 April 2009. <http://sres-associated.anu.edu.au/mensuration/author.htm>

Breiman, L., A. Cutler. 1995. Random Forests. 1 April 2009. <http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm>

Burns, P. 2008. The Statistical Bootstrap and Other Resampling Methods. Accessed 6 April 2009. <http://www.burns-stat.com/pages/Tutor/bootstrap resampling.html>

W. H. Cooke, Grala, Katarzyna, and Evans D., Collins. 2006. Assessment of pre- and post-Katraina fuel conditions as a component of fire potential modeling for southern Mississippi, USA. Journal of Forestry 105, 389-397.

Canham, C.D. & Loucks, O.L. 1984. Catastrophic wind-throw in the Presettlement forests of Wisconsin. *Ecology*, 65, 803–809.

Canham, C. D., M. J. Papaik, and E. F. Latty. 2001. Interspecific variation in susceptibility to windthrow as a function of tree size and storm severity for northern temperate tree species. *Can. J. For. Res*. 31:1-10.

Collins, C.A., D.W. Wilkinson, D.L. Evans. 2005. Multi-temporal analysis of Landsat data to determine forest age classes for the Mississippi statewide forest inventory - preliminary results. Proceedings of the Third International Workshop on the Analysis of Multi-temporal Remote Sensing Images. 2005 May 16-18. Biloxi, MS.

Curtis, J.D. 1943. Some observations wind damage. *J . For.* 41:877-882. [13] United States Department of Agriculture, 1956, Glossary of Terms Used in Forest Fire Control, Washington D.C.

Dale VH, Lugo AE, MacMahon J, Pickett STA. 1998. Ecosystem management in the context of large, infrequent disturbances. *Ecosystems* 1:546–57.

Difford, M. 2008 [R] Model R^2 and Partial R^2 Values. Accessed on 10 April 2009. <http://markmail.org/message/goqpsqusqsmamr4h#query:rsq.rpart+page:1+mid:ivcqpcaiixijl3r+state:results>.

Donoghue, D.N.M. "European Windstorms and the Role of Remote Sensing on Helping Assess Damage to Forestry." 4th International Workshop of Remote Sensing for Post-Disaster Response. Cambridge, Sept. 25-26 2006.

Evans, A. M., M. Tyrrell, A. Camp, C. Riely. 2007. Biotic and abiotic influences on severity of wind disturbance in forests of NW Pennsylvania. *Forest Ecology and Management*. 245(1-3): 44-53.

Everham, E.M.& Brokaw, NVL. 1996. Forest damage and recovery from catastrophic wind. Botanical Review. 62:113-185.

ESRI. 1998. ESRI Shapefile Technical Description. Accessed 14 April 2009. <http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>.

Faulkner, S.P. 2003. Ecological succession in Mississippi - patterns and processes. Handout, Department of Biological Sciences, Delta State University, 3 pp.

Foster, D.R. 1988a. Species and stand response to catastrophic wind in central New England, USA. *Journal of Ecology*, 76, 135–151.

Foster, D. R., & Boose, E. R. 1992. Patterns of forest damage resulting from catastrophic wind in Central New-England, USA. *Journal of Ecology, 80,* 79–98.

M.W. Gardner and S.R. Dorling, Artificial neural network: the multilayer perceptron: a review of applications in atmospheric sciences, *Atmos. Environ.* **32** (1998), pp. 2627–2636.

Garson, D. 2009. Logistic Regression. Accessed April 15 2009. <http://faculty.chass.ncsu.edu/garson/PA765/logistic.htm>.

Gillman, E.F. & Watson, D.G. 1994. Pinus taeda: Loblolly Pine. United States Forest Service. Southern Group of State Foresters. Fact Sheet ST-478. <http://hort.ufl.edu/trees/PINTAEA.pdf>

Gillman, E.F. & Watson, D.G. 1994. Pinus taeda: Loblolly Pine. United States Forest Service. Southern Group of State Foresters. Fact Sheet ST-478. <http://hort.ufl.edu/trees/PINPALA.pdf>.

