New Remote Sensing Methods for Detecting and Quantifying Forest Disturbance and Regeneration in the Eastern United States

A Dissertation Presented for the

Doctor of Philosphy

Degree

The University of Tennessee, Knoxville

Michael Joseph Hughes

August 2014

© by Michael Joseph Hughes, 2014
All Rights Reserved.

Acknowledgements

First and foremost, my thanks to Doug Kaylor for all of his support and insight, both as my husband and as the other half of the 'Forest Ecology Lab'. My thanks as well to my supervisors, Lou Gross and Dan Hayes, and to my comittee members, Paul Armswoth and Carol Harden, for their good advice throughout. Additionally, my appreciation to SCALE-IT and NASA for tuition and stipend support throughout graduate school, as well as funding this research. Finally, my thanks to those who made my graduate experience what it was: Jen Krauel, Josh Birkebak, Tyler Pannell, Jonathon Pruitt, Alex Pilote, and Anthony Walker for stimulating ecological conversation over many dinners; Cynthia Peterson for providing SCALE-IT and Harry Richards for his mentorship and friendship; the Armsworth Lab for adopting a stray; and all the other students in SCALE-IT and the department of Ecology and Evolutionary Biology who have been good friends these past five years.

Abstract

Forest disturbances, such as wildfires, the southern pine beetle, and the hemlock woolly adelgid, affect millions of hectares of forest each year in North America with significant implications for forest health and management. This dissertation presents new methods to quantify and monitor disturbance through time in the forests of the eastern United States using remotely sensed imagery from the Landsat family of satellites, detect clouds and cloud-shadow in imagery, generate composite images from the clear-sky regions of multiple images acquired at different times, delineate the extents of disturbance events, identify the years in which they occur, and label those events with an agent and severity. These methods operate at a 30×30 m spatial resolution and a yearly temporal resolution. Overall accuracy for cloud and cloud-shadow detection is 98.7% and is significantly better than a leading method. Overall accuracy for designating a specific space and time as disturbed, stable, or regenerating is 85%, and accuracy for labeling disturbance events with a causal agent ranges from 42% to 90%, depending on agent, with overall accuracy, excluding samples marked as 'uncertain', of 81%. Due to the high spatial resolution of the imagery and resulting output, these methods are valuable for managers interested in monitoring specific forested areas. Additionally, these methods enable the discovery and quantification of forest dynamics at larger spatial scales in a way other datasets cannot. Applying these methods over the entire extent of the eastern United States highlands reveals significant differences in disturbance frequency by ecoregion, from less than 1% of forested area per year in the Central Appalachians, to over 5% in the Piedmont. Yearly variations from these means are substantial, with disturbance frequency being twice as high as the mean in some years. Additionally, these analyses reveal that some disturbance agents, such as the southern pine beetle, exhibit periodic dynamics. Finally, although these methods are applied here to the problem of forest disturbance in the eastern United States, the core innovations are easily extended to other locations or even to other applications of landscape change, such as vegetation succession, shifting coastlines, or urbanization.

Table of Contents

Introduction **1**

1 Automated Detection of Cloud and Cloud Shadow **7**

1.1 Introduction 8

1.2 Background 10

1.3 Methods 12

1.3.1 Data Sets 12

1.3.2 Neural Network Classification 14

1.3.3 Spatial Post-processing 15

1.3.4 Classifier Assessment 16

1.3.5 Application: Obstruction-Free Summertime Composites 17

1.4 Results & Discussion 18

1.4.1 Network Selection 18

1.4.2 Comparison to FMask 20

1.4.3 Creating a multi-temporal composite 25

1.5 Conclusions 28

2 Natural Disturbances Increase Canopy Heterogeneity **30**

2.1 Introduction 31

2.2 Methods 33

2.2.1 Automated Creation of Patches 33

2.2.2 Study Area and Datasets 34

2.2.3 Statistical Analysis 36

2.3 Results 36

2.4 Discussion & Conclusions 38

3 Automated Detection of Forest Disturbance Events in Space and Time **40**
3.1 Introduction 41
3.2 Methods 44
3.2.1 Study Area 44
3.2.2 Data Processing 44
3.2.3 Accuracy Evaluation 54
3.2.4 Changescape 55
3.3 Results & Discussion 56
3.4 Conclusions 61

4 Algorithmic Attribution of Forest Disturbance Agents in Eastern Forests **62**
4.1 Introduction 63
4.2 Methods 65
4.2.1 Classifier Construction 65
4.2.2 Classifier Assessment 69
4.2.3 Estimation of Natural Forest Disturbance by Agent 69
4.3 Results & Discussion 70
4.3.1 Classifier Evaluation 70
4.3.2 Natural Disturbance by Agent in the Eastern US 74
4.4 Conclusions 76

Conclusion **79**

Bibliography **81**

Vita **93**

List of Tables

1.1 Network configuratations for cloud-detection neural networks 14
1.2 Multiway ANOVA of accuracy over the 12 evaluation subscenes for all cloud-detection neural networks, without applying post-processing spatial procedures. 19
1.3 Multiway ANOVA of accuracy over the 12 evaluation subscenes for all cloud-detection neural networks after applying spatial post-processing procedures. 21
1.4 Agreement over all 12 test sub-scenes for SPARCS and FMask compared to the evaluation masks. 21
1.5 Agreement over all 12 test subscenes for SPARCS and FMask compared to the evaluation masks. 26

2.1 ADS polygon count by disturbance type from each test area (A-L) used in analysis. 36

3.1 Agreement between VeRDET and expert segmentation. 58

4.1 Number of pixels used in training and evaluation by region and disturbance agent. 66
4.2 Multi-way ANOVA over the kappa scores from classifiers with different combinations of inputs calculated from classifying all pixels within ADS polygons from the 8 evaluation regions. 71
4.3 Effect size of input parameters on classification accuracy as measured by the difference in kappa calculated over samples from only the evaluation regions and from all regions. 71
4.4 Agreement matrix between disturbance agent classification and ADS labels. 74
4.5 Mean percent of forested area affected by disturbance agent. 76

List of Figures

1 G. E. Davis Lumber Company engine transporting lumber to the sawmill. Undated photograph 1912-1919; Bristol, TN. George Evan Davis collection, Archives of Appalachia. 2

2 Cotton grass and sphagnum bog. 2006; Dolly Sods Wilderness Area, Monongahela National Forest, Randolph Co. WV. Courtesy ForestWander.com 2

3 Informational sign during reforestation of the newly acquired Pisgah National Forest. 1915; Pisgah National Forest, NC. Courtesy United States Forest Service. 4

4 Standing dead Fraser Fir with young recruits. 2012; Clingman's Dome, Great Smoky Mountains National Park. Courtesy Doug Kaylor. 4

1.1 WRS2 path/row locations of imagery used in training (*black*) and testing (*red*) of the cloud detection neural networks. 13

1.2 Mean ranks of different methods for including spatial information in the network. Methods with the same letter are not significantly different by Tukey's honestly significant difference criterion. 20

1.3 Cloud and cloud-shadow image classification of a subscene from New South Wales using SPARCS and FMask 23

1.4 Cloud and cloud-shadow image classification of a subscene from Hidalgo, Mexico using SPARCS and FMask 24

1.5 Results from compositing clear-sky portions of four images in a region in eastern Tennessee. 27

2.1 Timeseries of NBR for two selected disturbances. 32

2.2 Locations of the twelve test areas with widespread disturbance as identified by US Forest Service Aerial Detection Surveys 35

2.3 Distributions of the five-year change in means and patch-variance within ADS polygons . . . 37
2.4 Estimates and confidence intervals of change in patch variance against change in vegetation index mean within ADS polygons. . . . 38

3.1 Study area showing mosaiced Landsat scenes for assessing VeRDET 45
3.2 2-dimensional TVR denoising at differing strengths subsumes more or less heterogeneity into patches. . . . 51
3.3 Dissimilarity scores for the 132 forested evaluation pixels. . . . 57
3.4 Segmented NDMI time-series (*dots*) for two different evaluation pixels from VeRDET (*dashed lines*) and the expert (*solid*). . . . 59
3.5 Changescape of eastern Tennessee / western North Carolina over 28 years. 60

4.1 Locations of regions used in training (*black*) and evaluation (*red*). . . . 65
4.2 Data thinning process within ADS polygons. . . . 68
4.3 Average kappa of disturbance agent classification networks with different numbers of hidden nodes. . . . 72
4.4 Percent of correct samples given maximum membership value. . . . 73
4.5 Classifier results for a known fire proceeded by a southern pine beetle outbreak. . . . 75
4.6 Percent of forest area affected by disturbance agent for selected ecoregions. 77

List of Attachments

A changescape of the eastern United States ..... `changescape-60.tif`

Introduction

In a fictionalized account of his mid-nineteenth century gentleman's hunting party, David Hunter Strother (1857), writing as Porte Crayon for Harper's Magazine, described a laurel brake in the primordial forest of Randolph County Virginia's Appalachian highlands:

> The hunters had been dodging the laurel-breaks all day. They seemed to dread the passage, and would frequently go miles around to avoid it. They had heard stories of men who had spent days in them, wandering in circles, and who had finally perished from starvation. ... With the horses the passage could not even be attempted without a previous clearing of the way by the ax-men. Upon consultation, it was considered necessary to cross the brake before them. ... They sank up to their knees in mud and water; they were throttled by the snake-like branches of the laurel, and were frequently obliged to resort to their hunting-knives to extricate a leg or an arm from its grasp. Ascending the stump of a riven hemlock, a striking picture presented itself. The laurel waved up and down as far as the eye could reach, like a green lake, with either shore walled by the massive forest, and out of its bed, rising singly or in groups of three or four, the tallest and most imposing of the fir species. The heads of our adventurers sometimes appeared hidden as they struggled through, and whether visible or invisible, the crackling of branches, the rustling of leaves, and the rolling fire of execrations marked their progress. All else was silent.

By the early twentieth century, the dense forests of the eastern United States that daunted early European settlers had been tamed by industrial logging and rail (Figure 1). What was not taken for lumber was harvested for charcoal to fuel iron furnaces. The slash – those small trees and limbs cut from larger trunks – that avoided the furnaces burned in prolonged fires sparked by lightning and rail cars. By 1920, the entirety of West Virginia had been logged. The soil of the temperate jungle recounted by Strother had burned hot and long enough to bake the clay into a bowl, transforming the forest into what is today a mountain top bog (Figure 2).

The passing of the Weeks Act in 1911 authorized the US federal government to purchase and preserve land, in part as a reaction to the lack of conservation by timber

Figure 1: G. E. Davis Lumber Company engine transporting lumber to the sawmill. Undated photograph 1912-1919; Bristol, TN. George Evan Davis collection, Archives of Appalachia.

Figure 2: Cotton grass and sphagnum bog. 2006; Dolly Sods Wilderness Area, Monongahela National Forest, Randolph Co. WV. Courtesy ForestWander.com

companies and also to maintain streams and waterways. These lands would become the national forests. However, because the eastern landscape was already under private ownership, most of the forest land purchased was, in the words of William Shands and Robert Healy, the lands nobody wanted (Shands and Healy, 1977). As a result, the national forests are, and have always been, fragmented with private inholdings and continual disturbance from extractive industries. The degree to which the preserved lands began as forests is illustrated by the 1915 spruce forest that would become the Pisgah (Figure 3).

In 1934 the Great Smoky Mountains National Park became one of the few regions of eastern forest to escape complete timbering. But even though locomotives and band saws would not enter the heart of the forest, the effects of human progress would still be felt in the form of invasive insects and disease. In the first half of the twentieth century, chestnut blight would nearly extirpate chestnut from the forest, killing over half of the hardwood trees in the Smoky Mountains (Whittaker, 1956). In the second half of the twentieth century, the balsam woolly adelgid would kill 80% of the high elevation conifers by decimating Fraser fir populations (Figure 4). Currently, the hemlock woolly adelgid is killing the dominant species of the low elevation valleys.

The Great Smoky Mountains are not unique in their history of invasion, and indeed benefit from active conservation programs. Today, hemlock woolly adelgid ranges northward to Maine. Invasive gypsy moth defoliates hardwoods in the northern Appalachians, while newly introduced species of anthracnose selectively remove *Cornus* species throughout their range. Additionally, in the Central Appalachians wildfires are burning an increasing number of ha each year (Lafon et al., 2005). Throughout, though, direct anthropogenic disturbance outpaces these so-called 'natural' disturbances (Powell et al., 2014). While old fields afforest, other regenerating forest stands are cut for new developments. And, of course, exploitative extractive industry in Appalachia did not stop with timbering; 6.8% of forest land in the Southern Appalachian region is estimated to have been lost to mountaintop removal mining (EPA, 2005).

The history of intense disturbance in the eastern United States created large unforested areas that were reforested through a mix of natural dispersal and human assistance. The initial deforestation reduced seed sources for the extracted trees species and changed local environmental and soil conditions. The forests that regrew are unlike the ones they replaced and may currently be undergoing some type of successional dynamic. The results are forest communities that are diverse and in flux.

The following chapters represent an attempt to quantify natural forest disturbances from fire, insects, and disease over recent decades in the central and southern Appalachian

Figure 3: Informational sign during reforestation of the newly acquired Pisgah National Forest. 1915; Pisgah National Forest, NC. Courtesy United States Forest Service.

Figure 4: Standing dead Fraser Fir with young recruits. 2012; Clingman's Dome, Great Smoky Mountains National Park. Courtesy Doug Kaylor.

highlands at a 30×30 m spatial resolution and a yearly temporal resolution using satellite imagery. Disturbance extents and severities are known for many individual disturbances in localized areas, but only within the last few years have overall disturbance estimates for the entire region been produced. These have also relied on satellite imagery. However, the mix of species, the fact that forest change occurs more or less naturally in the region, and the fact that many insects and diseases selectivity affect only some species and therefore only subtly change canopies, makes quantifying forest disturbance events using the spectral information from satellite imagery challenging. Because of this, no other methods systematically break down impacts into smaller regions or differentiate between disturbances caused by more than two agents as is done here.

