1-1-2014

Feasibility of Consistently Estimating Timber Volume through Landsat-based Remote Sensing Applications

Renaldo Josue Salazar Arroyo

Feasibility of consistently estimating timber volume through Landsat-based remote sensing applications

By

Renaldo Josue Salazar Arroyo

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Forest Resources
in the College of Forest Resources

Mississippi State, Mississippi

May 2014

Feasibility of consistently estimating timber volume through Landsat-based remote sensing applications

By

Renaldo Josue Salazar Arroyo

Approved:

Emily B. Schultz
(Co-Major Professor)

Thomas G. Matney
(Co-Major Professor)

David L. Evans
(Committee Member)

Zhaofei (Joseph) Fan
(Committee Member)

Andrew W. Ezell
(Graduate Coordinator)

George M. Hopper
Dean
College of Forest Resources

Name: Renaldo Josue Salazar Arroyo

Date of Degree: May 16, 2014

Institution: Mississippi State University

Major Field: Forest Resources

Major Professor: Emily B. Schultz and Thomas G. Matney

Title of Study: Feasibility of consistently estimating timber volume through Landsat-based remote sensing applications

Pages in Study: 65

Candidate for Degree of Doctor of Philosophy

The Mississippi Institute for Forest Inventory (MIFI) is the only cost-effective large-scale forest inventory system in the United States with sufficient precision for producing reliable volume/weight/biomass estimates for small working circle areas (procurement areas). When forest industry is recruited to Mississippi, proposed working circles may overlap existing boundaries of bordering states leaving a gap of inventory information, and a remote sensing-based system for augmenting missing ground inventory data is desirable. The feasibility of obtaining acceptable cubic foot volume estimates from a Landsat-derived volume estimation model (Wilkinson 2011) was assessed by: 1) an initial study to temporally validate Landsat-derived cubic foot volume outside bark to a pulpwood top estimates in comparison with MIFI ground truth inventory plot estimates at two separate time periods, and 2) re-developing a regression model based on remotely sensed imagery in combination with available MIFI plot data. Initial results failed to confirm the relationships shown in past research between radiance values and volume estimation. The complete lack of influence of radiance values in the model led to a re-assessment of volume estimation schemes. Data outlier trimming

manipulation was discovered to lead to false relationships with radiance values reported in past research. Two revised volume estimation models using age, average stand height, and trees per-acre and age and height alone as independent variables were found sufficient to explain variation of volume across the image. These results were used to develop a procedure for other remote sensing technologies that could produce data with sufficient precision for volume estimation where inventory data are sparse or non-existent.

DEDICATION

This work is dedicated to my parents who pushed me to stand out and never accept being like everyone else. They taught me to work hard, never be complacent, and always value the opinions of others. There was never a moment in my life where I felt anything other than total encouragement to follow my dreams and achieve my goals. They gave me unconditional love and support and I am forever grateful. I strive every day to be as good of a person as they were. They are my role models and their lessons are my daily motivation.

ACKNOWLEDGEMENTS

I would first like to thank my parents, Ronald Douglas Arroyo and Cecelia Salazar Arroyo, for teaching me the value of education. They always enrolled me in the best schools and provided me every opportunity to learn new things and see the world. I am who I am today because of them. I love them and I miss them every day.

I would like to thank my brother, Luiz Arroyo, for helping me every step of the way. Though he is my little brother, he has come up big for me in many occasions. I would also like to thank my grandparents, Diego Lucas Salazar and Dora Salazar, for all the time and effort they put in to help raise me. They worked hard all their lives and were inspirational. I would also like to thank my sisters, Janine Shapley and Julie Bousfield, who also pursued the field of education and serve as guides along my path. My sister, Christina Hubbard, is always a source of smiles, hugs, and encouragement. I also could not have completed a Ph.D. without the love and support of all of my family members, including Raleigh. You are there when I need you and always supportive.

I would like to acknowledge the work of my two major professors, Emily B. Schultz and Thomas G. Matney, for their effort in opening my eyes to all I could learn. I needed the push from Dr. Matney to try new things and needed the support from Dr. Schultz when I failed at those things. Together they make a great team. I would also like to thank the other members of my committee, David L. Evans and Zhaofei (Joseph) Fan, for their efforts in helping me fill in the gaps of my knowledge of spatial technologies.

Lastly, I want to acknowledge all of my friends that have helped me along the way in graduate school. Your smiles, jokes, potluck dinners, trips, positive attitudes, knowledge, and support have made my life easier every day of this whole experience. You are all important to me and I hope to be there for you as much as you were there for me.

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGEMENTS iii

LIST OF TABLES vii

LIST OF FIGURES viii

CHAPTER

I. INTRODUCTION 1

Background 1
Justification and Objective 4
Study Progression 5

II. LITERATURE REVIEW 6

Forest Inventory and Analysis (FIA) 6
Mississippi Institute for Forest Inventory (MIFI) 8
Remote Sensing in Forest Inventories 10
Remotely Sensed Volume Estimation 11
Shortcomings of Remote Sensing Related to Forest Volume Prediction 13
Probabilistic Problems of Change Detection in Long Sequences and Periods 13
Problems Associated with Data 14
Data Integrity 14
Data Preparation, Manipulation and Storage 15
Issues of Sampling Error 15
Registration/Georectification/Classification Issues (Effects on Accuracy of forest/non-forest) 15
Summary 16

III. INITIAL STUDY METHODS 19

Planning and Study Area 19
Data 21
MIFI Inventory Data 21
Landsat Thematic Mapper (TM) Data 22

Volume Estimates 27

IV. INITIAL STUDY RESULTS 29

Contribution of Radiance Band Information 30
Change Detection Time Interval 31
Outlier and Trimming Analysis 33

V. REASSESSMENT STUDY 35

Objectives 35
Methods....36
Study Area 36
Data 37
Verification of Initial Findings 39
Revised Models....40
Trimming Analysis 40

VI. REASSESSMENT STUDY RESULTS 42

Verification Results 42
Revised Regression Model Results....44
Data Trimming Results 46

VII. DISCUSSION 50

Failure of Radiance Values to Predict Volume....51
Change Detection and Forest Age Determination 51
Outlier Trimming Analysis 53
Volume Estimation Model 53
Remote Sensing Technologies 54

VIII. CONCLUSION 59

REFERENCES 62

LIST OF TABLES

1 Generalized power model index of fit (I^2), Mississippi Institute for Forest Inventory (MIFI) field inventory (±15% acceptable sampling range at the 95% confidence interval) cubic foot volume outside bark to a pulpwood top (CFVOBPW) per-acre estimations, Landsat-derived CFVOBPW per-acre estimations, and percent (%) difference comparisons. 30

2 Guidelines for manual correction of forest ages, based on Mississippi Institute for Forest Inventory (MIFI) ground plot estimates. 39

3 Data trimming iterations for Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region pine cover type data (n= 226) displaying precision changes in R^2, root mean square error (RMSE), estimated cubic foot volume outside bark to a pulpwood top per-acre ($\bar{Y}$), percent change in $\bar{Y}$ (% $\bar{Y}$), standard deviation (StDev), and removed and total plots. 47

4 Data trimming iterations for Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region hardwood cover type data (n= 219) displaying precision changes in R^2, root mean square error (RMSE), estimated cubic foot volume outside bark to a pulpwood top per-acre ($\bar{Y}$), percent change in $\bar{Y}$ (% $\bar{Y}$), standard deviation (StDev), and removed and total plots. 48

5 Data trimming iterations for Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region mixed cover type data (n= 66) displaying precision changes in R^2, root mean square error (RMSE), estimated cubic foot volume outside bark to a pulpwood top per-acre ($\bar{Y}$), percent change in $\bar{Y}$ (% $\bar{Y}$), standard deviation (StDev), and removed and total plots. 49

LIST OF FIGURES

1 The counties of Mississippi separated into five designated Mississippi Institute for Forest Inventory (MIFI) inventory regions. 2

2 Mississippi Institute for Forest Inventory (MIFI) Dynamic Reporter interface displaying a 50-mile radius working circle in the MIFI Central Inventory Region that partially overlaps Alabama. A generated report is also displayed. 3

3 The four county study area located in central Mississippi. 20

4 The four county study area represented in the three Landsat-derived layers (left to right: age, forest cover type, and radiance values) used in estimating cubic foot volume outside bark to a pulpwood top (CFVOBPW). 23

5 Age distribution for Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region pine data from temporal image differencing change detection methods in Wilkinson (2011). 24

6 A scatterplot of volume versus age of Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region pine data from Wilkinson (2011). 24

7 Age distribution for Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region pine data from corrected change detection with Microsoft Visual Studio® 2008 Visual C++ ® program. 26

8 A scatterplot of volume versus age of Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region pine data from corrected change detection and volume/age relationships. 27

9 Eighteen counties, highlighted in yellow, within the Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region. 36

10 Scatterplot of pine stratum cubic foot volume outside bark to a pulpwood top (CFVOBPW) versus Landsat band 5 (RB5) pixel radiance values for the Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region. 42

11 Scatterplot of pine stratum cubic foot volume outside bark to a pulpwood top (CFVOBPW) versus a band combination of Landsat band 4 pixel radiance values multiplied by Landsat 5 pixel radiance values (45) for the Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region. 43

12 Scatterplot of pine stratum cubic foot volume outside bark to a pulpwood top (CFVOBPW) versus Normalized Difference Vegetation Index (NDVI) for the Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region. 43

CHAPTER I

INTRODUCTION

Background

Forestry industry in Mississippi was directly responsible for 2.42% of the state's total employment, and forestry also provided 2.43 billion dollars of value-added in 2010 (Dahal et al. 2013). The total direct, indirect, and induced output of the forest products industry in Mississippi was 10.38 billion dollars (Dahal et al. 2013). Because of the importance of forests to the economy, precise and accurate periodic estimates of timber volume by forest type are necessary for assessing sustainability of this valuable resource and for attracting forest product industries by insuring that proposed mills can maintain a continuous supply of raw material for a planned period of time. Currently, the Mississippi Institute for Forest Inventory (MIFI) stratifies the state's forest land into three GIS strata (pine, hardwood, and mixed pine/hardwood) and randomly samples inventory plots in each stratum by county with a sampling error of ±15% at a minimum of a 90% confidence level (Parker et al. 2005). Mississippi is divided into five MIFI inventory regions (Figure 1) sampled on a rotating annual basis.

Figure 1 The counties of Mississippi separated into five designated Mississippi Institute for Forest Inventory (MIFI) inventory regions.

Inventory data are accessed by the MIFI Dynamic Reporter software (Matney and Schultz 2011) to produce reports for volume/biomass/carbon/weights for areas within the boundaries of the state by county, multi-county, MIFI region, polygon, or working circles (Figure 2).

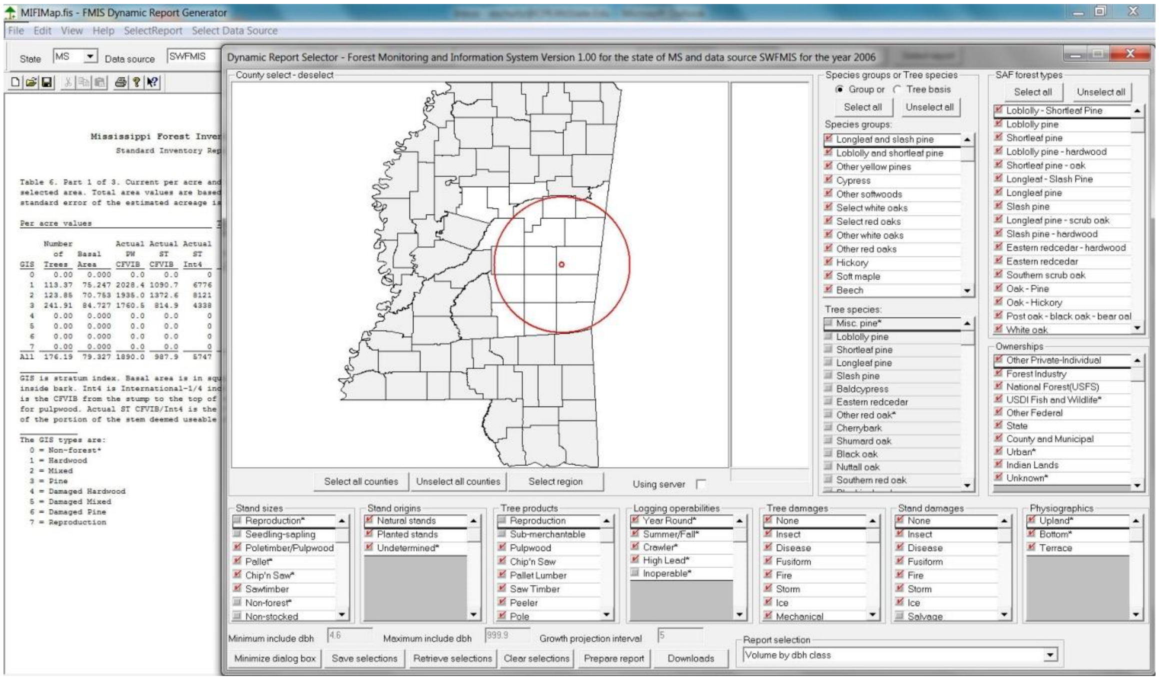

Figure 2 Mississippi Institute for Forest Inventory (MIFI) Dynamic Reporter interface displaying a 50-mile radius working circle in the MIFI Central Inventory Region that partially overlaps Alabama. A generated report is also displayed.

