Keywords: forest products, decision support, FIA, VCT, Landsat, county parcel data, machine learning, automated, forest age map
Spatiotemporal Informatics for Sustainable Forest Production Utilizing Forest Inventory and Remotely Sensed Data

Jobriath Scott Kauffman

(ABSTRACT)

The interrelationship between trees and humans is primordial. As pressures on natural resources grow and become more complex this innate connection drives an increased need for improved data and analytical techniques for assessing the status and trends of forests, trees, their products, and their services. Techniques for using readily available data such as the Forest Inventory and Analysis (FIA) database and output from forest disturbance detection algorithms derived from Landsat data, such as Vegetation Change Tracker (VCT), for estimating forest attributes across time from the state and inventory unit level down to the stand and pixel level are presented. Progressively more comprehensive harvest and parcel boundary records are incorporated appropriately. Quantification of attributes, including non-timber forest products and fine-scale age estimates, across the landscape both historically and into the future is emphasized. Spatial information on the distribution of forest resources by age-class provides knowledge of timber volume through time and across the landscape to support forest management for sustained production. In addition to monitoring forest resources in regards to their value as products for human consumption, their measurement facilitates analysis of the relationship of their spatial and temporal abundance to other resources such as water and wildlife.

This work received support from the USDA Forest Service Southern Research Station Forest Inventory and Analysis Program in Blacksburg, Virginia and the Center for Natural Resources Assessment and Decision Support at Virginia Tech.
In response to growing, complex pressures on natural resources, techniques for improving the timely estimation of the status and trends of forest resources across wide regions are presented. Methods for processing large amounts of readily available data, such as U.S. Forest Service Forest Inventory and Analysis data, and forest disturbance maps derived from remotely sensed data, are used for estimating forest attributes across time by region, state, county, and individually owned forest units. The inclusion of progressively more comprehensive harvest timing and parcel boundary records improves the accuracy of region-wide wall-to-wall and spatially precise forest age and harvest boundary maps at the individual stand level. In addition, region-wide quantification of tree-based non-timber forest products is facilitated by inventory measurements that are already used for timber products. These metrics can be used to measure forest resource supplies and attributes historically and for decision-support in maintaining sustainable production into the future. Monitoring and forecasting the abundance and attributes of these forest resources with spatial and temporal precision is also valuable for analysis of their relationship to other resources such as water and wildlife.
Dedication

This work is dedicated to Cole Scott Kauffman (October 21, 1953 - November 15, 2016), a giant among men. Thanks Dad! We did it! But there is still much to be learned from you about living, loving, and laughing. Thank the Lord for an extra portion of His grace when He gave me you. “Thanks be to God for his indescribable gift!” (2 Corinthians 9:15, NIV)
Acknowledgments

I would like to thank the following people and organizations for their impact on this research, my education, and on me as a person along this journey. The journey was long, so there are many people to thank. I thank God for placing each of you in my life.

Thanks to my advisor, Dr. Steve Prisley, for giving me the opportunity to work on this Ph.D. and to work at CeNRADS. It has been a wonderful experience. Thanks for being an outstanding teacher and mentor from my first GIS class with you through the end.

Thank you to many exceptional professors throughout all of my graduate experiences. First, much thanks to my committee members, Dr. John Coulston, Dr. Yang Shao, and Dr. Valerie Thomas, and to my co-author Dr. Jim Chamberlain. Your input was invaluable, and you challenged me to become better. I would also like to thank Dr. Evan Brooks, Dr. Kirsten deBeurs, and Dr. Randy Wynne. You are all first-class teachers and people.

I am very appreciative of the financial assistance received through CeNRADS to support this work from the USDA Forest Service and from our industry partners. Thank you.

It has been a pleasure working with many fine people at CeNRADS, including Delaney Beattie, Caitlin Carey, Shelia Crowe, Neil Crescenti, Josh Hammes, Scott Klopfer, Emily Smith-Mckenna, Andrew Phillips, Lola Roghair, Charlie Wade, Laura Wade, and Xiaozheng Yao. Thank you all. And thanks to many fellow graduate students, including Jill Derwin, Iris Fynn, Ranjith Gopalakrishnan, Allisyn Hudson-Dunn, Devita McCullough-Amal, Adam Oliphant, and Mike St. Germain. I learned a lot from you, and we suffered through much together. Thank you for your support and our conversations about this research.

Thank you friends, cousins, aunts, uncles, Pop Pop, Gramma, Nicole, Jason, and families for all of the love, support, prayers, and good times you have provided. Thanks Mom and Dad. You are amazing. You were there whenever I needed you, and you always set the sky as the limit. Your love and confidence in me got me here. We did it.

Finally, thank you from the bottom of my heart to Robin, my beautiful wife, and two super kids, Vivian and Joel. We did it. You all sacrificed so much. I am very proud to call you my family. The sky is the limit. Thanks for always sharing your lives, love, and fun times with me. I love you.
# Contents

1 Overview - Monitoring Timber and Non-Timber Forest Products and the Role of Land Change Science  
1.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1  
1.2 Information on forest resources and ecosystem services . . . . . . . . . . . . 6  
1.3 Context in Land Change Science . . . . . . . . . . . . . . . . . . . . . . . . 8  
1.4 Forest resource monitoring and remote sensing . . . . . . . . . . . . . . . . . . 14  
1.5 NTFPs and forest land use . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18  
1.6 Challenges and advances . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19  
1.7 References . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23  

2 Automated Estimation of Forest Stand Age Using Vegetation Change Tracker and Machine Learning  
2.1 Abstract . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28  
2.2 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29  
2.3 Methods . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35  
  2.3.1 2010 VCT disturbance map validation . . . . . . . . . . . . . . . . . 35  
  2.3.2 Evaluating the ability to detect clearcuts . . . . . . . . . . . . . . . . 35  
  2.3.3 Reclassifying forest disturbances by harvest method . . . . . . . . . . . 36  
  2.3.4 Calculating age . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 38  
2.4 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 39  
2.5 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 43  
2.6 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 48
3 Automated Harvest Detection and Delineation Using Vegetation Change Tracker and County Parcel Data

3.1 Introduction

3.1.1 Sources of forest stand age and volume information

3.1.2 Prior work and opportunities for refinement

3.1.3 Research objectives

3.2 Methods

3.2.1 Study area and data

3.2.2 Disturbance clump metrics

3.2.3 Sampling methods

3.2.4 Training, testing, and validation of machine learning algorithms

3.2.5 Stand-clearing disturbance maps, age maps, and partial harvest maps

3.2.6 Forest retention and conversion

3.3 Results

3.3.1 Reclassification accuracy

3.3.2 Comparison of VCT-derived age maps

3.3.3 Forest retention and conversion

3.4 Discussion

3.4.1 Input metrics

3.4.2 Bias and sampling

3.4.3 Forest retention and conversion

3.5 Conclusion

3.6 References

4 Monitoring Non-Timber Forest Products Using Forest Inventory Data: An Example with Slippery Elm Bark

4.1 Abstract

4.2 Management and policy implications
4.3 Introduction ......................................................... 100
  4.3.1 Non-timber forest products ................................. 101
  4.3.2 Forest Inventory and Analysis ............................... 102
4.4 Background ......................................................... 103
4.5 Methods ............................................................ 104
4.6 Results – slippery elm as example ............................. 111
4.7 Discussion .......................................................... 114
4.8 Conclusion .......................................................... 120
4.9 References .......................................................... 121

5 Conclusion ............................................................ 124
  5.1 Recap of the big picture ........................................... 124
  5.2 Procedure for fine scale forest inventory estimates .......... 125
  5.3 Additional research opportunities .............................. 128
  5.4 Conclusion .......................................................... 129
  5.5 References .......................................................... 130
List of Figures

1.1 NAIP aerial photo of portion of Dinwiddie County, Virginia with parcel bound-
aries in black and points denoting locations of FIA inventory plots. . . . . . . 3

2.1 NAIP aerial photo of portion of Dinwiddie County, Virginia with parcel bound-
aries in black and points denoting locations of FIA inventory plots. . . . . . 31

2.2 Portion of 2010 Virginia VCT map. . . . . . . . . . . . . . . . . . . . . . . . 34

2.3 Virginia VCT “age” map enhanced by reclassifying disturbed pixels as stand-
clearing or not. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 40

2.4 Estimates of forested acres in Virginia. . . . . . . . . . . . . . . . . . . . . 42

2.5 Forested acres in Virginia by age class, up to age 25. . . . . . . . . . . . . 43

2.6 Depiction of ability of enhanced VCT to conform to harvest boundaries. Top:
post-harvest NAIP aerial photography with photo-interpreted harvest bound-
aries; Bottom: Same image overlaid with semi-transparent enhanced VCT
“age” raster. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44

2.7 Depiction of conformance of enhanced VCT stand-clearing disturbances to
parcel boundaries. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 45

2.8 The enhanced VCT map reduces salt and pepper effects and more accurately
depicts stand age by reclassifying disturbances such as this 2010 thinning as
a partial disturbance before calculating years since most recent stand-clearing
disturbance. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 45

3.1 NAIP aerial photo of portion of Dinwiddie County, Virginia with parcel bound-
aries in black and points denoting locations of FIA inventory plots. . . . . . 57

3.2 Study area (shaded) includes portions of Virginia Counties within Landsat
path 15/row 34 with available parcel data . . . . . . . . . . . . . . . . . . . . 70
3.3 Reclassified peVCT disturbances: (Top) 2009 NAIP aerial photo with parcel boundaries, (Bottom) 2011 NAIP aerial photo with transparent disturbance clusters. (A) Violet = stand-clearing, (B) Green = partial, (C) Yellow = Non-harvest. 81

3.4 Examples (1986, 1991, 1996, and 2001) of the proportion of stand-clearing disturbances that are classified as forest in subsequent years 84

3.5 Proportion of 2011 coniferous forest and deciduous forest NLCD land covers that returned to a forest classification by years since 1991 stand-clearing disturbance 84

3.6 Classified disturbance clumps on left (blue = stand-clearing, purple = partial, orange = non-harvest); IFZ disturbance magnitude on right (brighter pixels equal higher magnitude) 86

3.7 Example calculations of input metrics. Gray and brown pixels are non-forest, light green and dark green pixels are forest, and red pixels are disturbed forest. 89

4.1 Geometric representation of the surface area of a truncated cone representing the bole of a tree. 109

4.2 Locations of FIA plots from the 2012 population evaluation group containing slippery elm trees 5 dbh or larger coded by count within plots and within species range map. 112

4.3 (a) Number of slippery elm in 2012, (b) percent change in number of slippery elm from 2007 to 2012, and (c) average yearly slippery elm mortality from 2007 to 2012 by FIA unit. 115

4.4 Estimated total surface area (million feet2) of live slippery elm trees 5 dbh or greater on forest land by diameter midpoint class from recent (2012) inventories 116

5.1 Process for partitioning VCT forest into harvest-sized clumps 126

5.2 Process for estimating volume of harvest clumps 127
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Error Matrix with percent accuracy rates for machine learning classification of disturbances larger than 1 acre as recorded by VCT.</td>
<td>40</td>
</tr>
<tr>
<td>3.1</td>
<td>Proportion of parcels containing a VDOF harvest record from each harvest class that also contain a VCT disturbance</td>
<td>67</td>
</tr>
<tr>
<td>3.2</td>
<td>Variables considered for classification algorithms</td>
<td>73</td>
</tr>
<tr>
<td>3.3</td>
<td>Sampling strata and types of disturbances and land use change</td>
<td>74</td>
</tr>
<tr>
<td>3.4</td>
<td>Models and overall accuracy rates for stand-clearing/non-stand-clearing classification of disturbance clumps with equal representation of harvest type strata</td>
<td>80</td>
</tr>
<tr>
<td>3.5</td>
<td>Models and overall accuracy rates for harvest/non-harvest classification of disturbance clumps with equal representation of harvest type strata</td>
<td>80</td>
</tr>
<tr>
<td>3.6</td>
<td>Agreement of most recent stand-clearing disturbance derived from each age map with time series of aerial photos</td>
<td>82</td>
</tr>
<tr>
<td>3.7</td>
<td>McNemar tests for differences in marginal proportions of successful determination of ‘most recent year of stand-clearing disturbance’</td>
<td>82</td>
</tr>
<tr>
<td>3.8</td>
<td>Metrics for clumps from Figure 3.6 by ID. Disturbance type is denoted in column one with: N = non-harvest, P = partial harvest, and S = stand-clearing disturbance.</td>
<td>87</td>
</tr>
<tr>
<td>4.1</td>
<td>Sample of trees found in FIA databases that are harvested for NTFPs</td>
<td>105</td>
</tr>
<tr>
<td>4.2</td>
<td>Summary metrics for NTFP tree species computed from FIA data.</td>
<td>106</td>
</tr>
<tr>
<td>4.3</td>
<td>Parameters (SE) and $df$ for linear regression lines modeling the relationship between $\sqrt{dbh}$ and $\sqrt{SA}$ of bark for boles of trees at least 5 in. dbh by species, using 2012 population evaluation groups.</td>
<td>110</td>
</tr>
</tbody>
</table>
4.4  Average annual net growth, mortality, and removals, and gross growth and net volume (in million cu. ft.) of live slippery elm trees 5 inches dbh or larger on forestland estimated from recent (2013) and previous inventories. Previous inventory values are approximately 5-year state prior inventories (e.g., state inventory closest to 2007). Does not include west Texas. . . . . . . . . . . . . . . . . . . . . . . . . . . . 113

4.5  Causes of mortality for slippery elm and for American elm co-located on plots. 117

4.6  Average annual 2013 net growth, mortality, and removals and 2007 and 2013 net volume (in million cu. ft.) with [sampling error percent] of trees 5 inches dbh or larger on forestland. Does not include west Texas. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 119
Chapter 1

Overview - Monitoring Timber and Non-Timber Forest Products and the Role of Land Change Science

1.1 Introduction

Renewable natural resources have the capacity to benefit this generation and the next, making their sustainable use a priority. A growing demand for these resources raises pressure on natural systems to produce and increases the value of informatics quantifying their abundance over time. Various ground measured and remotely sensed data sets can be combined in a complementary manner to enhance their value as important components in decision-making processes.

Distributions of forest by stand age and type at various spatial scales provide valuable information for optimizing forest production and sustainability. Two sources of forest stand age distribution, volume, and species information at the regional level are forest inventory plots and remotely sensed data. Each of these methods has its strengths and limitations.
Forest Inventory and Analysis

The three-phase sampling scheme of the US Forest Service Forest Inventory and Analysis (FIA) program, described by Bechtold and Patterson (2005), forms a comprehensive inventory of the nation’s forest resources. Precise estimates are attainable for forest-wide inventories at the county level and larger. A Phase 1 remote sensing component is designed to reduce variance through stratification, while Phase 2 consists of a systematic field measurement of randomly placed plots within a hexagonal grid system covering the continental United States. Each plot is assigned a land use code, and for those that meet the forest land use definition, an FIA field crew records each tree along with its species, status, and other attributes. Diameter at breast height (dbh) and height of each tree is measured, enabling estimation of cubic foot volume or other product and resource quantities. Plots can contain multiple ‘conditions’ which separate forest stands by ownership type, forest type, and age (USDA Forest Service 2014). FIA’s spatially and temporally comprehensive approach, which samples all of the nation’s forested lands over a periodic cycle is appropriate for monitoring status and trends of forest inventory (Reams et al. 1999; Bechtold and Patterson 2005). The Phase 3 forest health monitoring plots are a subset of Phase 2 plots and measure additional health indicator attributes yearly (Bechtold and Patterson 2005). Information from FIA on abundance, trends, and geographic distributions of products and resources can assist regional managers and policy makers by spatially monitoring their availability and sustainable use.

However, regional forest inventories such as FIA are not intended for stand-level estimation. Much forest management and planning occurs at the stand level by individual forest owners, especially in the Southeast. Therefore, modeling of forest resources sometimes requires information at this level. It is simply too costly to sample from every individual forest stand
at the regional level. Figure 1.1 shows a National Agriculture Imagery Program (NAIP) aerial photo of an approximately 30,000 acre portion of Dinwiddie County, Virginia, one-tenth the size of the entire county. The figure shows hundreds of separately managed areas of forest large enough for commercial timber harvesting, and five FIA plots. Individual management-based inventories requiring a sample of plots for every separately managed area for all of Virginia would be too costly for one organization to employ. Because comprehensive field measured data is not feasible for every individually managed unit, data on these units must be obtained through another source.

![Figure 1.1: NAIP aerial photo of portion of Dinwiddie County, Virginia with parcel boundaries in black and points denoting locations of FIA inventory plots.](image)

**Remote sensing**

Research to overcome difficulties of obtaining small area estimates of timber volume is of interest to forest managers. Maps of forest by major species groups derived from remotely
sensed Landsat data are common, including nation-wide land cover mapping such as the
30 meter pixel scale National Land Cover Database (NLCD) land cover map (Homer et al.
2015). Landsat continuously provides new multi-spectral images every eight to sixteen days
at a 30m pixel scale which facilitates the use of algorithms employing Landsat Time Series
Stacks (LTSS) to detect forest disturbances, including clearcuts (Goward et al 2006; Wulder
et al 2015). Vegetation Change Tracker (VCT) is one of these algorithms, with a suite of its
products freely available nationally for download (Huang et al 2010; Goward et al. 2016).
A simple way to calculate “age” of a forest is to measure the number of years since the last
clearcut. Additionally, LiDAR is capable of measuring individual tree heights and is also
becoming more common. Age estimates, in combination with forest type and stand height,
from remote sensing data at the stand or pixel level can make a major contribution toward
further refining volume estimates at this small scale. This remotely sensed data can be
obtained just as easily for a specific point as it can for an entire image, and can be combined
with other spatial data sources. For instance, counties collect data on parcel boundaries and
their features that is more and more available in on-line repositories housing collections of
counties available for downloading. Automated processing of these large amounts of data is
an obstruction that is becoming easier to overcome.

FIA has been developed mainly to estimate quantities and changes of timber products,
such as sawtimber and pulpwood, along with biomass which provides information about
carbon sequestration and potential bioenergy resources. This same data can and should be
used, in its current state or with some enhancements, to monitor non-timber forest products
(NTFPs), especially those derived from trees such as bark, nuts and fruits, sap, leaves and
needles, and boughs. Strides are being made toward this extended use of FIA to monitor
NTFPs, especially those derived from trees, but monitoring of these forest resources is not at the same level as timber (Kauffman et al. 2016). A diversity of species including white pine, sugar maple, pawpaw, common persimmon, butternut, black walnut, sweetgum, yellow poplar, black cherry, white oak, white willow, sassafras, slippery elm, and noble fir are valued for their ability to produce NTFPs and are inventoried by FIA.

The U.S. South leads the nation in timber harvesting where it is the most prevalent forest disturbance, especially in the Piedmont and Coastal Plain ecoregions (Fox et al. 2007, Cohen et al. 2016; Schleeweis et al. 2011; Pan et al. 2011). Automated methods for mapping these harvests are presented in Chapters 2 and 3. One component of these methods, VCT, is a disturbance detection algorithm that utilizes time series stacks of normalized (with respect to a sample of forested pixels) and integrated brightness values from appropriate bands of Landsat images (Huang et al 2010). Large deviations from the average forested pixel indicate likely disturbances which are often related to harvest in the U.S. South, but can also be related to other types of disturbance. The ability of VCT to detect and classify clearcuts and partial harvests such as thinnings and selection harvests is investigated here.

This introductory chapter provides an overview of the research components of this dissertation and also synthesizes appropriate papers on the role of Land Change Science (LCS) as it relates to forest disturbance and its ecological impacts. LCS, forest disturbance (detection, classification, location and timing), harvest delineation, forest stand dynamics and succession, and product volume estimation are integrally related. Harvesting of tree-derived non-timber forest products (NTFPs) is a unique subset of both non-stand-clearing forest disturbances and removal of products derived from trees that depends on forest land use other than plantation silviculture.
1.2 Information on forest resources and ecosystem services

Throughout its history the mission of the USDA Forest Service in general, along with FIA and its precursors, has included monitoring the nation’s forest resources, including timber and other forest products (LaBau et al. 2007). In recognition of the important role of forests for providing benefits such as timber products, non-timber products, beautiful scenery, wildlife habitat, recreation, and water quality improvement, congress delivered a mandate for the assessment of the nation’s renewable resources through the Forest and Rangeland Renewable Resources Planning Act of 1974 (Oswalt et al. 2012). This act requires comprehensive assessment of forest resources every ten years with updates at five-year intervals. Yielding to pressure from various interest groups opposed to timber harvesting, the act greatly influenced sampling strategies and broadened sampling focus to include more than just the timber benefits of forest ecosystems (LaBau et al. 2007).

Opposition to timber harvesting has all but vanquished it from public lands, leading to a dramatic decline in timber produced in the western U.S. over the last few decades. This, along with technological and silvicultural advances, has led to intensified timber production in the southern U.S. (Wear and Greis 2012). The southern forest that was mostly cutover forestland and degraded agricultural land several decades ago, has regrown. Today southern pine plantations are the most intensively managed forests in the world, rivaling southern hemisphere fast-growing exotic species (Fox et al. 2007). Due to the impact of industrial forestry productivity and plantation silviculture practices in the region leading to earlier financial maturity, forests in the southern U.S. have shorter average rotations than the rest
of North America, with more than 80 percent of forest stands less than 60 years old (Pan et al. 2011). This suggests that areas not in plantation forestry are still harvested frequently. However, the Appalachian mountain regions are harvested less frequently due to slope and access restrictions. Here, large areas of cleared forest are often the result of mining.