Chapter 1 describes a new automated algorithm to address the general problem of detecting clouds and their shadows in Landsat Thematic Mapper imagery. Since forests transpire and since low, wet, air masses are forced into higher, colder regions as they pass over the mountain peaks of the Appalachians, clouds and their accompanying shadows obstruct the majority of the views of the land surface from space. Existing methods leave substantial clouds and shadows undetected; these bright or dark deviations can appear to be disturbed areas. Additionally, most methods overestimate the extent of cloud and cloud-shadow objects in imagery in order to maximize the amount of contamination removed, which reduces the total number of clear-sky observations. The new method outperforms a leading existing method and is reliable enough that the output does not need to be manually inspected or corrected. Using this output, the clear-sky regions of several images acquired during the same summer are composited and used in the other analyses described.

In Chapter 2, an increase in canopy heterogeneity, as measured by several vegetation indices, is shown to follow disturbance events in eastern forests. This may be due to disturbance agents removing only some of the canopy biomass, possibly because of species-dependent selection. In addition, a new metric is introduced that uses satellite imagery to define patches of similar forest and also to quantify the amount of canopy heterogeneity within those patches directly from satellite imagery. This metric is then shown to provide additional information to discriminate between disturbance agents.

Chapter 3 provides details on a new, automated method to identify the locations, dates, and severity of both slow and acute forest disturbance events from a time-series of Landsat images. In an approach similar to LandTrendr (Kennedy et al., 2010), the method identifies break points in a time series of a vegetation index, and then fits a piecewise linear equation between the break points. In addition, the new method enforces spatial

cohesion of similar land cover patches and exposes a parameter that determines patch scale by limiting the level of heterogeneity allowed within a patch. In this way, disturbances with different characteristic spatial scales can be identified. Additionally, the identification of breakpoints in the time-series follows a similar mathematical framework as the patch creation, presenting the opportunity for future work to merge both the spatial and the temporal segmentation of disturbance processes into a single, unified step.

Chapter 4 exploits the results of chapter two to create a new method that labels disturbance events identified, using the method in chapter three, with one of several known agents. First, different types of information are explored for use in the classifier, then the best performing combination is selected. Overall, the classifier correctly labels approximately 70%-80% of the samples in the evaluation dataset, and broad patterns of disturbance outbreaks derived from the generated maps correspond with the known recent history.

Finally, a changescape of the central and southern Appalachians, showing disturbance severity and trend, is appended to this dissertation (`changescape-60.tif`). This file is in the geotiff format with a spatial resolution of 60 m and uses the same scheme as Figure 3.5.

Chapter 1

Automated Detection of Cloud and Cloud-shadow in Single-date Landsat Imagery Using Neural Networks and Spatial Post-processing

This chapter was first published in Remote Sensing as a research article by myself and Daniel Hayes (Hughes and Hayes, 2014). I formulated the questions, developed the method, performed the research, and was the primary author of the manuscript.

Abstract

The use of Landsat data to answer ecological questions is greatly increased by the effective removal of cloud and cloud shadow from satellite images. We develop a novel algorithm to identify and classify clouds and cloud shadow, SPARCS: Spatial Procedures for Automated Removal of Cloud and Shadow. The method uses a neural network approach to determine cloud, cloud shadow, water, snow/ice and clear sky classification memberships of each pixel in a Landsat scene. It then applies a series of spatial procedures to resolve pixels with ambiguous membership by using information, such as the membership values of neighboring pixels and an estimate of cloud shadow locations from cloud and solar geometry. In a comparison with FMask, a high-quality cloud and cloud shadow classification algorithm currently available, SPARCS performs favorably, with substantially lower omission errors for cloud shadow (8.0% and 3.2%), only slightly higher omission errors for clouds (0.9% and 1.3%, respectively) and fewer errors of commission (2.6% and 0.3%). Additionally, SPARCS provides a measure of uncertainty in its classification that can be exploited by other algorithms that require clear sky pixels. To illustrate this, we present an application that constructs obstruction-free composites of images acquired on different dates in support of a method for vegetation change detection.

1.1 Introduction

The Landsat archive provides an unprecedented opportunity to discover how our landscape has changed over the last 30 years. Much of the imagery, however, is contaminated with clouds and their associated shadows, particularly in the tropics and other forested areas with high transpiration (Ju and Roy, 2008). Therefore, the usefulness of this imagery for landscape change studies depends on reliably separating clear-sky regions from those obstructed by clouds and cloud-shadow. Because of the large number of scenes over multiple dates needed for such studies, accurate and reliable automated methods are essential for this task.

Significant work has been devoted to cloud and cloud-shadow identification. Many algorithms for cloud and cloud-shadow masking have been developed for other sensors,

particularly for AVHRR (Saunders and Kriebel, 1988; Derrien et al., 1993; Cihlar and Howarth, 1994; Simpson and Stitt, 1998) and MODIS (Ackerman et al., 1998; Gao and Kaufman, 1995; Luo et al., 2008). Some of these algorithms have then been adapted for use on Landsat data (see for example Oreopoulos et al. 2011). Since clouds are bright and cold, and cloud-shadows are darker than the surrounding landscape, a common approach is to apply a threshold to the spectral values (Martinuzzi et al., 2007) or some simple function of two or more spectral values (Choi, 2004; Oreopoulos et al., 2011). An early example is the Automatic Cloud Cover Assessment (ACCA) (Hollingsworth et al., 1996), which uses a series of successive thresholds over bands and band combinations to define a hierarchical set of rules for clouds. ACCA was not designed for precise spatial detection of clouds, however, but rather to estimate the percentage of cloud cover in a given Landsat scene. These methods operate over imagery acquired on a single date; multi-temporal methods can leverage additional data to detect clouds and their shadows (Goodwin et al., 2013; Kennedy et al., 2007). Since clouds are typically bright objects in a scene, and shadows necessary darken an area, the obstructions can be identified by looking for outliers from a reference scene. In an automated approach, however, where the reference scene must be selected algorithmically, this is only effective when a good cloud-detection method for single-date imagery is already in use, or when most images are cloud-free and therefore clouds and cloud shadows are outliers from the mean, an assumption that may not hold in areas with frequent cloud-cover.

Cloud shadows are more difficult to identify than clouds because the spectral information does not discriminate between shadows caused by clouds and shadows arising from other causes, such as terrain. Additionally, other dark land covers, such as dark vegetation or water bodies, have similar spectral signatures to shadows. To address this confusion, the cloud mask itself has been used to distinguish cloud shadows, using the known sensor and solar geometry to estimate where cloud shadows should occur given the location of clouds (Berendes et al., 1992; Simpson and Stitt, 1998; Hagolle et al., 2010; Huang et al., 2010; Martinuzzi et al., 2007; Zhu and Woodcock, 2012; Choi, 2004). Such approaches are sensitive to the height of clouds above the land surface, as this height is proportional to the two-dimensional distance of a cloud from its shadow, as seen in imagery. Cloud height can be estimated using the thermal band (Huang et al., 2010; Zhu and Woodcock, 2012), and combining this with a digital elevation model, as in the algorithm from Huang et al. (Huang et al., 2010), can further reduce error.

In this paper we develop a novel method for identifying clouds from Landsat TM and ETM+ imagery: Spatial Procedures for Automated Removal of Cloud and

Shadow (SPARCS). This development was motivated by our need for efficient and reliable cloud and cloud-shadow masking in a forest change detection application over highly heterogeneous land cover in the eastern U.S; existing methods were either too computationally intensive or missed many clouds or cloud-shadows which were detected as change. Four design objectives directed our development. First, the method should only use bands present on all sensors and avoid ancillary data sources in order to ensure that the method is applicable over the entire archive. Second, the method should be completely automated and free from operator input. Third, the method should be sufficiently computationally efficient to be applied over thousands of scenes. And finally, the method should provide a spatially-explicit measure of classifier certainty that can be propagated to the products relying on the resultant cloud and cloud-shadow masks, a feature we are unaware of in other cloud-detection algorithms.

To meet these goals, we use neural network classifiers (Haykin, 2008) to explore different methods of using spatial information contained in a single-date Landsat scene to address the cloud and cloud-shadow detection problem. These classifiers are trained using scenes with clouds and cloud-shadow labeled by human operators at USGS (Scaramuzza et al., 2012) and evaluated using additional manually labeled data. Using this evaluation, we choose a high-quality classifier to become the basis of SPARCS and apply a series of spatial post-processing procedures to resolve ambiguous pixels in the classifier outputs. We then compare SPARCS to a high-quality, commonly used method, FMask (Zhu and Woodcock, 2012). Like FMask, we also include a class for water and snow/ice, for completeness. Finally, we demonstrate the utility of our method using an example that exploits the classifier uncertainty provided by SPARCS to combine a multi-temporal image stack into an obstruction-free composite.

1.2 Background

Neural networks are non-linear supervised learning algorithms that can be trained to partition an input space into a set of classes. Neural networks work by learning h linear combinations of the input data, where h is determined by the operator, and passing each of these through a given non-linear thresholding function. These results are temporarily stored as hidden values. Then, the network repeats the process by taking c linear combinations of those hidden values, where c is the number of desired classes, and again passing them through a given non-linear thresholding function. These results are then interpreted as the input observation's membership in each output class and are wholly

dependent on the weights of the linear combinations at both stages. These memberships are continuous values between 0 and 1; if desired, a crisp classification can be performed by setting the highest membership value to 1 and all others to 0. The weights themselves are learned using an optimization procedure over a training dataset that contains observations labeled with their correct class. The correctness of the labels in this training dataset directly controls the quality of the resulting classifier. Additionally, a higher number of hidden values (h) allows more complex patterns in the input data to be learned, though it also increases the likelihood of learning spurious correlations in the training data and thereby reducing the generality of the classifier (Haykin, 2008).

Importantly, the input data must include non-ambiguous information about the desired output classes to be able to discriminate between observations. This, however, is not the case when attempting to identify cloud and cloud-shadow from aspatial Landsat pixel data over a wide range of land cover types, as the exact same spectral data can be associated with pixels of clouds, snow, or some other bright and cold feature, or associated with cloud-shadow, terrain shadow, water bodies, or some other dark terrain feature. In short, pixel data by itself is ambiguous. As such, an additional source of information is needed to resolve these cases, such as elevation data, which is useful to distinguish cloud shadow from terrain shadow, observations from multiple time periods to filter ephemeral values, or spatial relationships between pixels. We are most interested in harnessing spatial information because it is already present in the Landsat scene and because human classifiers can almost always visually identify clouds and cloud shadows within a scene, spatial information should be sufficient for discrimination. In addition to applying spatial adjustments to the classifier output in a post-processing step, we examine two simple methods for incorporating space into neural network inputs. The intuition behind both methods is that, by providing an estimate of 'average value' in a region, ambiguity caused by variation within objects could be reduced, creating a simpler problem for the neural network to learn. The first method is to simply calculate the mean spectral value within a neighborhood around each pixel. The second method uses the pixel values from the image after denoising using total variation regularization (TVR) (Rudin et al., 1992). TVR removes noise from an image, f, by finding a new image, u, that minimizes the functional:

$$\min_{u} \|u - f\|^2 + 2\alpha \, |\nabla u| \tag{1.1}$$

The first term is the sum of squared-errors between the new image and the original image; minimizing prevents the new image from diverging too far from the original. The second term measures the magnitude of differences between adjacent pixels; minimizing favors

outputs with similar adjacent values. Taken together, the method balances smoothing parts of the image with keeping original details. How this balance is struck is determined by α, which should be positive, with smaller values favoring more detail and larger values favoring more smoothing. Due to the nature of the gradient term (Rudin et al., 1992; Goldstein and Osher, 2009), the smoothing manifests as regions of the image with constant values and sudden discontinuous jumps at the edges. This constant value can be thought of as an average of the spectral values within the region.

1.3 Methods

1.3.1 Data Sets

Manually-generated cloud masks from the USGS LDCM Cloud Cover Assessment Data (Scaramuzza et al., 2012), were subset, refined, and used as a training dataset for several different neural network classifiers. The core-cloud and thin-cloud classes, which were separate in the USGS data, were combined into a single cloud class. Of the 157 scenes in the dataset, only the 18 scenes with more than 1% of pixels in both the cloud and cloud-shadow classes were considered. From these, twelve representative scenes (Figure 1.1) from different hemispheres and latitudes were selected for the training dataset; the other six scenes had inadequately accurate cloud and cloud shadow masks for training. To reduce the amount of data used while retaining land cover variability within scenes, four 1000×1000 pixel (30×30 km) regions, each separated by at least 2000 (60 km) pixels, were subset from each of these twelve scenes. These 48 subscenes were used for training only.

The 48 subscene masks were visually assessed for classification label accuracy. The USGS masks provided classes for clear-sky, cloud, and cloud-shadow. Additionally, water and snow/ice classes were added to the masks. To do so, the Normalized Difference Snow Index (NDSI) (Hall et al., 1998) and tassel-cap brightness (Crist, 1985) were used to locate potential water and snow/ice regions in the imagery. These proposed masks were then hand-edited to include regions missed by the thresholding and remove regions incorrectly included. The water and snow/ice masks were then combined with the USGS masks using the follow precedence rules. Pixels labeled as cloud in the USGS mask were unchanged. Pixels labeled clear-sky in the USGS masks and flagged as either water or snow/ice were changed to water or snow/ice, as appropriate. Pixels labeled cloud-shadow in the USGS masks and flagged as water were hand-set appropriately. Cloud-shadows over water were labeled as cloud-shadow where a distinction could be made.