Areas of polygons and working circles falling outside the state's boundaries are excluded from reports. When working circles, or polygons, representing future or existing mill location areas are analyzed that cross state boundaries, there is a lack of MIFI data creating gaps in information. Lower precision and plot density US Forest Service (USFS) Forest Inventory and Analysis (FIA) data can be used, when available, for approximately matching time frames. Otherwise, the required data must be filled in by: 1) conducting expensive inventories, 2) mirroring adjacent inventory areas, or 3) utilizing remote sensing methodologies. In the event that no appropriate time-matched ground data exist, procedures such as Landsat-based forest stratification by basal area (Schultz et al. 2006) and Landsat-derived stand structure estimation (Wilkinson 2011) showed promise for completing volume estimation gaps using only remotely sensed data. This study reports the results of attempting to extend previous research and evaluates the practicality of applying remote sensing methods to expanding volume estimates into excluded areas with sufficient accuracy to supplement the accepted standard of ground truth plot data.

Justification and Objective

Forest land comprises approximately 65% of the total area of Mississippi, roughly 19 million acres (Henderson et al. 2008) and is both economically and environmentally important to the state. For these reasons, timely and precise inventories are needed to: 1) attract new forest products industries, 2) retain existing forest products industries, 3) provide forest management information, and 4) monitor forest change and sustainability. Because a comprehensive inventory is necessary to produce accurate and precise estimates, methods are needed to augment expensive and time-consuming field-based inventories when additional data are required due to data gaps. Past research (Wilkinson 2011) suggests that the use of Landsat-derived data alone has potential for providing acceptable volumes when other more reliable inventory data is not available.

This study was conducted to spatially and temporally validate the feasibility of using Landsat-derived layers (age, cover type, and pixel radiance value) to estimate cubic foot volume outside bark to a pulpwood top (CFVOBPW) of GIS-based forest strata in central Mississippi. A Landsat-derived model for estimating CFVOBPW (Wilkinson 2011) was compared with field-based inventory (in-state MIFI ground truth plot data) data to determine if Landsat-derived estimates fell within an accepted standard of ground truth data. Once an acceptable standard of using the Landsat-derived volume estimation model for augmenting field-based inventories is established, forest resource industries can assess the costs/risks and benefits of utilizing more timely and less costly information for investment decisions.

Study Progression

Based on research literature, all methods described in Chapter III were derived for projecting CFVOBPW from three layers: age using temporal image differencing change detection, forest cover type, and pixel radiance values. Early in the analysis it was apparent that radiance data were not contributing to the precision of volume estimation equations in any way. Chapter IV discusses the initial results and why radiance values did not provide additional predictive power to volume estimation models. Chapter V presents revised models for predicting volume. Any of the revised models could be reasonably certain of being successful in applying established procedures for estimating height and stand density using other existing remotely sensed technology such as lidar, multispectral, or high resolution satellite imagery. The results of revised volume prediction models from remotely sensed estimates of age, average stand total height, and stand density are discussed in Chapter VI. Recommended procedures based on study results and suggested model improvements are the basis for the discussion in Chapter VII and conclusions in Chapter VIII.

CHAPTER II

LITERATURE REVIEW

Currently, data for two forest inventory systems are being implemented in Mississippi: the United States Forest Service (USFS) Forest Inventory and Analysis (FIA) and the Mississippi Forestry Commission (MFC) Mississippi Institute for Forest Inventory (MIFI). The USFS FIA inventory system uses remotely sensed data to stratify forest/non-forest acreage and determine field plot location, while the MIFI inventory uses it to produce a stratified sampling frame (USDA FS 2011). The use of remotely sensed data in forest inventories is also beneficial for displaying the expanse of inventoried land as well as its availability. However, problems do exist in using these data, such as issues with error generation, data continuity, accuracy, and the ability to stratify images into meaningful forest inventory strata.

Forest Inventory and Analysis (FIA)

Inventory data provided by FIA is widely used in many analyses and publications on topics such as remote sensing and forest inventory (McRoberts et al. 2002), urban sprawl and production (Barlow et al. 1998), fire probabilities (Munn et al. 2003), and forest products (Abt et al. 2000). FIA's congressional mandate is to provide national forest and range resource assessments, which include present and future potential of these lands for ensuring sustainable forestry practices (Avery and Burkhart 2002). FIA defines commercial timberland, pasture or ranchland with trees, plantations, unproductive forested land, and reserves as forestland (McRoberts et al. 2002). These forestland areas must be greater than or equal to 1 acre and maintain a minimum cumulative crown width of 120 feet (McRoberts et al. 2002).

Based on a systematic grid of permanent plots, FIA continuously samples all forested lands across the nation (Avery and Burkhart 2002). The USFS systematically samples forest and non-forest at an intensity of a maximum of 20% of the plots in each state per year, based on a nationally uniform cell grid of interlocking hexagons (USDA FS 2011). One plot is selected in each hexagon that encompasses approximately 6000 acres. In each standard field plot, a national set of ecological core measurements are collected by either federal, state, or contracted personnel (USDA FS 2011). FIA plots are systematically distributed as opposed to a random allocation (Crosby 2011). Avery and Burkhart (2002) noted that, while systematic sampling provides a very user-friendly sampling frame; this sampling design can have low accuracy if there is periodic variation. Avery and Burkhart (2002) also determined that estimating precision is problematic due to the fact that the sample is not random. Statistics gained from systematic samples are theoretically invalid because sample units do not have known probabilities of selection (Cochran 1977).

The USDA Forest Service Southern Research Station (SRS) has a total of 89,087 FIA inventory plots (35,653 forested) that they monitor throughout the states of Texas, Oklahoma, Arkansas, Louisiana, Mississippi, Alabama, Kentucky, Tennessee, Georgia, South Carolina, Florida, North Carolina, Virginia, and the territory of Puerto Rico (USDA FS 2011). The last periodic inventory of Mississippi started in 2009 and currently has 28% of the current subcycle completed (SRS FIA 2014). In each cycle, FIA measures a subcycle that is 20% of the original sample. FIA samples the original plots to gauge growth and mortality and the original sample is then adjusted for that subcycle's growth and mortality. The states that border Mississippi have cycles that began in 2009 (Louisiana and Tennessee), 2010 (Arkansas), or 2012 (Alabama) (SRS FIA 2014).

FIA was established to provide state level estimates but its plot density is too low to precisely estimate forest volumes for counties or the relatively small working circle radii currently utilized by forest industry. Depending on the size of the county and percentage of forest cover, FIA sampling error on a county level basis in forested areas is approximately ± 40% (assuming: 1) a fairly low coefficient of variation of 30, 2) a 95% confidence interval, and 3) eight to ten forested plots per county per 1-year cycle as has been recorded in Mississippi). Over a typical working circle, sampling error would be increased due to the low number of plots in a relatively small inventory area. FIA plot intensity is only acceptable when large regions or areas are examined, which is the purpose of these data.

Mississippi Institute for Forest Inventory (MIFI)

The lack of timely and accurate representation of the spatial distribution of timber resources at the county level in Mississippi created a need for the MIFI inventory system. A geospatially-based statewide database, beginning in Mississippi in 2004, was created to estimate forest volumes/biomass/carbon/weights by various categories. The objectives of the MIFI inventory are to: 1) conduct a forest inventory effective at the county level (±15% at a minimum of 90% confidence level) based on remote sensing, GIS, and GPS technologies, and 2) use computer technology to create original and derived data products available with online technologies enabling users to estimate timber supplies and monitor change in forest resources (Parker et al. 2005). Early inventories were completed at the 95% confidence level, but more recent inventories have been completed at the 90% confidence level due to lowered sampling intensities. MIFI's combination of remote sensing and forest inventory field estimates are the first of its kind at this large scale (Riggs et al. 2013).

The MIFI inventory utilizes an optimized stratified simple random sampling scheme (Cochran 1977) on an annual rotating basis among five regions. The two stage MIFI sampling process involves: 1) analysis of remotely sensed land cover classification and change along with statistical validation, and 2) detailed forest field estimates for each rotating region (Riggs et al. 2013). Remotely sensed data allow for the random allocation of sampling plots within forest GIS strata (pine, hardwood, and mixed pine/hardwood) creating the sampling frame for field-based estimates. One-fifth acre plots are assigned random locations based on forest type classification imagery (Riggs et al. 2013) and stand individual tree measurements are taken within 1/5-, 1/10-, and 1/20-acre concentric plots according to product classes sawtimber/pole/veneer, pulpwood, and non-merchantable, respectively (Riggs et al. 2013). One-hundredth acre subplots are used to measure reproduction for assessing future site stocking (Riggs et al. 2013). By combining both remote sensing technology and forest field inventory estimates, MIFI not only estimates how much volume is present but where that volume is located in Mississippi (Riggs et al. 2013). Imagery-based products from this combination of remote sensing and forest field inventory estimates provide an integration of forest age, land cover classification, and forest cover and composition (pine, hardwood, mixed pine-hardwood).

Remote Sensing in Forest Inventories

The National Space and Aeronautics Administration (NASA) Landsat program has low resolution satellite imagery available from 1972 to the present. The Landsat Thematic Mapper (TM) sensor system has been operational since 1982, and its higher spectral and spatial resolution imagery supplanted imagery from the previous sensor system that had been available since 1972, Landsat Multispectral Scanner (MSS) (Verbyla 1995). Seven Landsat satellites have become operational. However, during 2011-2013, Landsat 7's ETM^+ sensor had a scan line corrector (SLC) problem. Also during this time Landsat 8 had not launched, thus leaving Landsat 5 and its TM sensor as the only reliable source of Landsat data during this period. The use of Landsat imagery keeps costs low while providing a large volume of information (Heit and Shortreid 1991) and is well-suited for large-scale forest inventory.

Remote sensing is acknowledged as an integral tool for the implementation of statistically efficient large-scale forest inventories (Wilkinson 2011). Remote sensing has been used for national forest inventories to produce estimates including forest area, volume, condition, growth, mortality, removals, trends, and health (McRoberts and Tomppo 2007). It is a preferred method for forest classification and creation of forest sampling frame strata for large-scale inventories because it is far less expensive than ground observations and measurements (Schultz et al. 2006). Remotely sensed data can significantly reduce field travel costs; particularly eliminating time wasted by traveling to sites with no forest cover or forest land use (McRoberts and Tomppo 2007).

Remotely sensed satellite imagery allows different temporal resolutions due to time-lapse sequencing, which permits tracking of land cover changes over time and derivation of forest age (Collins et al. 2005). The ability to track current and temporal changes in forest stands provides a library of information that is easily accessible via spatial databases (Wilkinson 2011). Repeated time-series datasets provide vast quantities of information to base decision making over multiple time periods (Collins et al. 2005).

Forest classification is one of the most frequent uses for satellite data (Iverson et al. 1989). Remotely sensed satellite imagery can assist forestland owners and natural resource managers track forest age, type, composition, biomass, and productivity (Iverson et al. 1989). Ecological information such as landscape change, landscape patterns related to biological or physical phenomena, and the physiological processes of forest canopies can be gathered and/or modeled (Iverson et al. 1989). Mid-IR bands have proven to be indicators of maturity in forest canopies (Wilkinson 2011).

Other researchers have used Landsat TM imagery in conjunction with large scale forest inventory, but with limited results. Bauer et al. (1994) concluded that their estimates of acreage of different forest cover types at the survey level (88 ac) was 3% less than independent FIA estimates. At the county level, agreement between their results and FIA estimates ranged between -5 and 3.9%, and these differences are the result of differences in approach to sampling scheme as well as the complexity of the landscape.