In Virginia, more than half a million acres of forest have been lost since 1977, and this loss is expected to continue (VDOF 2015). The proportion of expected loss is not constant across forest types. Like much of the southern U.S., Virginia can be divided into three ecological sub-regions: Coastal Plain, Piedmont, and Mountain. Therefore, forestland management in Virginia reflects that of the entire region with hardwood silviculture prevailing throughout Virginia and high concentrations of pine plantations throughout the Coastal Plain and Piedmont. Sixty-two percent of the 15.9 million forested acres in Virginia are privately owned by individuals and families, while 19 percent are corporately owned and another slightly more than one percent is owned by forest product firms (VDOF 2015). Private ownership of land well-suited for pine plantation silviculture has led to the practice of a high rate of clearcut harvesting in the Piedmont and Coastal Plain. Due to increased demand for bioenergy, there is a possibility of increased area of planted pine and decreased area of natural pine for Alabama, Florida, and Georgia for five future land use scenarios as defined by the U.S. Forest Service (Abt et al. 2012). A recent Virginia forest assessment report observed high pressure on pulpwood (Prisley 2015). This could be due to some competition for pulpwood size timber coming from bioenergy facilities. All in all, growing global demand for wood products including bioenergy and pellet companies, declining supplies from other regions such as the beetle-killed forests in western North America, and shrinking forest land base in the southern U.S., all create increased pressure on remaining pine forest (Bragg et al. 2015).
1.3 Context in Land Change Science

Loss of forest cover and changes in land use due to mining, conversion to agriculture, urban expansion, low density housing, parcelization, fragmentation, and shifts in forest type establish a context for this research in LCS. A 2007 special issue of the Proceedings of the National Academy of Sciences (PNAS) presented the emergence of LCS as a fundamental component of environmental change and sustainability research (Turner et al. 2007).

LCS is described as a coupled human-natural system (Turner et al. 2007). Relationships and dynamics in the human subsystem, such as societal structures, policy and decision making, and economic markets effect change in human land use, land cover, and ultimately the biophysical subsystem. Conversely, structures and changes in the biophysical subsystem, such as environmental and ecological goods, services, and properties cause change in land cover, human land use, and consequently the human subsystem. Study within this field involves observation and monitoring of land use, land cover, the human subsystem, the biophysical subsystem, forces operating exogenously on the system, and the effects of the system on external processes. Thereby, LCS can be thought of as a complex subsystem of an even more complex system.

Interactions within the system occur at multiple scales, further increasing the complexity of the system. Ideologically, LCS views the environment in terms of its entire array of ecosystem goods and services, rather than focusing on individual, or even a particular suite, of resources (Turner et al. 2007). Despite the vast complexity of this system, reaching a level of understanding far below what is exact may be adequate and useful.

The 2007 special issue of PNAS provides a glimpse of the complexity of this human-natural
coupled system through several specific case studies revealing different facets of LCS. For example, Diaz et al. (2007) discuss the interaction of ecosystem services with land use and cover, focusing on the integration of various mechanisms by which functional diversity affects ecosystem properties that are directly relevant to ecosystem services. Irwin and Bockstael (2007) use parcel and land use data from Howard County, MD to show that low-density residential development is most strongly associated with fragmentation and suggest that fragmentation is known to constrain habitat, reduce the use of open space, and degrade natural resources. They further reveal that the National Land Cover Database (NLCD) land cover map is inadequate for use in determining the effects of land cover change on fragmentation because low-density housing is often classified as forest.

The body of LCS research incorporates various tools that are useful for analysis. Remotely sensed data has become an integral component of LCS research. Advances in air and spaceborne sensors enhance the ability to monitor many types of land use and land cover changes (such as forest disturbances), net primary productivity, flora, carbon sources and sinks, and biodiversity (Turner et al. 2007). Of particular significance to this research, Manson and Evans (2007) incorporate agent-based modeling (ABM) as a tool to simulate the relationship between human policies, household outlooks, and markets with reforestation in the Midwest U.S. as well as deforestation in southern Yucatán, Mexico. ABM models reproduce critical features of complex systems using component-level rules that can be programmed into computer simulations to model behaviors of individual actors that comprise the system (North and Macal 2007). Manson and Evans’ models incorporate parcel-level survey results and crop and timber prices. Chapters 2 and 3, herein, help facilitate higher spatial resolution input data for use with the Spatial Wood Supply Simulator (SWSS), an ABM for timber
supply-chain modeling (CeNRADS 2016). Remotely sensed data can be used to delineate potential harvest locations and estimate their age, forest type, and volumes in various product classes for use as input into the ABM. Agents in the model, such as paper mills, saw mills and bioenergy facilities, bid on these harvests to fulfill their demand for timber products. This is just one of many potential uses for data enhancements produced as a result of this research.

The history of LCS within the Forest Service includes an interagency research program entitled Resources Evaluation and Technique (RET) formed in the mid 1970’s (Labau et al. 2007). RET included representatives from the USDA Forest Service, Soil Conservation Service, and Economic Research Service, along with others from the U.S. Department of Interior Bureau of Land Management and Fish and Wildlife Service. RET played an important role in developing a National Land Classification System and a review of the state-of-the-art for remote sensing in natural resource inventories.

Land cover maps derived from remote sensing, when combined with information on land use, shed light on structure and dynamics within the LCS system. Land use is definitely influenced by cover. It is easier to convert cleared forest or agricultural lands to other agricultural uses, other forest uses, or develop it than it is to convert mature forest. Therefore, when addressing where land use change occurs, the focus should be placed on areas where tree cover is minimal, including recently cleared forests. However, cover is influenced by use as well. For instance, if humans want to use their land to build a shopping center or grow crops, they will remove all the trees, thereby changing the land cover. So it can be concluded that behavioral aspects of the human component of the system are driving land use change. It is possible that humans will begin to value building or farming without removing trees. This
Current forest definitions vary by the amount of tree cover required for an area to be classified as forest. In addition, some definitions of forest are based on cover while others are based on use. Differences in various land use and land cover definitions of forest can be significant. The following are various relevant definitions of forest:

1. FIA defines accessible forest land as, "Land within the population of interest that can be occupied safely and has at least 10 percent crown cover by live tally trees of any size or has had at least 10 percent canopy cover of live tally species in the past, based on the presence of stumps, snags, or other evidence. To qualify, the area must be at least 1.0 acre in size and 120.0 feet wide. Forest land includes transition zones, such as areas between forest and nonforest lands that meet the minimal tree stocking/cover and forest areas adjacent to urban and built-up lands. Roadside, streamside, and shelterbelt strips of trees must have a width of at least 120 feet and continuous length of at least 363 feet to qualify as forest land. Unimproved roads and trails, streams, and clearings in forest areas are classified as forest if they are less than 120 feet wide or less than an acre in size. Tree-covered areas in agricultural production settings, such as fruit orchards, or tree-covered areas in urban settings, such as city parks, are not considered forest land.” (USDA Forest Service 2014)

2. NLCD definitions of deciduous, evergreen, and mixed forests includes “areas dominated by trees generally greater than 5 meters tall, and greater than 20 percent of total vegetation cover.” (http://www.mrlc.gov/nlcd2011.php)
VCT uses forest training pixels identified from the forest peak of a histogram of top of atmosphere reflectance in several Landsat spectral bands where other known dark objects such as water and dark soils have been masked. Pixels with reflectance below this threshold are candidates for forest. These training pixels are checked for consistency with other land cover products. Pixels that fluctuate below and above this forest peak from year to year during the growing season are not considered forest. (Huang et al. 2008)

While FIA incorporates stratification based on forest/non-forest classes derived from remote sensing data, it relies most heavily on measurement of ground plots where use can be adequately ascertained by an expert field crew (USDA Forest Service 2014). NLCD relies on remotely sensed data as observed by sensors on board Landsat satellites and is therefore only capable of discerning land cover when additional map layers to define use are not available. In fact, it is known that the forest land use class as defined by FIA includes all major NLCD land cover classes (Coulston et al. 2013). The NLCD definition of forest does not include any areas with trees less than 5 meters tall. Therefore very young stands will not be included in the definition of forest. Thus young stands would be excluded from forest maps and forest age maps derived from NLCD. If an age map is created that is derived from forest defined by NLCD it might give the impression that there is very little young forest when it is NLCD’s definition of forest that is leaving these young forests out. NLCD and VCT have no restriction on the width and overall size of a forest stand. Therefore, they might include clumps of trees that are less than 120 feet wide or 1 acre total. However, eliminating these would most likely lead to only slight changes in a forest age distribution.

Coulston et al. (2013) lay out conditions under which forest cover and use are equivalent,
such as when a land use conversion like forest to agriculture is observable from a satellite, and when partial harvest regimes are not intense enough to decrease cover below the defined threshold. They also define circumstances when forest cover and use diverge, including short-term catastrophic loss of tree cover caused by hurricanes, fires, severe insect and disease outbreaks, and clearcut harvest regimes, and when land use is not observable with remote sensing, such as agroforestry, tree covered parks, and wooded residential areas.

Within this context of differing definitions of forest, it should be noted that foresters prefer land use definitions of forest because they are very interested not only in where forests are now but also in where there are plans for forests to be in the future for various reasons, including economic and ecological. Industrial timber operations continuously cut and regrow “forests”, but a pine plantation can have a very different land use than a natural deciduous, mixed, or even pine forest type, especially since LCS views a landscape for its entire array of ecosystem services, including biodiversity, recreation, aesthetics, and forest products other than timber. When a pure LCS view of land use is considered any land use map is hard-pressed to contain enough classes. In fact, most land use maps do not reflect the actuality that land often has multiple uses. With the need to continue to feed a growing population that takes up more space for housing and agriculture, there has been much debate over strategies in which agricultural production and conservation are integrated and co-occur versus those which segregate intensive agricultural, silvicultural and natural areas (Grau et al. 2013).

The disturbance regime for a given landscape is defined by its median area and timing of disturbances (Rogers 1996). On the ecological subregion scale, the Piedmont and Coastal Plain regions of Virginia and the South are dominated by clearcut harvest. Many scientists believe
we have already entered the anthropocene epoch of geological time in which understanding, predicting, and successfully managing ecological pattern, process, and change are limited by the understanding of how and why these items are reshaped by humans over the long term (Ellis 2015). Costanza et al. (2014) estimate the loss of ecosystem services from 1997 to 2011 due to land use change to be somewhere between 4.3 and 20.2 trillion dollars per year. Forest landholders gain future market dollars and perhaps a net gain in biomass by converting natural forests to intensely managed pine plantations, but what expense does this come with in loss of biodiversity and susceptibility to future invasive pests and disease outbreaks? A review of literature on the topic reveals that some taxa prefer early successional forest, while others do not, and by definition monoculture plantations have very low tree species richness (Bragg et al. 2015). Will landowners of forestland in the southern U.S. always value forests most for paper products and building materials? What will happen to intensively managed pine plantations when these products are no longer valued? Not much is known about the consequences of today’s intensive silvicultural practices in the southern U.S. on possible ecosystem services provided by tomorrow’s forests (Bragg et al. 2015). How should the future use or multiple uses of today’s plantations be weighed in today’s management practices?

1.4 Forest resource monitoring and remote sensing

Changes in land cover today may be driven by changes in land use from long ago (McEwan et al. 2011; Nowacki and Abrams 2015), from a time in which data is much less available than it is now. Remote sensing can be used to quantify change where there is not enough time and resources to send field crews, but it cannot collect data retrospectively. The longest
continuous earth monitoring system at forty-four years and counting, Landsat, is still paltry compared to all of the change in land use that has occurred throughout history. Despite this, and perhaps because of this, the opportunity to extend the Landsat record is valuable, shedding light on the entire complex coupled human-natural system which LCS endeavors to describe and understand at varying scales.

Remote sensing is an important tool for quantifying land use and land cover change. A review of best practices of change detection methods recommends Landsat for monitoring land use/land cover and disturbances and MODIS, with more frequent images at the cost of coarser spatial resolution, for monitoring changes in phenology (Willis 2015). A continuous record of these changes can shed light on system processes. Due to the complex nature of LCS, drivers of change are not always apparent or agreed upon (Nowacki and Abrams 2015; Pederson et al. 2015; Abrams and Nowacki 2015). Consistent measurements over time at an appropriate spatial scale and remeasurement period have enabled Landsat derived data to measure effects of natural and anthropomorphic processes on ecosystems, including stress or chronic loss, growth or increase, state change, change and resilience, cyclical change, and interactions and feedbacks (Kennedy et al. 2014). Beginning in 1972 with the launch of Landsat 1, the Landsat mission was most recently continued on February 11, 2013 with Landsat 8 (Roy et al. 2015) and is expected to be extended with Landsat 9 in 2023 (Wulder et al. 2015). This continuing archive held by the USGS allows for consistently processed, moderate spatial resolution, large area, long-term, terrestrial data records for resource management and global change studies and was made freely available in 2008 (Roy et al. 2015; Wulder et al. 2012). In 2010, when the USGS began its Landsat Global Archive Consolidation Mission, more data was held outside its archive than inside (Wulder et al. 2015).
Since then, the archive’s holdings have more than doubled with valuable past records. Its availability has facilitated the ability to create Landsat Time Series Stacks (LTSS) for an array of uses, including disturbance detection algorithms (Roy et al. 2015; Wulder et al. 2012).

These change detection algorithms, such as Land Trendr, VCT, Global Forest Watch (GFW), Exponentially Weighted Moving Average Change Detection (EWMACD), and others (Kennedy et al. 2010; Huang et al. 2010; Hansen et al. 2013; Brooks et al. 2014) are of immediate application to the research in this dissertation. Some of these algorithms have the ability to detect abrupt disturbance events, such as fire, wind, harvest, and urban development, while others are better suited for detecting more gradual agents of change, including insects, disease, and stress. The highly automated nature of these algorithms is valuable for rapid processing of large amounts of data in order to make prompt decisions. LTSS along with ensembles of these change detection algorithms and analysis of the shapes of time series plots of different measurements allow for attribution of various drivers to each change (Banskota et al. 2014; Kennedy et al. 2014; Cohen et al. 2016).

Cohen et al. (2016) analyzed forest disturbance rate by type over time using a pixel based random sample of human-interpreted Landsat time series in five geographic regions. In the southeast, harvest was the dominant disturbance agent over decline, fire, and other types of disturbances for the period from 1985 through 2012. Throughout the entire U.S., during this same time period, decline became the dominant disturbance type in 1996, switching places with fire. Schleeweis et al. (2011) relate canopy change observations and causal processes at multiple scales using satellite data and ancillary geospatial data for fires, hurricanes and tornadoes, suburbanization/urbanization, harvests, and insects and pathogens. They also
include thorough regional analyses in which a high harvest rate for the southeast is confirmed along with high rates of insect damage. Comparing rates from 1980-1990 with those from 2001-2005 shows a decrease in both insect-related disturbance and clearcut harvests. Pan et al. (2011) generated an age map for forested areas throughout the U.S. and Canada by integrating forest inventory data, remote sensing data, and supplementary fire data. In the U.S., Landsat and FIA data were used. A description of age structure and disturbance legacy is provided for regions throughout Canada and the U.S., and applications of the map for carbon studies are discussed.

Disturbance detection algorithms, such as VCT, can be used in conjunction with machine learning tools, such as Support Vector Machine (SVM) to reclassify disturbances by type (Zhao et al. 2015; Kauffman and Prisley 2016 - Chapter 2 herein). Kauffman and Prisley (2016) reclassify disturbances as stand-clearing or not in order to determine harvest boundaries and age by number of years since the last stand-clearing disturbance. Harvest boundaries and years since last stand-clearing disturbance can be related to quantities of ecosystem services, including, but not limited to, timber.

Hansen et al. (2010a) quantified gross forest cover loss (GFCL) by four biomes using MODIS stratification and subsequent primary Landsat data. This research was criticized for attempting to quantify forest health with GFCL alone and for comparing the southern U.S. to the rest of the world using a ‘cover’ definition of forest, rather than a ‘use’ definition of forest (Wernick et al. 2010; Reams et al. 2010). Wernick et al. (2010) show that a net increase in growing stock volume from FIA in the southern U.S. runs counter to the loss shown by Hansen et al. (2010a) for the same area. However, an increase in growing stock volume taken alone could also misrepresent forest health. It is possible for equal gains in growing stock
volume to coincide with quite different levels of overall forest health. An updated quantification of gross forest cover gain in conjunction with GFCL has since been made public on the GFW website and is updated continuously using Google Earth Engine and a 50 percent threshold tree crown cover density definition of forest (Hansen et al. 2013). This update shows high amounts of both loss and gain in the southern U.S. While multiple measures are used to quantify forest health, a ‘cover’ definition of forest is still employed. Reams et al. (2010) argue for the importance of field data in forest cover and use change monitoring in order to clearly assess use. Certainly, LCS is not at a point where it can rely entirely on remote sensing instruments for measuring land use change.

1.5 NTFPs and forest land use

NTFPs are commodities that can be bought, sold, and traded, or gathered directly for personal use. NTFPs are used for subsistence and are disproportionately important to marginalized groups such as impoverished rural families (Vaughan et al. 2013). Because of this, their value cannot fully be measured by markets. Despite this, they generate tens of billions of dollars of revenue annually (McClain and Jones 2005). A possible hyper-valuation of timber products through what might be a relatively ephemeral high-dollar amount in current markets could lead to loss of future NTFPs.

It can be argued that forests used for NTFPs should be monitored even though their market value may currently be much lower than forests used for timber production. Disturbance detection algorithms using LTSS can be used to monitor where and when forested stands are clearcut (Kauffman and Prisley 2016), yielding information about when land use change
might have occurred. Foresters are keenly aware of changes from forest to non-forest, and this same amount of care can be taken to determine when natural forest, which contains NTFPs such as slippery elm, is converted to planted pine, which systematically eliminates some species that produce NTFPs.

Techniques for estimating net forest land use change using remeasured FIA plots have been developed by VanDeusen et al. (2013), who showed statistically insignificant net change at the 0.025 level for the 34 states with remeasured plots at the time of analysis. These same procedures can be used to measure net loss of natural forest and other forest change, such as from forest suitable for NTFP production to forest that is not.

1.6 Challenges and advances

Because forests capable of producing NTFPs are used differently than pine plantations, when viewed from the LCS perspective there is value in differentiating them in land use maps. This research focuses on data-intensive methods for monitoring production capabilities of both types of forest. LCS goes beyond this focus to include the entire array of environmental and ecosystem services. This system is complex with interactions that depend on much more than the ability of forest to produce timber products or biomass for sequestering carbon, the focus of FIA.

Anthropogenic forest disturbances will continue to dominate private land as long as humans value what these forests are disturbed for. But humans are not immune to forces outside themselves, and sometimes their own desires have inadvertent side effects that return full force. Virginians and those throughout the South have been hit by free-trade policies that
have led to the offshoring of the U.S. furniture sector, mainly to China (Prestemon et al. 2015). In addition, the most recent recession has had negative effects on all Americans, but the collapse of U.S. housing construction has had an extreme effect on the southern U.S. where, for example, southern pine lumber production fell from 19.0 to 11.8 billion board feet between 2005 and 2009 (Hodges et al. 2011). Because of the economies of scale for the logger that come with access to larger machinery and economies of scale for the owner due to the ability to concentrate resources to a specific location, there is a trend toward larger harvest areas, but this is challenged by parcelization that drives its own trend toward smaller contiguous areas of forest under the same ownership (Moldenhauer and Bolding 2009; Kolis 2016; Hatcher et al. 2013). Owners of larger tracts profit more per acre because buyers of standing timber take not only the market situation but also harvest costs into consideration when making purchase offers (Kolis et al. 2014). While recent land cover changes can be the result of land use changes from long ago (McEwan et al. 2011), current human land use values combined with other current forces can dominate land cover changes. Indeed, humans value forest for uses other than growing timber, but there is difficulty in valuing these ecosystem services over monetary market values inherently attached to timber which can more easily be weighed, measured, and traded at a going rate (Turner et al. 2016).

Humans continue to struggle with detangling the complex drivers of land degradation (Turner et al. 2016). We do not know what will happen if we value less intense management of forest in the future or what we are sacrificing by valuing intense management now. How much of what drives land use is influenced by what we actually value, and how much is due to policies we have instituted that, for instance, facilitate global markets over local markets? Global markets require more packaging, and thus more wood consumption, while
local products require less packaging. The USDA Forest Service Forest Products web page (www.fs.fed.us/research/forest-products) states that “Countries with large steady quantities of industrial wood use are more likely to retain their forest base.” However, this is not evidence that there are not other ways to preserve forest. Furthermore, this statement does not define the characteristics of the forests that large steady quantities of industrial wood use preserves. Tidwell (2016) emphasizes the important role of the Forest Service in engaging the public to make decisions that will benefit the forests they value, but there is no apparent answer to how this can be accomplished.

This research investigates and presents methods for quantifying timber and non-timber forest products derived from trees throughout the United States with particular emphasis on Virginia. Spatially and temporally precise estimates of forest product quantities provide valuable information for decision support and sustainable management of forest resources. Ways to improve the ability to generate information for quantifying forest resources over large areas such as Virginia or the entire range of slippery elm (*Ulmus rubra* Muhl.) are presented here. Because these large areas require analysis of vast amounts of data, the ability to effectively combine readily available data from multiple sources is significant, allowing forest professionals to expedite the decision-making process.

Along with timber, NTFPs are also an important component of comprehensive forest management (Alexander et al. 2011). As pressures on these resources grow, techniques to produce scientifically sound information in a timely manner are becoming increasingly valuable. Therefore, the development, implementation, and thorough validation of decision support tools is central to this research.

In an effort to support sound decision-making about forest resources, the overarching objec-
Objective of this research is to develop and test analytical approaches to quantify forest characteristics across broad landscapes to enhance decision support. Specifically, this research:

(1) presents an overview of the literature on the relationship between forest dynamics and the harvesting of trees and their products.