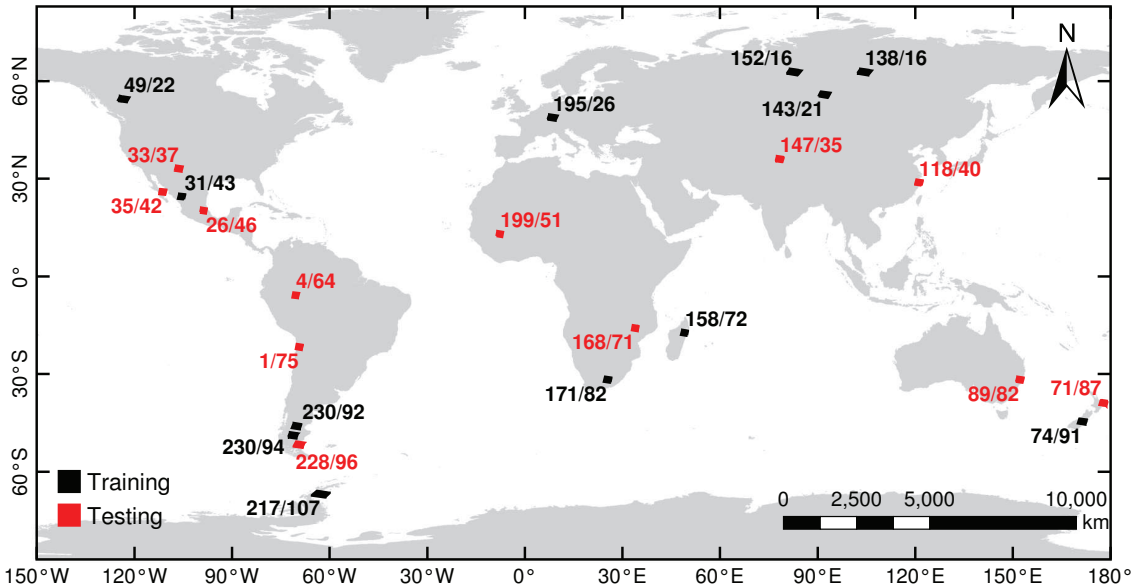

Figure 1.1: WRS2 path/row locations of imagery used in training (*black*) and testing (*red*) of the cloud detection neural networks.

Twelve additional scenes were selected for testing, and one 1000×1000 pixel (30×30 km) subscene was extracted from each. These 12 subscenes were then clustered over their spectral data using k-means clustering (MacQueen, 1967) with 100 cluster seeds, which sorts pixels into 100 groups with similar spectral values. Each resulting cluster was assigned to one of the five classes by an operator. Then, mislabeled image regions were hand-corrected using image editing software. These 12 subscenes were used only in classifier assessment, and not during training, in order to reduce the risk of neural network over-fitting during comparisons. All analyses comparing different networks and methods used only these testing scenes.

Landsat 7 ETM+ level 1T imagery and metadata for all scenes used in training and testing were acquired from the USGS archive. All scenes were from 2001, before the scan line corrector failure. The Landsat ETM+ data in bands 1-5 and 7 in each subscene were corrected to top-of-atmosphere reflectance (Moran et al., 1992; Chander et al., 2009) and then further corrected using dark-object subtraction (Chavez, 1996). The low-gain scaling of the thermal band (B_6) was converted to brightness temperature and then arbitrarily rescaled to values near the other bands to facilitate neural network learning:

$$\hat{B}_6 = B_6/100 - 2 \tag{1.2}$$

These values were used in all analyses.

Table 1.1: Network configuratations for cloud-detection neural networks. All configurations included aspatial ETM+ bands, but varied in network size (h) and the type of spatial inputs.

#	Network Size (h)	Spatial Averaging Method (calculated over Tassel Cap)	Intensity of Spatial Averaging
1	10	No Space	-
2	10	Local Average	5x5 Window
3	10	Local Average	9x9 Window
4	10	TVR	$\alpha = 0.05$
5	10	TVR	$\alpha = 0.10$
6	20	No Space	-
7	20	Local Average	5x5 Window
8	20	Local Average	9x9 Window
9	20	TVR	$\alpha = 0.05$
10	20	TVR	$\alpha = 0.10$
11	30	No Space	-
12	30	Local Average	5x5 Window
13	30	Local Average	9x9 Window
14	30	TVR	$\alpha = 0.05$
15	30	TVR	$\alpha = 0.10$

1.3.2 Neural Network Classification

A total of 15 neural network configurations were used to explore the role of classifier complexity and the inclusion of spatial information on classification accuracy (Table 1.1). Networks with 10, 20, and 30 hidden nodes were constructed using five different types of spatial inputs. In addition, all configurations included the aspatial spectral information (ETM+ bands 1-5,7) and rescaled brightness-temperature (ETM+ band 6) for each individual pixel. The first spatial input type added no spatial information and was used for baseline comparison. The second through fifth spatial input types added spatial information summarized from the first three components of the Tassel-Cap transformation (Crist, 1985). Spatial input types two and three added the local average in a region around each pixel in the three Tassel-Cap bands, using a 5x5 and 9x9 pixel neighborhood, respectively. Spatial input type four and five used tassel-cap pixel values after removing spatial noise using TVR (Rudin et al., 1992; Goldstein and Osher, 2009) with $\alpha = 0.05$, and $\alpha = 0.10$, respectively, which remove amounts of detail similar to the local averaging neighborhoods. The five types therefore represent a no-space baseline, plus two spatial-averaging methods each using two intensities.

Training data for the neural networks was randomly sampled from the 48 training subscenes after stratifying each subscene by class (cloud-shadow, cloud, water, snow/ice,

clear-sky). Where possible, 1500 pixels from each class were selected from each subscene. If a subscene did not have 1500 pixels of a class, all pixels of that class were selected. The aspatial spectral and brightness-temperature data, the 5x5 and 9x9 pixel averages in the Tassel-Cap indices, and the two TVR-denoised values of the Tassel-Cap indices were extracted for each pixel. This process was performed three times, with different stratified random samples selected each time, to generate three sets of training samples. A total of 166,639 samples were used for each network. Each network configuration was trained using each training set, generating a total of 45 networks consisting of the 15 configurations replicated three times.

Networks were trained using scaled conjugate gradient backpropagation by the MATLAB Neural Network Toolbox (Beale et al., 2013) `patternnet()` function.

1.3.3 Spatial Post-processing

Clouds and cloud-shadows are spatially-coherent objects in satellite imagery. SPARCS uses information on surrounding pixels to exploit this spatial coherency for the purposes of reducing classification error. In exploratory classifications using several different neural networks, error maps were generated to visually assess the spatial patterns of errors. From these observations, a series of six rules were developed to address spatially-definable error. Each of these rules operates over the continuous-valued membership images for each class. First, a 3x3 median filter is applied to the cloud and cloud-shadow membership images to reduce noise.

The second rule addresses confusion at water-land boundaries. Shallow water is often confused for clouds or snow/ice and the wet soil in the transition zone between water and land is often confused with cloud-shadow. To correct this, the cloud, cloud-shadow, and snow-ice membership of pixels within three pixels of large bodies of water are decreased.

The third rule uses sun and sensor geometry to identify areas of potential cloud-shadow using the cloud membership to reduce the significant ambiguity between hill-shade, wet ground, and cloud-shadow. Our approach closely resembles the method of Luo et al. (Luo et al., 2008) in that it defines a broad area of potential cloud shadow that accounts for a range of potential cloud heights, and then combines it with an estimate based on spectral values. Our method uses the cloud-shadow membership values for the initial estimate. The direction of the sun is first determined from scene metadata. Then, a copy of the cloud membership image is transposed away from the sun a distance determined by the sun elevation and a cloud height of 2250 m and then expanded (dilated) to include potential cloud heights from 1800 m to 2700 m above the ground, a height range

chosen to capture the dark shadows created by the optically thick cumulus clouds. This potential cloud shadow location is then further expanded and blurred using a 15x15 pixel filter to create a feasible zone of cloud-shadow. Finally, this estimate is multiplied with the cloud-shadow membership to increase cloud-shadow membership within the feasible zone and to reduce the cloud-shadow membership in areas outside of the feasible zone that likely represent terrain shadow erroneously classified as cloud-shadow.

The fourth rule addresses confusion between water and deep shadow, which can have equivalent spectral signatures. Pixels that have similar memberships in both cloud-shadow and water, meaning that the neural network classifier has identified them as ambiguous, are selected. Those that are also surrounded by pixels of high cloud-shadow membership then have their own cloud-shadow membership increased and their water membership decreased.

The fifth rule performs a similar function between clouds and snow/ice, which can have similar ambiguity, biasing membership toward clouds and away from snow/ice in pixels surrounded by clouds.

The final rule identifies pixels of high overall uncertainty and uses the membership of nearby pixels to predict the correct membership of the uncertain pixels. First, uncertainty is calculated as the variance between memberships, rescaled to be between 0 and 1. Then, a weighted average of nearby pixels is calculated for each membership class, with weights calculated as the product of each pixel's certainty and a Gaussian decay function over distance with $\sigma = 2$ pixels. Finally, new memberships are calculated as a linear combination of the original value and the spatial average, weighted using the pixel's uncertainty, such that more certain pixels retain their original value and uncertain pixels become more like the average value of pixels around them. This rule has the effect of homogenizing areas and removing noise.

1.3.4 Classifier Assessment

Each of the 45 neural networks was scored on each of the 12 evaluation subscenes based on the total classifier accuracy. For each pixel, the assigned class from the network was taken as the class with the maximum membership value over all classes. These classes were compared to the evaluation masks described in section 1.3.1. Because clouds and cloud-shadows are not discrete objects, there are many semi-obstructed pixels that form a transition zone between cloud or cloud-shadow and clear-sky. Since we are more concerned with identifying potential clouds and cloud shadows than with precisely

defining their extent, a three-pixel buffer around the areas labeled as cloud and cloud-shadow in the evaluation masks was constructed. Pixels within this buffer were scored as correct if they were labeled as either cloud or cloud-shadow, as appropriate, or the label in the evaluation mask. This reduced commission errors and increased overall accuracy for all methods, including FMask, by approximately 2%.

A multi-way ANOVA (Zar, 2010) was performed on the 540 accuracy scores to assess the contributions of classifier complexity and type of spectral information to accuracy, while accounting for the variations between subscenes. Networks with more inputs or more hidden nodes were not penalized for additional model complexity, since the purpose was to find the most effective classifier and not necessarily the most efficient.

After selecting a network to serve as the basis of our method, we then compared our method to FMask using the same classifier accuracy statistics derived from the testing data with buffered masks. We further compared the methods by calculating omission errors as the percentage of cloud or cloud-shadow pixels that were mislabeled as clear-sky and commission errors as the percentage of clear-sky pixels, outside of the buffered area, mislabeled as clouds or cloud-shadow. This was done because FMask has significant confusion between clouds and cloud-shadow, and we do not feel that including that confusion is relevant to the core question of whether the classifiers can separate clear-sky pixels from obstructions.

1.3.5 Application: Obstruction-Free Summertime Composites

Though SPARCS can provide a crisp classification wherein each pixel is labeled as exactly one class, by using the raw membership values, neural networks provide a measure of how certain the classifier is in assigning class membership. In this method we use these original membership values to create clear-sky composite images from a multi-temporal stack of Landsat TM scenes acquired within the same year. For the purposes of illustration, we selected four scenes acquired during late summer of 1990 from the same location in eastern Tennessee that had moderate cloudiness on visual inspection. Each of these scenes were classified using SPARCS to generate memberships in cloud, cloud-shadow, water, snow/ice, and clear-sky classes.

For each scene, the pixel memberships in the water and clear-sky classes were then combined to generate a clarity index (Q) for each pixel. Additionally, because including a marginally contaminated pixel is more harmful than excluding a marginally clear pixel, this index was squared.

$$Q = (m_W + m_L)^2 \tag{1.3}$$

where m_W is the water membership and m_L is the clear-sky membership. For the purposes of this example, snow and ice are considered as obstructions, as snow is seasonal in the area of interest. In high altitude or latitude areas where snow and ice are persistent features, the class should be included.

In order to reduce the influence of phenology, scenes are also weighted by a Gaussian decay function of day-of-year, such that scenes further away from a given date are weighted less than scenes near that date. For this example, we chose a late summer day (day of year 225) and used a standard deviation of 30 days:

$$w_j = \exp\left(-\left(\frac{d_j - 255}{60}\right)^2\right) \tag{1.4}$$

where d_j is the day of year that the j-th scene was acquired. Other days could be chosen, and comparison between composites weighted to different days could be fruitful to examine phenological effects, as long as care is taken that the decay function itself does not span significant phenological change.

Yearly summertime composites are then generated as a weighted average of all the summertime scenes each year, using the clarity index (Q) and the Gaussian-transformed distance from a target date (w_j)as weights. Each of the seven Landsat bands are computed independently:

$$A_b = \frac{\sum_{(j \in S)} B_{b,j} Q_j w_j}{\sum_j Q_j w_j} \tag{1.5}$$

where A_b is the composite image of band b, S is the set of selected scenes to be combined, $B_{b,j}$ is the image data of band b for scene j, and Q_j and w_j are the weights for scene j described above.

1.4 Results & Discussion

1.4.1 Network Selection

We trained 45 neural networks over 48 training subscenes and explored the effects of network size and type of spatial inputs on classification accuracy over 12 evaluation subscenes. The inclusion of spatial inputs into the neural network had no statistically significant impact on classifier accuracy over the evaluation dataset. (Table 1.2). Network size was significant at the 0.05 confidence level; in a post-hoc test using Tukey's Honestly Significant Difference (HSD) criterion (Kramer, 1956), statistically significant increases in total accuracy were seen in networks with 30 hidden nodes over those with 10 and

Table 1.2: Multiway ANOVA of accuracy over the 12 evaluation subscenes for all cloud-detection neural networks, without applying post-processing spatial procedures.

Source	Sum Sq.	df	Mean Sq.	F	Pr > F
Type of Space	0.004	4	0.001	1.51	0.198
Network Size (h)	0.004	2	0.002	3.18	0.043
Subscene	1.004	11	0.091	137.26	< 0.001
Space×Size	0.011	8	0.001	1.99	0.045
Space×Subscene	0.030	44	0.001	1.02	0.439
Size×Subscene	0.058	22	0.003	3.98	< 0.001
Error	0.298	448	0.001		
Total	1.408	539			

20 hidden nodes, which were similar (Figure 1.2). Absolute gains, though, were small; networks with 30 hidden nodes increased total accuracy by approximately 0.5%, or about a 15% decrease in error. The interaction term between network size and subscene was also significant, suggesting that the increase in accuracy is due to the ability of more complex networks to learn additional land-cover features, and that more complex networks are not simply overfitting training data but become more general classifiers.