Remotely Sensed Volume Estimation

Gemmell (1995) studied the use of Landsat TM data in a mixed coniferous forest in southeast British Columbia and found that, at a scale of 0.25 ha, sampling using Landsat TM data had limited use for predicting volume due to issues of spatial gaps, density, registration errors, and background reflectance. The author found that volume estimation may vary due to spatial scale, terrain, stand homogeneity, and other site characteristics. Because of these issues, the author suggested biophysical characteristic site assessment matched to the variable of prediction prior to the study might aid in selecting the correct method of image analysis. Trotter et al. (1997) stated that their multiple least squares regression model only explained 30% of the variance in timber volume (root mean square error = 100 m^3ha^{-1}). They concluded that Landsat TM bands 3 and 4 had the highest correlation with timber volume. The relationship between Landsat TM data and timber volume was significant but weak at the pixel scale, and they concluded, by spatially averaging regression or non-parametric line fitting results, that acceptable estimates were derived at the forest-stand scale (about 40 ha) (root mean square error = 41 and 46 m^3ha^{-1}) (Trotter et al. 1997).

Tokola and Heikkila (1997) studied a 2280 ha forest stand in eastern Finland and estimated total mean volume using satellite imagery, site quality maps, and field sample plot data. They found a relative error of 14% for 50 ha areas and a relative error of 17.4% for a regular farm forest holding (30 ha) (Tokola and Heikkila 1997). They also found that only the volume of the dominant species was within the 20% relative error they were seeking for acceptable accuracy (Tokola and Heikkila 1997). Overall, they found their results to be suitable to forested areas in terms of assessing total volume, but not for stand timber management planning due to a high relative error, 55.1%, in individual species estimates (Tokola and Heikkila 1997). Makela and Pekkarinen (2004) found restricted applicability when using low resolution satellite imagery, such as Landsat, to estimate stand parameters in small stands. Landsat TM data could be successful in strategic management planning at the regional level (Makela and Pekkarinen 2004). They concluded that, due to forest fragmentation and small sizes of ownership, the methods they used would be more suitable for larger homogenous stands; and if stand level accuracy is needed, aerial photographs or high resolution imagery would provide better data (Makela and Pekkarinen 2004).

Shortcomings of Remote Sensing Related to Forest Volume Prediction

Because accuracy and precision are essential to a good forest inventory, any derived imagery used in volume prediction must be held to acceptable standards of accuracy and precision. Problems were presented in past research that must be addressed.

Probabilistic Problems of Change Detection in Long Sequences and Periods

When using any remotely sensed derived estimate that encompasses multiple time periods, the probability that error will appear greatly increases with time. Assume that the probability of correctly classifying presence/absence, or forest/non-forest, on a single pixel is ρ; and, assuming independence of errors between pixels and time, the probability that a sequence of pixels will be correct is ρ^n, with *n* being equal to the number of years in the sequence. For example, if $\rho = 0.95$ and $n = 34$, the probability of a correct change detection is $\rho^{34} = 0.17$. This probability calculation demonstrates that the error probabilities increase exponentially with time.

For example, a change detection project uses 34 years of Landsat imagery data (Landsat imagery is assessed from 1972 to 2006) to determine forest age. If each image produced has a confidence level of 95% of correctly classifying pixels, then a 0.05 probability exists of misclassification. However, if all 34 years are used to classify pixels as forest or non-forest in the determination of forest age, then the probability of making at least one error is actually 0.83 (0.83= 1- ρ^{34}= 1- 0.17). This means the probability of producing change detection error in a long sequence is almost a certainty. From one time period to the next, the 0.05 statistical confidence level can be achieved, but the addition of multiple time periods leads to a significant increase in error and an almost certain misclassification.

Collins et al. (2005) found moderate results when assessing post classification comparison techniques and temporal image differencing of Landsat data to determine forest age classes for the MIFI inventory, and they concluded that their accuracy assessment methods were less definitive than the more robust ground and photo interpretative techniques. Collins et al. (2005) also noted that future studies could benefit from using their techniques as a base for change detection in Mississippi forests.

Problems Associated with Data

Data Integrity

To construct estimates applicable to a population, representative samples with known probabilities of selection must be obtained from the target population (Cochran 1977). Problems arise in studies when data, more specifically data outliers, are repeatedly trimmed from the representative sample at one project stage to the next leading to false relationships. This type of trimming can introduce bias into the dataset because the trimmed dataset is no longer a representative sample of its target population (i.e. the sample units probabilities are not known), and there may be an appearance of variation that may not reflect what is naturally occurring in the study. The maximum number of sample data points should always be included in the estimating equations. Only extreme outliers should be trimmed, and only one trimming should occur. Additional trimming of outliers and refitting of equations beyond the first pass begins to trim valid data points from the dataset, and the resulting equation is no longer representative of the originally drawn random sample.

Data Preparation, Manipulation and Storage

Procedures must be documented in detail to preserve data integrity and continuity over the span of a research or operational project and across data managers. Miscommunication or lack of information in projects and processes based on large quantities of stored images and their derived products, like in remote sensing and spatial analysis, can cause major problems at any stage of analysis. Regular maintenance and updating of detailed and ordered records allows the smooth transition of data from one scientist to the next (within or between project time frames).

Issues of Sampling Error

When FIA resamples their plots, they take a 20% subcycle of growth and mortality from the original sample, and then re-estimate the original sample. This is a sample of a sample. With bias already existing in the data, due to a systematic sampling scheme that does not allow for random sampling, sampling from the pre-existing sample only adds more bias to results and can cause results that do not reflect the actual existing forest structure.

Registration/Georectification/Classification Issues (Effects on Accuracy of forest/non-forest)

Remotely sensed imagery passes through levels of processing before ending up as a derived image containing data. The images are corrected and rectified for atmospheric degradation, and some images may also be re-projected, subsetted, and/or geometrically corrected by re-sampling (Campbell and Wynne 2011). Ground control points (GCPs) are established to assess and statistically adjust image spatial accuracy (Campbell and Wynne 2011). From that point, the image is then classified using spatial or geostatistical processes such as clustering. Errors are introduced whenever classification is involved (Campbell and Wynne 2011). Reflectance values from the sensor are converted from digital number (DN) values to radiance values to be related from one image to another (Campbell and Wynne 2011). Remotely sensed data thus provides a snapshot of spectral values at a specific spatial and temporal scale; and any method-produced errors introduced by the researcher can affect the usefulness of derived data (Jensen 2000).

Registration between images can lead to misclassification. Boundaries among different landscape level elements can cause edge effects, which lead to misregistration in the scanner (Campbell and Wynne 2011). These errors will be systematic, not random, and spatially related when occurring on edges or in patches of the image (Campbell and Wynne 2011). Ground truth observation is the only way to truly identify these pixels and give them an accurate classification and avoid error.

Because there is no way to determine when change actually occurred between two images in the sequence, midpoint values of dates between scenes where change occurred are used to compensate for this problem (Collins et al. 2005). Assumptions (like the midpoint designation) made by the user influence overall accuracy, because these assumptions affect how the change is detected (Collins et al. 2005).

Summary

Mississippi's two available forest inventory systems, FIA and MIFI, have different sampling schemes, intensities, and applications outside state boundaries. The

FIA inventory could potentially be utilized for analyses crossing political boundaries, and the MIFI inventory, only applicable within state boundaries, provides more precise estimates for smaller working circle areas.

In an effort to replace or supplement costly field data in large scale inventories, researchers have developed models for predicting inventory variables from remotely sensed data. Models have been developed for predicting forest volumes/biomass that can be used to stratify an image into meaningful strata and improve statistical efficiency. Efficiency results from the difference created among strata means and variances that lead to lowered sampling error (Cochran 1977). Remote sensing data have been used advantageously for predicting forested acreages for regional management planning; however, they have shown limited capabilities when applied to forest inventory estimation models. Data characteristics such as landscape and terrain fragmentation, spatial gaps, and stand density have all proved problematic in the development of forest volumes/biomass models from low-resolution Landsat TM imagery. Further research is needed to more fully explore the relationship between remotely sensed data and forest volume estimation models.

The extensibility provided by a solely Landsat TM based model (Wilkinson 2011) has the potential for both 1) offsetting the large cost of ground based large-scale forest inventories, and 2) having a model that falls within acceptable accuracy standards that can also be applicable over existing political boundaries. If such a model can be both spatially and temporally validated, there could be many opportunities for its cogent application in many other geospatial projects and methods. The high sampling intensity of the MIFI inventory provides an excellent platform for further investigating the relationship between plot and remotely sensed attributes. A MIFI inventory has a sampling error of less than 7.5% on a 50-mile radius working circle at a minimum 90% confidence level.

CHAPTER III

INITIAL STUDY METHODS

Planning and Study Area

Data gained from satellite imagery can potentially reduce the costs and time associated with ground-based forest inventories; however, for satellite imagery to provide reliable forest inventory data, it must produce consistent results both spatially and temporally (from one time period to the next). The basis for this initial study was to assess the feasibility of Landsat-derived imagery for producing consistent temporal volume estimates. A study was initiated to temporally validate Landsat-derived cubic foot volume outside bark to a pulpwood top (CFVOBPW) estimates by comparing them to MIFI field-based CFVOBPW estimates in two separate time periods. Model performance of a 1999 MIFI pilot study of four counties (Choctaw, Clay, Oktibbeha, and Winston) in central Mississippi was compared to 2006 MIFI Central Inventory Region and 2007 North Inventory Region data to establish a standard on which to evaluate temporal validity for Landsat-based estimations.

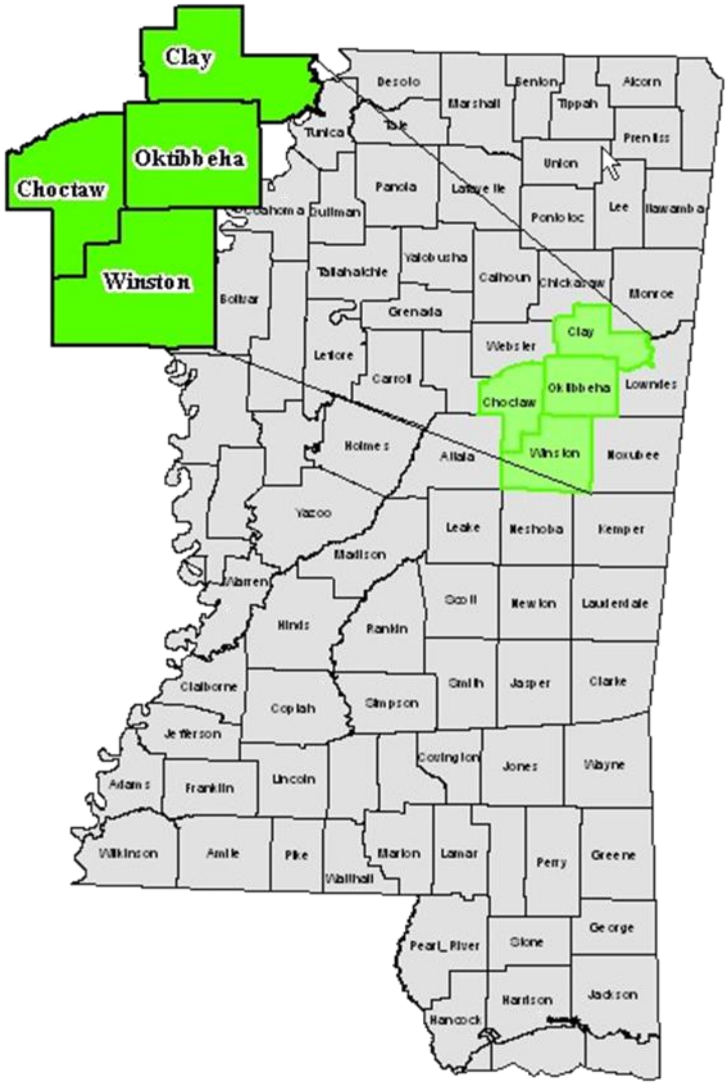

Figure 3 The four county study area located in central Mississippi.

Pixel radiance values, forest age, and forest cover data were derived from two Landsat Thematic Mapper (TM) images (1999 and 2006) that contained Choctaw, Clay, Oktibbeha, and Winston counties in central Mississippi (Figure 3). These data were then compared with field inventory data from MIFI inventories for 1999 and 2006 (Clay County 2007). These counties provide a temporal link in the data because both ground truth plot estimates and Landsat-based estimates were available for these counties in these specific years. The hypothesis of the initial phase of study was that Landsat-based estimates would fall within ± 20% of MIFI field-based ground truth estimates.