Gillman, E.F. & Watson, D.G. 1994. Pinus taeda: Loblolly Pine. United States Forest Service. Southern Group of State Foresters. Fact Sheet ST-478. <http://hort.ufl.edu/trees/PINELLA.pdf>.

Haanpä, S., S. Lehtonen, L. Peltonen, E. Talockaite. 2008. Impacts of wind storm Gudrun of 7th-9th January 2005 and measures taken in the Baltic Sea Region. Accessed Feb 2009. <http://www.astraproject.org/sites/download/ASTRA_WSS_report_final.pdf>.

Hook et al., 1991. D.D. Hook, M.A. Buford and T.M. Williams, Impact of Hurricane Hugo on the South Carolina coastal plain forest. *J. Coastal Res.* SI8 (1991), pp. 291–300.

Huang, C., L. Yang, C. Homer, M. Coan, R. Rykhus, Z. Zhang, B. Wylie, K. Hegge, Z. Zhu, A. Lister, M. Hoppus, R. Typcio, L. DeBlander, W. Cooke, R. McRoberts, D. Wendt, and D. Weyermann. 2001. Synergetic use of FIA Plot DATA and Landsat 7ETM+ Images for Large Area Forest Mapping. 35th Annual Midwest Forest Mensurationists Meeting and the Third Annual Forest Inventory and Analysis Symposium. October 17-19, 2001. <http://landcover.usgs.gov/pdf/synergistic.pdf>.

Ishizuka, M., Toyooka, H., Osawa, A., Kushima, H., Kanazawa, Y. & Sato, A. 1998 Secondary succession following catastrophic wind-throw in a boreal forest in Hokkaido, Japan: the timing of tree establishment. *Journal of Sustainable Forestry*, 6, 367–388.

Lawrence R, Bunn A, Powell S, Zambon M. 2004. Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote Sensing Environ* 90:331–336.

Leisch, F. 2003. The R foundation for Statistical Computing. Accessed 6 April 2009. http://www.r-project.org/.

Measells, M.K. & J. Henderson. 2008. 2005 Harvest of Forest Products. Mississippi State Extension Service. Department of Forestry. Accessed: Jan 7 2009. <http://msucares.com/forestry/economics/reports/2005_harvest_report.pdf>.

McCombs, J.W., S.D. Roberts, and D.L. Evans, 2003. Influence of fusing lidar and multispectral imagery on remotely sensed estimates of stand density and mean tree height in a managed loblolly pine plantation. Forest Science, 49(3), pp. 457-466.

McCormick, R. J., T. A. Brandner, and T. F. H. Allen. 2000. Toward a Theory of Meso scale Wildfire Modeling - A Complex Systems Approach Using Artificial Neural Networks. Pages 3 – 15 in: L. F. Neuenschwander, K. C. Ryan, G. E. Gollberg, and J. D. Greer, editors. Proceedings from The Joint Fire Science Conference and Workshop. University of Idaho and the International Association of Wildland Fire, Moscow, Idaho.

McMaster, K. J., 2005. Forest blowdown prediction: A correlation of remotely sensed contributing factors. *Northern Journal of Applied Forestry, 22,* 48–53.

McNairn, H., J. Shang, C. Champagne, E. Huffman, A. Smith, T. Fisette. 2004. A multi Sensor Approach to Inventorying Agricultural Land Use. International Symposium on Remote Sensing of Environement. June 20-24. <http://www.isprs.org/publications/related/ISRSE/html/papers/743.pdf>.

Miller, J., K. Borne, B. Thomas, Z. Huang, Y. Chi. 2005. Automated Wildfire Detection Through Artificial Neural Networks. NASA. 15 Mar 2009. <http://ntrs.nasa. gov/archive/nasa/casi.ntrs.nasa.gov/20050180456_2005177531.pdf >.