After post-processing the neural network output with spatial procedures, classification accuracy increased overall, from approximately 94.5% to approximately 97%. The post-processing procedures also exaggerated differences between methods of including spatial inputs to the neural network, which became statistically significant by the ANOVA F-test (Table 1.3). Tukey's HSD separated the TVR method with $\alpha = 0.05$ from the methods that used a mean over a local neighborhood, with the TVR method with $\alpha = 0.10$ and the method with no spatial inputs being intermediate between the groups (Figure 1.2). However, the difference between the means of methods to incorporate space within the neural network are within 1% of total classifier accuracy, much less than the increase gained by inclusion of spatial post-processing procedures. The spatial procedures preserved the patterns in accuracy between networks with different numbers of hidden nodes. Importantly, the application of spatial procedures greatly enhances the effectiveness of methods that have no spatial inputs to the neural network, suggesting that at least part of the classification rules learned by the networks that incorporated space are replicated by the post-processing spatial procedures. Given that calculating the spatial inputs, particularly TVR, is computationally expensive, there is no clear choice for an operational method. We selected a network with 30 hidden nodes and no spatial inputs to the neural network for further evaluation and to use in our cloud and cloud-shadow

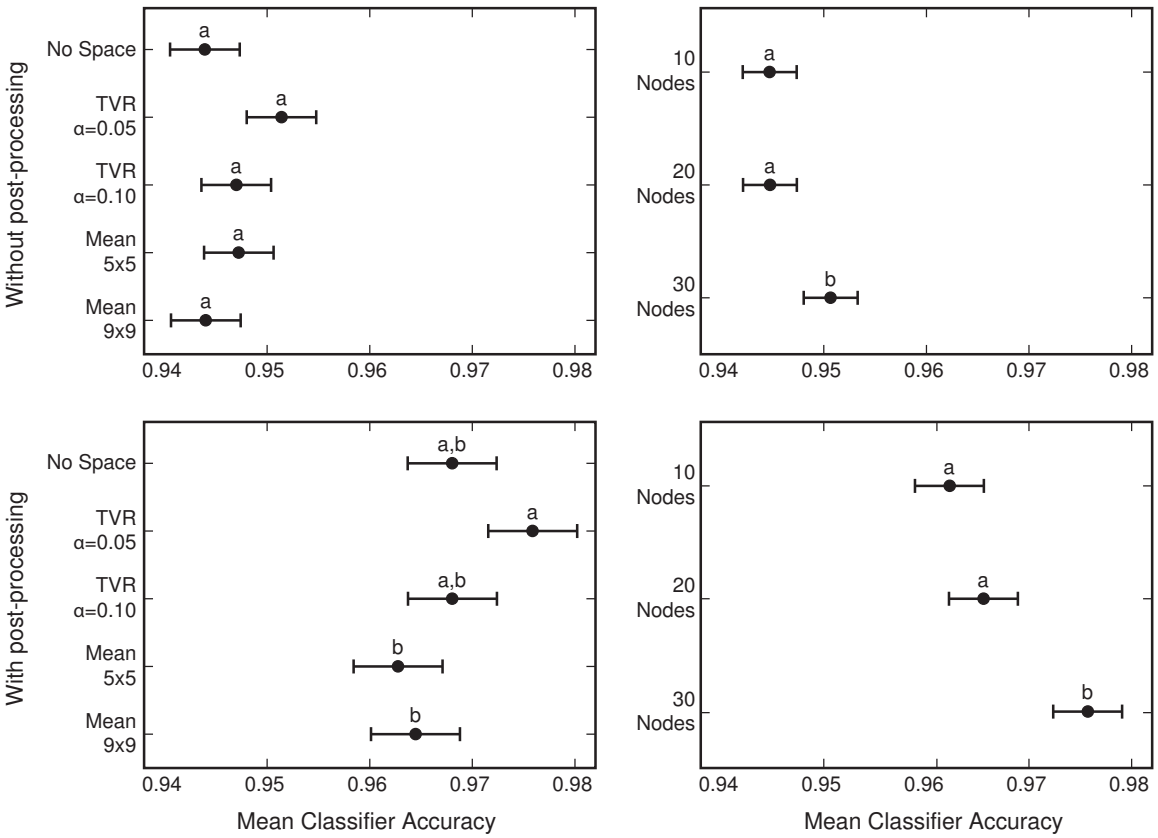

Figure 1.2: Mean ranks of different methods for including spatial information in the network. Methods with the same letter are not significantly different by Tukey's honestly significant difference criterion.

detection package, SPARCS: Spatial Procedures for the Automated Removal of Cloud and Shadow.

1.4.2 Comparison to FMask

SPARCS compares favorably with FMask over the 12 subscene evaluation dataset. The largest improvement is in correct identification of cloud shadow: SPARCS mislabels 3.2% of cloud-shadow pixels as clear-sky compared to FMask's 8.0%. Both methods perform well at identifying clouds, with FMask performing somewhat better by mislabeling 0.9% of cloud pixels as clear-sky compared to with SPARCS mislabeling 1.3%. Considering errors of commission, SPARCS performs substantially better by mislabeling 0.5% of clear-sky pixels as cloud-shadow and 0.2% as clouds, compared to FMask mislabeling 2.4% of clear-sky pixels as cloud-shadow and 2.8% as clouds.

Table 1.3: Multiway ANOVA of accuracy over the 12 evaluation subscenes for all cloud-detection neural networks after applying spatial post-processing procedures.

Source	Sum Sq.	df	Mean Sq.	F	Pr > F
Type of Space	0.011	4	0.003	2.53	0.040
Network Size (h)	0.018	2	0.009	8.17	< 0.001
Subscene	1.271	11	0.116	106.37	< 0.001
Space×Size	0.017	8	0.002	2.00	0.045
Space×Subscene	0.087	44	0.002	1.81	0.002
Size×Subscene	0.113	22	0.005	4.22	< 0.001
Error	0.487	448	0.001		
Total	2.004	539			

Table 1.4: Agreement over all 12 test sub-scenes for SPARCS and FMask compared to the evaluation masks.

Labeled as	Shadow	Cloud	Water	Snow/Ice	Clear
SPARCS					
Classed as Shadow	94.7%	1.0%	2.3%	1.7%	0.5%
Cloud	0.7%	97.2%	0.1%	0.9%	0.2%
Water	0.5%	0.0%	96.6%	1.0%	0.1%
Snow/Ice	0.9%	0.4%	0.0%	90.2%	0.0%
Clear	3.2%	1.3%	1.0%	6.2%	99.2%
FMask					
Classed as Shadow	69.9%	0.5%	0.6%	7.6%	2.4%
Cloud	20.9%	98.6%	0.3%	10.7%	2.8%
Water	1.0%	0.0%	96.6%	0.0%	0.0%
Snow/Ice	0.3%	0.0%	0.0%	72.4%	0.1%
Clear	8.0%	0.9%	2.4%	9.3%	94.7%

Spatial patterns of error are examined in Figures 1.4 and 1.3. A false color image of the scene mapping bands 5, 4 and 2 to red, green, and blue, respectively, is provided for reference (*top rows*). Classification output for SPARCS (*left*) and FMask (*right*) show agreement with the evaluation masks, with commission errors (*purple*) and omission errors (*red*) highlighted for clouds (*light colors*) and cloud-shadows (*dark colors*).

Figure 1.3 is of a scene with sparse mid-altitude clouds in the coastal region of New South Wales, Australia (WRS2 path/row 89/82). Both methods have strong cloud detection, though both miss some small, thin clouds in the northern portion of the image and thin clouds in the southwestern portion, as well as their respective shadows. Because the spectral signal of areas contaminated with thin clouds and their shadows is a mixture of cloud/cloud-shadow and the underlying landscape, they are especially difficult to detect,

as the resulting signal is ambiguous. Successful methods typically use multi-temporal image stacks, and then detect deviations from an average or consensus signal (Goodwin et al., 2013).

FMask predicts cloud-shadow by projecting the cloud mask onto the land-surface as a function of sun angle and topography without considering spectral information about shaded pixels. When this projection fails, this creates a pattern in the mask where the cloud-shadow mask is offset from the actual location of the cloud-shadow, as can be seen in several clouds in this image. SPARCS combines information about dark pixels from the neural network output with a similar, but much less precise, cloud projection approach to achieve more overall precision in cloud-shadow locations. But because of shadow/water ambiguity of dark pixels, this approach can causes over-shadowing in dark water near detected clouds, such as that in the eastern portion of the subscene.

A scene with dense, discrete clouds and cloud-shadows in Hidalgo, Mexico (WRS path/row 26/46) is presented in Figure 1.4. For both SPARCS and FMask, clouds are detected very well and most error occurs around the edges of cloud and cloud-shadow objects. In SPARCS, some bright landcover in the eastern portion of the image is misidentified as clouds with some spurious associated cloud-shadow. Additionally, some dark water is labeled as cloud-shadow. FMask exhibits some bright-landcover/cloud confusion as well, though less than SPARCS. Again, the cloud-shadow mask consistently fails to extend to the edges of cloud-shadow objects due to misprojection, resulting in substantial missed cloud-shadow while simultaneously labeling unshadowed areas as shadowed.

Both methods have a halo of commission error around cloud and cloud-shadow objects that results from design decisions to reduce contaminated pixels by expanding those masks slightly. This expansion approach creates a trade-off between commission and omission errors, with larger expansions capturing more obstructed pixels by sacrificing nearby clear-sky pixels. Much of this halo is ignored due to the three pixel buffer, described above, but some extends past that buffer and can be seen in the images. The larger halos in FMask represent the method's more aggressive efforts toward this goal.

Over all 12 evaluation subscences, SPARCS performs consistently better than FMask (Table 1.5). In only one subscene, which is predominately cloud cover over the Amazon rainforest, is the overall accuracy for SPARCS less than that for Fmask, and then only by 0.3%. In that scene SPARCS misses some thin clouds on the edge of the bulk of the cloud mass. FMask and SPARCS perform relatively well or poorly on the same subscenes, that is, when SPARCS performs well so does FMask, and vice versa, suggesting that in some

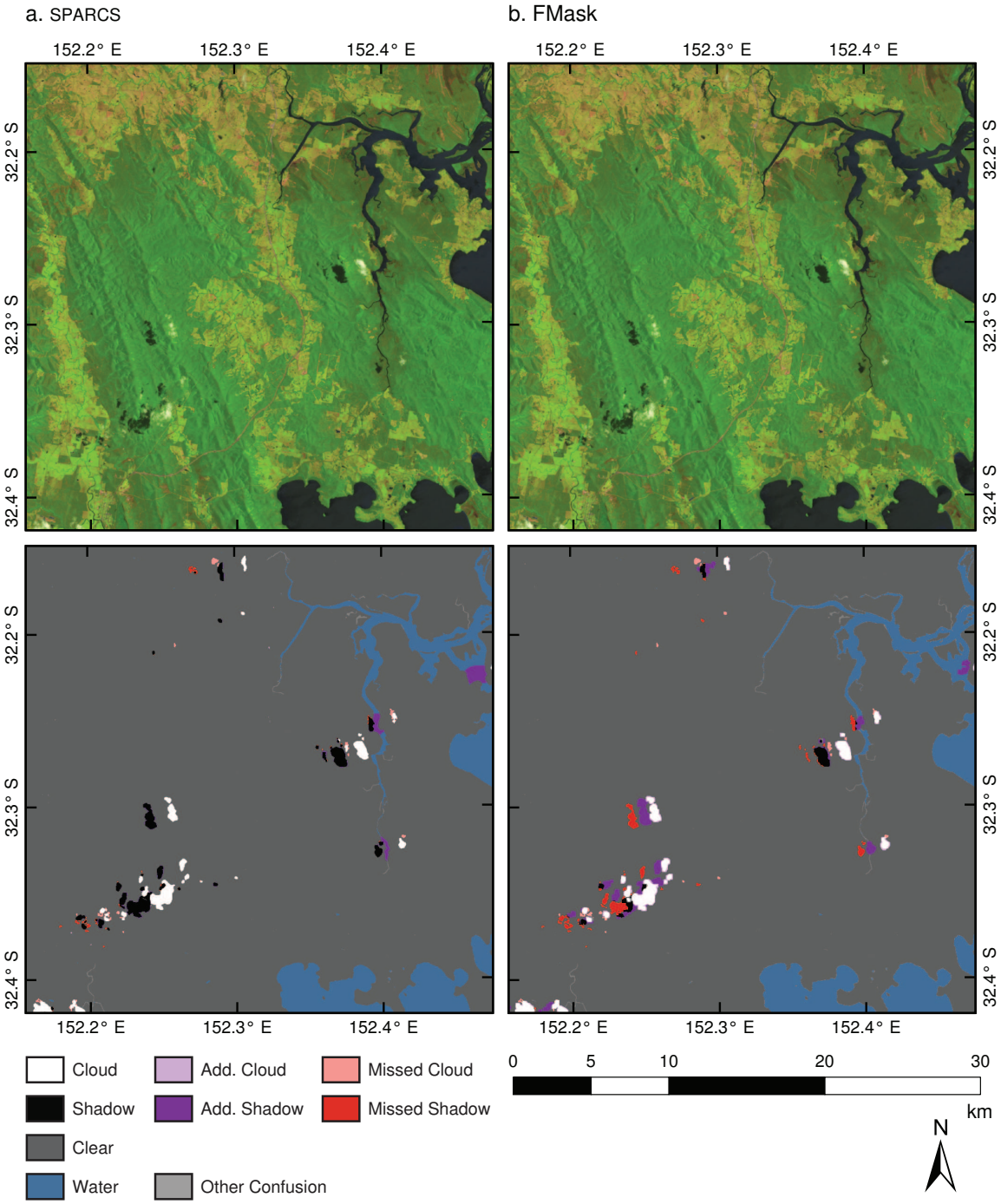

Figure 1.3: Image classification of a subscene from New South Wales, Australia (WRS2 path/row 89/82) acquired on April 21, 2001 using SPARCS (*left*) and FMask (*right*) with confusion between the classifications and evaluation masks highlighted.