(2) evaluates the use of VCT disturbance maps as a basis for estimating forest stand age and delineating harvest boundaries by re-classifying these disturbances as stand-clearing versus non-stand-clearing.

(3) improves upon age maps and harvest delineations in Chapter 2 by including parcel boundaries and shows that partial harvest disturbances, including thinnings, can be accurately classified.

(4) analyzes FIA data to monitor abundance and distribution of non-timber forest products (NTFPs) derived from trees.

Objective 1 is met in this overview chapter, synthesizing ideas and literature on the ecological implications of forest disturbances along with research on monitoring timber and non-timber forest products. Objectives 2, 3, and 4 are met in Chapters 2, 3, and 4 of this dissertation. These four objectives advance the ability to use readily available plot samples and remote sensing data to quantify forest products at multiple spatial scales and across various species. This allows for improved information which better enables decision-makers to plan for sustainable use of the resources required to produce forest products. In addition, historical and projected records of resource/product availability across the landscape will permit further analyses in relation to other resources such as water and wildlife. Spatial and temporal changes in quantities and trends of timber and non-timber forest products should
be considered as pieces of the complex, dynamic system described by LCS that includes ecological, social, political, and physical systems and feedback loops of which a very limited description has been provided here.

1.7 References


Nowacki, G.J., and M.D. Abrams. 2015. Is climate an important driver of post-European


USDA Forest Service. 2014. Forest Inventory and Analysis national core field guide Volume I: field data collection procedures for Phase 2 plots. USDA For. Serv., North Central Research Station FIA, St. Paul, MN.


Chapter 2

Automated Estimation of Forest Stand Age Using Vegetation Change Tracker and Machine Learning

Jobriath S. Kauffman and Stephen P. Prisley

Note: This paper was published in Mathematical and Computational Forestry and Natural-Resource Sciences

2.1 Abstract

The ability to automatically delineate forest stands and determine their age is useful for natural resources professionals. Two common approaches to estimating forest area and age-class distributions are inventory-based methods, such as Forest Inventory and Analysis (FIA), and remote sensing based methods. Vegetation Change Tracker (VCT) is an algorithm that uses time series stacks of Landsat images to identify forest disturbances. However, additional computation is required to identify type of disturbance. This paper evaluates the usefulness of machine learning tools, such as support vector machine (SVM), for reclassifying VCT
disturbances as stand-clearing disturbances or partial disturbances. Overall accuracy for a 2010 VCT disturbance map of the entire state of Virginia was determined to be 87 percent. 100 percent of 2010 Virginia clearcut harvests recorded in a reference dataset were classified as disturbances by VCT. Neighboring disturbed pixels, as classified by VCT, were clumped together and reclassified as stand-clearing disturbances or partial disturbances using SVM and variables for average disturbance magnitude and shape and size metrics of the clumped pixels, with an overall accuracy rate of 86 percent. The user’s and producer’s accuracy rates for stand-clearing disturbances were 88 percent and 95 percent respectively. In addition, an algorithm was developed in R for determining years since last stand-clearing disturbance for each pixel in a time series stack of reclassified VCT disturbance maps from 1984 to 2011. Neighboring pixels of the same age, in number of years since last stand-clearing disturbance, were clumped together and correspond, in general, to clearcut harvest boundaries.

2.2 Introduction

The fact that there is value in mapping forest stands is unquestioned. Distributions of forest by stand age and type at various spatial scales provide valuable information for optimizing forest production and sustainability. Both field measured forest inventories and remotely sensed data have been used to estimate forest area and age. Each of these methods has its strengths and limitations. Precise field measured inventory estimates from sample plots are possible over large areas but are costly for fine scale estimates across a large area due to the large number of sample plots required. Remotely sensed data can be obtained just as easily for a specific point as it can for an entire image. Automated processing of these large amounts of data is an obstruction that is becoming easier to overcome. While cost effective
methods for yearly, statewide, border-to-border mapping of forest by age and major species group have been elusive, they are nonetheless obtainable.

The three-phase sampling scheme of the US Forest Service Forest Inventory and Analysis (FIA) program, described by Bechtold and Patterson (2005), forms a comprehensive inventory of the nation’s forest resources. Precise estimates are attainable for forest-wide inventories at the county level and larger. A Phase 1 remote sensing component is designed to reduce variance through stratification, while Phase 2 consists of a systematic field measurement of randomly placed plots within a hexagonal grid system covering the continental United States. Each plot is assigned a land use code, and for those that meet the forest land use definition, an FIA field crew records each tree along with its species, status, and other attributes. Diameter at breast height (dbh) and height of each tree is measured, enabling estimation of cubic foot volume or other product and resource quantities. Plots can contain multiple ‘conditions’ which separate forest stands by ownership type, forest type, and age (USDA Forest Service 2014). FIA’s spatially and temporally comprehensive approach, which samples all of the nation’s forested lands over a periodic cycle is appropriate for monitoring status and trends of forest inventory (Reams et al. 1999; Bechtold and Patterson 2005). The Phase 3 forest health monitoring plots are a subset of Phase 2 plots and measure additional health indicator attributes yearly (Bechtold and Patterson 2005). Information from FIA on abundance, trends, and geographic distributions of products and resources can assist regional managers and policy makers by spatially monitoring their availability and sustainable use.

However, regional forest inventories such as FIA are not intended for stand-level estimation. Much forest management and planning occurs at the stand level by individual forest owners, especially in the Southeast. Therefore, modeling of forest resources sometimes requires
information at this level. It is simply too costly to sample from every individual forest stand at the regional level. Figure 1.1 shows a National Agriculture Imagery Program (NAIP) aerial photo of an approximately 30,000 acre portion of Dinwiddie County, Virginia, one-tenth the size of the entire county. The figure shows hundreds of separately managed areas of forest large enough for commercial timber harvesting, and five FIA plots. Individual management-based inventories requiring a sample of plots for every separately managed area for all of Virginia would be too costly for one organization to employ. Because comprehensive field measured data is not feasible for every individually managed unit, data on these units must be obtained through another source.

Figure 2.1: NAIP aerial photo of portion of Dinwiddie County, Virginia with parcel boundaries in black and points denoting locations of FIA inventory plots.

Timber product volume and biomass estimates at small scales are difficult to obtain but increasingly valuable for sustainable forest management. Maps of forest by major species groups are common, including nation-wide land cover mapping such as the 30 meter pixel
scale National Land Cover Database (NLCD) land cover map (Homer et al. 2015). Age estimates, in combination with forest type, from remote sensing data at the stand or pixel level can make a major contribution toward further refining volume estimates at this small scale.

A simple way to calculate “age” of a forest is to measure the number of years since the last clearcut. Algorithms using time series stacks of Landsat data, such as Vegetation Change Tracker (VCT), have proven to be reliable for detecting forest disturbances (Huang et al. 2010). After other known dark objects such as water and dark soils have been masked from Landsat images dating back to 1984 (Landsat 4), VCT uses forest training pixels identified from the forest peak in histograms of top of atmosphere reflectance in the near infrared and two short-wave infrared spectral bands (Huang et al. 2008, Huang et al. 2010). The means and standard deviations of these training pixels in the red and shortwave infrared bands are used to calculate an integrated forest z-score (IFZ) for each pixel in the image (Huang et al. 2010).

In time series of yearly height of season IFZ values, forested pixels will remain persistently below a threshold IFZ value, while non-forested pixels will remain above the threshold or fluctuate above and below it (Huang et al. 2010). Thus, a sudden increase in a pixel’s otherwise persistently low IFZ score indicates the timing of a forest disturbance within the time series. In this way, the VCT algorithm described by Huang et al. (2010) can be used to generate VCT products such as yearly disturbance maps at the 30 meter pixel level. The magnitude of these disturbances can also be calculated by finding the difference between a pixel’s average IFZ score (or other index) and its IFZ score for the disturbance year. Normalized difference vegetation index (NDVI) and normalized burn ratio index (NBRI),
and IFZ4 are also incorporated in the VCT algorithm and used to calculate similar measures of disturbance magnitude (Huang et al. 2010). NDVI measures photosynthetic capacity using the red and near-infrared bands. The calculation for NBRI is similar to NDVI but uses the near-infrared band and short-wave infrared band (band 7). Changes in NBRI can be used to measure burn severity. IFZ4 is calculated similarly to IFZ but uses only the near-infrared band. Maps of disturbance magnitudes of disturbed pixels measured each of these ways are in production for the contiguous United States.

Within secondary succession forests, especially those in which frequent harvest and regeneration occurs, groups of neighboring 30 meter pixels representing forest of the same age were most likely harvested together, or cleared by some other mechanism, at some point in the past. Clumping these neighboring pixels of the same age together can be used as a method for creating objects that conform to past harvest boundaries.

The Virginia Department of Forestry began keeping records of all harvests in Virginia in 2009 in order to facilitate inspections of best management practices (BMPs). The GPS point location of the first logging deck the BMP inspector comes to on the date of inspection is collected along with other harvest attributes. Typically, there are five or six thousand harvests per year in Virginia. Delineating these harvests with a GPS during inspection or by post-harvest photo interpretation is costly and time-consuming. Therefore, efforts to automate this process and extend the records back to 1984, the first available year of VCT disturbance maps, are worthwhile.

The combination of VCT disturbance maps dating back to 1984 with several recent years of comprehensive harvest records creates a convenient avenue to extend the record historically, back to 1984, accurately placing their occurrence spatially and temporally. This process can
be automated and used wherever VCT data is available, especially in areas dominated by harvest disturbances. Prior harvest locations are often the best indicators of where future harvests will occur. Therefore, in addition to extending harvest records historically, data of this type can be used to help forecast the locations of future harvests. Past and future harvest locations yield valuable information for many uses, including timber procurement, climate modeling, water resource modeling, and wildlife habitat analysis.

![Figure 2.2: Portion of 2010 Virginia VCT map.](image)

This research demonstrates procedures for using yearly VCT disturbance maps to create clumps of neighboring pixels that were disturbed collectively. These clumps can then be spatially linked to recent harvest records, and metrics related to shape, size, and average disturbance magnitude of each clump can be used to train machine learning tools used for classifying disturbances as either stand-clearing or partial. Throughout this paper the term “enhanced VCT” will be used when referring to VCT disturbance maps that have been reclassified to include only stand-clearing disturbances. Raster stacks of stand-clearing disturbance maps can be used to subsequently create an “age” map by calculating the number
of years since the last stand-clearing disturbance. The process for creating a map of this type for Virginia will be described in the next section. VCT data products are anticipated to be available nationally, creating opportunities to repeat these methods anywhere in the contiguous United States, especially in areas dominated by harvest disturbances.

2.3 Methods

2.3.1 2010 VCT disturbance map validation

VCT disturbance maps, and disturbance magnitude maps created from the VCT algorithm, were obtained for all of Virginia. An accuracy assessment of the 2010 VCT disturbance map for Virginia was performed using before and after aerial photography, 2008 National Agriculture Imagery Program (NAIP) and 2012 NAIP, respectively. Figure 2.2 depicts a portion of the 2010 VCT disturbance map. VCT yearly disturbance maps are classified into six groups: persisting non-forest, non-forest after a disturbance, persisting forest, forest after a disturbance, forest disturbed in the current year, and water. 100 sample points within each class were randomly chosen. For simplicity, the two non-forest groups (200 points total) and the two forest groups (200 points total) were combined together. This assessment can be used to validate the accuracy of using the map to identify each of these classes, including forest and forest disturbances.

2.3.2 Evaluating the ability to detect clearcuts

In conjunction with the 2010 VCT disturbance map accuracy assessment, the ability of VCT to detect clearcut harvest was evaluated. VCT mapped all sample points that were clearcuts
as disturbances. In order to add weight to this assessment, the Virginia harvest records were used to look for possible errors outside of the sample used for validation in which VCT classified a clearcut harvest as something other than a disturbance. All harvest locations that did not intersect a VCT disturbance were inspected using the before and after aerial photography in order to find out if there was perhaps a stand-clearing disturbance that VCT missed. If VCT is doing reasonably well at detecting stand-clearing disturbances, the next step would be to reclassify all VCT disturbances as stand-clearing or not.

2.3.3 Reclassifying forest disturbances by harvest method

It is known that VCT detects yearly forest disturbances using a Landsat time series stack of one scene per year during the height of growing season (Huang et al. 2010). Sometimes more than one scene is used to allow for more cloud free pixels. The latest scene used for Virginia in 2009 was taken on September 14th, while the earliest scene taken for Virginia in 2010 was on June 3rd. Therefore, point locations of Virginia timber harvest data records from the Virginia Department of Forestry for inspections between these dates were intersected with the 2010 VCT disturbance maps covering Virginia. This encompasses all of the land area in Virginia within Landsat path/row scenes 14/34, 14/35, 15/33, 15/34, 15/35, 16/33, 16/34, 16/35, 17/33, 17/34, 17/35, 18/34, 18/35, 19/34, and 19/35.

Cells classified as disturbances within each VCT disturbance map were evaluated for adjacency to neighboring disturbed cells with the Queen’s case of 8 directions (right, left, above, below, or diagonal) using the ‘clump’ function in the R ‘raster’ package (Hijmans 2015). In this manner, connected groups of neighboring disturbed pixels were assumed to be the result of the same forest disturbance. Therefore, they were clumped together and given a common
identity. Since VCT detects both stand clearing disturbances and partial disturbances, the ability to calculate stand “age” must be facilitated by reclassifying VCT disturbances as stand-clearing or not. Classification of VCT disturbances by type using machine learning tools, including SVM, is appropriate and has proven to be effective (Zhao et al. 2015).

Average VCT disturbance magnitudes as measured by IFZ, NDVI, NBRI, and IFZ4 were obtained for each disturbance clump. In addition, various shape and size metrics were calculated for each clump using the ‘PatchStat’ function in the ‘SDMtools’ R package (VanDerWal et al. 2014). Clearcut harvests tend to have higher disturbance magnitudes, larger areas, and less complex shapes as they conform to more linear parcel boundaries.

There were 1170 VCT disturbance clumps from 2010 that intersected with one of the harvest site point locations in the date range specified above. Half of the disturbance clumps were used for training and half for validation of three machine learning tools used to reclassify VCT disturbance clumps as stand-clearing or partial disturbances. Support vector machine (SVM; ?), k nearest neighbor (kNN; Meyer et al. 2015), and the ‘rpart’ (Therneau et al. 2015) R package classification algorithms were trained using a somewhat arbitrary selection of variables, including disturbance magnitudes measured three ways (IFZ, NDVI, and NBRI), area, and fractal dimension index of the training clumps. Fractal dimension index is a measure of shape complexity. The kNN algorithm classifies a point in the feature space according to the majority class of a predefined number of nearest neighbors. SVM is also a supervised classification algorithm that maximizes the separation between two classes with a hyperplane. The ‘rpart’ package is an implementation of the Classification and Regression Trees (CART) algorithm, following Breiman et al. (1984) in most details. CART and rpart recursively split the feature space by finding the value of a variable that separates
the training sample data into classes that minimize incorrect classification. In an effort to maximize automation of the reclassification process, no further effort was made to calibrate the models or select the most advantageous features for improving accuracy. For the kNN algorithm, k=9 was arbitrarily chosen for the number of nearest neighbors. Kernel methods for producing non-linear classifiers are possible with SVM, but the default ‘linear’ kernel was used for simplicity. Each of the three trained models was used to classify each member of the validation set as a stand-clearing disturbance or a partial disturbance, and the results were compared to the actual harvest data.

The trained machine learning classification tools were used to reclassify each clump in each disturbance map as a stand-clearing disturbance or not based on the majority class of the three models. This process was automated and repeated for the 15 Landsat scenes for the entire study area, each year for the 28 years from 1984 to 2011, using R and its “raster” package (R Core Team 2015, Hijmans 2015). The resulting maps of stand-clearing disturbances are what has been defined above as enhanced VCT disturbance maps.

### 2.3.4 Calculating age

R was also used to create a time series stack of the enhanced VCT maps in order to calculate the “age” of each pixel by determining the number of years since the last stand-clearing disturbance. Neighboring pixels of the same “age” using the Queen’s case of 8 directions were clumped together and the “Eliminate” filter function in Erdas Imagine (Geosystems 2013) was used to clean edges and eliminate small clumps less than five pixels. Thus, clumps smaller than one acre were removed, and those pixels were back-filled with information from the surrounding clumps.
2.4 Results

The overall accuracy rate of the 2010 VCT disturbance map was estimated to be 87 percent. According to this assessment, an estimated 100 percent of the 2010 Virginia stand-clearing disturbances were classified as disturbances by VCT. Furthermore, clearcuts from 2010 harvest records were overlaid with 2010 VCT disturbances. This was done to insure that there was not a clearcut harvest at the location of the harvest record that VCT did not pick up. It was confirmed that VCT did not miss any clearcut harvests. GPS point locations of harvests are taken at the first logging deck that the forestry official comes to when inspecting the logging site. Logging decks are most often on the edge of the harvest site and are sometimes outside of the harvest site altogether, so it is not expected that all harvest point locations will be within the actual boundaries of the harvest.

The greatest overall reclassification accuracy rate was achieved with SVM at 86 percent (Table 2.1). SVM correctly classified 95 percent of clearcuts but misclassified 42.5 percent of partial harvests. The results for kNN and rpart were comparable.

An enhanced VCT “age” map for all of Virginia was created (Figure 2.3). This map shows age as the number of years since the last stand-clearing disturbance regardless of whether the cleared forest returned to forest or remained non-forest. The total number of acres that VCT considered persisting forest or disturbed forest for at least one year in the timespan from 1984 to 2011 is 17.6 million (Figure 2.4). This compares with 16.2 million acres if post-disturbance non-forest pixels as identified in the most recent VCT disturbance map are removed. Therefore approximately 8 percent of the pixels that are given a forest “age” in the enhanced VCT “age” map have either converted to non-forest, or have not regenerated
Table 2.1: Error Matrix with percent accuracy rates for machine learning classification of disturbances larger than 1 acre as recorded by VCT.

<table>
<thead>
<tr>
<th>Method</th>
<th>Classification Data</th>
<th>Error Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-Clear-Cut</td>
<td>77 23 0.77</td>
</tr>
<tr>
<td></td>
<td>Clear-Cut</td>
<td>57 428 0.882</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.575 0.949 0.863</td>
</tr>
<tr>
<td>rpart</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-Clear-Cut</td>
<td>76 30 0.717</td>
</tr>
<tr>
<td></td>
<td>Clear-Cut</td>
<td>58 421 0.879</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.567 0.93 0.850</td>
</tr>
<tr>
<td>kNN</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-Clear-Cut</td>
<td>64 34 0.653</td>
</tr>
<tr>
<td></td>
<td>Clear-Cut</td>
<td>70 417 0.856</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.478 0.925 0.822</td>
</tr>
</tbody>
</table>

to the point that VCT can discern it is forest, with IFZ values remaining above the forest threshold.

Figure 2.3: Virginia VCT “age” map enhanced by reclassifying disturbed pixels as stand-clearing or not.
The enhanced VCT estimate of forest acres, after removing post-disturbance non-forest pixels, is less than 2 percent higher than the FIA point estimate (Miles 2016). It is also comparable to forest area estimates using the 2011 NLCD land cover map (Homer et al. 2015). The NLCD “shrub/scrub” class includes true shrubs and young trees less than 5 meters tall. Much of what is classified as shrub in Virginia is actually early successional forest or trees stunted from environmental conditions. Therefore two NLCD totals are shown in Figure 2.4. The first includes the deciduous forest, evergreen forest, mixed forest, and woody wetlands classes without shrub. The second also includes the “shrub/scrub” class. The NLCD estimates straddle the lower and upper 95 percent confidence bounds of the FIA estimate, respectively.

Figure 2.5 compares forest area estimates for the enhanced VCT with FIA in Virginia by age class (Miles 2016). These estimates generally coincide, but special note should be made where enhanced VCT estimates fall outside the confidence bounds of FIA. In addition to low area estimates when excluding the pixels that no longer look like forest in the 0-5, 6-10, and 21-25 year age classes, and the high estimates when including these pixels in the 0-5 and 11-15 year age classes, the enhanced VCT also overestimates forest acres when compared with FIA in the over 25 age group, even when post-disturbance non-forest pixels are excluded. The estimate without post-disturbance non-forest pixels in this age group is 12.5 million and rises to 12.7 million when these pixels are included. Possible reasons for these differences will be discussed in the next section.

Figure 2.6 shows the tendency of clumps of neighboring reclassified VCT disturbances of the same “age” to conform to actual harvest boundaries. These clumps also frequently conform to parcel boundaries (Figure 2.7). Figure 2.8 shows how the resulting enhanced VCT gets rid
of isolated pixels and provides a more realistic determination of stand “age” by reclassifying disturbed pixels in a 2010 thinning as a partial disturbance, giving the accurate number of years since the most recent stand-clearing disturbance.
2.5 Discussion

This study demonstrates both the value of comprehensive harvest records for use in training and validating machine learning models, as well as the ability to successfully utilize shape metrics in an automated manner to classify disturbance clumps by harvest method. It is the intent that similar automated methods will be used to reclassify VCT disturbances and
create forest age maps and harvest boundaries in other areas of interest.

In an effort to maximize automation for ease and efficiency in creating future age maps and harvest boundaries, the full capacity of the machine learning tools for correctly classifying disturbances as clearcut or partial harvests was not realized. Features were chosen without
Figure 2.7: Depiction of conformance of enhanced VCT stand-clearing disturbances to parcel boundaries.