Nix, L.E.; Hook, D.D.; Williams, J.G.; Blaricom, D.V. 1996. Assessment of hurricane damage to the Santee Experiment Forest and the Francis Marion National Forest with a geographic land research and management related to the storm. Gen. Tech. Rep. SRS-5. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 552 p.

NOAA. "Chronological List of All Hurricanes which Affected the Continental United States: 1851-2005." May 2006. NOAA/HRD/AOML 10 Dec. 2008. <http://www.aoml.noaa.gov/hrd/hurdat/ushurrlist.htm>.

NOAA. "NHC Archive of Hurricane Seasons." March 2009. NOAA/NHC. <http://www.nhc.noaa.gov/pastall.shtml>.

Ormsby, J. P., Choudhury, B. J., & Owe, M. 1987. Vegetation spatial variability and its effect on vegetation indices. *International Journal of Remote Sensing, 8,* 1301–1306.

Özesmi, S.L. and U. Özesmi. 1999 An artificial neural network approach to spatial habitat modeling with interspecific interaction, *Ecological Modeling* 116, pp. 15–31.

Parker, R.C., P.A. Glass, H.A. Londo, D.L. Evans, K.L. Belli, t.G. Matney, E.B. Schultz. 2005. Mississippi's forest inventory pilor program: Use of computer and spatial technologies in large area inventories. Forest and Wildlife research Center, Bulletin FO 274, Mississippi State University. 43pp.

Peltola, H. M. 2006. Mechanical stability of trees under static loads. American Journal of Botany 93: 1501–1511.

Peterson, C.J. & Pickett, S.T.A. 1990. Microsite and elevational influences on early forest regeneration after catastrophic wind-throw. *Journal of Vegetation Science*, 1, 657–662.

Peterson, C. J. 2000b. Damage and recovery of tree species after two different tornadoes in the same old growth forest: a comparison of infrequent wind disturbances. *For. Ecol. Manage* 135:237–252.

Peterson, C.J. 2007. Consistent influence of tree diameter and species on damage in nine Eastern North America Tornado blowdowns. *Forest Ecology and Management*, 250: 96-108.

Petty, J.A. and Swain, C., 1985. Factors influencing stem breakage of conifers in high winds. *Forestry* 58 1, pp. 75–85.

Powell, M. D., S. H. Houston, L. R. Amat, and N Morisseau-Leroy, 1998: The HRD real time hurricane wind analysis system. *J. Wind Engineer. and Indust. Aerodyn.* 77&78, 53-64.

Powell MD, Bowman D, Gilhousen D, Murillo S, Carrasco N, St. Fleur R. 2004. Tropical cyclone winds at landfall: the ASOSCMAN wind exposure documentation project. Bull Am. Meteorol Soc 85:845–51.

Purevdorj, T., Tateishi, R., Ishiyama, T., & Honda, Y. (1998). Relationships between percent vegetation cover and vegetation indices. *International Journal of Remote Sensing, 19*, 3519–3535. doi:10.1080/014311698213795.

Putz, F. E., Coley, P. D., Lu, K., Montalvo, A., & Aiello, A. 1983. Uprooting and Snapping of Trees-Structural Determinants and Ecological Consequences. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere, 13,* 1011–1020. doi:10.1139/x83-133.

Ramsey, E.W.; Chappell, D.K.; Baldwin, D.G. 1997. AVHRR imagery used to Identify Hurricane Damage in a Forested Wetland of Louisiana. 63: 293-297.

Ripley, B. 2009. The Rpart Package. Accessed 26 March 2009. <http://cran.r project.org/web/packages/rpart/rpart.pdf>.

Rodgers JC III, Gamble DG, McCay DH, Phipps S. 2006. Tropical Cyclone Signals within Tree Ring Chronologies from Weeks Bay National Estuary and Research Reserve. *J Coast Res.* 22:1320–1329.

Roy, D. P., Kennedy, P., and Folving, S. 1997. Combination of normalized difference vegetation index and surface temperature for regional scale European forest cover mapping using AVHRR data. *International Journal of Remote Sensing*. 18, 1189–1195.