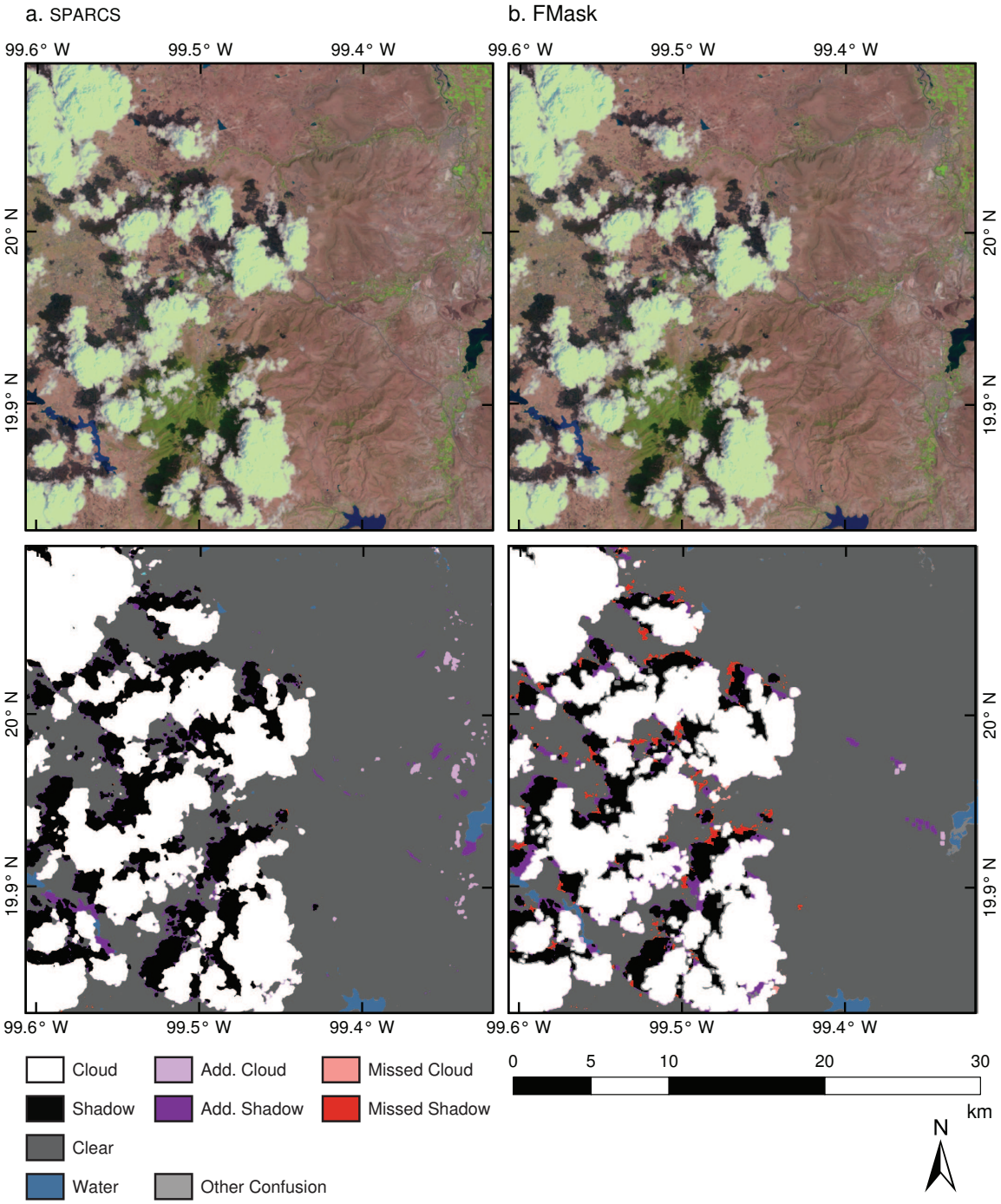

Figure 1.4: Image classification of a subscene from Hidalgo, Mexico (WRS2 path/row 26/46) acquired on February 1, 2001 using SPARCS (*left*) and FMask (*right*) with confusion between the classifications and evaluation masks highlighted.

subscenes separating clouds and cloud-shadow from clear-sky is simply a more difficult problem than in others.

In forest disturbance identification, mislabeling cloud and cloud-shadow as clear-sky (omission) is generally a larger problem than mislabeling clear areas as clouds or cloud-shadow (commission), because dark or bright spots can be mistaken for ephemeral disturbance (Huang et al., 2010). Commission errors become important, however, in areas with frequent cloud cover, such as the tropics or mountainous regions, where actual clear-sky views are rare. In these cases, proper identification of clear pixels is necessary to increase the number of observations and maximize the likelihood of comparing images with similar days of year and phenology. We believe SPARCS provides a balance between these competing objectives. Although SPARCS consistently misses more cloud cover than FMask, it does so at the gain of substantially reducing commission error in otherwise obstructed scenes, and so increases the likelihood of observing rare clear-sky pixels in cloudy areas.

1.4.3 Creating a multi-temporal composite

To illustrate the usefulness of our method, we generated a composite image from Landsat 5 TM images acquired over eastern Tennessee on four dates in the summer of 1990 (Figure 1.5). Though SPARCS was developed using data from Landsat 7 ETM+, since it does not use the panchromatic band it can be easily applied to the TM archive. Two of the images in this example have significant cloud cover. One scene, from July 6, is mostly clear-sky, though a few cloud and cloud-shadow pairs are detected and removed from the composite. A scene from August 16 contains several thin clouds. Scene weights shown are the sum of the clear-sky and water memberships from SPARCS output, multiplied by a deceasing function of distance from August 1, which is used to account for shifts in phenology as images are acquired further away from the target day. Additionally, the algorithm provides the total of the weights used (Figure 1.5, *bottom right*), which can be useful in calculations and analyses further down a processing pipeline to determine, for example, that insufficient data was available for an area during a certain year, or that a seemingly anomalous area actually has significant support.

In the scene from August 16, large portions of the thin cloud are assigned intermediate weights due to classifier uncertainty. By squaring these values, our algorithm trusts these areas substantially less than the same areas in clear images. By providing a continuous measure of uncertainty, though, different algorithms and operators can choose their own thresholds for how conservative they wish to be with data inclusion. The classifier

Table 1.5: Agreement over all 12 test subscenes for SPARCS and FMask compared to the evaluation masks.

	Missed Shadow	Missed Cloud	Over Shadow	Over Cloud	Overall
Jammu and Kashmir, India: pr 147/35 (36.4°N, 78.8°E). February 20, 2001.					
SPARCS	19.2%	10.1%	1.0%	0.6%	97.2%
FMask	6.9%	2.5%	7.2%	7.2%	86.8%
New Mexico, USA: pr 33/37 (33.5°N, 105.9°W). February 11, 2001.					
SPARCS	2.9%	0.2%	0.1%	0.0%	99.6%
FMask	3.4%	0.0%	4.7%	4.4%	92.4%
Zhejiang, China: pr 118/40 (28.6°N, 120.4°E). March 11, 2001.					
SPARCS	0.4%	0.1%	0.7%	0.2%	99.4%
FMask	5.0%	0.3%	3.3%	4.2%	94.9%
Baja California Sur, Mexico: pr 35/42 (25.5°N, 111.1°W). March 22, 2001.					
SPARCS	2.5%	0.0%	0.2%	0.0%	99.8%
FMask	17.1%	0.4%	0.2%	0.2%	99.2%
Hidalgo, Mexico: pr 26/46 (20.0°N, 99.5°W). February 1, 2001.					
SPARCS	0.2%	0.0%	1.0%	0.7%	98.9%
FMask	9.1%	0.5%	2.1%	0.5%	97.4%
Koulikoro, Mali: pr 199/51 (13.3°N, 7.3°W). January 30, 2001.					
SPARCS	0.0%	0.0%	0.0%	0.0%	100.0%
FMask	0.0%	0.0%	0.6%	0.3%	99.1%
Amazona, Brazil: pr 4/64 (5.2°S, 70.9°W). March 13, 2001.					
SPARCS	9.5%	1.9%	2.2%	0.8%	97.6%
FMask	9.4%	1.3%	8.4%	1.0%	97.9%
Tete, Mozambique: pr 168/71 (16.2°S, 33.3°E). April 10, 2001.					
SPARCS	0.0%	0.0%	0.0%	0.0%	100.0%
FMask	0.0%	0.0%	0.0%	0.0%	100.0%
Antofagasta, Chile: pr 1/75 (21.4°S, 68.9°W). January 20, 2001.					
SPARCS	17.7%	18.0%	2.5%	0.0%	96.4%
FMask	24.8%	7.4%	7.2%	13.5%	79.8%
New South Wales, Australia: pr 89/82 (32.3°S, 152.3°E). April 21, 2001.					
SPARCS	6.9%	11.2%	0.2%	0.0%	99.6%
FMask	59.3%	9.7%	0.5%	0.1%	98.9%
Hawke's Bay, New Zealand: pr 71/87 (39.2°S, 177.1°E). April 12, 2001.					
SPARCS	9.1%	0.8%	1.1%	0.2%	98.6%
FMask	12.7%	2.1%	0.1%	0.1%	99.8%
Santa Cruz, Argentina: pr 228/96 (51.9°S, 70.5°W). January 11, 2001.					
SPARCS	3.9%	1.4%	1.0%	2.5%	97.8%
FMask	6.8%	0.4%	5.0%	2.5%	97.1%

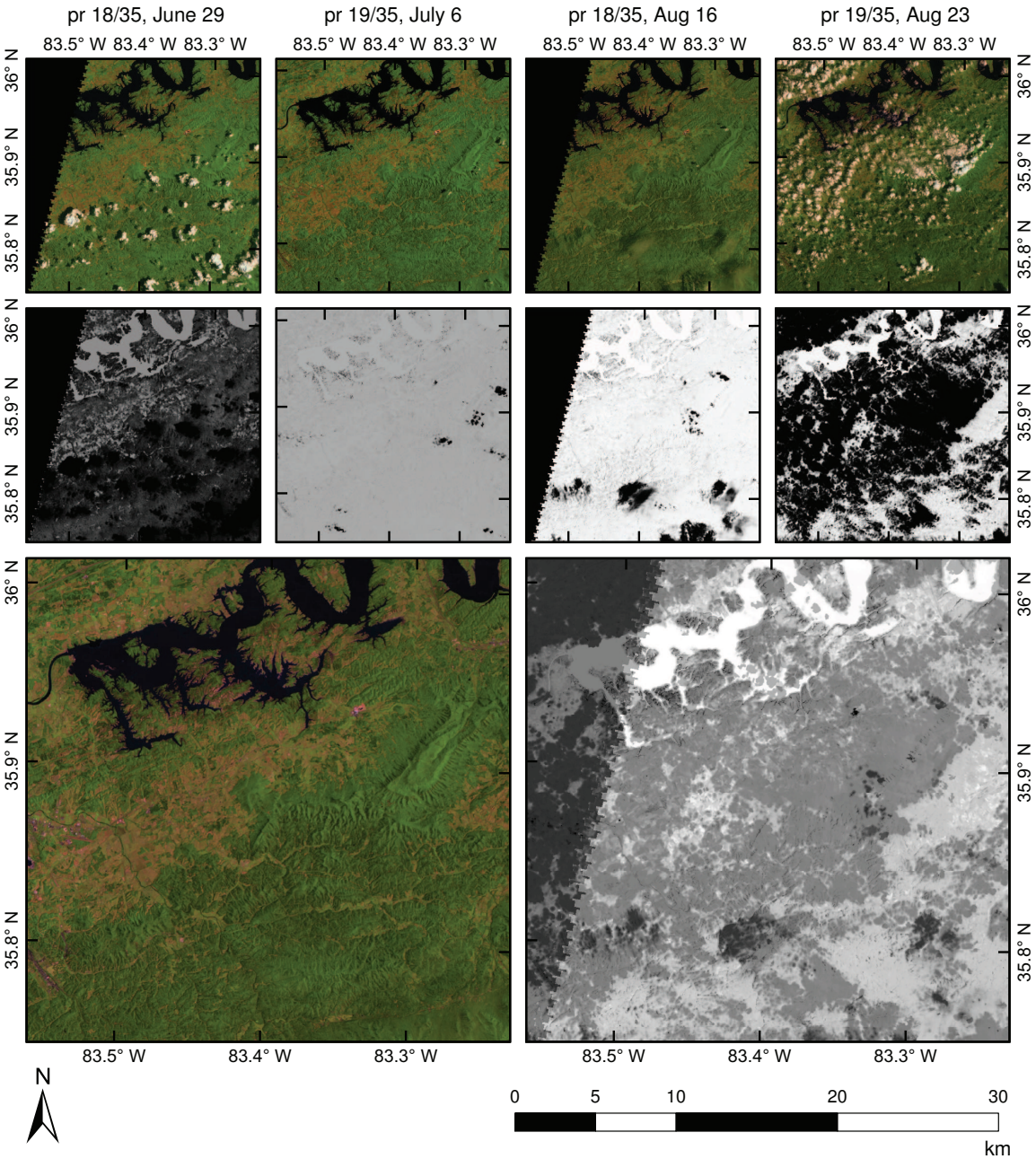

Figure 1.5: A 30×30 km region in eastern Tennessee from four Landsat 5 TM scenes acquired during the summer of 1990 (*top row*) and their respective weights from zero to one (*black* to *white*) from SPARCS (*second row*), where weights of zero signify contaminated or unusable pixels. The results from compositing (*bottom left*) and the sum of the weights used to determine each pixel (*bottom right*) are included.

is also consistently uncertain about several areas around Douglas Lake, the body of water in the northern part of the images. But because the algorithm takes a weighted mean of each image on a pixel-by-pixel basis, only the relative weight through time is important, allowing them to be combined successfully. They are, however, somewhat more susceptible to contamination by clouds or cloud-shadow, as the relative difference between the weights of contaminated and uncontaminated pixels is less.