Figure 2.8: The enhanced VCT map reduces salt and pepper effects and more accurately depicts stand age by reclassifying disturbances such as this 2010 thinning as a partial disturbance before calculating years since most recent stand-clearing disturbance.

testing to see which set might produce the highest accuracy rates. Also, algorithm parameters were not tuned to maximum accuracy. Some strategies, that would not hinder automation while improving accuracy, can and will be implemented. Four disturbance magnitudes and
various shape and size metrics were calculated for each disturbance clump. One simple method for selecting from these variables would be to exclude one variable from each pair of highly correlated variables while keeping all of the remaining variables. The results here show that good accuracy rates can be obtained without tuning model parameters. While the other algorithms required no parameter specification, k=9 was arbitrarily chosen for kNN. Perhaps an arbitrary parameter specification, such as k=9, could be used for kNN in the future, or kNN could be replaced by another algorithm that does not require parameter specification.

It is possible that the method used to sample clumps of disturbed pixels is biased toward clearcuts over partial harvests. Partial harvests often appear as multiple disjoint clumps in the same parcel rather than one solid clump like most clearcuts. Thus, it seems that there is a smaller chance that the logging deck of a partial harvest will intersect a VCT disturbance clump. The actual proportion of disturbances that are stand-clearing is likely to be somewhere between the proportion observed in the sample and 50 percent. It is unlikely to be below 50 percent because most partial harvests are followed by a clearcut, but it is impossible for a stand to be partially harvested after it has already been cleared. An improved sampling process can overcome this deficiency by intersecting harvests with parcels and then parcels with disturbances. Otherwise, correcting for the bias in this sample indicates that if VCT detects all stand-clearing disturbances and an equal area of partial disturbances, the overall map accuracy rate could be as low as 76 percent. On the other hand, by including parcel data, it is likely that the number of disjoint clumps in a parcel could be a useful feature to help to distinguish partial harvests and improve accuracy. Partial harvests are more likely to appear as several disjoint clumps within the same parcel, while
clearcuts most often appear as one contiguous clump.

In addition, the model used to classify harvest disturbances as stand clearing or not is also used for other disturbances detected by VCT. This would indicate that the sample is biased toward harvest disturbances. The concern here seems to be minimal because harvests are the overwhelming disturbance type by area in Virginia. Other disturbances are likely to be so small that they will not affect stand age. Nonetheless, opportunities exist to reclassify VCT disturbances using auxiliary GIS data. These include conformance with parcel boundaries, shape metrics of disturbance clumps, LiDAR forest structure metrics, and variables such as distance to road and slope. Harvests are unlikely to occur in areas that are far from a road or with steep slope.

It should be noted that because a percentage of the pixels with an associated age have converted to non-forest and the number of years until regeneration begins depends on forest type and management intensity, there is room for improvement in the enhanced VCT “age” map. An analysis of the percentage of pixels that were classified as cleared, in regards to the methods of this paper, and return to forest in one-year increments after disturbance, according to VCT, is underway. Percentages can be calculated by forest type and/or ecoregion in order to shed light on differences in time to detect regeneration using VCT.

When considering the area estimates in Figure 2.4 and Figure 2.5 it is important to note that there are differences in how forest is defined by VCT, FIA, and NLCD (Huang et al. 2010, Bechtold and Patterson 2005, Homer et al 2015). VCT uses a threshold cutoff of an integrated measure of number of standard deviations above the mean a pixel’s brightness in the red and two shortwave infrared Landsat bands is compared to the average of a forest sample. FIA imposes area, width, stocking, and use restrictions which are not accounted for
with VCT. The figures for NLCD do not include developed open space which can include vegetation planted in developed areas for recreation, erosion control, or aesthetic purposes. In addition the shrub/scrub class can include forest areas where there are young trees in an early successional stage that are less than 5 meters tall. However, this class and the woody wetlands class can also include true shrubs, of which there is very little in Virginia. It is reasonable to believe that the total forest acres estimate for the enhanced VCT exceeds FIA because there are less restrictions on what it defines as forest.

Furthermore, higher estimates for enhanced VCT over FIA are not consistent across age classes when excluding VCT classified post-disturbance non-forest. This reveals a time lag after a disturbance before VCT can detect a return to forest. It is conceivable that the reclassified enhanced VCT stand-clearing disturbance maps and the enhanced VCT age map derived from them can be intelligently combined with information from NLCD or a similar land cover map to arrive at a more accurate age map. For instance, a disproportionately high number of pixels that are classified as forest by VCT but grassland by NLCD are in the 0-5 age group. Thus, VCT forest pixels estimated to be older according to enhanced VCT and are also classified as grassland according to NLCD are less likely to be forest with the correct age estimate. Perhaps some of these pixels were incorrectly classified by VCT as forested and some are actually young forest with incorrect older age estimates.

2.6 Conclusion

While the enhanced VCT product is a good proxy for “age” and generally conforms to harvest boundaries, some work remains. Despite the overwhelming presence of secondary
succession forest in Virginia due to clearcut harvesting practices in the south, a majority of
the forest in Virginia has not been disturbed since 1984. Therefore its age cannot be precisely
determined, and large clumps of undisturbed forest cannot be broken up into stand-sized
pieces using reclassified VCT disturbances. These stand-sized pieces would be a good basis
for modeling future change. It may be possible to use variables such as LiDAR height and
structure metrics along with other remotely sensed and auxiliary GIS data to create these
stand-sized units.

Additional work needs to be done to generate identities for unique harvests rather than
unique clusters. Partially forested parcels may include disjoint clumps of harvested pixels
even if the harvest was a clearcut. Therefore parcel data should be combined with VCT data
to identify disjoint clearcut harvest clumps that are most likely part of the same clearcut
harvest. This could be important for modeling efforts in which only harvests that meet a
minimum area threshold are considered. For instance, small harvests may be excluded when
modeling commercial wood supply. Further processing should be done to combine adjacent
clumps that differ by one VCT year because a single harvest spanned consecutive VCT years.

Even without addressing these imperfections, reclassifying clumps of disturbed pixels in
yearly VCT maps as stand-clearing disturbances or not based on average disturbance mag-
nitude, shape, and size results in a good proxy for “age” and objects that conform to harvest
boundaries. The net result is an historical record of harvest boundaries that can also be
used to predict when and where future harvests will occur. A decades long historical record
begins to shed light on the impact of variables such as policy change, social and cultural
values, and ownership demographics on harvesting practices, although their overall impact
may not be known for many more decades. Estimates of biomass or timber volume across
time provide valuable data for procurement foresters, landscape ecologists, climate scientists,
water resource experts, and many others.

Acknowledgments

We are very grateful to Chengquan Huang for generously sharing the data on which this
work is based. Special thanks to the anonymous reviewers whose dedication and insightful
feedback significantly improved our manuscript. Thanks also to the Center for Natural
Resources Assessment and Decision Support and its partners for their support of this project.

2.7 References

Program - national sampling design and estimation procedures. USDA For. Serv. Gen.

trees. Wadsworth.

Hijmans, R.J. 2015. raster: Geographic data analysis and modeling. R package version 2.4-

Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold,
N.D., Wickham, J.D., and Megown, K. 2015. Completion of the 2011 National Land Cover
Database for the conterminous United States-Representing a decade of land cover change

machines for land cover classification. International Journal of Remote Sensing. 23(4):725-
749.

Use of a dark object concept and support vector machines to automate forest cover change


Chapter 3

Automated Harvest Detection and Delineation Using Vegetation Change Tracker and County Parcel Data

3.1 Introduction

A forest stand can be defined as a contiguous group of trees that are homogeneous in condition, species composition, age distribution, and ownership. Stage of stand development is a major component of forest growth and yield modeling and is closely related to stand age (Avery and Burkhart 2002). Extensive information on the distribution of forest resources both spatially and by age-class provides the required knowledge of timber volume through time and across the landscape to allow for a sustained production forest management philosophy (Bettinger et al. 2010). One definition of forest stand age is the number of years since regeneration. Regeneration is often preceded by a clearcut harvest. Therefore, the precise locations and timing of clearcut harvests is of considerable interest.

Opposition to timber harvesting has all but vanquished it from public lands in the United
States, leading to a dramatic decline in timber produced in the western U.S. over the last few decades. This, along with technological and silvicultural advances, has led to intensified timber production in the southern U.S. (Wear and Greis 2012). The southern forest, that was mostly cutover forestland and degraded agricultural land several decades ago, has regrown. Today southern pine plantations are the most intensively managed forests in the world, rivaling southern hemisphere fast-growing exotic species (Fox et al. 2007). Due to the impact of industrial forestry productivity and plantation silviculture practices in the region leading to earlier financial maturity, forests in the southern U.S. have shorter average rotations than the rest of North America, with more than 80 percent of forest stands less than 60 years old (Pan et al. 2011). This suggests that areas not in plantation forestry are still harvested frequently. However, the Appalachian mountain regions are harvested less frequently due to the predominance of hardwood forests, slope and access restrictions, and less amenability to plantation silviculture. Here, large areas of cleared forest are often the result of mining.

Cohen et al. (2016) analyzed forest disturbance rate by type over time using remotely sensed data in five geographic regions. In the southeast, harvest was the dominant disturbance agent over decline, fire, and other types of disturbances for the period from 1985 through 2012. Schleeweis et al. (2011) relate canopy change observations and causal processes at multiple scales using satellite data and ancillary geospatial data for fires, hurricanes and tornadoes, suburbanization/urbanization, harvests, and insects and pathogens. They also include thorough regional analyses in which a high harvest rate for the southeast is confirmed.

The study of forest disturbances and harvest patterns can be thought of as a subset of Land Change Science (LCS). Land Change Science is described as a coupled human-natural
system (Turner et al. 2007). Relationships and dynamics in the human subsystem, such as societal structures, policy and decision-making, and economic markets effect change in human land use, land cover, and ultimately the biophysical subsystem. Conversely, structures and changes in the biophysical subsystem, such as environmental and ecological goods, services, and properties cause change in land cover, human land use, and consequently the human subsystem. Study within this field involves observation and monitoring of land use, land cover, the human subsystem, the biophysical subsystem, forces operating exogenously on the system, and the effects of the system on external processes (Turner et al. 2007). Thereby, LCS can be thought of as a complex subsystem of an even more complex system.

A 2007 special issue of the Proceedings of the National Academy of Sciences (PNAS) presented the emergence of LCS as a fundamental component of environmental change and sustainability research (Turner et al. 2007). Loss of forest cover and changes in land use due to mining, conversion to agriculture, urban expansion, low density housing, parcelization, fragmentation, and shifts in forest type establish a context for the distribution of forest spatially and temporally by age-class, ownership, and type in LCS.

The body of LCS research incorporates various tools that are useful for analysis. Remotely sensed data has become an integral component of LCS research. Land cover maps derived from remote sensing, when combined with information on land use, shed light on structure and dynamics within the LCS system. Advances in air and space-borne sensors enhance the ability to monitor many types of land use and land cover changes, including those resulting from forest disturbances. This facilitates global assessments in changes in net primary productivity, flora, carbon sources and sinks, water and food supply, and wildlife
biodiversity (Turner et al. 2007).

Therefore, stand age mapping can be thought of as a component of both forest resource monitoring and LCS. Its applications include growth and yield estimates at the stand level, and spatially and temporally precise wildlife habitat, water quality, and carbon stock assessment. In addition to timing and location of clearcut harvests, information on partial harvest timing and location can facilitate further insight into the dynamics of forest resources. Tools and data sources for stand age mapping are especially useful in Virginia and throughout the southern US where timber harvesting remains frequent.

3.1.1 Sources of forest stand age and volume information

Two sources of forest stand age distribution and volume information are forest inventory plots and remotely sensed data.

Forest Inventory and Analysis

The three-phase sampling scheme of the US Forest Service Forest Inventory and Analysis (FIA) program, described by Bechtold and Patterson (2005), forms a comprehensive inventory of the nation’s forest resources. Precise estimates are attainable for forest-wide inventories at the county level and larger. A Phase 1 remote sensing component is designed to reduce variance through stratification, while Phase 2 consists of a systematic field measurement of randomly placed plots within a hexagonal grid system covering the continental United States. Each plot is assigned a land use code, and for those that meet the forest land use definition, an FIA field crew records each tree along with its species, status, and other
attributes. Diameter at breast height (dbh) and height of each tree is measured, enabling estimation of cubic foot volume or other product and resource quantities. Plots can contain multiple ‘conditions’ which separate forest stands by ownership type, forest type, and age (USDA Forest Service 2014). FIA’s spatially and temporally comprehensive approach, which samples all of the nation’s forested lands over a periodic cycle is appropriate for monitoring status and trends of forest inventory (Reams et al. 1999; Bechtold and Patterson 2005). The Phase 3 forest health monitoring plots are a subset of Phase 2 plots and measure additional health indicator attributes yearly (Bechtold and Patterson 2005). Information from FIA on abundance, trends, and geographic distributions of products and resources can assist regional managers and policy makers by spatially monitoring their availability and sustainable use.

However, regional forest inventories such as FIA are not intended for stand-level estimation. Much forest management and planning occurs at the stand level by individual forest owners, especially in the Southeast. Therefore, modeling of forest resources sometimes requires information at this level. It is simply too costly to sample from every individual forest stand at the regional level. Figure 1.1 shows a National Agriculture Imagery Program (NAIP) aerial photo of an approximately 30,000 acre portion of Dinwiddie County, Virginia, one-tenth the size of the entire county. The figure shows hundreds of separately managed areas of forest large enough for commercial timber harvesting, and five FIA plots. Individual management-based inventories requiring a sample of plots for every separately managed area for all of Virginia would be too costly for one organization to employ. Because comprehensive field measured data is not feasible for every individually managed unit, data on these units must be obtained through another source.
Remote sensing

Remote sensing is an important tool for quantifying land use and land cover change. A review of best practices of change detection methods recommends Landsat, with its 30m pixel size, for monitoring land use/land cover and disturbances (Willis 2015). Beginning in 1972 with the launch of Landsat 1, the Landsat mission was most recently continued on February 11, 2013 with Landsat 8 (Roy et al. 2015) and is expected to be extended with Landsat 9 in 2023 (Wulder et al. 2015). This continuing archive held by the USGS allows for consistently processed, moderate spatial resolution, large area, long-term, terrestrial data records for resource management and global change studies and was made freely available in 2008 (Roy et al. 2015; Wulder et al. 2012). In 2010, when the USGS began its Landsat Global Archive Consolidation Mission, more data was held outside its archive than inside.
Since then, the archive’s holdings have more than doubled with valuable past records. Its availability has facilitated the ability to create Landsat Time Series Stacks (LTSS) for an array of uses, including disturbance detection algorithms (Roy et al. 2015; Wulder et al. 2012).

These disturbance detection algorithms, such as Land Trendr, Vegetation Change Tracker (VCT), Global Forest Watch (GFW), Exponentially Weighted Moving Average Change Detection (EWMACD), and others (Kennedy et al. 2010; Huang et al. 2010; Hansen et al. 2013; Brooks et al. 2014) are of immediate application to the research here. The highly automated nature of these algorithms is valuable for rapid processing of large amounts of data. Some have the ability to detect abrupt disturbance events, such as fire, wind, harvest, conversion to agriculture, and urban development, while others are better suited for detecting more gradual agents of change, including insects, disease, and stress.

Hansen et al. (2013) quantify gross forest cover gain and loss in their publicly available maps on the GFW website (http://www.globalforestwatch.org/). These maps are updated continuously using Google Earth Engine and a 50 percent threshold tree crown cover density definition of forest (Hansen et al. 2013). They show high amounts of both loss and gain in the southern U.S. as their forest ‘cover’ definition of forest detects loss due to harvests and gain when forest regrows. Because remote sensing is limited to a ‘cover’ definition of forest, Reams et al. (2010) argue for the importance of field data in forest cover and use change monitoring in order to clearly assess use. Certainly, LCS is not at a point where it can rely entirely on remote sensing instruments for measuring all land use change.

However, some Landsat-based disturbance detection algorithms can be relied on to pre-
dominately detect all disturbances in a landscape that cause removal or mortality of most trees within a stand (Kauffman and Prisley 2016). We define these types of disturbances as stand-clearing. Some causes of stand-clearing disturbances are clearcut harvest removals, wide-spread mortality due to wind or fire, and conversion from forest to agriculture and developed areas. Most stand-clearing disturbances that do not result in conversion to non-forest create even-aged stands (Bettinger et al. 2010). Even-aged stands are those for which ages of all trees in the stand are generally within 20 percent of the average, while uneven-aged stands are those with at least two distinct age ranges (Bettinger et al. 2010). For even-aged stands there are many definitions of stand age, all related to time since regeneration, but for uneven-aged stands age must be calculated by some manner of average total age of the individual trees in the stand (Bettinger et al. 2010). It can be argued that even if stands resulting from a stand-clearing disturbance are converted to uneven-aged stands through selection harvest management practices, their age will be closely related to time since regeneration during the early years of stand development.

Disturbance clumps derived from VCT disturbance maps are aggregates of disturbed pixels, as defined in Kauffman and Prisley (2016). While their boundaries may not match exactly with stands delineated from aerial photography, due to image registration and pixel size, the alignment is quite good for clearcut harvests. The match is more difficult with non-stand-clearing disturbances, such as thinnings, in which all of the pixels in the disturbed area may not be detected because at least part of the canopy often remains in a 30 meter Landsat pixel. Also, disturbance clumps may not match stands delineated by aerial photography because it is often advantageous for forest owners and managers to harvest neighboring stands, of different type, simultaneously. Furthermore, natural disturbances caused by fire and wind
have no regard for forest type, stand age, or management boundaries.

Maps derived from disturbance detection algorithms, such as the VCT, can be used to calculate a good proxy for age at the pixel or stand level by measuring the number of years since the last stand-clearing disturbance if forest regeneration occurs in those areas (Huang et al 2010, Kauffman and Prisley 2016). This facilitates age and volume estimates at a stand-sized scale across a large area.

3.1.2 Prior work and opportunities for refinement

Age and harvest boundaries from Vegetation Change Tracker

Detection of forest disturbances at the pixel level via the VCT algorithm relies on two steps. First, yearly time series values of integrated forest z-scores (IFZ) are calculated for each pixel after masking out other known dark surfaces such as water and dark soils from Landsat images dating back to 1984 (Huang et al. 2008, Huang et al. 2010). These IFZ pixel values can be thought of as measures of forest likelihood. They are calculated as normalized deviations from the mean reflectance of forest training pixels from a local image window in the red and shortwave infrared bands (bands 3, 5, and 7 from Landsat 5 and Landsat 7). Next, the yearly time series of IFZ values for each pixel is analyzed to determine when pixels with typically high forest likelihood suddenly exhibit low forest likelihood. In years when this occurs, the pixel is classified as disturbed forest. The VCT algorithm creates maps of disturbed forest pixels for each year since 1984 (Huang et al. 2010).

Kauffman and Prisley (2016) used three supervised machine learning algorithms for auto-
mated estimation of forest stand age and harvest delineation. Connected groups of disturbed pixels in the eight-pixel queen’s case neighborhood, from the time series of yearly VCT disturbance maps, were assumed to be the result of the same forest disturbance and were clumped together and given a common identity. The clumps were subsequently reclassified, using the machine learning tools, as stand-clearing disturbances or not, based on several metrics calculated for each clump. The variables used for reclassification included shape and size metrics common in landscape ecology and calculated using the SDMtools package in R (VanDerWal et al. 2014). Also included were the average disturbance magnitudes of pixels within each disturbance clump, measured four ways, as changes in values of IFZ, normalized difference vegetation index (NDVI), normalized burn ratio (NBR), and IFZ4, which is similar to IFZ but alternatively uses the single Landsat band 4 rather than an integrated z-score from Landsat 5 and Landsat 7 bands 3, 5, and 7.

Each VCT disturbance map was reclassified in this way, resulting in a time series stack of reclassified stand-clearing disturbance maps. From these maps the number of years since the most recent stand-clearing disturbance could be calculated for each pixel as a proxy for age. The result was the enhanced VCT (eVCT) age map product defined by the authors.

The resulting map demonstrates the additional benefit of using it as a mechanism to delineate harvest boundaries and partition the forestland of Virginia into spatially precise forest units or objects at the harvest-sized level. These forest units provide an appropriate scale for which each can be assigned volume estimates in four timber product classes (hardwood pulpwood, softwood pulpwood, hardwood sawtimber, and softwood sawtimber) for modeling and forecasting timber supply. The use of R for this task was advantageous because it allowed
for automated processing of large amounts of data, specifically 28 disturbance maps, one for each year from 1984 through 2011 for an area equivalent to one Landsat scene. Fifteen different Landsat scenes are required to cover all of Virginia.

**Tuning and comparing algorithms**

Supervised machine learning algorithms used in the above reclassification process included support vector machine (SVM), rpart, and k nearest neighbors (kNN). The ‘rpart’ package in R is an implementation of the CART algorithm and follows Breiman et al. (1984) in most details (Therneau et al. 2015). Kauffman and Prisley (2016) demonstrated that these tools in conjunction with shape, size, and disturbance magnitude metrics are valuable components for successfully reclassifying disturbance clumps as stand-clearing or not. Other algorithms that have not previously been tested with the available data include logistic regression and random forest (R Core Team 2015, Liaw and Wiener 2002). Proper specification of a logistic regression model requires that model assumptions can be met but may be less computationally intensive than other tools. Random forest may outperform other models in terms of accuracy, but could also be limited by computational speed in order to generate an ensemble of predictions for each disturbance clump considering the number of trees in the random forest. If computational speed is not an issue, an ensemble approach using multiple tools may be appropriate if it leads to improved results.