Data

MIFI Inventory Data

CFVOBPW was selected for comparing MIFI inventory and Landsat-based volume estimates because it was the MIFI inventory target/design variable and exhibited the least amount of variation among volume estimates. MIFI CFVOBPW estimates were obtained for four Mississippi counties (Choctaw, Clay, Oktibbeha, and Winston) inventoried in the 1999 pilot study (Parker et al. 2005) and re-inventoried in the 2006 MIFI Central Inventory Region (for Choctaw, Oktibbeha, and Winston counties) (Glass 2007) and the 2007 MIFI North Inventory Region (for Clay County) (Glass 2008). Choctaw, Oktibbeha, and Winston counties were placed in the Central Inventory Region while Clay County was placed in the North Inventory Region as it was inventoried in a different annual rotation year.

MIFI CFVOBPW estimates for the four counties and two time periods were obtained from Mississippi Forest Inventory Dynamic Reporter Version 7 software (Matney and Schultz 2011) available through the Mississippi State University Forest and Wildlife Research Center (FWRC) and downloadable from www.timbercruise.com (Download Center-Desktop-Other Solutions-Mississippi Institute of Forest Inventory Dynamic Inventory Reporter). Dynamic Reporter forest type and species group report selections were made using criteria similar to those employed by MIFI in determining GIS forest cover type sampling strata. The pine stratum was specified as consisting of all species groups within all Society of American Foresters (SAF) forest types that were solely composed of pine. The hardwood stratum was specified as consisting of all species groups within all SAF forest types that were hardwood only. There were too few MIFI plots (n = 15) containing mixed pine and hardwood SAF forest types for a confident comparison of Landsat-derived and MIFI volume estimates (Wilkinson 2011). A minimum merchantable GIS age of 10 was used for both pine and hardwood, and non-merchantable plots (non-forest, sub-merchantable, and regeneration) were excluded.

Landsat Thematic Mapper (TM) Data

Techniques used for deriving volume data from Landsat TM imagery were modified from Wilkinson (2011). Both 1999 and 2006 image data sets were 30-meter resolution in the Mississippi Transverse Mercator (MSTM) projection. Age, forest type, and radiance layers were derived from the imagery and subset to the four-county study area (Figure 4). These derived image layers were used as independent variables in hardwood and pine stratum models.

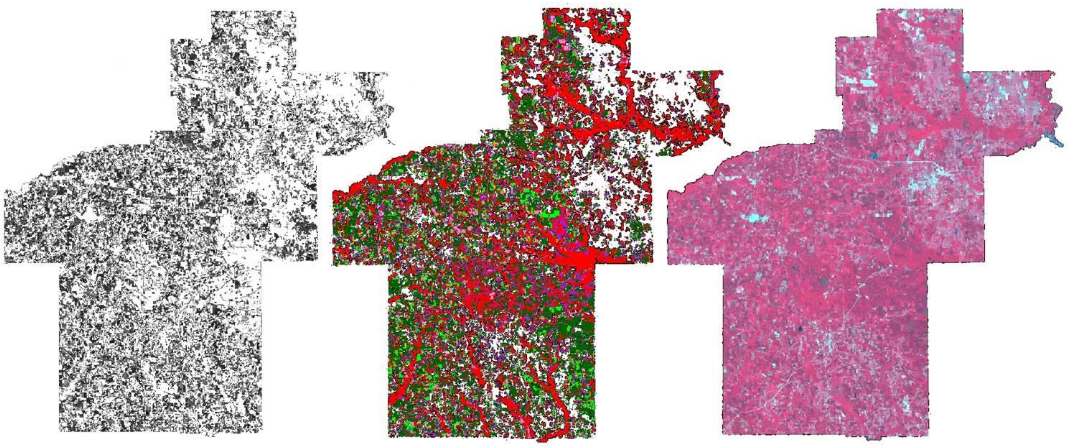

Figure 4 The four county study area represented in the three Landsat-derived layers (left to right: age, forest cover type, and radiance values) used in estimating cubic foot volume outside bark to a pulpwood top (CFVOBPW).

Forest and non-forest cover type distinctions were made using approximately 300 clusters in unsupervised classification (Collins et al. 2005). Age layers for 1999 and 2006 were constructed from a sequence of 14 times, representing every third year, from 1972 to 2006. Initially, age layers created in ERDAS Imagine® 8.7 by Wilkinson (2011) were used to determine GIS ages assigned to pixels. Examination of the age distribution from the Wilkinson (2011) image showed a disproportionate number of young ages (skewed to the left) and missing age gaps (Figure 5). Further examination of plot volume values versus assigned GIS ages showed that a larger proportion of young stands had higher volumes than would be possible (Figure 6).

Figure 5 Age distribution for Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region pine data from temporal image differencing change detection methods in Wilkinson (2011).

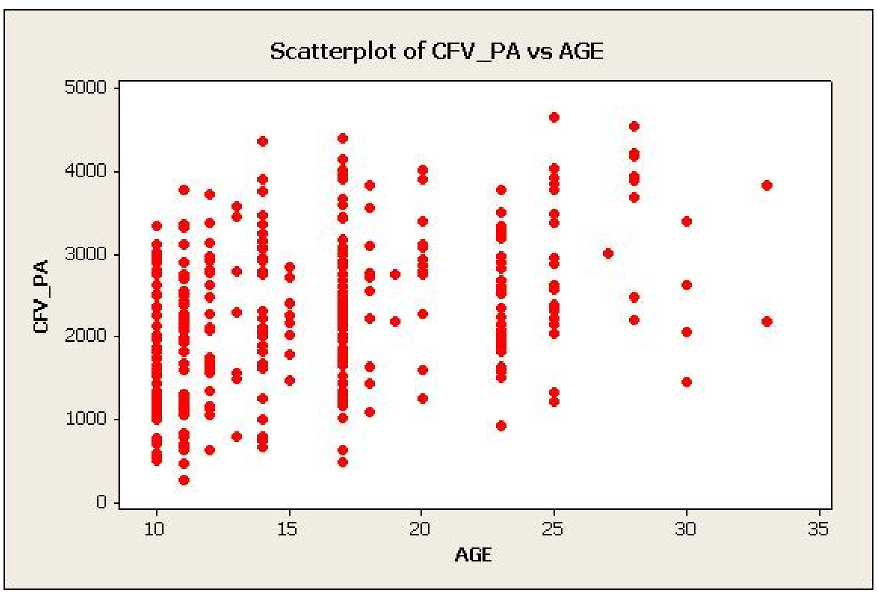

Figure 6 A scatterplot of volume versus age of Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region pine data from Wilkinson (2011).

These findings suggested that false non-forest/harvest designations were being made in the change detection system and were generating false young stand ages. It was concluded that the algorithm used to generate age layers was flawed, and not being able to determine the exact algorithm used, a new one was created with Microsoft Visual Studio® 2008 Visual C++ ® available from www.timbercruise.com/Utility/AgeLayerAlgorithm.zip.

The forest age predicting algorithm involved comparing forest/non-forest values for the same pixel over a time ordered sequence of multiple images. The ERDAS Imagine® 2010 forest/non-forest layer was converted to standard binary raster format for high speed processing. This backward searching and forward processing algorithm constructed an array of forest/non-forest values for each pixel calculated over 2- to 3-year intervals from 1972 to 2012. The image year for each forest/non-forest pixel value was stored in a separate array. Beginning at the most recent time interval, the algorithm searched backward to find the first forested pixel value that was followed in time by a non-forested pixel value. The interval in years (2 or 3 years) between forested and non-forested pixels was calculated from the separate array of image years and divided by two. Since the exact time of harvest (forest disappearance) within the time interval was unknown, dividing by two gave an average year that was designated as pixel base age 0. One year was then added to the age to avoid having plots with age 0. The algorithm then processed forward in time and added years to the base age until it reached the most recent time interval. If a non-forested condition was never found during the backward search, then age was indeterminate and, therefore, set to 40^+ years. GIS age distribution (Figure

7) and GIS age-volume comparisons (Figure 8) were improved dramatically by the new algorithm.

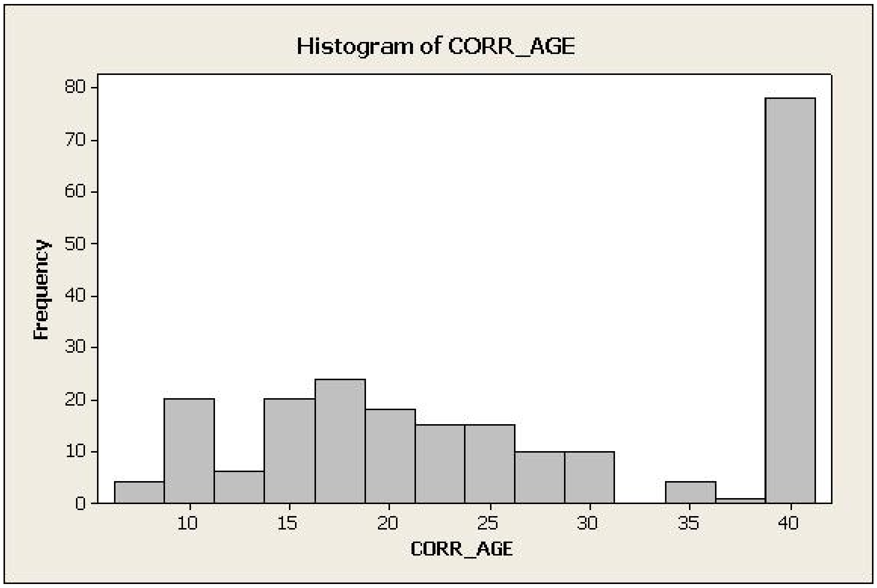

Figure 7 Age distribution for Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region pine data from corrected change detection with Microsoft Visual Studio® 2008 Visual C++ ® program.

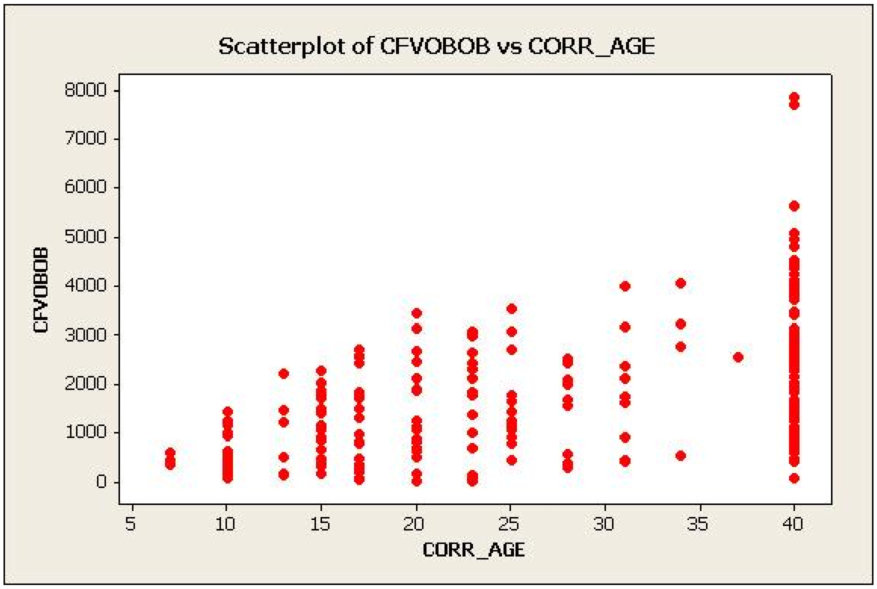

Figure 8 A scatterplot of volume versus age of Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region pine data from corrected change detection and volume/age relationships.

Hardwood and pine forest type pixel classifications were made interpretatively using an 80% or greater hardwood or pine rule (Collins et al. 2005). Any forested pixel less than 80% was separated into the mixed category. The radiance layer was generated by converting pixel reflectance values to pixel radiance values with a procedure described in Chander et al. (2007) using ERDAS Imagine's® Model Maker function.

Volume Estimates

Once Landsat-derived age, GIS forest cover type, and pixel radiance value layers were created, hardwood and pine CFVOBPW per-acre models (Equations 1 and 2) were constructed using techniques derived from Schultz et al. (2006).

$$\text{Pine CFVOBPW} = 22822.8 - 23600.6AGE^{-0.3476} - 5089.6R1^{-0.3449}R2^{0.2654}R3^{-0.0510}R4^{0.0756}R5^{0.2832}AGE^{0.1747} \quad (1)$$

$$\text{Hardwood CFVOBPW} = 316973 - 319999 AGE^{-0.00544} - 13053.8 R2^{1.6721} R3^{-4.1671} R4^{-2.4279} R5^{4.8113} AGE^{1.4272} \quad (2)$$

Where: *CFVOBPW* = estimated cubic foot volume outside bark to a pulpwood top; *AGE* = age determined from a forest/non-forest change detection temporal sequence beginning in 1972; *R1* = Landsat band 1 radiance; *R2* = Landsat band 2 radiance; *R3* = Landsat band 3 radiance; *R4* = Landsat band 4 radiance; *R5* = Landsat band 5 radiance.