1.5 Conclusions

We presented a neural network approach to detect cloud and cloud-shadow, as well as water and snow/ice, in Landsat TM and ETM+ imagery, SPARCS: Spatial Procedures for the Automated Removal of Clouds and Shadow. SPARCS uses only single date imagery, does not rely on ancillary datasets, and outperforms another high quality method that operates on similar constraints with overall accuracy of 98.8% compared to 95.3%. Additionally, it is completely automated, does not require specifying new parameters for different scenes, and classification of a Landsat scene completes in under 5 minutes on a desktop computer using an AMD Athalon II 3.1 GHz dual-core processor with 8 Gb of RAM, meeting our design goals.

Unlike other cloud and cloud-shadow detection algorithms, SPARCS is a fuzzy classifier; crisp classification can be achieved by labeling each pixel as the highest-valued membership class and then using the variance among class memberships as an accompanying measure of uncertainty. Knowing uncertainty allows spatial analyses that utilize SPARCS-generated cloud masks to create more accurate spatial products that have more robust estimates of error.

We explored the inclusion of spatial information as an input to the neural network classifier and found limited support for their inclusion. No method summarizing spatial information increased overall accuracy by more than 0.5%. However, a post-processing stage using expert-defined rules increased accuracy by 3.5%. Of particular usefulness is the rule to differentiate between cloud-shadows and terrain shadow that combines predicted cloud-shadow locations from solar geometry and cloud locations with neural network output in the cloud-shadow and water classes. Inclusions of data from larger spatial areas summarized using total variation regularized denoising (with $\alpha > 0.1$) or using a log-polar representation of the local neighborhood (Javier Traver and Bernardino, 2010), which has promise in the field of robotic vision, as neural network inputs may be useful avenues of future research. However, these methods are computationally intensive. We believe that

a multi-stage method that first classifies several single-date scenes of the same location using a method such as the one described here, and then uses those classifications in a second-stage multi-date classifier to resolve cloud and cloud-shadow within each single-date scene is the most promising way forward.

Chapter 2

Natural Disturbances Increase Canopy Heterogeneity in Species Rich Forests

This chapter was submitted to Landscape Ecology in May 2014 as a short communication by myself and Daniel Hayes. I formulated the question, performed the research, and was the primary author of the manuscript.

Abstract

Forest insects, such as the southern pine beetle and the hemlock woolly adelgid, have affected millions of hectares of forest in North America with significant implications for forest health and management. Unlike in stand clearing events, such as clear-cuts and severe fires, the more selective insect disturbances leave the canopy partially intact in the species rich forests of the eastern United States. These more subtle, gradual, and spatially heterogeneous insect disturbances make detecting small inter-annual changes in remotely sensed vegetation indices difficult. However, because these disturbances discriminate for certain host species, and because eastern forests are patchy mosaics, we hypothesize that these types of disturbance events will increase spatial heterogeneity in eastern forest canopies. Here, we introduce an automated method to produce a metric of forest canopy heterogeneity from common vegetation indices derived from satellite imagery. We then use known disturbance locations mapped by the USDA Forest Service Aerial Detection Survey to demonstrate that this metric increases after forest disturbance events caused by several widespread agents in the eastern United States. Finally, we show that this metric provides information about forest disturbance agent independent from the vegetation indices themselves, and is therefore a worthwhile measure to include in applications for mapping and characterizing forest disturbances.

2.1 Introduction

Different disturbance types vary in severity and frequency in ecosystems, but can also vary in how selective they are in affecting individual trees within a forest (Frolking et al., 2009). For example, the hemlock woolly adelgid only attacks hemlock trees (Orwig et al., 2012), whereas the Asian gypsy moth is a more cosmopolitan parasitoid that affects many deciduous species (Townsend et al., 2012). Further, wildfires can affect a range of individuals with varying impacts depending on species and size (Garren, 1943; Schimmel and Granstrom, 1996).

Forest insects, such as the southern pine beetle and the hemlock woolly adelgid, have affected millions of hectares of forest in North America with significant implications for

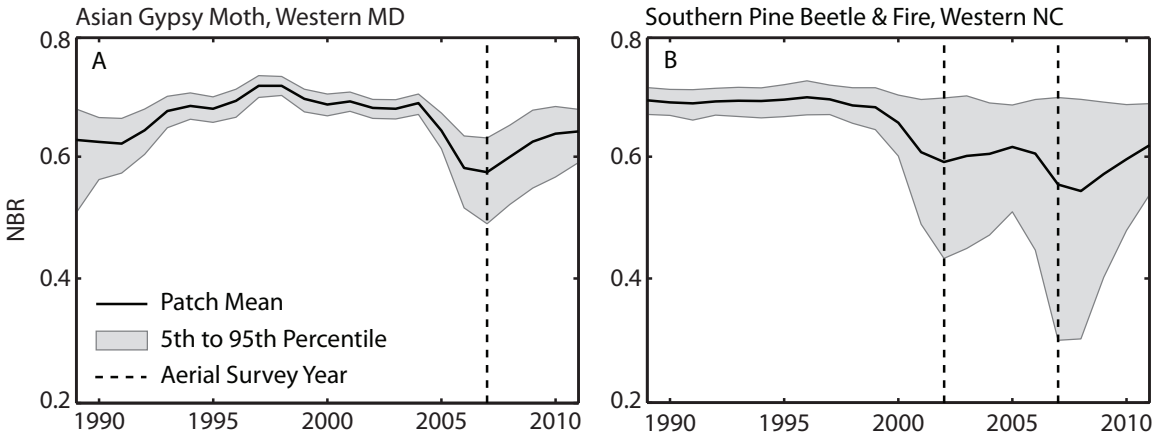

Figure 2.1: Normalized burn ratio (NBR) calculated from the historical Landsat image record for areas in the eastern U.S. and detected as disturbance events in the US Forest Servive Aerial Detection Surveys (ADS). In two example ADS polygons, one disturbed by Asian gypsy moth in western Maryland (A) and one disturbed by southern pine beetle followed by fire in western North Carolina (B), the mean of pixel-wise NBR values (*solid lines*) within the polygons decreases prior to the detection of the disturbance event (*dashed line*), but since the effect is spatially heterogeneous the range of NBR values (*shaded region*) within the polygons increases.

forest health and the carbon cycle (Hicke et al., 2012). However, unlike stand clearing events such as clear-cuts and severe fires, detecting the gradual, more subtle impacts of forest insects and disease using remotely sensed imagery remains challenging (Holmgren and Thuresson, 1998; Goodwin et al., 2008; Frolking et al., 2009; Oumar and Mutanga, 2011; Hansen and Loveland, 2012), though see Townsend et al. (2012) for progress in this area.

In general, forest cover can be measured using vegetation indices derived from remotely sensed imagery, and changes in these indices can be used to detect forest change events (Coppin et al., 2004). When a selective disturbance occurs in a species-rich forest, such as the deciduous forests of the eastern United States, the average value over a large area in a vegetation index will decrease. However, the underlying heterogeneity in the forest will result in a spatially heterogeneous change to canopy cover, causing the vegetation index to decrease in some areas and not in others. This heterogeneous response can be detected by measuring the change in the spatial variance of a vegetation index within a local neighborhood before and after a disturbance event (Figure 2.1).

The neighborhood within which to calculate spatial variance could be expressed as a moving window over the landscape grid, similar to the surface metrics approach used by McGarigal et al. (2009). However, because the value of the local variance is dependent on all of the values in the window, the extent of changes will have fuzzy boundaries and can change based on changes in neighboring vegetation. One can imagine, for example, the

local variance around a forest pixel changing due to changes in crops within the window. To avoid this, we use the longstanding notion in landscape ecology of patches (Turner et al., 2001), here defined as a contiguous area of similar land cover. However, unlike most patch-based approaches that consider patches homogenous areas and then examine the between-patch relationships and structure, we use the patch as a local boundary within which to analyze changes in the variance of a vegetation index.

Since for most disturbance detection applications predefined patches are not available, we first describe a simple image segmentation method to generate patches directly from satellite imagery without needing *a priori* patch extents. Then, we calculate the spatial variance of three vegetation indices within those patches (patch-variance), and show that mean patch-variance within the boundaries of known natural disturbances in the eastern United States increases after those known disturbance events. Further, we demonstrate that this measure of increased heterogeneity provides additional, independent information from the vegetation indices themselves, and is therefore a worthwhile measure to include in applications for mapping and characterizing forest disturbances.

2.2 Methods

2.2.1 Automated Creation of Patches

The method begins with a greyscale image representing a vegetation index, here derived from Landsat Thematic Mapper imagery. Patches are constructed from similar-valued regions within the vegetation index image using the assumption that different land cover types have different responses in the vegetation index. We then apply a denoising algorithm that removes local variation while preserving object edges, the two-dimensional total variation regularization (TVR) (Rudin et al., 1992; Goldstein and Osher, 2009). The result is a mosaic with patches of same-valued pixels.

To identify and label those patches such that each pixel unambiguously belongs to exactly one patch, we apply a zero-crossing algorithm to the Laplacian of Gaussian (LoG) of the denoised image (Marr and Hildreth, 1980). The LoG is a smoothed second-derivative operator over space. In LoG-filtered images, step-like edges appear as a change in sign, which the zero-crossing algorithm then identifies. This process guarantees closed contours and these enclosed areas become patches. Pixels on the edges of patches, where the LoG crossed zero, are assigned to the neighboring patch with the most similar pixel value. Finally, the method merges very small patches with total areas under 20 pixels into neighboring patches.

Patch-variance for each pixel is calculated as the variance in the vegetation index of all pixels that are both within a given patch and within a maximum radius of the pixel. The maximum radius constraint enforces locality since some patches, such as rivers, lakes, or large areas of uninterrupted forest can be quite large. Here, we calculate patch-variance using the same vegetation index used to construct the patches themselves, but this need not be the case. Since patches are constructed from relatively similar-valued regions, this means that patch-variance will likely be smaller than the variances calculated from independently derived patches. This is a conservative choice to illustrate the general applicability of patch-variance when data is limited to a single vegetation index that must be used for both tasks.

Both the TVR and LoG have parameters that control the strength of denoising (α) and Gaussian smoothing (σ), respectively. We selected an α of 0.05 for TVR and a σ of 1.5 for LoG. Higher values, representing more smoothing and less detail, create fewer, larger patches, and smaller values create more, smaller patches. There are no objectively correct values for these parameters, though they do determine the scale at which patches are created. For forest disturbances, we find that values of α between 0.02 and 0.10 and of σ between 1.0 and 3.0 are capable of generating patches that delineate the extents of disturbed areas in the eastern United States while still generalizing heterogeneous regions, but additional research in this area is needed to fully detail the size and types of disturbances that correspond to different parameter values. Additionally, we used a value of 9 pixels (270 m) for the maximum radius from the pixel of interest. This was suitable since most patches were smaller than this radius. We explored values from 5 (150 m) to 15 (450 m), which produced similar results. Larger values should be selected if larger patches are created from higher values of α and σ.

2.2.2 Study Area and Datasets

We selected twelve 45×45 km test areas in the eastern United States that had at least one, and often more than one, widespread disturbance agent recorded in the US Forest Service Aerial Detection Survey (ADS) between the years of 1997 and 2011 (Figure 2.2). Five disturbance agents met our two requirements by occurring in more than one test area and being represented by a minimum of 25 polygons: fire, southern pine beetle, hemlock woolly adelgid, Asian gypsy moth, and beech bark disease. A total of 898 ADS polygons representing these five disturbance agents were present in the twelve test areas; all were used in analysis (Table 2.1). In addition, we constructed 25 control polygons at a random location and year within each test area, for a total of 300 control polygons across all test

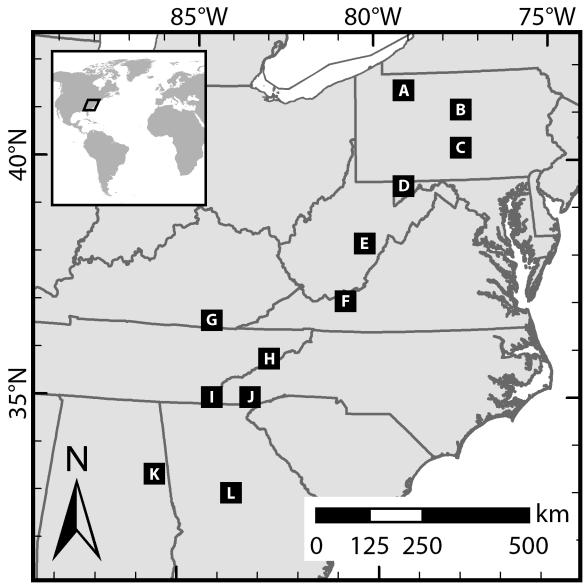

Figure 2.2: Locations of the twelve 45×45 km areas of interest in the eastern United States, representing a variety of forest types and disturbance agents. All disturbances in the US Forest Servive Aerial Detection Survey dataset within these regions were analyzed.

areas. Control polygons were squares with edge lengths randomly selected between 300 and 600 m, which is consistent with the size of patches automatically generated by the method presented in this paper, but is toward the smaller end of ADS polygon sizes. Control polygons were selected from the entire test area, including both disturbed and undisturbed areas.

Landsat images from multiple dates in each test region were declouded and combined to generate a single summertime composite image for each year between 1984 and 2011 using SPARCS (Hughes and Hayes, 2014). Three vegetation indices were calculated over each composite image for each year: Tassel Cap Angle (TCA, Powell et al. 2010), the normalized difference vegetation index (NDV Tucker 1979) and the normalized difference burn ratio (NBR van Wagtendonk et al. 2004). In addition, patch-variance was calculated over each composite for each year and vegetation index.