The accuracy of kNN can be improved by choosing an optimal number of neighbors, k. Kauffman and Prisley (2016) chose k=9. Instead, classification could be repeated for varying numbers of neighbors, and the value of k that yields the highest accuracy rate could be chosen.
A similar grid search approach could be used to optimize the multiple parameters of SVM.

While the kNN algorithm is easy to understand and simple to employ in R, it comes with two major disadvantages. A Euclidean distance in the feature space must be calculated between every disturbance clump that must be classified and every disturbance clump used for training. This becomes even more computationally intensive when the number of variables increases (Richards 2013). Also, it is biased toward the class which is more prevalent in the training sample, especially when the number of nearest neighbors, k, is larger. On the edge of a cluster where it may approach or overlap another cluster, the more prevalent can win out just because there are more of them overall.

Further effort to calibrate parameters of SVM and rpart could also result in improved accuracy. One of the advantages of SVM over rpart is that it allows for non-linear partitioning of the data through choice of the kernel function. This would give it some advantage over rpart in which each branch of a tree can only partition the dataset perpendicularly to one of the axes. However, the specification for SVM used a linear kernel, so this potential advantage could not be fully exploited. On the other hand, an advantage of rpart is that it can perform well with a large number of variables without any danger of breaking any model assumptions, such as the uncorrelated dependent variable assumption required for logistic regression.

Kauffman and Prisley (2016) did not include comparison of models with various feature selections. A key characteristic of random forest is that the importance of each feature over a forest of classification trees can be quantified and ranked. This can be used in a feature selection process for the random forest algorithm or another supervised machine learning tool.
SVM allows for different kernels including linear, polynomial, radial basis, and sigmoid (Meyer et al. 2015). Polynomial and radial basis function kernels are the two most common in remote sensing and are known to work well (Richards 2013). If a function that best separates features in Kauffman and Prisley (2016) is not linear, the polynomial kernel function is a likely candidate. It is appropriate to tune parameters specific to the kernel function via a trial-and-error or grid searching technique which narrows the spacing in the vicinity of the optimal parameters (Richards 2013).

The confusion matrix in Kauffman and Prisley (2016) shows good classification performance for the stand-clearing disturbance class (94.9 percent), however, the omission error for the non-stand-clearing disturbance class was high (42.5 percent). This means that there are a substantial number of non-stand-clearing clumps that are being classified as stand-clearing disturbances. The effect is that these misclassified disturbance clumps were possibly cleared further in the past than the age map shows, maybe never, depending on whether the misclassified non-stand-clearing disturbance occurred before or after the latest actual stand-clearing disturbance. In other words, there are a good number of forested pixels that are actually older than what is shown on the map, and this is more likely for pixels that were disturbed more recently.

There are several possible opportunities to improve classification accuracy for the non-stand-clearing disturbance class. First, other classification algorithms may have better success than those already chosen, including an ensemble of algorithms. Another opportunity to improve classification accuracy is to more properly calibrate the parameters for the algorithms that have already been tried. Also, perhaps there are variables that were not previously included.
that would improve accuracy. Finally, improved sampling techniques may more appropriately represent the non-stand-clearing class.

**Sampling procedures and parcel data**

A large portion of disturbances conform to property boundaries and seem far more likely to be anthropogenic (i.e., harvests or deliberate clearing). Therefore, we suspected that the inclusion of parcel boundaries could aid significantly in three areas: (1) improved sampling methods, (2) expansion of the array of quality features for use in the reclassification algorithm, (3) enhancing the ability to also distinguish between partial harvests and non-harvest disturbances/errors in the disturbance map using these same features.

The Virginia Department of Forestry (VDOF) began keeping records of all harvests in Virginia in 2009 in conjunction with inspections of best management practices (BMPs). The GPS point location of the first logging deck that the BMP inspector encounters on the date of inspection is collected along with other harvest attributes, including harvest type. Harvests are classified into three types, “clearcut”, “partial harvest (including thinning)”, and “biomass”. Biomass harvests are clearcuts that remove more woody material than a normal roundwood harvest. Typically, there are five or six thousand harvests per year in Virginia. Only harvest records that undoubtedly correspond to 2010 VCT disturbance maps for Virginia were used for eVCT. This includes harvests after the latest Landsat scene used for the 2009 VCT map (September 14th) and before the earliest Landsat scene used for the 2010 VCT disturbance map (June 3rd). For eVCT these harvest records were used to select the disturbance clumps that were used to calibrate, test, and validate the supervised machine
learning classification algorithms. It is possible that this method for selecting the sample clumps of disturbed pixels was biased toward stand-clearing disturbances over non-stand-clearing disturbances. Partial harvests often appear as multiple disjoint clumps in the same parcel rather than one solid clump like most clearcuts. Thus, it seems that there is a smaller chance that the logging deck of a partial harvest will intersect a VCT disturbance clump. An improved sampling process can overcome this deficiency by intersecting harvests with parcels and then parcels with disturbances.

In addition, it was noticed that harvest boundaries often conform to parcel boundaries. Forest landowners cannot legally harvest what they do not own. Additional metrics based on parcel boundaries are reasonable candidates for expanding the array of features used to reclassify disturbance clumps. For instance, it is likely that the number of disjoint clumps in a parcel could be a useful feature to help distinguish between stand-clearing and partial harvests thereby improving accuracy. After all, partial harvests are more likely to appear as several disjoint clumps within the same parcel, while clearcuts most often appear as one contiguous clump.

Preliminary analysis of the intersection of parcel data with a 2010 VCT disturbance map for Landsat path 15/row 34 and VDOF harvest records that correspond to the latest Landsat scene used for the 2009 VCT map (September 14th) and the earliest Landsat scene used for the 2010 VCT disturbance map (June 3rd) reveals an interesting and noteworthy fact. Table 3.1 illustrates that parcels that contain point locations of partial harvests (including thinnings) are nearly as likely to also contain a VCT classified forest disturbance as parcels that contain point locations of clearcut harvests, and they are over fifteen times more likely
Table 3.1: Proportion of parcels containing a VDOF harvest record from each harvest class that also contain a VCT disturbance

<table>
<thead>
<tr>
<th>Clearcut</th>
<th>Partial</th>
<th>Unknown/multiple</th>
<th>None</th>
<th>All Parcels</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.954</td>
<td>0.938</td>
<td>0.957</td>
<td>0.062</td>
<td>0.063</td>
</tr>
</tbody>
</table>

to contain a VCT classified forest disturbance than parcels where no VDOF harvest was recorded within the September 14th to June 3rd time frame. It should be noted that parcels where “no harvest” occurred may actually contain harvests outside these most likely dates of correspondence to yearly Landsat images used for VCT. Therefore the proportion of parcels that contain “no harvest” and a VCT disturbance is inflated in the table. The revelation that most parcels that contain a partial harvest record also contain a VCT disturbance prompts the question “is it possible to develop tools to accurately classify VCT disturbances as partial harvests and create historic maps of partial harvests?”

3.1.3 Research objectives

The primary focus of this work is to further enhance reclassification of VCT disturbances as stand-clearing disturbances, partial harvests, or small non-harvest related disturbances, thereby increasing the value of reclassified VCT data products. Note that for stand age mapping it is only necessary to distinguish between stand-clearing disturbances and non-stand-clearing disturbances. Therefore, this was the main classification objective. It is also valuable to map partial harvests. Our strategy for accomplishing this is to add a second stage of classification that separates harvest disturbances from non-harvest disturbances. Disturbances classified as non-stand-clearing in the main classification stage and as harvests in the second classification stage can be mapped as partial harvests. To this end, we evaluate
the following hypotheses:

1. A combination of adding metrics derived from parcel boundaries, modifying the sampling procedure, including other potential machine learning tools, and including a more thorough process for tuning and comparing possible machine learning tools will improve classification rates of stand-clearing versus non-stand-clearing disturbances.

2. Higher reclassification accuracy rates will lead to a more accurate historical record of yearly maps of stand-clearing disturbances from 1984 through 2011 which can also be used to create more accurate forest stand age maps and include boundaries of past clearcut harvests and other stand-clearing disturbances.

3. The same sampling procedures, metrics, and machine learning tools that facilitate improved accuracy rates for distinguishing between stand-clearing and non-stand-clearing disturbances will also work well for reclassifying disturbances as harvest-related or not and allow for the partitioning of disturbances into stand-clearing disturbances, partial harvests, and small non-harvest related disturbances.

4. As an example of further applications using data derived from reclassified VCT data, reclassified VCT disturbance maps combined with forest and non-forest classes from original VCT disturbance maps can be used to create informative charts of the proportion of pixels that return to a forested state after a stand-clearing disturbance by number of years since the disturbance.
3.2 Methods

The research involved compiling remote sensing and field data for a portion of the southern piedmont and coastal plain. This data included VCT disturbance maps, VDOF harvest records, parcel boundary data, and Google Earth imagery.

3.2.1 Study area and data

Landsat path 15/row 34 was chosen as the study area because it is the single scene that incorporates most of the Virginia southern piedmont and coastal plain ecoregions which contain some of the most actively managed forests in Virginia. The research involved compiling remote sensing derived VCT data, parcel boundary data, and field measured VDOF harvest records for the study area. The harvest records were used to validate the accuracy of reclassifying VCT disturbances as stand-clearing disturbances, partial harvests, and non-harvest disturbances. In addition, Google Earth (http://earth.google.com/) was used to separately analyze the accuracy of stand age maps derived from a time series of stand-clearing disturbance maps.

VCT remote sensing data

The VCT algorithm has proven to be a reliable method for detecting forest disturbances using time series stacks of Landsat data (Huang et al. 2010, Kauffman and Prisley 2016). VCT yearly disturbance maps and disturbance magnitude maps were obtained for Landsat path 15/row 34. The disturbance maps are classified into six groups: persisting non-forest, non-
forest after a disturbance, persisting forest, forest after a disturbance, forest disturbed in the current year, and water. Clumps of neighboring disturbed pixels obtained from these maps will be classified as stand-clearing disturbances, partial harvests, or non-harvest disturbances less than 5 acres. Time series of yearly VCT disturbance magnitude (IFZ, NDVI, NBR, and IFZ4) maps of disturbed pixels were evaluated for potential use as input variables in algorithms that facilitate this reclassification.

**VDOF harvest field data**

VDOF harvest records that undoubtedly correspond to 2010 VCT disturbances, as described in the introduction, and that are located within the study area were used to calibrate, test, and validate potential classification algorithms.
Parcel data

Parcel data from within the study area was obtained for 44 counties and municipalities from the Virginia Geographic Information Network repository of county level parcel data. A portion of these counties is only partially contained in Landsat path 15/row 34. Six county parcel shapefiles were not readily available at the time of download. Therefore, the study area for disturbance type classification consists of the portions of the counties within path 15/row 34 with available parcel data as pictured in Figure 3.2.

Google Earth imagery

Google Earth was used to determine the approximate year of the most recent stand-clearing disturbance for a random selection of 389 VCT-classified forested points. Google Earth typically supplies a time series of aerial photos throughout the study area beginning approximately in 1994. After 1994 there is usually a gap until the next available image, often around 2003. After 2003 more frequent photos are available.

3.2.2 Disturbance clump metrics

As in Kauffman and Prisley’s (2016) eVCT methods, connected groups of disturbed pixels in the eight-pixel queen’s case neighborhood were assumed to be the result of the same forest disturbance and were clumped together and given a common identity. However, as opposed to eVCT, there was no minimum size clump. For each disturbance clump the average IFZ, NDVI, NBR, and IFZ4 disturbance magnitudes of the pixels comprising the clump were calculated as in the methods of Kauffman and Prisley (2016). In addition, Kauffman and
Prisley’s (2016) methods were used to calculate shape and size metrics with the ‘PatchStat’ function in the ‘SDMtools’ R package (VanDerWal et al. 2014).

Various additional metrics were calculated beyond the methods of Kauffman and Prisley (2016). First, it was noticed that many pixels on the edge of a forest/non-forest boundary are misidentified as disturbed due to misregistration. Therefore, a count of forested pixels in the eight-pixel queen’s case neighborhood of each pixel in a disturbance clump was obtained and an average count for all the pixels in the entire clump was aggregated for each clump. All additional metrics were calculated by intersecting the VCT disturbance map with parcel boundaries in a county-by-county fashion in order to obtain counts and ratios of pixels, disturbed pixels, and disturbance clumps. A complete list of metrics can be seen in Table 3.2.

3.2.3 Sampling methods

Tract size plays an important role in landowner willingness to sell and economies of scale for loggers, with tracts smaller than 20 acres traditionally thought to be frequently cost prohibitive (Row 1978, Cubbage 1982, Greene et al. 1997). According to VDOF harvest records, less than one percent of harvested acres occur on tracts less than five acres. In order to focus on areas where harvests are likely to occur in the future, sampled VCT disturbances classified as harvests were limited to those five acres or more. Perhaps this threshold could be raised to suit other preferences. However, there is evidence that with increased parcelization loggers may be willing to harvest on smaller tracts rather than pay a premium on other costs, such as transportation (Moldenhauer and Bolding 2009).
<table>
<thead>
<tr>
<th>METRIC</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 IFZ.distbMagn</td>
<td>change magnitude of IFZ x 10</td>
</tr>
<tr>
<td>2 NDVI.distbMagn</td>
<td>change magnitude of NDVI x 100 + 100</td>
</tr>
<tr>
<td>3 NBR.distbMagn</td>
<td>change magnitude of NBR x 100 + 100</td>
</tr>
<tr>
<td>4 IFZ4.distbMagn</td>
<td>change magnitude of band 4 forest index x 10</td>
</tr>
<tr>
<td>5 core.area.index</td>
<td>quantifies core area as a percentage of clump area</td>
</tr>
<tr>
<td>6 n.cell</td>
<td>the number of cells for each clump</td>
</tr>
<tr>
<td>7 n.core.cell</td>
<td>the number of cells in the core area without the edge area</td>
</tr>
<tr>
<td>8 n.edges.internal</td>
<td>the number of internal cell edges of the clump</td>
</tr>
<tr>
<td>9 n.edges.perimeter</td>
<td>the number of outer perimeter cell edges of the clump</td>
</tr>
<tr>
<td>10 perim.area.ratio</td>
<td>number of outer perimeter cell edges per cell: n.edges.perimeter/n.cell</td>
</tr>
<tr>
<td>11 shape.index</td>
<td>clump perimeter divided by min perimeter of maximally compact patch</td>
</tr>
<tr>
<td>12 mean.forest.cell.count</td>
<td>mean count of forested cells in 9-cell neighborhoods of clump cells</td>
</tr>
<tr>
<td>13 n.cell.parcel</td>
<td>the total number of cells in a parcel</td>
</tr>
<tr>
<td>14 n.distb.parcel</td>
<td>the total number of disturbed cells in a parcel</td>
</tr>
<tr>
<td>15 n.clumps</td>
<td>the total number of clumps in a parcel</td>
</tr>
<tr>
<td>16 mean.clump.size</td>
<td>average number of disturbed pixels per clump: n.distb.parcel/n.clumps</td>
</tr>
<tr>
<td>17 clumps.p.distb.cell</td>
<td>clumps per disturbed pixel: n.clumps/n.distb.parcel</td>
</tr>
<tr>
<td>18 clumps.p.cell.parcel</td>
<td>clumps per parcel pixel: n.clumps/n.cell.parcel</td>
</tr>
<tr>
<td>19 clumps.p.distb.parcel</td>
<td>clumps per disturbed pixel per pixel: (n.clumps * n.cell.parcel)/n.distb.parcel</td>
</tr>
<tr>
<td>20 distb.proportion</td>
<td>proportion of pixels disturbed: n.distb.parcel/n.cell.parcel</td>
</tr>
<tr>
<td>21 n.cell.distb.parcel</td>
<td>the number of cells of a disturbance clump within a parcel</td>
</tr>
<tr>
<td>22 parcel.cell.per.total</td>
<td>proportion of disturbance clump within a parcel: n.cell.distb.parcel/n.cell</td>
</tr>
</tbody>
</table>
The parcel and harvest data were utilized to select a sample, stratified by harvest type, of VCT disturbance clumps from the 2010 VCT disturbance map (Table 3.3). First, 200 clearcut/biomass harvests at least five acres in size and 200 partial harvests at least five acres in size were randomly selected from the VDOF harvest records. Each harvest point was matched and moved, if necessary, to the interior of the largest disturbance clump representing the harvest in the parcel in which the majority of the harvest occurred. Additionally, 200 parcels that did not contain a harvest were randomly selected. A point was created, within the largest disturbance clump less than five acres, for each of the randomly chosen parcels that did not include a harvest point. Six partial harvest locations did not coincide with a VCT disturbance, therefore these harvests were not represented in the sample with a disturbance clump. Clumps that were five acres or larger and were within the boundaries of a parcel with no harvest were considered to be stand-clearing disturbances and were not included in the non-harvest sample.

<table>
<thead>
<tr>
<th>Disturbance/Harvest Type</th>
<th>n</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stand-clearing harvest ≥ 5 acres</td>
<td>200</td>
<td>industrial clearcut harvests and biomass harvests</td>
</tr>
<tr>
<td>Partial harvest ≥ 5 acres</td>
<td>200</td>
<td>industrial thinnings and selection harvests</td>
</tr>
<tr>
<td>Stand-clearing harvest &lt; 5 acres</td>
<td>0</td>
<td>small clearcut harvests and biomass harvests</td>
</tr>
<tr>
<td>Partial harvest &lt; 5 acres</td>
<td>0</td>
<td>small thinnings and selection harvests</td>
</tr>
<tr>
<td>Stand-clearing non-harvest ≥ 5 acres</td>
<td>0</td>
<td>conv. to high intens. dev./ag., high intens. wind/fire</td>
</tr>
<tr>
<td>Partial non-harvest ≥ 5 acres</td>
<td>0</td>
<td>insect/disease, wind/fire damage</td>
</tr>
<tr>
<td>Stand-clearing non-harvest &lt; 5 acres</td>
<td>200</td>
<td>conversion to developed/agricultural error</td>
</tr>
<tr>
<td>Partial non-harvest &lt; 5 acres</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In this manner, there were three types of disturbance events that were left out of the sample. There were both clearcut and partial harvest records for harvests less than five acres in size.
Maps of boundaries and ages of these disturbances were not deemed critical for our purposes. Certainly this sampling scheme could be altered to suit other needs. Stand-clearing non-harvests of at least five acres in size were also not included in the sample. While a correct year of most recent stand-clearing disturbance can be used to correctly estimate stand age even for disturbed areas attributed to causes other than harvest, we are most interested in estimating ages for forest units defined by ownership rather than intense fire, wind, or other catastrophic natural events.

It should be considered that when classifying by disturbance clump, equal representation of each strata under-represents the overwhelming number of non-harvest disturbance clumps. However, these non-harvest clumps are much smaller in area. Nonetheless, an area-proportional stratification also can not be exact because partial harvest disturbance clumps or groups of clumps rarely cover the entire harvest area.

### 3.2.4 Training, testing, and validation of machine learning algorithms

The sample was randomly divided into training and testing subsamples of two-thirds and one-third of the sample respectively. Ten-fold cross validation with 3 repeats was used on the training sample to calibrate several models with the ‘caret’ R package (Kuhn et al. 2016). These included random forest, SVM, kNN, rpart, and logistic regression. A variable importance plot derived from random forest was created for all variables in Table 3.2 (Liaw and Wiener 2002). For each model type, variables were entered into the model by rank in the variable importance plot, as long as there was less than a 0.50 correlation coefficient.
with respect to each of the previously entered variables, until there was no improvement in classification accuracy for the testing subsample. Models using all of the variables in Table 3.2 were also considered for random forest and rpart. Additionally, model parameters for SVM and kNN were calibrated for peak accuracy with the testing subsample. Finally, the test sample was used to validate accuracy rates for the best model. This entire process was followed to arrive at a model for distinguishing between stand-clearing and non-stand-clearing disturbances. This will be called the ‘stand-clearing model’. To allow for the possibility of the use of a different machine learning tool and selection of variables and model parameters, this process was repeated for determining the model for distinguishing between harvests and non-harvests. This will be called the ‘harvest model’.

3.2.5 Stand-clearing disturbance maps, age maps, and partial harvest maps

New stand-clearing disturbance maps were used to generate an age map for the study area which was compared to the eVCT age map and the unenhanced VCT map of year of most recent disturbance. Maps of partial harvests were also created.

Map generation

The machine learning tool, selection of features, and calibration of parameters that produced the greatest accuracy rates for stand-clearing versus non-stand-clearing disturbance clumps (the stand-clearing model) was used to classify all VCT disturbance clumps for path 15/row 34. This resulted in a time series stack of reclassified disturbance maps similar to the eVCT
stand-clearing disturbance maps of Kauffman and Prisley (2016). Because this iteration of enhanced VCT maps relies on parcel boundary data it will be referred to here as parcel-enhanced VCT (peVCT). The algorithm used by Kauffman and Prisley (2016) to determine the number of years since the last stand clearing disturbance was again used to create a forest stand age map.

Additionally, the optimal selection of machine learning tool, features, and parameter calibration for distinguishing between harvest and non-harvest disturbances (the harvest model) was used to create a time series stack of non-harvest disturbances. Disturbances that were neither classified as non-harvest nor stand-clearing in either of the time series stacks were considered to be partial harvests, creating a third time series stack of mutually exclusive disturbance events.

**Map comparison**

Once the peVCT age map was created, it was compared with the original VCT ‘most recent year of disturbance’ and the eVCT age map to determine how each performs relative to the others. To this end, these maps were individually compared to ‘most recent year of disturbance’ as determined from Google Earth at 389 randomly selected VCT-classified forested points to identify obvious discrepancies. Proportions of agreement with Google Earth were calculated for all 389 points because some VCT classified persisting forest pixels may have actually been disturbed since 1984. Proportions of agreement were also calculated for only the 205 points where there was a VCT disturbance. All three of these maps are identical for points identified as persisting forest by VCT. Thus, McNemar’s test was used to test
for differences in proportions of agreement with aerial photos from Google Earth for pairs of maps at the 205 points where a VCT disturbance occurred. McNemar’s test is used to test for differences in proportions for dependent matched pairs (Campbell and Wynne 2011, McNemar 1947).