The models are generalized power models with radiance and age as the independent variables and an additional power term involving age only to correct for the over/under prediction of the image model. Both power models were fitted with SAS® Version 8.0 non-linear (NLIN) procedure using the Gauss-Newton iteration with CFVOBPW per-acre as the dependent variable. The power models achieved higher index of fit (1 minus the quantity of the error sum of squares divided by the total sum of squares), I^2= 0.18 for pine and I^2= 0.16 for hardwood, than other models tested. Final models were incorporated into a Microsoft Visual Studio® 2008 Visual C++ ® program to calculate average per-acre volume for the hardwood and pine GIS forest cover types. The sample estimates from the MIFI inventory were compared to the estimates produced by applying Equation 1 and 2 on the imagery to calculate volume by GIS forest cover type of pine or hardwood.

CHAPTER IV

INITIAL STUDY RESULTS

MIFI 1999 CFVOBPW per-acre estimates were 2,607 and 2,380 for hardwood and pine cover classes, respectively, while Landsat-derived per-acre volume estimates were 3,808 and 1,978, respectively (Table 1). The Landsat hardwood estimate exceeded the hypothesized acceptable ± 20% sampling error while the Landsat pine cover class fell within the ± 20% sampling error. MIFI 2006 CFVOBPW per-acre estimates were 2,525 for hardwood and 2,064 for pine cover classes. Landsat per-acre volume estimates were 2,477 and 2,091 for hardwood and pine cover classes, respectively. Both Landsat estimates were within 2 percentage points of the MIFI estimates, well within the hypothesized ± 20% sampling error.

Table 1 Generalized power model index of fit (I^2), Mississippi Institute for Forest Inventory (MIFI) field inventory (±15% acceptable sampling range at the 95% confidence interval) cubic foot volume outside bark to a pulpwood top (CFVOBPW) per-acre estimations, Landsat-derived CFVOBPW per-acre estimations, and percent (%) difference comparisons.

Forest cover /year	I^2	MIFI inventory	Landsat-derived	% difference
		-------CFVOBPW ac^{-1}-------		
Hardwood	0.16			
1999		2607±521	3808	46.1
2006/7		2525±505	2477	1.9
Pine	0.18			
1999		2380±476	1978	16.9
2006/7		2064±413	2091	1.3

Because the hardwood cover class CFVOBPW estimation could not be temporally validated, a detailed analysis of the prediction system and formulation of new methods was necessary. Contribution of radiance band information, change detection time interval, and outlier/trimming analysis were explored as possible weaknesses in the prediction system.

Contribution of Radiance Band Information

Besides age, radiance band values were the only other model inputs considered in examining the failure of CFVOBPW temporal validation. Past research has shown positive relationships between volume estimation and band radiance values that were not achieved in this study. Power models used by Schultz et al. (2006) produced coefficients of variation (R^2) for predicting basal area from pixel radiance greater than or equal to 29.5 for all three GIS cover types. The standard error of estimate was examined, and a non-linear asymptotic z-test on the parameters was carried out to determine the degree of contribution that radiance values made in the prediction of CFVOBPW.

Inclusion of the radiance value variables in the model did not reduce the standard error of estimate. Their contribution to the model was less than 0.5% of the total variation. None of the radiance band parameters were found significantly different from zero by a non-linear asymptotic z-test. Thus, it was concluded that radiance values were not contributing to the prediction of CFVOBPW and formulation of a new model based on variables (age, height, and stand density) known to contribute to volume estimation and obtainable through additional remotely sensed technologies was deemed desirable.

Change Detection Time Interval

Because age was the most important variable in the prediction of CFVOBPW, each step in its calculation was examined beginning with the Landsat images and change detection algorithm. Errors in the GIS age layer algorithm were found and addressed early in the study resulting in the creation of a Microsoft Visual Studio® 2008 Visual C++® age calculation algorithm available from www.timbercruise.com/Utility/AgeLayerAlgorithm.zip. In the course of creating this algorithm, issues with the time interval of data used in change detection became evident.

The probability of making an error in pixel forest/non-forest classification, and thus the calculation of forest age, is exponentially increased by the number of sequential images used in the calculation $(1 - \rho^n)$, as previously discussed with ρ = the probability of correctly classifying presence/absence, or forest/non-forest, on a single pixel and, assuming independence of errors between pixels and time, the probability that a sequence of pixels will be correct is ρ^n, with n being equal to the number of years in the sequence.

However, if classification errors in the sequence could be identified using a computer algorithm, automated error correction could be possible. For example, if a series of 1's represent the forested condition and 0's represent non-forested, an annual series that contained designations like 1110111 might be construed as an error in the fourth sequence. This type of sequence could be reasonably assumed to include an error and be automatically corrected, since a forested condition using Landsat interpreted data would not be detected one year after a harvest. Based on experience, it takes approximately five to six years for a harvested forest to be recognized as new forest. However, this type of automatic correction could not be employed for this study because there were 2- to 3-year gaps between Landsat images used to construct forest age layers. This meant, if a clear-cut harvest occurred immediately following a recorded Landsat image, five years could elapse before the next forested (code of 1) image was recorded which is enough time for a new forest to be detected. The 2- to 3-year image time interval in the database occurred both due to the time/cost of classifying images for the whole state and the ability to acquire cloud-free days in the leaf-on months in Mississippi.

Change detection via temporal image differencing produced a volume result that fell outside the 90-95% confidence level targeted by MIFI. Using a multiple image change detection sequence over a 34-year time period produces a probability of 0.83 of making at least one error per pixel sequence. Assuming each age layer is independent of all other layers, an error correction analysis spanning five to six years between Landsat images that are classified on a 2- to 3-year basis is adequate time to allow harvest (forest disappearance) and regeneration detection over multiple time intervals or multiple locations. Therefore, a distinction cannot be made between a legitimate harvest/regeneration and a classification error. Because of the coarse nature of the time interval (two to three years), detection and correction of forest/non-forest classification errors was not possible. Annual intervals between image data classification should be constructed to allow the development of proper error correction routines and thus increase the precision of volume estimates.

Outlier and Trimming Analysis

During the course of the detailed analysis of the prediction system, an attempt was made to reproduce the methods and results by Wilkinson (2011) and Schultz et al. (2006). It was observed that the Wilkinson (2011) MIFI plot dataset had plot observations trimmed (outliers eliminated from the dataset) from the original data set based on age and cover type classification (337 of 1485) before model analysis occurred. Trimming appeared to result in false regression relationships and biased estimates. A large number of plots also exhibited erroneous volume/age relationships. Adding to these problems of missing or incorrect plot observations was the distribution of plot points. Wilkinson's (2011) model analysis contained 416 pine plots out of 543 total plots, which is 77% of the total number plots. The GIS cover layer only displays 49% of the forested pixels as pine, which means an inordinate number of pine plots were selected. This result was also found for hardwood and mixed plot selection with only 18% of hardwood plots (37% of total forested pixels) and 5% of mixed plots (15% of forested pixels) used in model analysis. To further substantiate this observation, a data trimming analysis was designed as part of the next study phase.

Study focus shifted from temporal and spatial validation to a refinement of data processing and integrity and the development of a revised volume estimation model to produce precise results within an acceptable range of ground truth standards. The results of the initial study showed that the CFVOBPW model driver was age and that radiance values did not improve the model significantly. Stand age, height, and density were investigated as model inputs because they are all known to contribute to volume estimation and have the potential for prediction from other forms of remotely sensed imagery such as lidar, multispectral, and high resolution imagery.

CHAPTER V

REASSESSMENT STUDY

Objectives

Initial study results revealed the failure of temporal validation of a Landsat-derived model to estimate CFVOBPW from age, cover type, and pixel radiance value layers in a four-county area in central Mississippi. Radiance band values did not contribute to the model. However, to eliminate the case that model failure could have been due to some unknown characteristic in the four-county data, the same Landsat-derived model was used to predict volume for the entire 18-county MIFI Central Inventory Region.

Secondly, formulation of a new model was conducted to establish the best possible relationship between remotely sensed image data and volume. The objective was to develop a volume prediction model from change detection derived forest age and height and trees per-acre inputs that could be derived from other remotely sensed data such as lidar, high resolution, or multi-spectral imagery.

A third objective was to examine the effects of data trimming on precision and applicability that were observed during the initial study.

Methods

Study Area

Data from the MIFI Central Inventory Region 2006 inventory (Figure 9) were used as observed/ground truth data for all model development and testing.

Figure 9 Eighteen counties, highlighted in yellow, within the Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region.

The 18 MIFI Central Inventory Region counties (Attala, Carroll, Choctaw, Clarke, Jasper, Kemper, Lauderdale, Leake, Lowndes, Montgomery, Neshoba, Newton, Noxubee, Oktibbeha, Scott, Smith, Webster, and Winston) cover 6.9 million acres, of which 4.8 million were classified as having a designated forest cover type (Wilkinson 2011). Of these 4.8 million acres classified as forest, roughly 49% are classified as pine, 37% as hardwood, and 14% as mixed cover type (Wilkinson 2011). This area contains both private and public land, with four large parcels of public land including national forest (Bienville and Tombigbee National Forests), wildlife refuge (Sam D. Hamilton Noxubee Wildlife Refuge), and state owned lands (Mississippi State University John W. Starr Memorial Forest) (Wilkinson 2011).

Data

Based on results from the initial study, all MIFI plot data with improbable GIS age-volume relationship pairs were excluded from analysis. Five hundred and eleven MIFI Central Inventory Region plots were used in three separate analyses: 1) the verification of the initial results of CFVOBPW estimation models from Landsat-derived age, cover type, and band radiance layers, 2) the development of revised CFVOBPW estimation models from inputs of GIS age and MIFI plot data (height and trees per-acre) that could be derived from other sources of remotely sensed data such as lidar, multispectral, or high resolution imagery, and 3) the effect of data trimming on precision and prediction.

Landsat 5 Thematic Mapper (TM) imagery was downloaded from the US Geological Survey (USGS) Glovis interface (http://glovis.usgs.gov) and the layers were stacked and partitioned to the study area using ERDAS Imagine® 2010. Landsat bands were obtained from a single image for path 22, row 37 with an acquisition date of May 2006. The projection and datum for this image was Mississippi Traverse Mercator (MSTM) and North American Datum 1983 (NAD83). This one image contained the entire MIFI Central Inventory Region from which three image layers were then created: forest age, forest cover, and pixel radiance value.

Imagery was processed in an identical manner to that outlined in the initial study. Ground truth plot data for individual tree estimates and forest cover data were obtained from randomly located 1/5-acre MIFI inventory plots (Riggs et al. 2013).

Landsat-derived forest age classification of the MIFI Central Inventory Region was estimated in 2- or 3-year intervals over all operational Landsat years (1972-present) using temporal image differencing similar to Collins et al. (2005) but modified in the initial study by a Microsoft Visual Studio® 2008 Visual C++ ® available from www.timbercruise.com/Utility/AgeLayerAlgorithm.zip.

Fifty-three of 511 total Landsat-derived change detected ages corresponding to MIFI ground truth plot data were error corrected. These errors were identified by comparing MIFI ground truth volumes to ages of stands with similar volumes and average diameters (Table 2). The sequence of forest/non-forest designations for MIFI plot pixels identified as errors was examined for logical consistency with ground truth volumes. Sequence errors were manually corrected.

Table 2 Guidelines for manual correction of forest ages, based on Mississippi Institute for Forest Inventory (MIFI) ground plot estimates.

Age	Volume	Expected Range
	--------------------CFVOBPW ac^{-1}	-------------------
10	1000	1000-1499
15	2000	1500-2499
20	3000	2500-3499
30	3500	3500-3999
40	4000	>4000

Verification of Initial Findings

Initial study temporal validation results showed that radiance band values did not contribute to volume estimation, and to verify this result, the scale of the study area was increased from multi-county (4 counties) to regional (18-county MIFI Central Inventory Region). Minitab®16 statistical software was used to evaluate the relationship of radiance values from individual bands and age (independent variables) with MIFI CFVOBPW (the dependent variable) and select the best performing model. Estimated volume was plotted against MIFI ground truth plot data as well as pixel radiance values in all individual bands, multiple interaction terms, and various standard indices that one might assume would have an effect on volume pixel radiance value data to display possible trends. Linear regressions were evaluated according to the combination of variables providing the best R-squared (R^2) value. Variables were eliminated from models based on lack of contribution in the regression equation.