Within each ADS and control polygon, mean vegetation indices and patch-variances for each year between 1992 and 2011 were extracted and labeled with the disturbance year and agent. The changes in the mean and patch-variance were then calculated between the year of the survey and the year five years prior. The five year interval was chosen to ensure gradual declines prior to aerial detection were captured. No effort, though, was made to avoid other disturbances that may have affected the forest plot five years prior, such as a bark beetle outbreak preceding a fire (e.g. Figure 1B).

Table 2.1: ADS polygon count by disturbance type from each test area (A-L) used in analysis.

	Fire	SPB	HWA	AGM	BBD	Total
A					247	247
B	9			47		56
C				42		42
D				72		72
E			17	4	50	71
F		21	1			22
G	1	48				49
H	12	113				125
I	2	52	8			64
J	3	19	43			65
K	3	27				30
L	1	56				57
Total	31	336	69	165	297	898

2.2.3 Statistical Analysis

Changes in vegetation index mean and patch-variance between the year of disturbance and five years prior were compared among the five disturbance types identified by the ADS polygons and the control polygons using the non-parametric Kruskall-Wallis test (Corder and Foreman, 2009). Tukey's Honestly Significant Difference (HSD) was used to determine if the mean and patch-variance in areas affected by each disturbance type were different from the control polygons (Mathworks, 2013). Patch-variance was log-transformed prior to comparison to meet the normality assumption of ANOVA.

Because changes in patch-variance of each vegetation index co-vary with changes in the mean of that index, it is possible that changes in patch-variance provide no additional information about disturbed areas, making an additional, comparatively computationally intensive metric unneeded. To determine if different disturbance types respond differently to changes in patch-variance and mean vegetation index, the change in patch-variance within each disturbance type was regressed against its corresponding change in mean vegetation index and the resulting regression lines were compared using analysis of covariance (Mathworks, 2013).

2.3 Results

For all three vegetation indices examined, the mean value within disturbed ADS polygons was, on average, lower in the year of the survey than five years prior, indicating lower

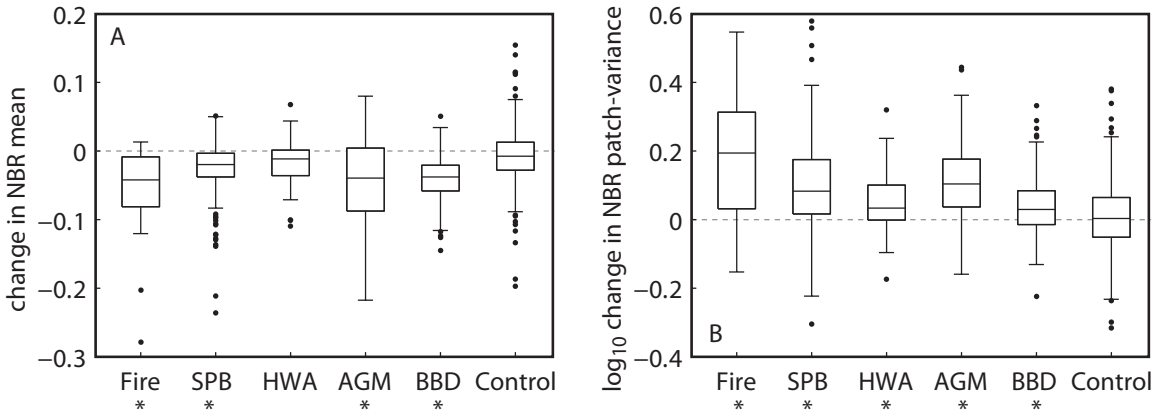

Figure 2.3: Distributions of the five-year change in means (A) and log-transformed patch-variances (B) in NBR values within ADS polygons of different disturbance agents (fire, southern pine beetle (SPB), hemlock woolly adelgid (HWA), Asian gypsy moth (AGM), beech bark disease (BBD), and random controls). Disturbance agents marked with * have mean rank order that are statistically different than the control by Tukey's HSD. Results for NDVI and TCA are similar.

vegetation cover for each disturbance agent (Figure 2.3A). Additionally, all disturbance agents had distributions different than the control polygons by Tukey's HSD at the 0.05 confidence level, except for hemlock woolly adelgid. Variances within each disturbance type, however, are large, and many individual ADS polygons had mean vegetation indices higher in the year of survey than five years prior.

In addition, for all three vegetation indices, the patch-variance within disturbed ADS polygons was higher in the year of the survey than five years prior, indicating an increase in heterogeneity for each disturbance agent (Figure 2.3B). In patch-variance, all disturbance agents had distributions significantly different than the control polygons by Tukey's HSD at the 0.05 confidence level, including hemlock woolly adelgid. Again, all disturbance agents exhibited a wide range of values with many individual polygons becoming less heterogeneous after a disturbance.

Changes in vegetation index mean and patch-variance are negatively correlated (NBR r = -0.66, NDVI r = -0.53, TCA r = -0.52). Regression lines between change in mean and change in patch-variance for each disturbance agent had significantly different slopes by ANCOVA F-test. All slopes were negative, meaning that heterogeneity increases when the average vegetation index decreases in all disturbance types. Relationships between different disturbance agent's slopes, however, were different between vegetation indices (Figure 2.4). In NBR, fire and southern pine beetle had slopes that were significantly more negative than Asian gypsy moth and beech bark disease by Tukey's HSD at the 0.05 confidence level. The relationship in hemlock woolly adelgid was intermediate between

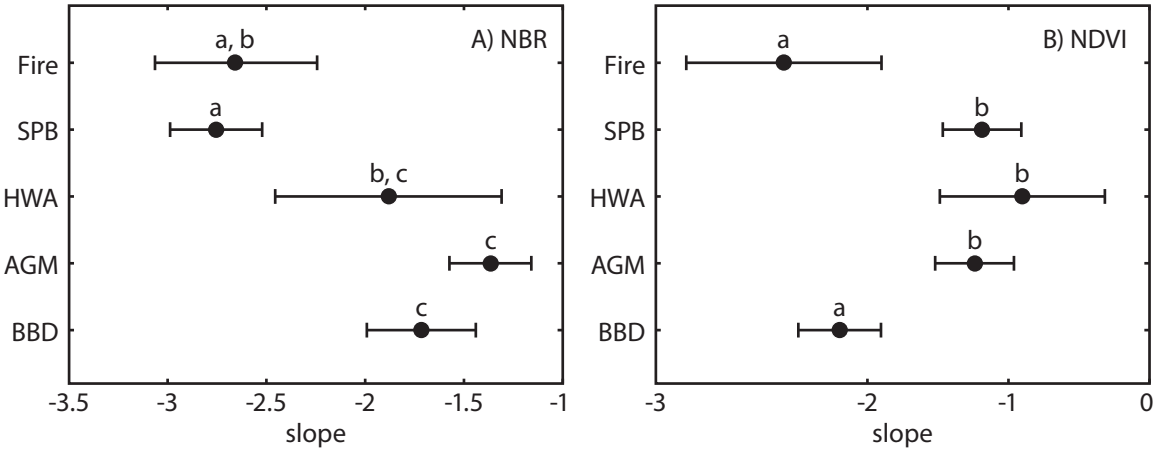

Figure 2.4: Estimates and confidence intervals of the NBR and NDVI regression slopes of change in patch-variances regressed against change in mean vegetation index within ADS polygons representing different disturbance agents (fire, southern pine beetle (SPB), hemlock woolly adelgid (HWA), Asian gypsy moth (AGM), and beech bark disease (BBD)). Disturbance agents marked by the same letter do not have statistically different slopes by Tukey's HSD. Regression slopes for TCA followed the same pattern as those for NDVI.

the two groups. In NDVI and TCA, fire and beech bark disease had significantly more negative slopes than southern pine beetle, hemlock woolly adelgid, and Asian gypsy moth by Tukey's HSD.

2.4 Discussion & Conclusions

Patch-variance increases in areas affected by common disturbance agents providing additional evidence that natural disturbances increase canopy heterogeneity in species-rich forests and thus maintain complexity in forest structure and composition (Bianchini et al., 2001; Lundquist and Beatty, 2002; Turner, 2005; Kneeshaw and Prévost, 2007; Frolking et al., 2009). Different disturbance agents interact with eastern forest ecosystems in different ways, leading to variable increases in heterogeneity for similar decreases in overall canopy cover caused by different disturbance agents. These differences are detectable and potentially exploitable by automated methods to label different agents of forest disturbance. In addition, even though various vegetation indices are highly correlated, the interactions between the mean and patch-variance in different vegetation indices are different for different disturbance types, indicating that a multivariate approach will be useful for this task.

However, the magnitude, and even direction, of the changes in patch-variance can be quite different between ADS polygons representing the same disturbance types. Some

of the variation in these distributions arises from using a fixed 5-year window from the survey date as an 'undisturbed' reference that, in some cases, is itself disturbed. Additional variation comes from the quality of the ADS polygons themselves, which have been shown to have limited accuracy in delineating disturbance extents (Meddens et al., 2012) and often lump multiple small disturbances into a single large polygon (McConnel et al., 2000). This results in polygons that include a mixture of disturbed and undisturbed forest. In an accuracy assessment of ADS polygons in the mountain west, Johnson and Ross (2008) found that 43% of sites within ADS polygons were in fact not disturbed, and 27% of sites within ADS polygons had no disturbance within 500 m. In our study areas, some of the largest polygons are undergoing overall regeneration while still having pockets of disturbance. A benefit of the automated patch-generation method presented in this paper is that those small patches can be individually discovered.

Vegetation indices do saturate at high canopy cover (Glenn et al., 2008). Therefore vegetation indices would indicate that a forest with very high canopy cover is uniformly dense, even if it had high heterogeneity. A disturbance that affects this kind of canopy by uniformly decreasing cover can cause an increase in patch-variance by lowering the heterogeneous canopy below the saturation threshold and into a region where differences can be measured. In these cases, a measure of spatial heterogeneity can be a more sensitive detector of canopy change than the mean, and may help explain why hemlock woolly adelgid polygons were not significantly different from control patches when comparing means, but were when comparing patch-variance.

Here, we have described one metric to characterize heterogeneity: the variance in vegetation indices within a patch. Patch-variance, though, is a member of a class of metrics to describe image texture (du Buf et al., 1990), and similar measures are currently being explored to detect and map biological invasions (Ge et al., 2006; Bradley, 2013). Additional studies of the applicability of texture analysis to the related problems of identifying landscape patches and labeling those patches with disturbance agents, as appropriate, are warranted.

Chapter 3

Automated Detection of Forest Disturbance Events in Space and Time

Abstract

Forest disturbances drive successional changes, release carbon stored in the living forest carbon sink, and alter nutrient cycling. Detecting these important changes using remotely sensed methods allows for a complete census of vegetation change over the study area, unlike in traditional plot-based methods. However, in the species-rich and structurally complex forests of the eastern United States, disturbance events caused by low-intensity fires or species-specific insects and disease are often partial, and therefore difficult to detect using satellite based methods that rely only on total vegetation amount. Here, a set of new algorithms, collectively called VeRDET, is presented that uses a novel patch-based approach to incorporate spatial information from the Landsat Thematic Mapper sensor to detect disturbance, stable, and regeneration periods in a time-series of imagery. VeRDET utilizes SPARCS to identify cloud and cloud-shadow in individual Landsat images, generates a yearly clear-sky composite of those images, calculates a vegetation index using that composite, spatially segments the vegetation index into patches using total variation regularized denoising, and then temporally segments the time-series of each pixel in the patches into a piecewise linear function. For each pixel, the slopes of the segments in the piecewise linear function are interpreted as disturbed, stable, or regenerating. Four vegetation indices are explored: tassel-cap Angle, NDVI, NBR, and NDMI. In an evaluation subset of pixels that were manually interpreted by an expert, NDMI was found to best match the expert interpretation. Using NDMI, VeRDET is closer to the expert than 80% of randomly generated interpretations 89% of the time. Additionally, using NDMI, VeRDET correctly labeled 85% of the years in the evaluation sample as either undergoing disturbance, stable, or regenerating. These results are similar to LandTrendr, a similar disturbance detection algorithm, which correctly labeled 86% of years when operating in the less complex forests of the Pacific Northwest.

3.1 Introduction

Forest disturbances maintain and drive successional changes in species composition (White et al., 2011), are known to impact carbon storage and the strength of forest carbon sinks (Goward et al., 2008; Williams et al., 2013) hydrologic dynamics (Bleby et al., 2010; Pugh and Gordon, 2013), and nutrient cycling and retention (Aber et al., 2002; Bernal et al., 2012). Forest disturbance and regeneration processes occur in patches at different spatial scales ranging from single tree-fall gaps to stand-clearing forest fires. In addition, temporal scales vary; tree stress and mortality may happen slowly over several years, or changes

may be relatively quick with a reduction in forest cover occurring over a single season. Importantly, disturbance from insect and pathogen outbreaks, drought, and fire may be increasing in frequency and severity due to changes in climate (Dale et al., 2001; Weed et al., 2013). Accurate methods that map historical trends in forest change are necessary for monitoring and to establish baseline levels of forest change.

Disturbance and regeneration are particularly challenging to map and quantify in forests of the eastern US. These forests are species rich and may lack clear successional trajectories (Eyer, 1980). In addition, gradual declines in live biomass from prolonged climate stress, pollution, or low-intensity insect outbreaks are difficult to detect without intensive monitoring effort.

Remotely sensed imagery is well-suited for detecting and monitoring forest disturbances, typically by contrasting spectral reflectance at two or more wavelengths to create some vegetation index. Nearly all vegetation indices rely on the fact that green vegetation reflects brightly in the near infrared region of the electromagnetic spectrum, and compares measurements from that region to a region with lower reflectance. For example, the normalized difference vegetation index (NDVI) compares the near infrared to the red, which is highly absorbed by chlorophyll.