A most likely year of last disturbance was also determined for each sample point by choosing the last disturbance year of the map (VCT, eVCT, or peVCT) that agreed with the possible year range from Google Earth. It is expected that the reclassified eVCT and peVCT age maps are more likely to agree with the aerial photography than the original VCT map, so they were given preference. For instance, when these maps disagreed but both showed no discrepancy with the aerial photography, the difference was split. If the original VCT also agreed it was ignored. The year of last disturbance for the original VCT was used only if it was the only one that agreed with Google Earth. If there was no agreement with Google Earth, the midpoint of the year range in Google Earth was used. When the most likely year of the most recent stand-clearing disturbance was prior to the time series stack of VCT maps, 1983 was used.

The absolute and signed differences in year of last disturbance between Google Earth and each of eVCT and peVCT was calculated. A paired t-test was performed to test for differences in the absolute differences between VCT and eVCT. Additional t-tests were performed individually on the signed differences to see if there was a statistically significant difference between eVCT/peVCT and the Google imagery.
3.2.6 Forest retention and conversion

Yearly maps of stand-clearing disturbances were intersected with subsequent years of forest/non-forest maps derived from VCT. In this manner, each stand-clearing disturbance pixel could be followed in subsequent years to see if it had returned to a VCT forest classification or not. Charts and graphs of the proportion of stand-clearing disturbance pixels, as classified by peVCT, that are subsequently classified as forest by number of years since they were cleared were created for each stand-clearing disturbance map year. These charts were explored and analyzed in order to discover relevant interesting patterns.

3.3 Results

3.3.1 Reclassification accuracy

The stand-clearing models yielded overall accuracy rates for distinguishing between stand-clearing and non-stand-clearing disturbances (Table 3.4). A reduced random forest model performed the best with an overall accuracy rate of approximately 92 percent, slightly better than a full model. This reduced model included IFZ.distbMagn, n.edges.internal, peri.area.ratio, mean.forest.cell.count, and mean.clump.size with the number of variables sampled at each split (mtry) set to 5. Overall accuracies for the models ranged between 87 percent and 92 percent, supporting the use of any of the models for reclassification.

Maintaining these equal-representation strata in the sample and using the restricted random forest model yielded producer’s accuracy rates of 0.8484 for stand-clearing disturbances and 0.9535 for non-stand-clearing disturbances with the test sample.
Table 3.4: Models and overall accuracy rates for stand-clearing/non-stand-clearing classification of disturbance clumps with equal representation of harvest type strata

<table>
<thead>
<tr>
<th>Method</th>
<th>Variables</th>
<th>Parameters</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>rpart</td>
<td>all</td>
<td>cp = 0.03488372</td>
<td>0.876</td>
<td>0.724</td>
</tr>
<tr>
<td>kNN</td>
<td>1, 8, 10, 12</td>
<td>k = 9</td>
<td>0.871</td>
<td>0.719</td>
</tr>
<tr>
<td>logistic reg</td>
<td>1, 8, 10, 12, 16, 19</td>
<td>k = 9</td>
<td>0.888</td>
<td>0.751</td>
</tr>
<tr>
<td>random forest</td>
<td>1, 8, 10, 12, 16</td>
<td>mtry = 4</td>
<td>0.919</td>
<td>0.818</td>
</tr>
<tr>
<td>SVM (polynomial)</td>
<td>1, 8, 10, 12, 16</td>
<td>degree = 2, scale = 0.1, C = 1</td>
<td>0.903</td>
<td>0.786</td>
</tr>
</tbody>
</table>

The harvest model with the best accuracy rate was a full random forest model (with mtry set at 4) at approximately 94 percent, slightly better than a reduced model (see Table 3.5).

The producer's accuracy rates were 0.8750 for non-harvests and 0.9583 for harvests with the test sample.

Table 3.5: Models and overall accuracy rates for harvest/non-harvest classification of disturbance clumps with equal representation of harvest type strata

<table>
<thead>
<tr>
<th>Method</th>
<th>Variables</th>
<th>Parameters</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>rpart</td>
<td>all</td>
<td>cp = 0.008</td>
<td>0.920</td>
<td>0.817</td>
</tr>
<tr>
<td>kNN</td>
<td>1, 8, 12, 19</td>
<td>k = 5</td>
<td>0.877</td>
<td>0.720</td>
</tr>
<tr>
<td>logistic reg</td>
<td>1, 8, 12, 19</td>
<td>k = 5</td>
<td>0.914</td>
<td>0.806</td>
</tr>
<tr>
<td>random forest</td>
<td>all</td>
<td>mtry = 2</td>
<td>0.937</td>
<td>0.856</td>
</tr>
<tr>
<td>SVM (radial)</td>
<td>1, 8, 12, 19</td>
<td>C = 1</td>
<td>0.888</td>
<td>0.734</td>
</tr>
</tbody>
</table>

The intersection of the stand-clearing disturbance map and the harvest map creates a map with stand-clearing, partial harvest, and non-harvest disturbance categories. Figure 3.3 illustrates the success of this map for classifying disturbances into these categories. The figure shows a 2009 National Agriculture Imagery Program (NAIP) photo before the 2010 disturbance map as well as the 2010 VCT disturbances color-coded according to the categories and overlaid on a 2011 NAIP photo.
Figure 3.3: Reclassified peVCT disturbances: (Top) 2009 NAIP aerial photo with parcel boundaries, (Bottom) 2011 NAIP aerial photo with transparent disturbance clusters. (A) Violet = stand-clearing, (B) Green = partial, (C) Yellow = Non-harvest.
3.3.2 Comparison of VCT-derived age maps

The proportions of successful determination of ‘most recent year of stand-clearing disturbance’ as compared to Google Earth aerial photos are shown in Table 3.6. These agreement rates were determined for all 389 forested points in the ‘All VCT forest’ column, and a separate agreement rate was determined for the 205 points where a VCT disturbance occurred in the ‘Disturbance only’ column. Table 3.7 shows that McNemar tests for differences in proportions indicate a statistically significant difference in proportions of successful determination of ‘most recent year of stand-clearing disturbance’ between the unenhanced VCT-derived age map and both the eVCT and peVCT enhanced age maps. There is no statistically significant difference between eVCT and peVCT.

Table 3.6: Agreement of most recent stand-clearing disturbance derived from each age map with time series of aerial photos

<table>
<thead>
<tr>
<th>Map</th>
<th>All VCT forest</th>
<th>Disturbances only</th>
</tr>
</thead>
<tbody>
<tr>
<td>VCT</td>
<td>87.66%</td>
<td>76.59%</td>
</tr>
<tr>
<td>eVCT</td>
<td>92.29%</td>
<td>85.37%</td>
</tr>
<tr>
<td>peVCT</td>
<td>93.83%</td>
<td>88.29%</td>
</tr>
</tbody>
</table>

Table 3.7: McNemar tests for differences in marginal proportions of successful determination of ‘most recent year of stand-clearing disturbance’

<table>
<thead>
<tr>
<th>Map pair</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VCT/eVCT</td>
<td>3.67</td>
<td>0.000242</td>
</tr>
<tr>
<td>VCT/peVCT</td>
<td>3.21</td>
<td>0.001328</td>
</tr>
<tr>
<td>eVCT/peVCT</td>
<td>0.53</td>
<td>0.596112</td>
</tr>
</tbody>
</table>

The average absolute difference between the year of most recent stand-clearing disturbance estimated from eVCT and the most likely year of stand-clearing disturbance is 2.53 years. This is slightly more than the average absolute difference between the year of most recent...
stand-clearing disturbance as determined by peVCT and the most likely year of stand-clearing disturbance, at 2.25 years. The difference between these absolute differences is not statistically significant ($t = 0.53$, $n = 205$), indicating there is not a case to say that peVCT is more accurate than eVCT.

When looking at the signed differences eVCT appears to underestimate the number of years since the last stand-clearing disturbance, while peVCT appears to overestimate it. The average signed difference between the eVCT estimated year of most recent stand-clearing disturbance and the most likely year of most recent stand-clearing disturbance is 1.91 years. This same difference, replacing eVCT with peVCT, is -1.61 years. Both of these are statistically different than zero with $t = 4.45$ and $t = -3.75$, respectively, for a sample of 205 VCT disturbance points.

### 3.3.3 Forest retention and conversion

Figure 3.4 shows the proportion of stand-clearing disturbance pixels, as classified by peVCT, that are subsequently classified as forest by years since they were cleared for a sample of disturbance years (1986, 1991, 1996, and 2001). Time series of these plots could be used to identify trends in rates of land cover change by answering questions such as ‘is there a trend in the maximum proportion of pixels that return to a forest classification?’

Figure 3.5 demonstrates a similar comparison of 2011 NLCD classified deciduous forest and coniferous forest that was cleared in 1991 according to peVCT.
Figure 3.4: Examples (1986, 1991, 1996, and 2001) of the proportion of stand-clearing disturbances that are classified as forest in subsequent years.

Figure 3.5: Proportion of 2011 coniferous forest and deciduous forest NLCD land covers that returned to a forest classification by years since 1991 stand-clearing disturbance.
3.4 Discussion

3.4.1 Input metrics

Referring to Figure 3.2, metrics 1 through 4 are measures of disturbance magnitude, 5 through 11 are measures of shape and size, 12 measures the average proximity to the forest edge, and 13 through 22 are counts of disturbed pixels and total pixels in a parcel, disturbance clumps in a parcel, and their ratios. Many of these variables measure similar things and are highly correlated. We will not discuss all of them in depth here but will highlight those that were included in the reduced stand-clearing and harvest models. Variables 1, 8, 10, 12, 16, and 19 were included in at least one of the reduced stand-clearing models. These same metrics, minus 10 and 16, were used in the reduced harvest models.

Disturbance magnitude

According to a random forest variable importance plot for accuracy, metric 1, the average disturbance magnitude for all pixels in a disturbance clump as measured by IFZ, is the most important variable in both the stand-clearing model and the harvest model. It measures the magnitude of the increase in IFZ value from the previous year to the year of disturbance. In general, the more trees removed from a pixel the greater the IFZ score due to more light being reflected from bare earth. Figure 3.6 shows that stand-clearing disturbances generally have the highest disturbance magnitude due to more trees being removed from within each pixel. This can also be seen in Table 3.8.
Figure 3.6: Classified disturbance clumps on left (blue = stand-clearing, purple = partial, orange = non-harvest); IFZ disturbance magnitude on right (brighter pixels equal higher magnitude)
Table 3.8: Metrics for clumps from Figure 3.6 by ID. Disturbance type is denoted in column one with: N = non-harvest, P = partial harvest, and S = stand-clearing disturbance.

<table>
<thead>
<tr>
<th>Clump ID</th>
<th>n.edges.internal</th>
<th>perim.area.ratio</th>
<th>IFZ.dstbMagn</th>
<th>mean.forest.cell.ct</th>
<th>mean.clump.size</th>
<th>clumps.p.distb.parcel</th>
</tr>
</thead>
<tbody>
<tr>
<td>42662 N</td>
<td>0</td>
<td>4</td>
<td>33</td>
<td>8</td>
<td>73.875</td>
<td>0.000003</td>
</tr>
<tr>
<td>42648 N</td>
<td>0</td>
<td>4</td>
<td>20</td>
<td>5</td>
<td>73.875</td>
<td>0.000003</td>
</tr>
<tr>
<td>42614 N</td>
<td>0</td>
<td>4</td>
<td>35</td>
<td>7</td>
<td>73.875</td>
<td>0.000003</td>
</tr>
<tr>
<td>42567 N</td>
<td>0</td>
<td>4</td>
<td>115</td>
<td>9</td>
<td>73.875</td>
<td>0.000003</td>
</tr>
<tr>
<td>42371 N</td>
<td>0</td>
<td>4</td>
<td>36</td>
<td>8</td>
<td>86</td>
<td>0.000007</td>
</tr>
<tr>
<td>42363 P</td>
<td>1148</td>
<td>0.5833333</td>
<td>51.67857</td>
<td>8.636905</td>
<td>86</td>
<td>0.000007</td>
</tr>
<tr>
<td>42355 P</td>
<td>0</td>
<td>4</td>
<td>40</td>
<td>1</td>
<td>86</td>
<td>0.000007</td>
</tr>
<tr>
<td>42319 P</td>
<td>60</td>
<td>1.272727</td>
<td>52.72727</td>
<td>8.49091</td>
<td>86</td>
<td>0.000007</td>
</tr>
<tr>
<td>42250 P</td>
<td>200</td>
<td>1.142857</td>
<td>37.61429</td>
<td>8.2</td>
<td>86</td>
<td>0.000007</td>
</tr>
<tr>
<td>42043 P</td>
<td>0</td>
<td>4</td>
<td>15</td>
<td>3</td>
<td>10</td>
<td>0.000448</td>
</tr>
<tr>
<td>37703 P</td>
<td>40</td>
<td>1.777778</td>
<td>21.61111</td>
<td>8</td>
<td>6.25</td>
<td>0.000310</td>
</tr>
<tr>
<td>37687 P</td>
<td>0</td>
<td>12</td>
<td>12</td>
<td>8</td>
<td>6.25</td>
<td>0.000310</td>
</tr>
<tr>
<td>37623 P</td>
<td>12</td>
<td>2</td>
<td>41.33333</td>
<td>9</td>
<td>6.25</td>
<td>0.000310</td>
</tr>
<tr>
<td>37608 P</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>9</td>
<td>6.25</td>
<td>0.000310</td>
</tr>
<tr>
<td>42497 S</td>
<td>2200</td>
<td>0.302521</td>
<td>114.4639</td>
<td>8.796639</td>
<td>73.875</td>
<td>0.000003</td>
</tr>
<tr>
<td>42003 S</td>
<td>680</td>
<td>0.4210526</td>
<td>134.9211</td>
<td>8.584211</td>
<td>156</td>
<td>0.000036</td>
</tr>
</tbody>
</table>

Parcel metrics

- The mean clump size, metric 16, is simply the number of disturbed pixels in a parcel divided by the number of disjoint disturbance clumps. Disturbance clumps associated with partial harvests are usually smaller because they normally cover only a portion of the harvest area.

- Metric 19 is the number of disjoint disturbance clumps within a parcel boundary divided by the number of disturbed pixels, and multiplied by the number of pixels in the parcel, as shown in Table 3.2. This value is higher for a greater number of clumps, for larger parcels, and for fewer disturbed pixels in the parcel. Many thinnings occur on large parcels and have both a high number of clumps in the parcel with a low number of disturbed pixels. This is because partial harvests only remove some of the trees. Therefore, many 30 meter pixels throughout the harvest area remain undisturbed or with only a portion of the pixel disturbed, thereby going undetected by VCT.
Metrics calculated from cell counts

Example calculations of metrics 8, 10, and 12 can be seen in Figure 3.7.

- Metric 8, the number of edges of internal cells was also important in both models. It can be considered both a measure of shape and size. The number of internal edges, those that are not perimeter edges, is calculated for each cell, and the sum of the individual cells determines the total. Thus, internal edges are always counted twice because two adjacent cells always share an edge. Thicker clumps with smoother edges have higher values of this metric.

- Metric 10 is the ratio of perimeter edges of the clump to the area of the clump. It is not shown here, but perimeter edges consist of any edge incident with an undisturbed pixel, including those possibly in the interior of the clump. This is a measure of shape complexity. Clearcuts depicted in Figure 3.6 and quantified in Table 3.8 have lower values of this metric, while clumps belonging to thinnings have higher values. Square clumps (including a single pixel) have the highest value. In addition to increasing with shape complexity, values of this metric decrease with size when shape remains constant.

- The mean number of forest pixels, including the center pixel and any other disturbed pixels, in a nine cell neighborhood of each pixel of the disturbance clump is calculated for metric 12. Pixels on the edge of a forest will have low values of this metric.

Machine learning tools do not consider the reasons a variable might be important for classification. Recall that variables were entered into the reduced models one at a time, in order
of importance according to the random forest variable importance plot, while eliminating variables correlated higher than 0.5 with a previously entered variable. Chance was at play in selecting the samples, so another run could produce different results. Alos, it is possible that entering variables according to another criterion, such as t-scores for logistic regression, could improve accuracy for some of the reduced models.

### 3.4.2 Bias and sampling

The non-stand-clearing disturbance producer’s accuracy rates for peVCT (95.4%) is a substantial improvement over Kauffman and Prisley’s (2016) results for eVCT (57.5%). This is due to several interacting factors. First, non-harvest disturbances were not represented in Kauffman and Prisley (2016). Non-harvest disturbances are likely much easier to distinguish.
from stand-clearing disturbances than are partial harvests. Second, the sampling method for this paper excludes harvests that are less than five acres from the partial harvest and stand-clearing harvest strata, while all non-harvest clumps are limited to less than five acres. Reducing this threshold may negatively affect accuracy rates but may be more suitable for applications where age is important at a smaller scale. The lack of a hard size requirement in Kauffman and Prisley (2016) may have been compensated for by the requirement that only harvest points that intersect disturbance clumps were selected for the sample. This method may make it more likely for larger disturbance clumps to intersect a harvest point than smaller clumps. Recall that for the improved sampling method in this work, sample harvest points are moved to the largest portion of a disturbance clump that lies within a single parcel and within the boundaries of the harvest. Finally, some of the improvement in the producer’s accuracy for non-stand-clearing disturbances can be attributed to the inclusion of parcel metrics. However, of the ten most important variables for classification, as determined by a variable importance plot generated from the random forest model, only \textit{mean.clump.size} at number 10 is derived using parcel boundaries. The omission of \textit{mean.clump.size} only slightly reduces the overall accuracy and the non-stand-clearing producer’s accuracy.

Both of the enhanced VCT products, eVCT and peVCT, reclassify VCT disturbance clumps as stand-clearing or not. Two types of error are possible when reclassifying a disturbance clump. The reclassification algorithm could classify a disturbance clump as not stand-clearing when it really is. This type of error could be called a type one error. A type two error would then occur when the algorithm classifies a disturbance as stand-clearing when it really is not. Thus, age maps derived from the raw VCT product, in which there is no classification of disturbances as stand-clearing or not, can only include type two errors.
VCT does not distinguish between stand-clearing and non-stand-clearing disturbances, so they must all be treated similarly, as stand-clearing disturbances. Recall that eVCT appears to underestimate the number of years since the most recent stand-clearing disturbance while peVCT appears to overestimate it. Therefore, it seems to follow that eVCT makes more type two errors while peVCT makes more type one errors.

Perhaps omitting the non-harvest stratum from the sample or weighting it proportionally by area will result in a calibration that is less biased toward making type two errors. Furthermore, the size of the ‘stand-clearing’ and ‘partial harvest’ strata could be made proportional to the yearly average of these in the harvest record according to either area or numbers. However, this would not totally omit bias because stand-clearing disturbances that are not related to harvest are not included in the harvest record. Continued research in this area could lead to further improvement of forest age maps.

### 3.4.3 Forest retention and conversion

VCT recognizes regenerating areas of deciduous forest more quickly than regenerating coniferous with 67 percent of cleared forests returning to a deciduous forest classification within just four years compared to 40 percent for coniferous forest. However, a slightly higher percentage of pixels classified as coniferous in 2011 eventually return to forest than those classified as deciduous.

Figure 3.5 also shows an eventual obvious decrease in the percent of previously cleared forested pixels that are classified as coniferous forest once again after a maximum at 14 years post-disturbance. This pattern is absent from cleared forest that returns to deciduous
forest. One explanation of this is that a portion of coniferous forest is in managed pine stands with shorter rotations and industrial thinnings. VCT classifies some of the pixels in these thinned areas as disturbances. The percentage of forested pixels continues to decrease as perhaps more thinning occurs and stands are eventually cleared. It would be interesting to see how planted pine forest compares to natural pine. Perhaps possible differences like earlier return to forest and a more pronounced dip in percentage of pixels classified as forest could be used in an automated fashion to help to identify pine plantations.

Unfortunately, if a stand was just cut, VCT cannot predict the number of years it will take to return to forest again. In fact, some cleared forest will be converted to other uses. Information in addition to VCT is needed to conclusively determine when this occurs. However, it can be seen from Figure 3.4 that within a few years a majority of cleared pixels have returned to forest, eliminating the possibility that a land use change occurred at the last clearing for those pixels. This could be combined with other methods for determining land cover in order to hone in more accurately on land use. For instance, three years after a stand-clearing disturbance nearly half of these pixels have been classified by VCT as forest once again, so they can be assumed to be in forest land use. Additionally, from all pixels that were cleared in 2001 and appeared as non-forest in 2011, 109 were randomly chosen and followed through a time series of aerial photos from Google Earth. Seventy-nine percent of the pixels that converted from forested to developed or agricultural did so within the first three years. In fact, 77 percent converted in the first 2 years, 90 percent in the first five, and 97 percent converted by year 7. Perhaps a classification method complementary to VCT and tuned to identify agricultural and developed land uses could identify these land cover transitions. Pixels where the combined classifications contradict or are inconclusive could
be classified as in transition. These pixels that are in transition could be treated specially in measures of net forest loss where there is controversy over using a land cover definition of forest (Wernick et al. 2010; Reams et al. 2010).