Non-linear regression power models, similar to those used to estimate volume during temporal validation (Equations 1 and 2), were also evaluated using SAS® NLIN.

Revised Models

Revised models for predicting volume were based on the inputs of MIFI ground truth plot observations of height and stand density (trees per-acre) together with Landsat-derived age. These variables and their interactions are known to have the strongest relationship to volume estimation as has been reported in growth and yield study literature (Avery and Burkhart 2002, Burkhart and Tome 2012, and Spurr 1952). Stand height, trees per-acre, and age can be estimated consistently from remotely sensed data (Evans et al. 2006, McCombs et al. 2003). MIFI input data were used for model construction but could be replaced in future studies with the same variables derived from remotely sensed data calibrated with ground truth values from plot data (McCombs et al. 2003).

Landsat-derived age and MIFI volume data were sorted to evaluate data consistency in volume/age relationships and highlight possible misclassifications. Plots with uncorrectable forest/non-forest age sequencing were eliminated from the study (7% of total plots). Minitab®16 statistical software was used to create and evaluate regression models for estimating volume using combinations of stand age, height, trees per-acre, and their interactions.

Trimming Analysis

Results of previous studies utilizing Landsat imagery data in forest stand volume estimation were favorable (Wilkinson 2011, Schultz et al. 2006), but were not consistent with the initial findings of this study. A detailed analysis of why initial findings did not substantiate previous results, led to examination of the data that had been trimmed. Trimming appeared to have affected the resulting regression models.

A regression fitted to a dataset, returns the mean values of the independent variables; i.e. if every plot in the dataset were predicted, the original dataset plot means would be returned. If plot values are eliminated at random from the dataset, there should be little effect on the mean; however, if plots values are eliminated based on any non-random procedure, the means will be affected. If enough data points are trimmed because they are outliers (points with undue influence), then predicted plot means will be shifted away from the unbiased estimate. To examine to what degree regression estimations are affected by trimming outlier points to different levels, Minitab® regression outlier analysis was carried out in a repetitive regression-trimming process. Minitab® flags outliers during regression analysis based on a combination of leverage, Cook's D, and DFITS tests (Minitab, Inc. 2009). After each regression analysis, any data points identified as outliers were eliminated from the dataset and the regression was repeated until no more outliers were identified. Number of trimmed points, R-squared (R^2), root mean square error (RMSE), average CFVOBPW per-acre ($\bar{Y}$), percent difference in average CFVOBPW, and standard deviation were recorded after each regression in the iteration. Separate analyses were conducted for each of the three GIS strata.

CHAPTER VI

REASSESSMENT STUDY RESULTS

Verification Results

No relationships between linear regression Landsat band pixel radiance value inputs (or combinations of band pixel radiance or vegetative indexes based off band pixel radiance input variables) and CFVOBPW were apparent from data scatterplots (Figures 10, 11, and 12). Initial 4-county study results were verified in finding no added value in including radiance band variables in regressions predicting volume.

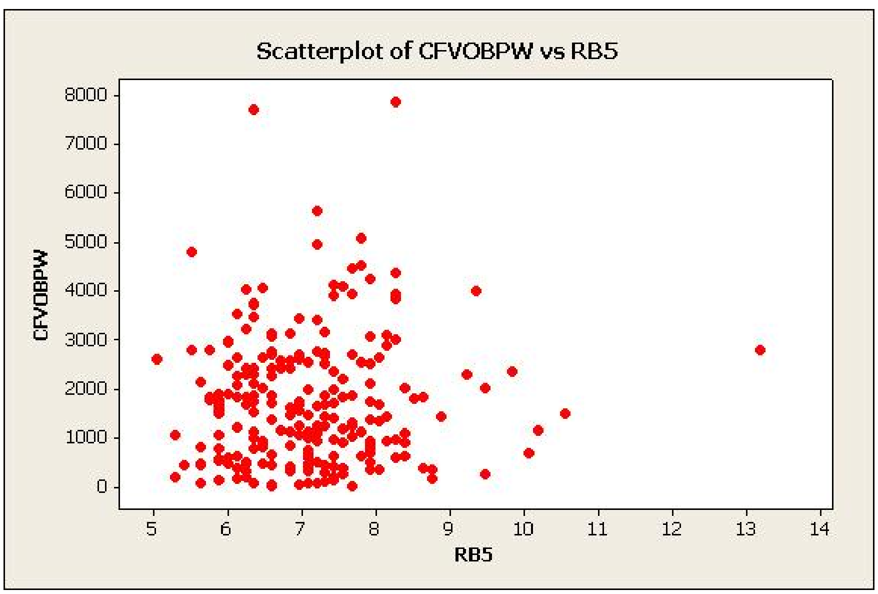

Figure 10 Scatterplot of pine stratum cubic foot volume outside bark to a pulpwood top (CFVOBPW) versus Landsat band 5 (RB5) pixel radiance values for the Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region.

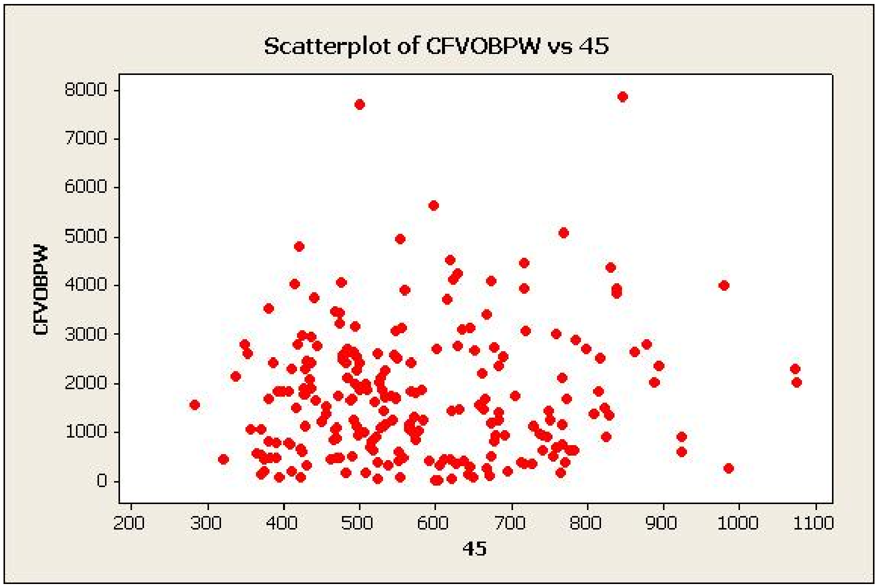

Figure 11 Scatterplot of pine stratum cubic foot volume outside bark to a pulpwood top (CFVOBPW) versus a band combination of Landsat band 4 pixel radiance values multiplied by Landsat 5 pixel radiance values (45) for the Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region.

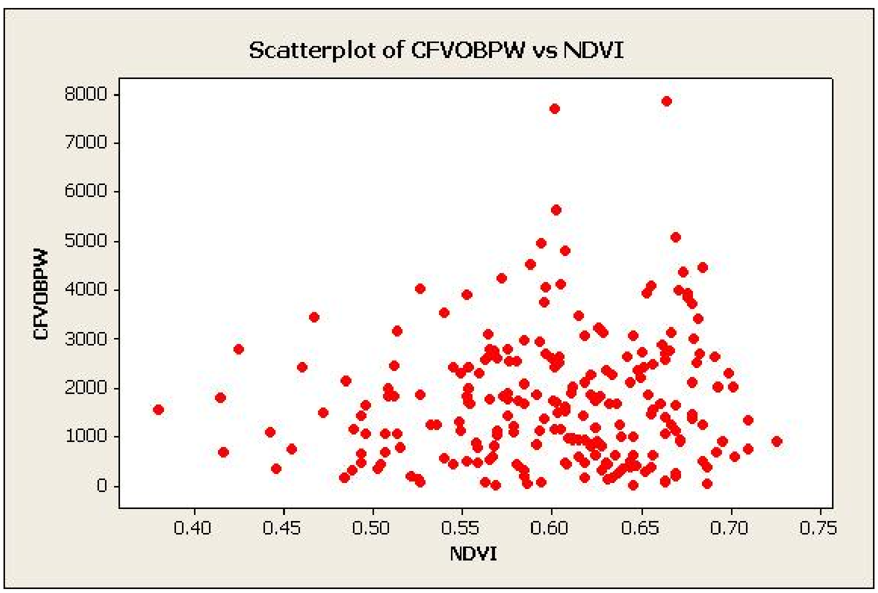

Figure 12 Scatterplot of pine stratum cubic foot volume outside bark to a pulpwood top (CFVOBPW) versus Normalized Difference Vegetation Index (NDVI) for the Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region.

Because Landsat pixel radiance value individual band layers, interactions, or standard indices displayed no significant influence on volume estimation models, attention was turned to developing models with alternative input variables known to have strong relationships to volume and that could also be replaced in future studies with data derived from remote sensing technologies.

Revised Regression Model Results

Revised models for predicting volume were based on the inputs of MIFI ground truth plot observations of height and stand density (trees per-acre) together with Landsat-derived age. The best linear regression equation (Equation 3) for CFVOBPW estimation of the pine cover type was a combination of height, age, and stand density (trees per-acre) which produced an R-squared value of 58.5% and a standard error of prediction of 0.698.

$$LNV = -5.52 + 2.39LNH + 0.551LNA + 0.428LNTPA \tag{3}$$

Where: *LNV* = the natural log of estimated cubic foot volume outside bark to a pulpwood top, *LNH* = the natural log of the total average height of the plot, *LNA* = the natural log of the age of the plot, and *LNTPA* = the natural log of trees per-acre per plot.

This regression model explained a majority (over one-half) of the variation in the data and contains all three variables identified as important in growth and yield study literature (Avery and Burkhart 2002, Burkhart and Tome 2012, and Spurr 1952). A second linear regression model (Equation 4), with height and age as input variables, also produced a significant result where R-squared explained over one-half of the variation in the inference population (51.7%), with a standard error of prediction of 0.751, but less than the 3-variable model.

$$LNV = -1.49 + 1.91LNH + 0.627LNA \tag{4}$$

Where: *LNV* = the natural log of estimated cubic foot volume outside bark to a pulpwood top, *LNH* = the natural log of the total average height of the plot and *LNA* = the natural log of the age of the plot.

In an additional attempt to utilize radiance band information in predicting volume, a model with inputs consisting of MIFI height, GIS age, and Landsat imagery Band 5 pixel radiance values was developed for the pine cover type. This model explained over one-half of the variation in volume (52.2%); but Band 5 radiance data was not statistically significant at the 0.05 level. This was the only regression model with a variable involving pixel radiance values that showed any potential for contributing to volume estimation.

Radiance bands also did not contribute to hardwood cover type volume estimation. The same combinations of MIFI and GIS age variables were significant in the hardwood regression models (Equations 5 and 6) as for the pine cover type.

$$LNV = -2.40 + 1.83LNH + 0.365LNA + 0.350LNTPA \tag{5}$$

$$LNV = 0.957 + 1.42LNH + 0.415LNA \tag{6}$$

These regression volume estimation models produced lower precision (R-squared) than with the pine models (42.8% for Equation 5 and 35.9% for Equation 6), but produced lower standard errors of prediction of 0.665 and 0.702 respectively. This result can most likely be attributed to greater variation in the size and shape of individual hardwood trees compared to pine trees.

Results for the mixed cover type were comparable to the hardwood cover type. Linear regression volume estimation models produced lower R-squared values (41.1% for Equation 7 and 32.1% for Equation 8) than those for the pine cover type but were very similar to the hardwood cover type.

$$LNV = -4.34 + 1.82 LNH + 0.740 LNA + 0.449 LNTPA \quad (7)$$

$$LNV = -0.07 + 1.47 LNH + 0.648 LNA \quad (8)$$

The mixed cover type was expected to produce the lowest R-squared as a consequence of methods employed for grouping pixels (Collins et al. 2005) and because of the smaller number of observed plots. The standard error of prediction was highest in the mixed cover type (0.795 for Equation 7 and 0.847 for Equation 8). The mixed cover type contained all forested pixels that did not fit the pine or hardwood cover type requirements, thus leaving an assortment of pixels with higher variation.

Data Trimming Results

The linear regression volume estimation model with input variables height and age (Equations 4, 6, and 8) was selected to examine gains in precision when outlier plots are removed from the inference population.