Rosson, J. F., Jr. 1995. Forest resources of Louisiana, 1991. F. S. U.S. Department of Agriculture, Southern Forest Experiment Station. *Resour. Bull. SO-1,* 92. New Orleans, Louisiana.

Rounds, S.A. 2002. Development of a Neural Network Model for Dissolved Oxygen in the Tualatin River, Oregon. Second Federal Interagency Hydrologic Modeling Conference. USGS. July 29 - August 1, 2002.

Sadar, S. A., Hayes, D. J., Hapinstall, J. A., Coan N, M., and Soza, C., 2001, Forest change monitoring of a remote biosphere reserve. *International Journal of Remote Sensing*, 22, 1937–1950.

Salford Systems. 2009. Accessed 19 March 2009. <http://www.salford systems.com/1211.php>.

Schmoldt, D.L. 2001. Application of artificial intelligence to risk analysis for forested ecosystems. Risk Analysis in Forest Management, 49–74.

Schroff, E. 2008. "December 2007 Storm." Accessed 4 January 2009. <http://www.dnr.wa.gov/BusinessPermits/Topics/ForestPracticesApplications/Pages/fp_storm_damage.aspx>.

Schultz, E.B., T.G. Matney, D.L. Evans, I. Fujisaki. 2006. A landsat stand basal area classification suitable for automating stratification of forest into statistically efficient strata. 6 pages in S. Land, T. Blaschke, E. Schpfer, editors, First International Conference on Object-based Image Analysis. Salzburg University, Austria. July 4-5, 2006.

Sewell, J. 2002. Bill to Create Mississippi Institite for Forest Inventory Signed by Governor Musgrove. 2 April 2009. <http://mdah.state.ms.us/arrec/digital_archives/musgrove/pdfs/5029.pdf>.

Silva, A.P. (2003). Neural networks application to spatial interpolation of climate variables. STSM on the Framework of COST 719 from 6-10 October 2003 at Vienna – Austria.

Stathers, R.J., T.P. Rollerson, and S.J. Mitchell. Windthrow handbook for British Columbia forests. B.C. Min. For., Victoria, B.C. Working Paper 9401.

Stergiou, C., D. Siganos. 1996. Neural Networks. 1 April 2009. <http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html.>

Touliatos, P. and E. Roth. 1971. Hurricanes and trees: Ten lessons from Camille. *J. For.* 285-289.

USDA Department of Agriculture. 1975. Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys. Agriculture Handbook No. 436.

Sheffield, Raymond M.; Thompson, Michael T. 1992. Hurricane Hugo Effects on South Carolina's Forest Resource. Res. Pap. SE-284. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 60 p.

USDA Forest Service (2007). The forest inventory and analysis database:Database description and users guide version 2.1. National Forest Inventory and Analysis Program, U.S. Department of Agriculture, Forest Service, Southern Research Station.

USDA Forest Service. Southern Research Station. 2006. Accessed 3 March 2009. <http://www.srs.fs.usda.gov/pubs/su/010/threats.htm>.

Walker, L. R., J. Voltzow, J. D. Ackerman, D. S. Fernandez, and N. Fletcher. 1992. Immediate impact of Hurricane Hugo on a Puerto Rican rain forest. *Ecology* 73:691-694.

Wang, F., & Xu, Y. J. (2008). Hurricane Katrina-induced forest damage in relation to ecological factors at landscape scale. *Environmental Monitoring and Assessment*. doi:10.1007/s10661-008-0500-6.

Weaver, P.L. 1989. Forest changes after hurricanes in Puerto Rico's Luquillo mountains. *Interciencia* 14, pp. 181–192.

Wesley, D. A., Poulos, G. S.,Meyers, M. P., Snook, J. S., and Judson, A., 1998, Observations and forcing mechanisms during the October 1997 Front Range blizzard and forest destruction. in *Preprints, Eighth Conference on Mountain Meteorology, 3–7 August 1998, Flagstaff, AZ, American Meteorological Society* (Boston, Massachusetts), pp. 25–30.