3.5 Conclusion

Spatially precise locations and temporally precise timing of timber harvests provide valuable information to landowners and those in the forest products industry. This type of information is also valuable for monitoring land use and land cover change, wildlife habitat, water resources, carbon sequestration, and other ecosystem services. Although there was not a statistically significant difference in eVCT versus peVCT stand age maps, peVCT correctly identified the year of most recent stand-clearing disturbance more often. More refinement of sampling methods and development of metrics based on location of disturbed pixels in relation to parcel boundaries could further improve the ability to correctly classify stand-clearing versus non-stand-clearing disturbances. The possibility that neighbors sometimes cooperate and the fact that a single landowner often owns adjoining or nearby parcels was not taken into account. Therefore, additional information on ownership could also improve classification accuracy. The ability to delineate and track forest stands for a myriad of purposes is valuable. One such example of the importance of stand-clearing disturbance maps, in addition to facilitating the creation of age maps, was demonstrated through the ability to track regeneration patterns of cleared forest stands across time and by forest type. Metrics based on parcel boundaries also proved important in classification algorithms for distinguishing between harvest and non-harvest related disturbances. The use of parcel data
in conjunction with disturbance algorithms such as VCT is useful for harvest detection and delineation. Continued research into the utilization of parcel boundary data is warranted.

3.6 References


USDA Forest Service. 2014. Forest Inventory and Analysis national core field guide Volume I: field data collection procedures for Phase 2 plots. USDA For. Serv., North Central Research Station FIA, St. Paul, MN.


Chapter 4

Monitoring Non-Timber Forest Products Using Forest Inventory Data: An Example with Slippery Elm Bark

Jobriath S. Kauffman, Stephen P. Prisley, James L. Chamberlain

Note: This paper was published in Journal of Forestry

4.1 Abstract

The US Forest Service, Forest Inventory and Analysis (FIA) program collects data on a wealth of variables related to trees in forests. Some of these trees produce non-timber forest products (NTFPs; e.g., fruit, bark, sap) that are harvested for culinary, decorative, building, and medicinal purposes. At least 11 tree species inventoried by FIA are valued for their bark. For example, slippery elm (*Ulmus rubra* Muhl.) is included in FIA forest inventories, and the
bark is used for its medicinal values. Despite widespread use of NTFPs, little quantitative information about abundance, distribution, and harvest is available to support sustainable management. Methods for using the FIA database to monitor and explain the situation regarding selected NTFPs are presented. The focus is on using FIA data to assess for: (1) geographic distribution, (2) abundance, (3) applicable metrics (e.g., square feet of bark), and (4) change over time.

### 4.2 Management and policy implications

The commercial value of non-timber forest products is significant and not fully assessed (Alexander et al. 2011). Development of awareness that some NTFPs originate from trees and are already included in FIA is critical to incorporating these products into forest monitoring policy. For example, slippery elm, and at least 10 other tree species are harvested for their bark, which is typically not considered in management decisions. Use of readily available data from FIA, and graphical and tabular analytics that can be replicated across various species and geographic areas, can provide valuable insights to both managers and policy makers by spatially monitoring availability and sustainable use of these resources for this valuable NTFP. Using FIA to improve monitoring of understory species harvested for non-timber products is more limited. Protocols have been developed to collect data on some understory vegetation but they are limited in scope and use. Policy directives to enhance these measurements would help quantify understory vegetation species and allow for better monitoring of non-timber forest products. An automated process for analyzing and summarizing FIA data on NTFPs in common trade units is a desirable and feasible goal. Research
funds to support these efforts could markedly improve monitoring of NTFPs.

4.3 Introduction

The forests of the United States provide an abundance of resources, products, recreational opportunities, and ecosystem services that benefit the people living in and near them, as well as around the world (Oswalt et al. 2012). Among these are NTFPs—plants and fungi harvested from forests used for a diversity of purposes. The ability to quantify the spatial and temporal distribution and abundance of NTFPs is important for monitoring the resource, assessing sustainable use, and answering broad-scale research questions.

The American Herbal Products Association solicits and aggregates data on quantities of medicinal plants that companies purchase as raw materials (Dentali and Zimmermann 2012). The amounts of NTFPs harvested for commercial purposes alone is significant, including products coming from trees. For instance, more than 300,000 pounds of slippery elm bark were harvested, annually (2006-2010), from U.S. forests (Dentali and Zimmermann 2012). However, more data are needed to estimate total NTFP harvests to include non-commercial harvests. This includes a lack of information on harvests from many state forests and little knowledge about the prevalence of illegal poaching (Frey and Chamberlain 2015). In light of this need, we explore ways to quantify various aspects of tree species that provide non-timber products using the FIA database.
4.3.1 Non-timber forest products

NTFPs have been significant to the culture and commerce of the United States since before the country was founded. They are integral to subsistence economies, as well as to peoples' health, food security and spiritual livelihoods (Emery and Pierce 2005, McClain et al. 2008). They expand the scope and scale of the forest products industry to include culinary, medicinal, decorative, landscaping and nursery markets. More people harvest non-timber products from U.S. forests for non-commercial benefits than they do for commercial (Alexander et al. 2011). The commercial value of these forest products, however, is significant and not fully assessed in forest management planning (Alexander et al. 2011).

NTFPs are collected for recreational, commercial, and subsistence purposes (Vaughan et al. 2013). Throughout the history of the United States, people of diverse backgrounds and cultures have derived their livelihood from NTFPs (Chamberlain et al. 1998, Emery 2002). However, lack of trust often inhibits collectors from sharing their knowledge with others (Vaughan et al. 2013). For example, Native Americans have special knowledge of NTFPs that has been passed through generations. Combining traditional ecological knowledge with science-based knowledge is providing valuable insights into ways to improve management for these products (Emery et al. 2014, Hummel and Lake 2015).

NTFPs originate from fungi and plants, including forest mosses, lichens, herbs, shrubs, vines, and understory and overstory trees. A general perception is that NTFPs come from plants other than trees. There are, however, many NTFPs that originate from trees which are included in the FIA database. Iconic NTFPs from trees include the sap extracted from sugar maple (*Acer saccharum* Marsh.), the bark of paper birch (*Betula papyrifera* Marsh.),
and the fruit of black walnut \((\textit{Juglans nigra} \text{ L.})\). Additionally, Christmas ornamentals from the boughs of species such as noble fir \((\textit{Abies procera} \text{ Rehder.})\) are an important part of the floral and decorative NTFP segment (Blatner et al. 2009). Current understanding about these forest trees as producing NTFPs is lacking. Gleaning from knowledge sources such as FIA databases may provide important insights into the dynamics of NTFPs from trees.

### 4.3.2 Forest Inventory and Analysis

The FIA program of the US Forest Service is the primary source of information on status and trends of the nation’s forest resources (Reams et al. 1999). By sampling all of the nation’s forested lands over a periodic cycle, FIA provides the most comprehensive field inventory conducted today (Bechtold and Patterson 2005). Since NTFPs are an important subset of these renewable resources, efforts to improve information about these forest products by utilizing and enhancing FIA is warranted.

FIA’s spatially and temporally comprehensive sampling approach is appropriate for monitoring tree products in forests. Generally, within each subplot, trees are identified by species and status (living or dead) and measured for height, diameter, damage, and cause of death (USDA Forest Service 2014). These measurements can be used to monitor the status of NTFPs that come from tree species. Using expansion factors which rely on trees per acre and number of acres of a forest condition within a plot, tree measurements at the plot level allow for number and volume estimates that can be aggregated to any collection of plots, such as counties, FIA units, sub-state regions, states, and groups of states (Bechtold and Patterson 2005).
Measurement of understory plants that produce NTFPs is much less comprehensive. While the sampling protocol and FIA data structure include provision for percent canopy cover measurement of understory plants, these measurements are available only for states from the Rockies west, including Alaska, Arizona, California, Colorado, Idaho, Montana, New Mexico, Nevada, Oregon, Utah, Washington, and Wyoming.

4.4 Background

Several studies have made use of FIA data to monitor and assess the status of selected NTFPs. Products that have been examined using FIA data include bark from paper birch (Emery et al. 2014), pine nuts from pinyon pines (*Pinus edulis* Engelm., *P. monophylla* Torr. and Frem., *P. monophylla* var. fallax Torr.; Shaw et al. 2005) and maple syrup from sugar maple (Farrell 2012).

These studies support the use of FIA data to document the spatial distribution and dynamics of NTFPs from trees. For example, the ability to detect trend and magnitude of low levels of change caused by drought, insects, and disease on pinyon mortality has been attributed to FIA’s annual inventory sampling design with yearly panels free of geographic bias (Shaw et al. 2005). Combining FIA data on the diameter at breast height (dbh) of maple trees and the distance of each plot to the nearest road, Farrell (2012) was able to estimate the regional production potential of sugar maple stands. Supplementing FIA data with additional knowledge can improve resource assessments and the ability to manage for paper birch bark (Emery et al. 2014). Because paper birch bark is used to make various items, the number and measurements of trees that meet minimum size requirements are needed. Emery et al.
(2014) estimated total bark surface area for trees with at least a five-inch dbh as well as for trees with at least an 11-inch dbh, which are more suitable for making items such as canoes. Change over time in these resources can be quantified through analysis of data from re-measured plots.

The examples here illustrate the usefulness of leveraging FIA data to quantify abundance, spatial distribution and change over time of NTFPs derived from tree species. Despite the accessibility of FIA data, summary analyses of this type have been lacking for many NTFP species. The following demonstrates approaches and methods for analyzing any tree species in the database with specific example calculations and results.

### 4.5 Methods

Using FIA databases and expert knowledge of NTFPs, an initial 19 tree species (Table 4.1) that are harvested for their non-timber values were identified by US Forest Service personnel and categorized according to use. The majority of the species in this list are valued for their bark and used in herbal medicines, and a large number are enjoyed for their culinary benefits. This list of species is not comprehensive and can be expanded to any pertinent NTFP tree species.

All of these tree species were analyzed across their entire range using 2007 and 2012/2013 population evaluation groups. On occasion, a population evaluation group for the desired year was not available for a given state. Most often this was because measurement of that year’s panel extended well into the next year. In these unusual cases, the next closest available year was used.
Table 4.1: Sample of trees found in FIA databases that are harvested for NTFPs

<table>
<thead>
<tr>
<th>Location</th>
<th>FIA code</th>
<th>Common name</th>
<th>Scientific name</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>East/West</td>
<td>375</td>
<td>Paper birch</td>
<td><em>Betula papyrifera</em></td>
<td>Bark, Decorative</td>
</tr>
<tr>
<td>East</td>
<td>129</td>
<td>White pine</td>
<td><em>Pinus strobus</em></td>
<td>Bark, Medicine</td>
</tr>
<tr>
<td>East</td>
<td>601</td>
<td>Butternut</td>
<td><em>Juglans cinerea</em></td>
<td>Bark, Medicine</td>
</tr>
<tr>
<td>East</td>
<td>611</td>
<td>Sweetgum</td>
<td><em>Liquidambar styraciflua</em></td>
<td>Bark, Medicine</td>
</tr>
<tr>
<td>East</td>
<td>762</td>
<td>Black cherry</td>
<td><em>Prunus serotina</em></td>
<td>Bark, Medicine</td>
</tr>
<tr>
<td>East</td>
<td>802</td>
<td>White oak</td>
<td><em>Quercus alba</em></td>
<td>Bark, Medicine</td>
</tr>
<tr>
<td>East</td>
<td>931</td>
<td>Sassafras</td>
<td><em>Sassafras albidum</em></td>
<td>Bark, medicine</td>
</tr>
<tr>
<td>East/West</td>
<td>927</td>
<td>White willow</td>
<td><em>Salix alba</em></td>
<td>Bark, Medicine</td>
</tr>
<tr>
<td>West</td>
<td>231</td>
<td>Pacific yew</td>
<td><em>Tazus brevifolia</em></td>
<td>Bark, Medicine</td>
</tr>
<tr>
<td>East</td>
<td>975</td>
<td>Slippery elm</td>
<td><em>Ulmus rubra</em></td>
<td>Bark, Medicine</td>
</tr>
<tr>
<td>East</td>
<td>621</td>
<td>Yellow-poplar</td>
<td><em>Liriodendron tulipifera</em></td>
<td>Bark, Siding</td>
</tr>
<tr>
<td>East</td>
<td>367</td>
<td>Pawpaw</td>
<td><em>Asimina triloba</em></td>
<td>Fruit, Edible</td>
</tr>
<tr>
<td>East</td>
<td>521</td>
<td>Common persimmon</td>
<td><em>Diospyros virginiana</em></td>
<td>Fruit, Edible</td>
</tr>
<tr>
<td>East/West</td>
<td>602</td>
<td>Black walnut</td>
<td><em>Juglans nigra</em></td>
<td>Medicine</td>
</tr>
<tr>
<td>East/West</td>
<td>561</td>
<td>Gingko</td>
<td><em>Gingko biloba</em></td>
<td>Leaves, Medicine</td>
</tr>
<tr>
<td>East</td>
<td>318</td>
<td>Sugar maple</td>
<td><em>Acer saccharum</em></td>
<td>Sap, Edible</td>
</tr>
<tr>
<td>West</td>
<td>106</td>
<td>Two-needle pinyon</td>
<td><em>Pinus edulis</em></td>
<td>Seeds, Edible</td>
</tr>
<tr>
<td>West</td>
<td>133</td>
<td>Singleleaf pinyon</td>
<td><em>Pinus monophylla</em></td>
<td>Seeds, Edible</td>
</tr>
<tr>
<td>West</td>
<td>143</td>
<td>Arizona pinyon pine</td>
<td><em>Pinus monophylla var. fallax</em></td>
<td>Seeds, Edible</td>
</tr>
</tbody>
</table>
Summary metrics were defined for each species to assess abundance, map spatial distribution, and estimate change over time (Table 4.2). Abundance was quantified and mapped by total number of trees, number of trees by diameter class, plot locations with live trees, and number of trees by FIA unit. Changes in abundance were presented by examining differences in periodic remeasurements and percent change in number of trees overall, in the number of trees by diameter class, by FIA unit, net growth, removals, and mortality (where available), and mortality by FIA unit (where available).

Table 4.2: Summary metrics for NTFP tree species computed from FIA data.

<table>
<thead>
<tr>
<th>Type of Measurement</th>
<th>Summary Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abundance</td>
<td>Number of Trees</td>
</tr>
<tr>
<td></td>
<td>Number of Trees by Diameter Class</td>
</tr>
<tr>
<td></td>
<td>Surface Area</td>
</tr>
<tr>
<td></td>
<td>Surface Area by Diameter Class</td>
</tr>
<tr>
<td>Spatial Distribution</td>
<td>Plot Locations</td>
</tr>
<tr>
<td></td>
<td>Number of Trees by FIA Unit</td>
</tr>
<tr>
<td></td>
<td>Surface Area by FIA Unit</td>
</tr>
<tr>
<td>Trends</td>
<td>Number of Trees and Percent Change (2007 to 2012)</td>
</tr>
<tr>
<td></td>
<td>Number of Trees by Diameter Class (2007 and 2012)</td>
</tr>
<tr>
<td></td>
<td>Surface Area and Percent Change (2007 to 2012)</td>
</tr>
<tr>
<td></td>
<td>Net Growth, Mortality, Removals, Gross Growth, and Volume (2007 and 2012)</td>
</tr>
<tr>
<td>Trends by Location</td>
<td>Percent Change in Number of Trees by FIA Unit</td>
</tr>
<tr>
<td></td>
<td>Percent Change in Surface Area by FIA Unit</td>
</tr>
<tr>
<td></td>
<td>Mortality by FIA Unit</td>
</tr>
<tr>
<td></td>
<td>Net Growth, Removals, and Mortality by State</td>
</tr>
</tbody>
</table>

Summary metrics were most easily obtained using EVALIDator, an online reporting tool that can be used to summarize FIA data (http://apps.fs.fed.us/Evalidator/evalidator.jsp; Miles 2014). EVALIDator cannot be used to report plot-level details but can limit results to a given species. For example, to obtain number of trees by diameter class for each FIA unit for a certain species, the user makes a series of selections from EVALIDator. The user first specifies the attribute to be summarized (number of all live trees on forestland), then
selects the appropriate population evaluation group(s), followed by a row, column, and/or page factor on which to summarize (such as FIA Unit and diameter class), and finally the user may limit the query to a certain tree species or in some other way if desired (such as dbh at least five inches).

Other queries, including approximate plot locations (LAT/LON) done at the plot level, must rely on database management tools such as Microsoft Access. Plot coordinates are approximated to within 1.0 mile of the exact plot location, and up to 20 percent of the private plot coordinates are swapped with another similar private plot within the same county (O'Connell et al. 2015). Entire state Access databases can be downloaded from the FIA website (http://www.fia.fs.fed.us/tools-data/). Access databases contain built-in queries for state-level estimates by population evaluation group. Experienced users can alter these built-in queries to summarize estimates at the unit, county, or plot level. For example, a query for number of all live trees on forestland at the state level could be altered to obtain number of trees on a plot by changing the query to include a plot identifier. By adding fields for population evaluation group, tree diameter, or species code, the query could be altered to limit results to certain measurement years and trees with at least a five-inch dbh and of a certain species. Furthermore, fields for latitude and longitude could be added to a query to obtain the approximate geographic coordinates of each plot in the result.

FIA data from multiple states is required when querying for results across an entire species’ range. In these cases it may be easier to use other database management software such as Microsoft SQL Server. The individual tables representing the entire nation needed to perform the desired queries can also be downloaded from the FIA website. Some of these
nationwide tables are too large for Access to handle, and performing the same query in Access repetitively for each state can be cumbersome. Thus, maps of point locations of plots containing selected species were created using SQL Server. Writing scripts of code in SQL or another language is also useful for automating analyses of new FIA data in successive years.

Calculating surface area

For trees from which bark is harvested, the bole surface area is of particular interest for estimating product quantity. Eleven tree species were identified that are valued for their bark (Table 4.1). FIA defines the bole of a tree as the portion from a one-foot stump to a four-inch diameter top (O’Connell et al. 2015). Accepting this definition, a truncated cone can be used as an appropriate representation of the bole. Equation 4.1 is the standard formula for calculating the surface area (SA) of a truncated cone in square feet.

\[
SA = \pi \times \left( \frac{r_1}{12} + \frac{r_3}{12} \right) \times s
\]  

(4.1)

In equation 2.1, \( r_3 \) denotes the radius at the top of the bole (2 inches, or one half of the 4 inch diameter), \( r_1 \) represents the radius in inches at a one-foot stump, and \( s \) symbolizes the slant height in feet (shortest distance between edges of the top and bottom of the bole). The bole of a tree with a specified dbh and total height can be represented by Figure 4.1. In this diagram, \( r_1 \) and \( r_3 \) are derived from dbh (two times \( r_2 \)), total height (\( h_4 \)), and similar triangles. This results in Equation 4.2 for surface area:

\[
SA = \pi \times \left( \frac{dbh}{24} \cdot \frac{h_4 - 1}{h_4 - 4.5} + \frac{1}{6} \right) \times \sqrt{\left[ \left( \frac{h_4 - 1}{6} \cdot \frac{h_4 - 4.5}{r_2/12} \right) - 1 \right]^2 + \left[ \left( \frac{dbh}{24} \cdot \frac{h_4 - 1}{h_4 - 4.5} - \frac{1}{6} \right) \right]^2}
\]  

(4.2)
where $SA$ is the surface area in square feet, $dbh$ is the diameter at breast height in inches and $h_4$ is the total height (actual length plus any missing broken piece; O’Connell et al. 2015). Records for live trees at least five inches dbh on forestland can be obtained from the FIA database for the desired population evaluation group. Total surface area for the measurement unit can be estimated by multiplying the surface area for each tree in the sample (as calculated by Equation 4.2) by the number of trees per acre for the plot/condition containing the tree and the number of acres represented by the plot/condition, then summing over all plot/conditions (O’Connell et al. 2015).

With FIA data on thousands of trees sampled across their range for most species, regression equations provide a good means for estimating bark SA directly from dbh. This eliminates the need for total height measurements and simplifies the process for estimating total SA for future calculations. We calculated SA using Equation 4.2 and dbh and total height from each live tree of a selected species at least five inches dbh prior to performing the regression. Table
4.3 provides the parameters of linear equations for each of the 11 bark species resulting from regressing $\sqrt{SA}$ on $\sqrt{dbh}$. These square root transformations help to maintain the ordinary least squares regression assumption of equal variance across the range of the explanatory variable, providing for better estimates.

Table 4.3: Parameters (SE) and $df$ for linear regression lines modeling the relationship between $\sqrt{dbh}$ and $\sqrt{SA}$ of bark for boles of trees at least 5 in. dbh by species, using 2012 population evaluation groups.