Minitab® identified a total of 38 outlier plots from 226 total pine plots, or 17% of the data after 5 regression-trimming iterations. Iterations continued until no outlier plots remained, and the final linear regression model (Equation 9) produced a significant improvement in precision (R^2=65.6%) over the regression with untrimmed data (R^2=51.7) and a significant increase in CFVOBPW per-acre estimates after 5 iterations (11.9% increase) (Table 3).

$$LNV = 0.452 + 1.45LNH + 0.582LNA \quad (9)$$

Table 3 Data trimming iterations for Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region pine cover type data (n= 226) displaying precision changes in R^2, root mean square error (RMSE), estimated cubic foot volume outside bark to a pulpwood top per-acre ($\bar{Y}$), percent change in $\bar{Y}$ (% $\bar{Y}$), standard deviation (StDev), and removed and total plots.

Iteration	R^2	RMSE	$\bar{Y}$	% $\bar{Y}$	StDev	Removed	Total
1	51.7	0.751	1732	0.0	1328	10	216
2	60.3	0.570	1767	2.0	1253	13	203
3	61.6	0.486	1847	6.7	1241	7	196
4	63.1	0.447	1888	9.0	1235	7	189
5	65.6	0.416	1937	11.9	1228	0	189
						Total=37	

A total of 219 hardwood cover type plots were available for analysis, of which a total of 50 were identified as outliers from iterative Minitab® outlier analysis, eliminating 23% of the original data. Regression Equation 10 resulted from the final iteration.

$$LNV = 1.76 + 1.38LNH + 0.265LNA \quad (10)$$

Trimming of outlier plot data eliminated almost one-quarter of the hardwood cover type plots from the total study area, which resulted in a significant precision increase (19.7%) in R-squared from 35.9% to 55.6% as well as a significant increase in CFVOBPW per-acre (Table 4).

Table 4 Data trimming iterations for Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region hardwood cover type data (n= 219) displaying precision changes in R^2, root mean square error (RMSE), estimated cubic foot volume outside bark to a pulpwood top per-acre ($\bar{Y}$), percent change in $\bar{Y}$ (% $\bar{Y}$), standard deviation (StDev), and removed and total plots.

Iteration	R^2	RMSE	$\bar{Y}$	% $\bar{Y}$	StDev	Removed	Total
1	35.9	0.702	2191	0.0	1551	14	205
2	44.2	0.545	2242	2.3	1467	15	190
3	49.9	0.462	2333	6.5	1427	9	181
4	54.2	0.412	2328	6.3	1346	5	176
5	56.2	0.392	2324	6.1	1324	2	174
6	55.6	0.390	2343	6.9	1319	0	174
						Total=45	

The mixed cover type had the lowest number of initial plots (66) and was expected to show a significant increase in precision when outlier plot data were removed. Out of 66 plots, Minitab® flagged 12 as outliers or 19% of the total. The final regression model (Equation 11) produced a statistically significant increase in R-squared (25.9%), from 32.1% to 58%.

$$LNV = 1.25 + 0.819LNH + 1.04LNA \tag{11}$$

The regression equation developed from the trimmed dataset explains almost twice as much volume variation as from the untrimmed dataset. The mixed cover type produced the most significant gains in precision and percent difference in CFVOBPW per-acre estimates (Table 5).

Table 5 Data trimming iterations for Mississippi Institute for Forest Inventory (MIFI) Central Inventory Region mixed cover type data (n= 66) displaying precision changes in R^2, root mean square error (RMSE), estimated cubic foot volume outside bark to a pulpwood top per-acre ($\bar{Y}$), percent change in $\bar{Y}$ (% $\bar{Y}$), standard deviation (StDev), and removed and total plots.

Iteration	R^2	RMSE	$\bar{Y}$	% $\bar{Y}$	StDev	Removed	Total
1	32.1	0.847	2020	0.0	1393	5	61
2	46.5	0.662	2163	7.1	1351	4	57
3	54.0	0.537	2294	13.6	1300	1	56
4	52.1	0.518	2330	15.3	1282	2	54
5	58.0	0.481	2388	18.2	1269	0	54
						Total=12	

CHAPTER VII

DISCUSSION

Landsat-derived models based on age and pixel radiance value layers failed to estimate hardwood and pine CFVOBPW per-acre with sufficient precision ($I^2 = 0.16$ for hardwood and $I^2 = 0.18$ for pine) to discriminate volume differences among pixels during temporal validation. Pixel radiance values did not contribute to the model. Because other researchers reported successful use of models that included pixel radiance values in estimating per-acre volumes, negative results obtained for a MIFI 4-county area of Mississippi were verified for a much larger 18-county Central Inventory Region area. Detailed analyses were conducted on the calculation of age (the only other model input besides radiance values) that included the development of a Microsoft Visual Studio® 2008 Visual C++ ® change detection algorithm as opposed to the algorithm used by Collins et al. (2005), Schultz et al. (2006), and Wilkinson (2011). A data outlier/trimming study found significant differences in volume estimation per-acre. A CFVOBPW estimation model was constructed from GIS age and MIFI plot variables that could conceivably be estimated from remotely sensed technologies such as lidar with sufficient precision to substitute for missing data in a MIFI inventory analysis.

Failure of Radiance Values to Predict Volume

Inputs to initial study models (Equations 1 and 2) that failed to predict CFVOBPW were composed of Landsat radiance bands 1 – 5 and change detection derived age. None of the six individual band pixel radiance value layers, interactions, or standard indices contributed to the prediction of volume. Band 4 was expected to exhibit the most influence because it has been reported as a predictor of vegetative leaf cover (Campbell and Wynne 2011) and possibly related to tree size. Band 5 has been reported as an indicator of vegetative moisture content (Campbell and Wynne 2011) and could, perhaps, be a relative measure of tree vigor. However, the reflectance of energy from trees crowns as recorded by the Landsat TM sensor appears unrelated to per-acre tree volume (Figures 10 – 11). The case might be made that derived measures of crown size could contribute to volume estimation on an individual tree basis, but there was no expected link between these radiance band values and crown size or volume on a per-acre basis.

Change Detection and Forest Age Determination

Identification of age classification problems proved to be challenging. Issues associated with length of the interval between Landsat images, the change detection algorithm itself, and outlier trimmings were all investigated. Other possible problems such as errors caused by data clumping procedures (kernel, kriging or IDW, neighborhood functions) that can mask non-existent trends in pixel radiance values were avoided in this study.

Calculation of the probability of classification error accumulated over a 34-year change detection period demonstrated that forest age estimation could be a significant source of CFVOBPW model error. The change detection process could be improved by annual Landsat imagery forest/non-forest classification as opposed to 2- to 3-year periods. Two- to three-year periods left data gaps too large to properly identify errors that could be automatically corrected. If an error were detected for a system that only classified age every two to three years, one to five years could elapse between when an actual harvest occurred and when it was detected. At that point, no determination could be made as to whether it was a legitimate error or a harvest. Five years is sufficient time for a harvested area to be detected as regeneration. The long intervals between change detection evaluations prevent the development of proper error correction routines that could greatly increase the precision of volume estimates.

It is often difficult to discover the exact methods of routines that are wholly or partially constructed in previous research or from within utilities contained in software packages such as ERDAS Imagine®. Datasets derived from these sources should always be analyzed for statistical and biological consistency before adaptation. Age distributions from previous research could not be validated by known MIFI age distributions and age-volume relationships (Figures 5 – 8); therefore, an alternative algorithm was constructed whose outputs are consistent with ground truth.

When a problem arises, the tendency is to remove plots composed of young stands with high volume estimates. Erroneous age-volume relationships are apparent in young stands but cannot be delineated in older stands. Older stands could legitimately have high or low volumes and, without an error correction routine or ground truth plot inventory data, there is no appropriate basis for separating bad estimates from good ones.

Incorrectly removing stands with high volume can cause considerable bias in volume estimates.

Outlier Trimming Analysis

Manipulation of data by trimming outlier plots showed statistically significant increases in R-squared precision as well as significant increases in average volume per-acre. It appeared that past researchers could have trimmed data sufficiently to produce artificial relationships between radiance values and volume. Excluding data points also created regression estimations that at the means of the independent variables do not predict the average plot volumes and are, therefore, biased predictors of volume for the derived image. This result was confirmed by the trend in significant gains in precision observed by trimming data to different levels for all GIS cover types. Problems in the determination of independent variables must be addressed rather than trimming outlier plots. The trimming analysis in Tables 3, 4, and 5 reveals the extent to which results can be affected once outliers are removed. The number of outliers, and thus the need for trimming outliers, may be greatly reduced once issues in calculating age from forest/non-forest change detection methods are resolved.

Volume Estimation Model

Because models (Equations 1 and 2) based on radiance values failed to estimate volume, an approach was taken to build a model based on proven predictor variables identified from growth and yield studies that were also estimable from other remotely sensed technologies. Successful volume models were constructed for pine (R^2 = 58.5%), hardwood (R^2 = 42.8%), and mixed cover types (R^2 = 41.1%) using age, height, and trees per-acre as inputs. The linear regression models explained over half the variation in the pine data while only eliminating a minimum number of plots that were not correctable (7%). Using the maximum number of data points possible is essential for forest inventories and their associated models because the questions being answered by an inventory are “what is out there”, “how much of it is out there”, and “where is it” (Riggs et al. 2013). In order to obtain an unbiased estimate, the diversity of sites, stand densities, and species for the entire area under consideration must be fully represented in the data and used to allocate field inventories or create prediction models that meet stated precision goals.

The volume prediction equations (Equations 3, 5, and 7) based on age, height, and trees per-acre derived from remotely sensed data can serve as the basis for future research involving technologies such as lidar, multispectral, or high resolution imagery.

Remote Sensing Technologies

Much unexplained variance exists in remotely sensed imagery. The coarse resolution of Landsat TM data (30-meter) could lead to classification error and problems with precision (Gemmell 1995, Makela and Pekkarinen 2004). An issue with resolution arises with natural stand-level clumping of trees. At 30-meter resolution, the scale is beyond the individual canopy and the pixels capture groups of trees. Because in nature trees tend to clump together, the resolution of the Landsat imagery may be picking up pockets of hardwood or pockets of pine, but not sufficiently capturing or representing a mixed cover type that exists in the forest. This stand-level clumping causes an underestimation of the mixed cover type component, which may be more representative of certain areas than a pine or hardwood cover type designation.

The level of aggregation of pixel reflectance at the Landsat 30-meter resolution could cause a lack of relationship in the model. The resolution also becomes problematic when trying to compare open stands versus dense, closed canopy stands. The same volume could exist in both stands, but the pixel radiance values could be vastly different. These resolution scale issues lead to problems in precision and classification accuracy.

Multiple image sequencing provides its own set of unique problems such as the variation interjected by changes in atmospheric conditions when a satellite sensor takes images of the same area at two different time periods. Weather, scanner malfunctions, atmospheric influence, misclassification, and algorithm issues are examples of issues involved in image classification. High resolution imagery, such as 1-meter National Agriculture Imaging Program (NAIP), provides a more detailed (1-meter as opposed to 30-meter resolution of Landsat imagery) view of the landscape and individual attributes in images but with greater variability due to multiple days of acquisition. The use of this imagery could possibly help clarify the land cover under consideration, but higher resolution also brings its own level of complexity. An increase in resolution could lead to more misclassification of pixels due to spectral mixing and place more pixels in the mixed forest cover type (ex. pixel size being smaller than some hardwood canopies thus putting an individual tree in multiple pixels). Object-based classification would need to be conducted to identify the individual attributes of the image correctly. The advantages of high resolution imagery must be weighed against the disadvantages of increased data storage requirements, processing time, and greater processing capabilities.

Combining lidar data, which provides a three dimensional view of the structure of trees in a stand or plot, with improved Landsat-derived age layers and MIFI ground truth plot data could possibly provide a solution to accuracy and precision problems. With height, age, and stand density (trees per-acre) already explaining more than half of the variance, more precise height and stand density measurements gained from high accuracy lidar data can aid in model refinement. Lidar data has been used in estimation of individual tree and average stand height (Evans et al. 2006). Additional research into understanding shadowing from canopies, and the effect of only receiving partial reflectance from the lower parts of the stem (Campbell and Wynne 2011) could lead to more precise estimates of three dimensional volumes. Much of Mississippi, including the entire MIFI Central Inventory Region, currently has some form of lidar data available; however, most existing lidar data are not at the resolution necessary for individual tree measurements. Individual stem density combined with stand height (derived from elevation of ground and canopy) could provide reliable volume estimates over large areas (Evans et al. 2006). McCombs et al. (2003) reported that density can be estimated using lidar in combination with multispectral imagery for individual stem measurements. Lidar information alone has not been shown to provide precise volume estimates, but local ground truth plots, or similar data available through MIFI, can be used to calibrate data for better precision than field estimates alone (Evans et al. 2006). Parker and Evans (2009) also discussed some disadvantages and inherent bias in lidar data. Some bias is introduced because of the inability of the lidar sensor to consistently hit the terminal leader, which is more problematic in hardwoods than in conifers. In addition, midstory and understory trees may not register when hidden beneath a dominant canopy causing tree count bias.

The two primary considerations prohibiting lidar use in current large-scale forest inventory are cost and data density. Hummel et al. (2011) compared traditional stand inventories with lidar-based inventory at the stand level (30,000 ac), and they found that the cost and accuracy was not significantly different between the two inventory methods. Lidar-based inventory was close to three dollars per-acre at the stand level, but this number could be reduced to a more feasible estimate because there is an inverse relationship between number of acres and cost to acquire and process lidar data (Hummel et al. 2011). To provide timely and precise inventory estimates, lidar data would need to be refreshed every 5 years. This makes lidar-based forest inventory cost prohibitive until costs of acquiring and processing lidar can be reduced. Data density could become a problem with lidar-based inventory due to the large amount of data storage necessary for lidar returns. Hummel et al. (2011) reported a mean density of ground returns of 1.44 points/m^2, which is sufficient for western conifers but would be a higher density in southeastern forests. This larger density of ground returns translates to an extremely large amount of data when multiplied over a large-scale forest inventory in the southeastern United States. Appropriate storage space would also be required for archiving multiple years of large-scale forest inventory data. If costs can be reduced and proper storage and processing capabilities made available, lidar-based forest inventory could provide a remote sensing-based system for augmenting missing ground inventory data with sufficient precision.

Remotely sensed height and trees per-acre estimates derived from lidar data combined with improved algorithms for Landsat-derived ages that are calibrated and verified with existing MIFI ground truth data have the potential for significant gains in the precision of forest volume prediction.

CHAPTER VIII

CONCLUSION

The Mississippi Institute for Forest Inventory (MIFI) combines the use of remotely sensed imagery and field inventory plot measurements to provide the state with accurate, precise, and up-to-date forest inventory volume estimates. The state and its forest product industry investors can obtain detailed breakdowns of available volume and where that volume is located. USFS Forest Inventory and Analysis (FIA) inventory data are also available in Mississippi, but these data are sparse and contain potential bias because they are obtained from a systematic sampling scheme and not selected through a random sample of known probability (Cochran 1977). Remote sensing-based volume estimation would be desirable to supplement MIFI data when proposed working circles may overlap bordering state boundaries leaving a gap of inventory information or where there are missing ground inventory data. The feasibility of obtaining acceptable cubic foot volume estimates from a Landsat-derived volume estimation model (Wilkinson 2011) was assessed.

Landsat-derived volume estimation models proposed by Wilkinson (2011) could not be temporally validated. The relationship between band radiance values and volume estimation was not significant, contrary to results from previous studies. The failure of radiance values to make a significant contribution to volume estimation prompted further research into other possible methods of precisely estimating volume from remotely sensed data. It was shown that, if growth and yield model independent variable requirements were used as inputs, 58% or more of volume variation in the inference population could be explained. Revised linear regression models included remotely sensed independent variable of age and MIFI ground truth plot independent variables of height and trees per-acre, and these independent variables were found sufficient to produce variation of volume across the image. Results supported abandoning development of volumetric relationships with radiance values and suggest focusing on models based on age obtained from refined change detection methods, height from lidar imagery, and density measures from a combination of lidar and multispectral or high resolution imagery. Future research should focus on developing low cost and practical methods of obtaining accurate and precise values of these variables from remotely sensed imagery.

Exploration into the various sources of independent variable error in the models pointed to problems with temporal image differencing change detection used to calculate forest age, issues with misclassification of GIS cover type, and the lack of accounting for stand density and height in previous models. It was shown that the probability of generating errors in a forest/non-forest change detection sequence is exponentially increased by the number of sequential images used in the calculation. Also, an analysis of the time interval between images in change detection demonstrated that annual intervals between image data classification must exist to allow development of proper error correction routines and thus increase the precision of volume estimates.

The dramatic change in estimated CFVOBPW per-acre volume observed in the analysis of data trimming of randomly sampled points demonstrated that trimming should be avoided. A modest amount of data trimming resulted in unacceptably biased estimates. When a regression is developed from a random sample, instead of trimming observations with bad independent variables, an effort should be made to readdress and correct the bad independent variables.

These results were used to develop a procedure that would produce sufficient precision for volume estimation where field inventory data are sparse or non-existent. Further research is required to determine if this procedure can be used with existing technologies on an image in two successive time periods to determine relative growth.

REFERENCES

Abt, R.C., F.W. Cubbage, and G. Pacheco. 2000. Southern forest resource assessment using the subregional timber supply (SRTS) model. Forest Products Journal, 50(4): 25-33.

Avery, T.E. and H.E. Burkhart. 2002. Forest Measurements, 5th edition. McGraw Hill, Boston.456 pp.

Barlow, S.A., I.A. Munn, D.A. Cleaves, and D.L. Evans. 1998. The effect of urban sprawl on timber harvesting: A look at two southern states. Journal of Forestry, 96(12): 10-14.

Bauer, M.E., T.E. Burk, A.R. Ek, P.R. Coppin, S.D. Lime, T.A. Walsh, D.K. Walters, W. Befort, and D.F. Heinzen. 1994. Satellite Inventory of Minnesota Forest Resources. Photogrammetric Engineering and Remote Sensing, 60(3): 287-298.

Burkhart, H.E. and M. Tome. 2012. Modeling Forest Trees and Stands. Springer, New York. 457 pp.

Campbell, J.B. and R.H. Wynne. 2011. Introduction to Remote Sensing, Fifth Edition. The Guildford Press, New York. 667 pp.

Chander, G., B.L. Markham, and J.A. Barsi. 2007. Revised Landsat-5 Thematic Mapper radiometric calibration. IEEE Geoscience and Remote Sensing Letters. 4(3): 490-494.

Cochran, W.G. 1977. Sampling Techniques, Third Edition. Wiley and Sons, New York. 428 pp.

Collins, C.A., D.W. Wilkinson, and D.L. Evans. 2005. Multi-temporal analysis of Landsat data to determine forest age classes for the Mississippi statewide forest inventory - preliminary results. P. 10-14 in Proceedings of the 3rd international workshop on the analysis of multi-temporal remote sensing images, Biloxi, MS, May 16-18, 2005. IEEE. [On CD-ROM].

Crosby, M.K. 2011. Consequences of GIS Classification Errors on bias and Variance of Forest Inventory Estimates: Effects on a Mill Location. Published dissertation, Mississippi State University. 50 pp.

Dahal, R.P., I.A. Munn, and J.E. Henderson. 2013. Forestry in Mississippi: the impact of the industry on the Mississippi economy—an input-output analysis. Res. Bull. FO 438. Mississippi State, MS: Forest and Wildlife Research Center. 22 p.

Evans, D.L, S.D. Roberts, and R.C. Parker. 2006. LiDAR – A new tool for forest measurements? The Forestry Chronicle. 82(2): 211-218.

Gemmell, F.M. 1995. Effects of forest cover, terrain, and scale on timber volume estimation with thematic mapper data in a Rocky Mountain site. Remote Sensing of Environment, 51: 291-305.

Glass, P. 2007. State of Mississippi central district forest inventory. Jackson, MS: Mississippi Institute for Forest Inventory. 24 p. http://www.mifi.ms.gov/documents/MIFI_Central_Report.pdf.

Glass, P. 2008. State of Mississippi north district forest inventory. Jackson, MS: Mississippi Institute for Forest Inventory. 24 p. http://www.mifi.ms.gov/documents/MIFI_North_Report.pdf.

Heit, M. and A. Shortreid (Eds.). 1991. GIS Applications in Natural Resources. GIS World, Fort Collins. 381 pp.

Henderson, J.E., I.A. Munn, G. Perez-Verdin, and D.L. Grebner. 2008. Forestry in Mississippi: the impact of the forest products industry on the post-Katrina Mississippi economy-an input-output analysis. Forest and Wildlife Research Center, Research Bulletin FO 374, Mississippi State University. 31 pp.

Hummel, S., A.T. Hudak, E.H. Uebler, M.J. Falkowski, and K.A. Megown. 2011. A comparison of accuracy and cost of lidar versus stand exam data for landscape management on the Malheur National Forest. Journal of Forestry, 109: 267-273.

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

Jensen, J.R. 2000. Remote Sensing of the Environment, An Earth Resource Perspective. Prentice Hall, Upper Saddle River. 544 pp.

Jones, T.L., E.B. Schultz, T.G. Matney, D.L. Grebner, D.L. Evans, C.A. Collins, and P. Glass. 2010. A forest product/bioenergy mill location and decision support system based on county level forest inventory and geo-spatial information. Chapter 7. In: Gan, J., S. Grado, and I. Munn, eds., Global Change and Forestry: Economic and Policy Impacts and Responses. Nova Science, NY. pp. 119-133.

Makela, H. and A. Pekkarinen. 2004. Estimation of forest stand volumes by Landsat TM imagery and stand-level field-inventory data. Forest Ecology and Management, 196: 245-255.

Matney, T.G., and E.B. Schultz. 2011. Mississippi Forest Inventory Dynamic Reporter Version 7. Mississippi State, MS: Forest and Wildlife Research Center. http://www.mifi.ms.gov. [Date accessed: January 7, 2013]

McCauley, C.K. and J.P. Caufield. 1990. Using mixed-integer programming to determine the optimal location for an oriented strandboard plant in Alabama. Forest Products Journal, 40(2): 39-44.

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): 457-466.

McRoberts, R.E. and E.O. Tomppo. 2007. Remote sensing support for national forest inventories. Remote Sensing of Environment, 110: 412-419.

McRoberts, R.E., D.G. Wendt, M.D. Nelson, and M.H. Hansen. 2002. Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates. Remote Sensing of the Environment, 81: 36-44.

Minitab, Inc. 2009. Minitab Statistical Software, release 16 for Windows, State College, Pennsylvania. Minitab® is a registered trademark of Minitab, Inc. Portions of the input and output contained in this publication/book are printed with permission of Minitab, Inc.

Munn, I.A., Y. Zhai, and D.L. Evans. 2003. Modeling forest fire probabilities in the south central United States using FIA data. Southern Journal of Applied Forestry, 27(1): 11-17.

Parker, R.C., and D.L. Evans. 2009. LiDAR Forest Inventory with Single-Tree, Double-, and Single-Phase Procedures. International Journal of Forestry Research. 2009:1-6.

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

Riggs, A., W. Tucker, W., and P. Glass. 2013. 2012-2013 Forest Inventory Southwest Region, Mississippi. Jackson, MS: Mississippi Institute for Forest Inventory. 20 p.

Schultz, E.B., T.G. Matney, D.L. Evans, and I. Fujisaki. 2006. A Landsat stand basal area classification suitable for automating stratification of forest into statistically efficient strata. Six pages in Proceedings of the 1st international conference on object-based image analysis, July 4-5, 2006. S. Lang, T. Blaschke, and E. Schopfer (eds.). International Society for Photogrammetry and Remote Sensing (ISPRS) Vol. No. XXXVI - 4/C42 (ISSN 1682-1777). Salzburg, Austria.

Southern Research Station, Forest Inventory and Analysis (SRS FIA). 2014. The Inventory, Issue 32 January 2014. Available at: http://srsfia2.fs.fed.us/.

Spurr, S.H. 1952. Forest Inventory. Wiley and Sons, New York. 476 pp.

Tokola, T. and J. Heikkila. 1997. Improving satellite based forest inventory by using a priori site quality information. Silva Fennica 1(31): 67-78.

Trotter, C.M., J.R. Dymond, and C.J. Goulding. 1997. Estimation of timber volume in a coniferous plantation forest using Landsat TM. International Journal of Remote Sensing, 18(10): 2209-2223.

USDA Forest Service (USDA FS). 2011. USDA Forest Service Forest Inventory and Analysis (FIA) Fact Sheet Series. Accessed on August 12, 2011. Available from: http://fia.fs.fed.us/library/fact-sheets.

Verbyla, D.L. 1995. Satellite Remote Sensing of Natural Resources. Lewis Publishers, Boca Raton. 198 pp.

Wilkinson, D.W. 2011. Landsat-derived stand structure estimation for optimizing stratified forest inventories. Published dissertation, Mississippi State University. 42 pp.