<table>
<thead>
<tr>
<th>Species</th>
<th>FIA code</th>
<th>a</th>
<th>b</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper birch</td>
<td>375</td>
<td>-6.541 [0.025]</td>
<td>4.661 [0.009]</td>
<td>25062</td>
</tr>
<tr>
<td>White pine</td>
<td>129</td>
<td>-7.835 [0.021]</td>
<td>5.039 [0.006]</td>
<td>30749</td>
</tr>
<tr>
<td>Butternut</td>
<td>601</td>
<td>-6.440 [0.223]</td>
<td>4.553 [0.071]</td>
<td>323</td>
</tr>
<tr>
<td>Sweetgum</td>
<td>611</td>
<td>-8.223 [0.014]</td>
<td>5.402 [0.005]</td>
<td>54711</td>
</tr>
<tr>
<td>Black cherry</td>
<td>762</td>
<td>-7.474 [0.210]</td>
<td>5.020 [0.007]</td>
<td>30540</td>
</tr>
<tr>
<td>White oak</td>
<td>802</td>
<td>-7.292 [0.017]</td>
<td>4.974 [0.005]</td>
<td>52250</td>
</tr>
<tr>
<td>Sassafras</td>
<td>931</td>
<td>-6.971 [0.051]</td>
<td>4.781 [0.018]</td>
<td>6800</td>
</tr>
<tr>
<td>White willow</td>
<td>927</td>
<td>-4.118 [0.715]</td>
<td>3.523 [0.220]</td>
<td>13</td>
</tr>
<tr>
<td>Pacific yew</td>
<td>231</td>
<td>-2.906 [0.395]</td>
<td>2.664 [0.156]</td>
<td>25</td>
</tr>
<tr>
<td>Slippery elm</td>
<td>975</td>
<td>-7.018 [0.046]</td>
<td>4.820 [0.016]</td>
<td>6102</td>
</tr>
<tr>
<td>Yellow-poplar</td>
<td>621</td>
<td>-8.270 [0.018]</td>
<td>5.504 [0.005]</td>
<td>36241</td>
</tr>
</tbody>
</table>

Establishing a relationship between SA and dbh allows for surface area estimates requiring only dbh of trees in the sample. If using FIA data the process for estimating total SA for a species across a region is to perform an EVALIDator query for number of trees by diameter class in that region. This eliminates the need for the manager to query the database to retrieve dbh and total height for each tree in the appropriate sample and then expand the estimate on the appropriate plots to all forest area in the region of interest. EVALIDator insures that the correct trees are used in the sample for estimating the number of trees by diameter class in the desired region. By substituting the midpoint of each diameter class
and the appropriate parameters from Table 4.3 into the following formula,

\[ SA = \left( a + b\sqrt{dbh} \right)^2 \]  \hspace{1cm} (4.3)

SA within each class can be estimated by multiplying by the number of trees in each class obtained from EVALIDator. Summing the results of each class yields the overall total surface area.

### 4.6 Results – slippery elm as example

The approach described was used to obtain estimates for all tree species in the study, and results from slippery elm are presented here as an example. The inner bark of slippery elm trees is valued as one of the most common herbal remedies and as an ingredient in throat lozenges, nutritional supplements, and for other medicinal purposes (Pengelly and Bennett 2011). Because of the medicinal value of its bark, slippery elm was chosen as the first tree species to analyze employing these methods.

The native range of slippery elm is most of the eastern United States, and the distribution of FIA plots containing slippery elm closely matches this range (US Geological Survey 1999; Figure 4.2). An EVALIDator query for the 2013 population evaluation groups of states across its range estimates 1011 million ± 55 million slippery elm trees, of which 20.5 percent or 207 million (± 8 million) were at least 5 inches dbh.

Annual net growth, removals, mortality, and net volume were estimated for 2007 and 2013 population evaluation groups, and 95 percent confidence bounds were provided (Table 4.4). Estimated average annual net growth decreased from 13 million cubic feet to 0.59 million
Figure 4.2: Locations of FIA plots from the 2012 population evaluation group containing slippery elm trees 5 dbh or larger coded by count within plots and within species range map.
cubic feet with 95 percent confidence intervals containing negative net growth (mortality exceeds growth) for both sample years. Estimated average annual mortality increased from 57 million cubic feet to 69 million cubic feet. Estimated average annual removals increased from 9.6 million cubic feet in 2007 to 16 million cubic feet in 2013, approximately 71 percent.

Table 4.4: Average annual net growth, mortality, and removals, and gross growth and net volume (in million cu. ft.) of live slippery elm trees 5 inches dbh or larger on forestland estimated from recent (2013) and previous inventories. Previous inventory values are approximately 5-year state prior inventories (e.g., state inventory closest to 2007). Does not include west Texas.

<table>
<thead>
<tr>
<th>Parameter (95% C.I.)</th>
<th>2007 (million cu. ft.)</th>
<th>2013 (million cu. ft.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Average annual net growth</td>
<td>13 [-1.6,28]</td>
<td>0.59 [-9.1,10.3]</td>
</tr>
<tr>
<td>Average annual mortality</td>
<td>57 [42,72]</td>
<td>69 [60,78]</td>
</tr>
<tr>
<td>Average annual removals</td>
<td>9.6 [5.5,14]</td>
<td>16 [11,20]</td>
</tr>
<tr>
<td>Average annual gross growth</td>
<td>70 [40,100]</td>
<td>70 [51,88]</td>
</tr>
<tr>
<td>Net volume of live trees</td>
<td>1890 [1776, 2005]</td>
<td>1699 [1600, 1798]</td>
</tr>
</tbody>
</table>

* Negative net growth values are usually due to mortality but can also occur on live trees that have a net loss in volume because of damage, rot, broken top, or other causes.

The low and declining net growth of slippery elm corresponds to a decreasing number of trees. The number of trees at least 5-inch dbh declined from an estimated 231 million ± 10 million in 2007 to 207 million ± 9 million trees in 2013. These 95 percent confidence intervals do not overlap, which suggests a statistically significant difference, and represents an approximate 10.4 percent decrease in number of trees at least 5-inch dbh.

Figure 4.3 illustrates comparisons in number of trees and mortality by FIA unit. The map on the top (a) shows the number of slippery elm in 2012 by FIA unit, the middle map (b) shows percent change in number of slippery elm from 2007 to 2012, and the map on the
bottom (c) shows average yearly slippery elm mortality from 2007 to 2012 in millions of cubic feet. These maps can help managers and scientists identify areas with higher mortality and greater percent change that may warrant further attention as opposed to areas that have higher mortality merely due to an abundance of slippery elm. Units with a high number of trees, large percent decreases in number of trees, and high mortality may merit closer investigation and higher priority for management or interventions.

Figure 4.4 illustrates the distribution of slippery elm bole surface area by diameter class, calculated by substituting the midpoint of each diameter class and the slippery elm parameters from Table 4.3 into Equation 4.3 and then multiplying by diameter class frequencies from an EVALIDator query for number of slippery elm by diameter class (right-hand bars). Clearly, a majority of surface area is in small diameter classes. Summing across all diameter classes, we estimate the total area of slippery elm bark in 2012 at 10.97 billion square feet. Diameter class estimates when using total height and dbh for each tree in the database are shown in the left-hand bars for comparison purposes.

4.7 Discussion

Far more slippery elm volume is being harvested than is being grown (Table 4.4), indicating high harvest pressure. More cause for concern is the fact that mortality is seriously outpacing removals (Table 4.4), resulting in a statistically significant drop in number of live trees. Therefore, further information related to possible causation was extracted from the FIA database. Comparing the mortality rate (mortality as a percent of total volume) of slippery elm with the three most common tree species associated with slippery elm (i.e.,
Figure 4.3: (a) Number of slippery elm in 2012, (b) percent change in number of slippery elm from 2007 to 2012, and (c) average yearly slippery elm mortality from 2007 to 2012 by FIA unit.
Figure 4.4: Estimated total surface area (million feet²) of live slippery elm trees 5 dbh or greater on forest land by diameter midpoint class from recent (2012) inventories
most commonly co-occurring on plots) shows that the mortality rates for slippery elm and American elm (*Ulmus americana* L.) are greater than 4 percent (4.05 percent and 4.64 percent respectively) on plots containing at least one slippery elm, much higher than for sugar maple and sweetgum (0.69 percent and 0.86 percent respectively). Most likely, this is due to the susceptibility of both species to Dutch elm disease (*Ophiostoma ulmi* (Buisman) Nannf. and *Ophiostoma ulmi* Brasier).

However, overharvesting or improper harvesting of slippery elm bark is a possible explanation for the increased mortality. Table 4.5 shows the causes of mortality for slippery elm and American elm on plots with at least one slippery elm tree recorded in the FIA database. The percentage of deaths due to disease was lower for slippery elm than it was for American elm, while the percentage of deaths in the ‘unknown, not sure, other’ category was greater. This category includes death from human activity not related to silvicultural or land-clearing activity. This could include improper removal of bark and provides a basis for additional investigation.

<table>
<thead>
<tr>
<th>Cause of mortality</th>
<th>American Elm</th>
<th>slippery elm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insect</td>
<td>53</td>
<td>19</td>
</tr>
<tr>
<td>Disease</td>
<td>2358</td>
<td>482</td>
</tr>
<tr>
<td>Fire</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Animal</td>
<td>41</td>
<td>6</td>
</tr>
<tr>
<td>Weather</td>
<td>242</td>
<td>106</td>
</tr>
<tr>
<td>Vegetation</td>
<td>416</td>
<td>267</td>
</tr>
<tr>
<td>Unknown/not sure/other*</td>
<td>1286</td>
<td>478</td>
</tr>
<tr>
<td>Silvicultural or landclearing activity</td>
<td>852</td>
<td>274</td>
</tr>
<tr>
<td>Total</td>
<td>5256</td>
<td>1637</td>
</tr>
</tbody>
</table>
Broad analyses as described above and summarized in Table 4.2 have been completed for all 19 tree species in Table 4.1. This shows that FIA data can be used to monitor NTFPs from many different tree species. Rather than include all of these results here, total net growth, removals, mortality, 2007 net volume, and 2013 net volume on forestland for trees at least 5 inches dbh with their associated sampling error percent from EVALIDator are summarized in Table 4.6. Calculating growth-removals ratio, mortality as a percent of net volume, and percent decrease in volume as a means to monitor status of a resource may be appropriate for species with low percent standard errors for the above estimates. For example, the growth/removals ratios (0.04 vs. 1.99), mortality as a percent of volume (4.06 vs. 0.72), and percent change in volume (-11.24 percent vs. 3.91 percent) of slippery elm (former) versus sugar maple (latter) can be compared. Slippery elm’s lower growth removals ratio, higher mortality as a percent of volume and negative percent change in volume indicate that it is of more concern than sugar maple. However, the low growth/removals ratio for slippery elm should be taken with caution due to the high sampling error percent for net growth.

Surface area models other than the simple truncated cone used here (Figure 4.11, Equation 4.2) are common, such as the two stacked conical frustum model used by Emery et al. (2014). In their study, stump diameter to dbh ratios and taper estimates were available. However, these estimates are not contained in the FIA database and are not readily available for many species. For the purposes of estimating change, consistent use of any good surface area model is appropriate.

Adding stump diameter measurements (two times $r_1$ in Figure 4.1) for each tree in the FIA database along with additional research to provide taper estimates for species valued for
Table 4.6: Average annual 2013 net growth, mortality, and removals and 2007 and 2013 net volume (in million cu. ft.) with [sampling error percent] of trees 5 inches dbh or larger on forestland. Does not include west Texas.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Average annual net growth</th>
<th>Average annual mortality</th>
<th>Average annual removals</th>
<th>2007 volume</th>
<th>2013 volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper birch</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>6047 [1.96]</td>
<td>5246 [2.02]</td>
</tr>
<tr>
<td>White willow</td>
<td>.132 [186]</td>
<td>.281 [63.1]</td>
<td>x</td>
<td>25.6 [49.2]</td>
<td>26.4 [59.8]</td>
</tr>
<tr>
<td>Pacific yew</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>40.0 [14.7]</td>
<td>40.9 [12.6]</td>
</tr>
<tr>
<td>Pawpaw</td>
<td>0.36 [40.5]</td>
<td>.082 [47.1]</td>
<td>0.20 [65.2]</td>
<td>3.10 [31.0]</td>
<td>2.77 [42.2]</td>
</tr>
<tr>
<td>Gingko</td>
<td>.004 [90.0]</td>
<td>x</td>
<td>x</td>
<td>1.37 [72.5]</td>
<td>.782 [99.4]</td>
</tr>
<tr>
<td>Two-needle pinyon</td>
<td>-22.5 [-]</td>
<td>119 [-]</td>
<td>*</td>
<td>7423 [2.36]</td>
<td>7473 [2.12]</td>
</tr>
<tr>
<td>Singleleaf pinyon</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>4167 [5.07]</td>
<td>3853 [2.86]</td>
</tr>
</tbody>
</table>

x none observed in sample
* not measured for significant portion of range
- total sampling error percent for multiple states in Interior West not available with EVALIDator
their bark would enhance the database to allow for alternative surface area calculations. In fact, including measurements for heights at four-inch tops of boles ($h_3$ in Figure 4.1) or height and diameter at some other height of the upper portion of the bole would facilitate research on taper. This would allow for models of the bole which more closely represent the shapes of individual species and may result in better surface area estimates for bark species.

4.8 Conclusion

The FIA database is a useful information source for monitoring tree species that are valued for NTFPs. The FIA sampling design provides good estimates over large areas and, in general, regularly re-measures plots. The database is improving with more comprehensive net growth, mortality, and removals estimates as western states complete their first cycles of annual inventory.

FIA data alone provide a good start for estimating quantities of bark products for tree species by calculating surface area. However, different bark components, such as inner bark versus outer bark, are valued depending on species. Furthermore, bark products are rarely traded in square feet. Additional research is needed to convert estimates made from tree measurements (e.g. bark surface area) to quantities relevant for trading NTFPs (e.g. dry weight of inner bark).

Our sample analysis of slippery elm along with completed analyses of the other species in Table 4.1 and examples of others studies involving NTFPs and FIA shows that FIA data can be augmented for various NTFP species and types of products, including syrup, nuts, and bark (Emery et al. 2014, Farrell 2012, Shaw et al. 2005). While procedures for estimating
quantities of some NTFPs have not incorporated FIA, procedures such as the one used by Blatner et al. (2005) to estimate noble fir bough weight could be adapted to incorporate FIA or utilized if additional measurements are included with FIA.

Access to reliable spatial and temporal data is vital to most effectively manage forestlands for NTFPs. FIA data can be used as an important component of a process that monitors NTFPs and manages forests in a manner inclusive of them. Tree data of this type and amount can alert the professional forester when a species is in jeopardy. Expert knowledge can refine measurements and techniques specific to each species. The use of FIA data for analysis of NTFPs can be replicated across species and wide geographic areas in an automated manner. However, using FIA to understand understory species is more limited. Subplot data on vegetation is currently constrained to western states and gives percent canopy cover data on subplots containing the resource. Expanding to all states and enhancing vegetation measurements could help quantify NTFP products derived from understory species.

4.9 References


USDA Forest Service. 2014. Forest Inventory and Analysis national core field guide Volume I: field data collection procedures for Phase 2 plots. USDA For. Serv., North Central Research Station FIA, St. Paul, MN.


Chapter 5

Conclusion

5.1 Recap of the big picture

Distributions of forest by stand age and type at various spatial scales provide valuable information for optimizing forest production and sustainability. Two sources of forest stand age distribution, volume, and species information are forest inventory plots and remotely sensed data. Plot and tree data from the US Forest Service Forest Inventory and Analysis (FIA) program are an integral component of large area forest inventories, providing precise field measured estimates from sample plots for areas equal to or larger than several counties (Bechtold and Patterson 2005). FIA has been developed mainly to estimate quantities and changes of timber products, such as sawtimber and pulpwood, along with biomass which provides information about carbon sequestration and potential bioenergy resources. However, a diversity of species including white pine, sugar maple, pawpaw, common persimmon, butternut, black walnut, sweetgum, yellow poplar, black cherry, white oak, white willow, sassafras, slippery elm, and noble fir are valued for their ability to produce NTFPs and are inventoried by FIA. Chapter 4 illustrates that the use of FIA data in its current state or
with some enhancements can and should be extended to monitor non-timber forest products (NTFPs), especially those derived from trees such as bark, nuts and fruits, sap, leaves and needles, and boughs.

Precise field measured inventory estimates of both timber and NTFPs from sample plots are possible over large areas but are costly for fine scale estimates across a large area due to the large number of sample plots required. Results from chapters 2 and 3 take steps to overcome difficulties of obtaining small area estimates of timber volume across large areas by identifying stand-clearing clumps of forest disturbances from yearly Vegetation Change Tracker (VCT) disturbance maps based on average disturbance magnitude, shape, size, and parcel metrics. This results in the ability to calculate stand age for objects that conform to harvest boundaries resulting in state sized or larger stand-level age maps.

### 5.2 Procedure for fine scale forest inventory estimates

Age and harvest boundary maps such as those illustrated in chapters 2 and 3 are valuable for producing detailed forest inventory information at a fine spatial scale, like the forest stand or parcel, across a broad spatial domain like a state or larger region. For instance, information like volume and biomass (tons per acre) by broad species group (hardwood/softwood), and relative quantities of different types of products (pulpwood/sawtimber) can be produced. Massive quantities of digital spatial data are available throughout the US for producing this information.

This digital spatial data includes VCT. As demonstrated by this research it does a good job of identifying forest, and within land cover identified as forest it does not miss stand-clearing
disturbances such as clearcut harvests and intense fires. There is evidence that this is the case in Virginia (Kauffman and Prisley 2016). In addition, an adequate sample of forest plots is available through FIA at the state level, for states the size of Virginia, from perhaps just one inventory year for the reference data needed to calibrate and validate algorithms for reclassifying VCT disturbances as stand-clearing, thinning, or small non-harvest disturbances. Smaller states may need to add more plots by including additional inventory years or extending the sample into neighboring states. With the help of Forest Service officials it is possible to extract values from enhanced VCT rasters and other layers derived from remotely sensed and auxiliary data that coincide with actual sample plot locations. These other layers include land cover and land use data for forest type, LiDAR metrics for determining height and density, and perhaps soils data. A flow chart illustrating this information generation process can be seen in Figures 5.1 and 5.2.

Three key uncertainties in this process are reclassifying disturbances as stand-clearing, thinning or small non-harvest disturbances, estimating the actual age of forest stands as opposed
to years since clearing, and including the effects of varying management intensities on volume. Chapters 2 and 3 show that machine learning tools such as random forest, SVM and CART can accurately reclassify VCT disturbance clumps using metrics for size, shape, average disturbance intensity, and parcel derived metrics (Kauffman and Prisley 2016). However, some bias is present in both the eVCT sample from Kauffman and Prisley (2016) and the peVCT sample in Chapter 3. The result is that, on average, eVCT seems to underestimate forest stand age due to an imbalance toward errors in which non-stand-clearing disturbances are not reclassified as such. On the other hand, peVCT may have a propensity for an imbalance in errors toward those that reclassify stand-clearing disturbances as non-stand-clearing disturbances when they really are not, resulting in average overestimation of age. Other sampling procedures that more accurately represent disturbance types, perhaps proportionally to area, could rectify this problem.
It is possible that the actual age of a stand can be more properly estimated using the number of years since a stand-clearing disturbance was first classified as forest by VCT. However, there will be many pixels that never return to forest because of land cover change or because they were last cleared too close to the year of the last VCT map. Additionally, maps of thinnings produced using procedures for classifying partial harvests in Chapter 3 can be used in conjunction with other datasets or variables such as number of years to return to forest, shape metrics of the disturbance clump, proximity of the disturbance to a road, and slope to facilitate estimation of management intensity. An additional pine plantation forest type could be incorporated into the forest type input for Figure 5.2. Separate regression models could be calibrated based on deciduous, mixed, natural pine, and planted pine forest types, or other methods such as nearest neighbor imputation could be used. Further research in these areas can minimize these uncertainties.

5.3 Additional research opportunities

Further research on accurately combining disjoint disturbance clumps under one harvest identity is valuable for applications in which total harvest area is important such as input data for the Spatial Wood Supply Simulator (SWSS), an agent-based model for timber supply-chain modeling (CeNRADS 2016). Clumps of neighboring pixels with the same age as determined by using Vegetation Change Tracker (VCT) data and the algorithms developed in Chapter 2 provide a good start for defining harvest boundaries. However, the discussion and conclusion of Chapter 2 outline several obstacles and difficulties with directly using the age proxy maps for this purpose. For instance, some harvests are spatially disjoint because
forest in general or the type of forest of interest to the harvester is spatially disjoint within the ownership parcel.

LiDAR and Landsat data have been used in conjunction with plot data, and auxiliary geospatial data to estimate forest inventory and timber volume at small scales (Goerndt et al 2012, Hudak et al 2007, Parker and Evans 2004). In addition, Lidar has been used to delineate forest stands (Lepännen et al 2008). Additional research can help to overcome the obstacles to delineating stand boundaries as outlined in Chapter 2. Age maps derived from VCT may not result in stand sized clumps for contiguous clumps of forest that have not been disturbed since 1984. Likewise, neighboring stands that were cleared the same year appear as one clump. However, it is possible that LiDAR could detect stand-level differences in both average height and variation in height for contiguous chunks of persisting forest and neighboring stands cleared the same year. A comparison of forest stand objects created using LiDAR metrics in conjunction with methods such as those from Lepännen et al. (2008) with forest stand objects created by clustering equal “age” pixels as in Chapters 2 and 3 would be useful. Both of these resulting sets of forest stand objects could be compared to forest stands delineated by experts using aerial photography. While currently not the case, evolving technology and a rise in value may quite possibly lead to comprehensive and frequent LiDAR coverage.

5.4 Conclusion

This research paves a path and take strides in quantification of abundance of forest tree products across time and at fine spatial scales over broad areas. In the process it provides a
basis for tools that can be used to better understand forest disturbance and land use change
dynamics. Forest dynamics and spatial and temporal changes in forest products parallel
changes in other resources related to forests such as water and wildlife, providing knowledge
and tools that can be used for monitoring an array of natural resources.

5.5 References

Program - national sampling design and estimation procedures. USDA For. Serv. Gen.

Center for Natural Resources Assessment and Decision Support (CeNRADS). 2016. Wis-
consin Wood Supply Assessment. 67 p.

forest attributes using ground data and remote sensed auxiliary information. Forest Science.
59(5):536-548.

neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data.


delineation of forest stands from LiDAR data. In: ISPRS Conf. on GEOBIA 2008.

Parker, R.C., and D.L. Evans. 2004. An application of LiDAR in a double-sample forest