Methods to detect forest disturbance from remotely sensed imagery vary substantially in complexity. The most straightforward methods compare two images, typically acquired before and after a known disturbance event, to assess the extent and intensity of the event. Differences between images can be located by calculating differences (Masek et al., 2008), mapping different times to different RGB channels in a color image (Wilson and Sader, 2002), or by considering the direction of change in multispectral space (García-Haro et al., 2001). In these methods the magnitude of the deviation indicates the amount of change, either explicitly through statistical methods or, particularly for the RGB-mapping methods, implicitly through visual inspection of the resulting image. The defining character, though, is that these methods find differences in imagery from dates selected by the operator, and so do not provide additional information on when the disturbance event occurred.

Similar methods to those above, but operating over time-series stacks of imagery, are made possible by the large amounts of available data from MODIS (Lunetta et al., 2006) and from the opening of the Landsat archive (Vogelmann et al., 2009). These methods are capable of providing both the location and times of disturbance events, typically to create historical baselines of disturbed areas or for ongoing monitoring. For example, the Vegetation Change Tracker (VCT) (Huang et al., 2010), finds disturbances by detecting deviations in the vegetation index time-series and labeling these years as disturbed. A

few methods examine the entire time-series as a whole. LandTrendr (Kennedy et al., 2010) finds disturbances by fitting a piecewise linear trend to the vegetation index time-series and labeling those segments with negative slope as disturbances. Both of these methods use Landsat TM/ETM+ imagery, which is well suited for fine-scale spatial monitoring due to its 30x30 m spatial resolution while providing a temporal precision of one or two years. A similar approach takes advantage of the higher temporal resolution of MODIS by fitting a sine curve to the phenological signal and then fitting a piecewise linear trendline to the residual to detect disturbances with subannual precision (de Jong et al., 2012).

Here, I present a new method, VeRDET (Vegetation Regeneration and Disturbance Estimates through Time), which detects variable-length disturbance, stable, and regeneration periods from a yearly time-series at 30×30 m spatial resolution, and which is suitable for the monitoring of individual stands and, when spatially aggregated, for establishing historical frequencies of forest change within ecoregions. Like LandTrendr (Kennedy et al., 2010), VeRDET fits a piecewise linear trend model such that the segments of the trend line are allowed to be of varying length and each determines a period of disturbance, regeneration, or stability, depending on slope. However, whereas LandTrendr treats each pixel independently, VeRDET first constructs patches of similar land cover from the satellite imagery within each year and assigns them a common value. It then uses these patch-level values to track each pixel through time. In this way, pixels that represent a forest stand or agricultural plot are tied together, even though they may have some internal heterogeneity. This patch-creation step can accept a parameter that describes the amount of heterogeneity to be subsumed into the patch, which effectively describes the spatial scale of generated patches. Additionally, VeRDET finds breakpoints in the time-series not by a heuristic search using deviations, but by evolving a global function that accepts a desired noise threshold as a parameter. Similar to the parameter in the patch-creation step, this parameter effectively describes the temporal scale of interest.

For evaluation, VeRDET was run over four different vegetation indices, which were compared for accuracy against human-operator created evaluation data. In addition, a new method for evaluating the disturbance time-series output is introduced. The method accounts for both accuracy in breakpoint identification and the severity of the disturbance by comparing differences in the output between all points in the time-series from their evaluation values. These differences are then normalized by comparing them to differences from an appropriate null model. Finally, the results from this method are themselves compared to the results from established agreement table methods.

3.2 Methods

3.2.1 Study Area

The study area spans approximately 25,000 sq km in the southeastern United States, including portions of northeast Tennessee, northwest North Carolina, southwest Virginia, and a small part of southeast Kentucky (Figure 3.1). In addition, it incorporates parts of six WRS-2 path/rows, the tiling system used by the Landsat program. This area was chosen for its mix of species-rich deciduous and coniferous forests in the Ridge and Valley and Blue Ridge ecoregions (Omernik 1987), and for its multiple land uses, including protected forest, managed forests, agricultural, grazing, residential, and urban areas. Additionally, the study area was nearby, which facilitated field validation by on site visits and discussion with local managers during development.

3.2.2 Data Processing

3.2.2.1 Acquisition and Preprocessing

All Landsat TM 4 and TM 5 imagery acquired between May 1 and September 31 from 1984 to 2011 that intersected the study area and had an estimated cloud cover of less than 50% were downloaded from the USGS website. All images downloaded were Level 1 products geometrically corrected to subpixel accuracy using ground control points. Landsat TM sensors detect electromagnetic radiation in six spectral bands corresponding to blue, green, and red visible light, one near-infrared, and two short-wave infrared bands, as well as one thermal infrared band. Image data is provided for each band as 8-bit digital numbers between 0 and 255.

Linear scaling coefficients were calculated from scene metadata and mean solar irradiances from Chander et al. (2009) to rescale digital numbers for each band to top of atmosphere reflectance.

3.2.2.2 Dark Object Subtraction

To correct for atmospheric scattering and better approximate surface reflectance, an automated dark object correction that takes into account different relative scattering of different wavelengths (Chavez, 1988, 1996), with some modification, was applied to each band of each image. First, the darkest and brightest values of pixels in each band are found by selecting values corresponding to the 0.01 and 99.99 percentiles of all pixels in the band, respectively. Because these radiometric values are derived from discrete digital numbers,

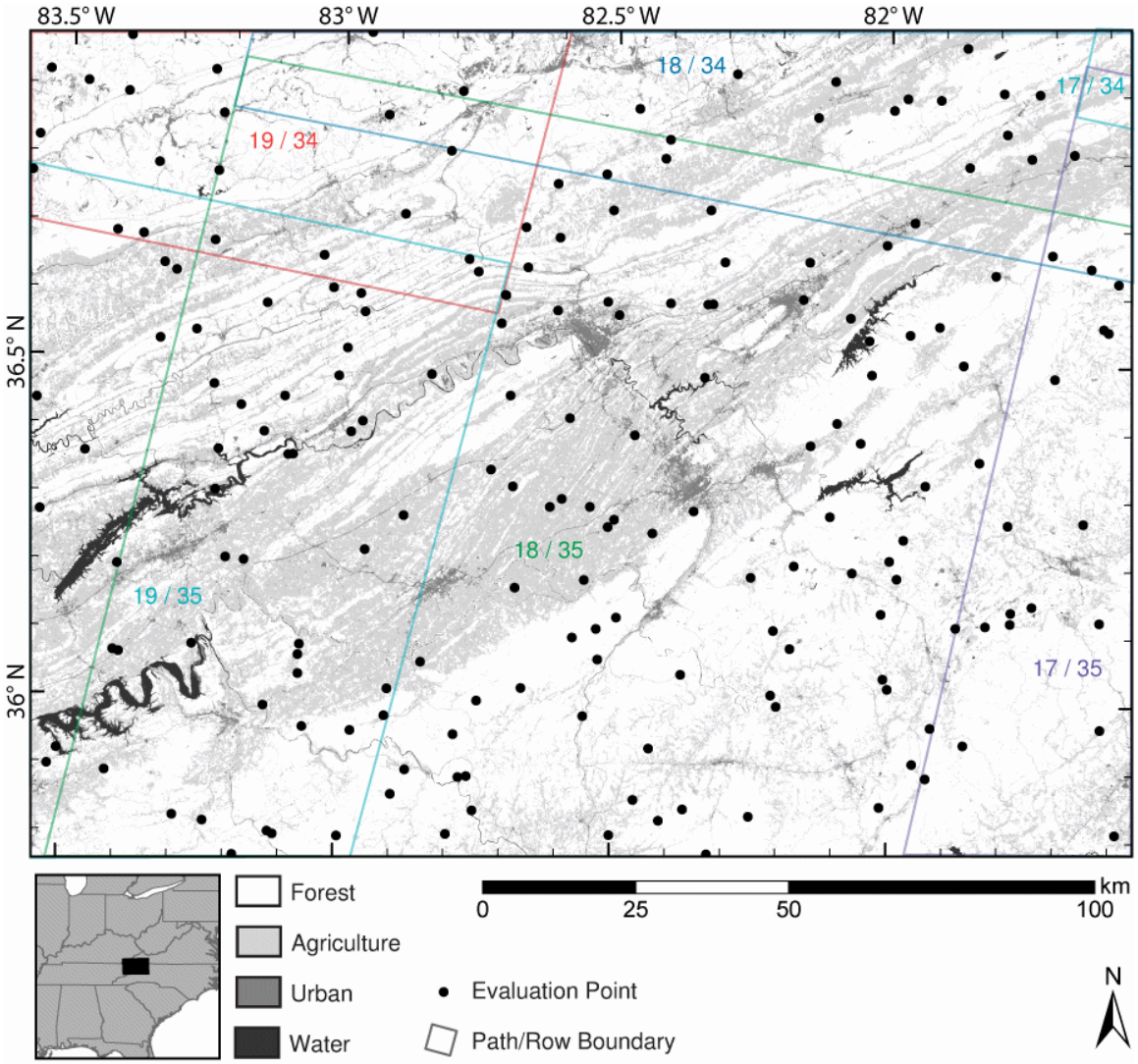

Figure 3.1: Study area showing evaluation points (*black dots*) and mosaiced Landsat scenes (path/rows as *colored outlines*) over the Ridge and Valley ecoregion, characterized by linear folds and agriculture in eastern Tennessee, and the Blue Ridge ecoregion, characterized by uninterrupted forests in western North Carolina.

simply selecting the n-th number in the sorted spectral intensities, where n corresponds to the 0.01 percentile for example, can make the method sensitive to the chosen percentile due to discontinuities between adjacent values. To prevent this, the sorted spectral intensities are linearly interpolated such that the central position in a run of same-valued numbers retains its value, and the numbers between central positions transition linearly between them.

The value of the darkest pixels in each band is an estimate of the amount of atmospheric scattering affecting light at that band's wavelength. However, the relative atmospheric scattering of different wavelengths of electromagnetic radiation should scale as a negative power of wavelength:

$$S(\lambda) = C\lambda^{\alpha}$$

where $S(\lambda)$ is the amount of scattering at wavelength λ, C is a base amount of scattering, and $alpha$ is a measure of particle size. In an atmosphere with only Rayleigh scattering, where light is scattered only from interactions with atomic particles much smaller than the wavelengths, relative scattering scales with λ^{-4}. In an atmosphere with significant Mie scattering, where light is scattered by larger particles such as smoke, dust, pollen, or pollution, relative scattering scales approximately with λ^{-1}, with scaling at λ^{0} representing complete scattering, such as from thin clouds (Chavez, 1988). Real atmospheres exhibit scattering somewhere between these extremes.

To ensure that the atmospheric scattering correction follows a realistic model, while accounting for different atmospheric conditions at different times and places, the dark-object values for each band are log transformed and regressed against the log-transformed geometrically-weighted band centers for each of the six Landsat TM spectral bands in order to estimate C and α in the equation above. The dark object offsets are then calculated from this model. In addition, high values for C, and values of $\alpha > -0.5$ can be used as quality flags for the scene in order to avoid those with large amounts of atmospheric particulate.

The rescaling coefficients from the radiometric normalization step are then updated such that reflectances equal to the dark object offset are zero, reflectances equal to the bright object value remain the same, and intermediate values are linearly interpolated between these end points. Finally, rescaled reflectances that are less than 0.001 are set to the very physically improbable value of 0.001. This prevents impossible negative values while preserving the value of 0 for pixels flagged as having no data.

N := *number of pixels*

```
for each Band
    dark(b)    := percentile(Band, 0.01)
end

[intercept, slope] := regress(log(Band_Centers), log(dark))

for each Band
    offset := exp(intercept + slope*Band\_Center)
    bright := percentile(Band, 99.99)
    gain := bright / (bright - offset)

    Band := gain*(Band-offset)
    Band := max(Band, 0.001)
end
```

3.2.2.3 Tile Creation

Imagery from USGS is provided in Universal Transverse Mercator (UTM) projection local to the scene. Like most mountain systems, clouds are common in the eastern forest highlands. In order to maximize the amount of usable imagery, all images were transformed into a single common projection. UTM zone 17N was chosen as this base projection as it covers the study area with minimal distortion and many scenes were already in this projection.

A tile system was established based on UTM zone 17N, wherein the projected area is divided into a grid of 45×45 km (1500×1500 pixels) squares with the origin at (0E, 0N). Tiles of imagery are then clipped out of each scene such that each tile corresponds to a square in the grid, plus a 1.5 km (50 pixel) buffer along each side, for a final tile size of 48×48 km (1600×1600) pixels. The buffer is used to mitigate any edge effects in processing.

Together, the reprojection and the tile clipping allow imagery of the same location from adjacent WRS-2 path/rows to be combined, doubling the temporal resolution for about 50% of the study area. In addition, the tiles have a much small file size than full Landsat scenes, allowing many tiles to be read into computer memory at the same time during the temporal processing steps described belw.

3.2.2.4 Cloud Detection

Clouds and cloud shadows block the view of the earth's surface. Pixels contaminated with clouds and their shadows, as well as with any lingering or early snow or ice, are identified using the neural network and spatial post-processing approach in SPARCS (Hughes and Hayes, 2014). SPARCS provides continuous-valued memberships of cloud, cloud-shadow, water, snow/ice, and clear-sky classes. The memberships of the clear-sky and water classes are summed to represent the likelihood of an unobstructed pixel (q), such that q is near zero for pixels contaminated with cloud or cloud-shadow and near one for good obstruction-free pixels.

3.2.2.5 Obstruction-free Summertime Composites

Multiple images from the same year are combined by taking a weighted average to construct an obstruction-free summertime composite for each year in a method similar to that described in Hughes and Hayes (2014). The weights used in the average are a function of both q, from the cloud-detection, and of the distance from a target day in order to reduce noise introduced via phenology, w. The cloud-detection weights are squared to reduce the contribution of uncertain pixels for the composite, a choice that is appropriate when there are many images. For this study, the 200th day of the year (June 18/19) was selected for the target day and the weight was calculated as:

$$w_t = \exp\left(-\left(\frac{d_t - 200}{45}\right)^4\right)$$

where d_t is the day of year the given scene t is acquired. A composite of each Landsat band are then constructed as:

$$C_{y,b} = \frac{\sum_{t=1}^{N} q_t^2 w_t B_{b,t}}{\sum_{t=1}^{N} q_t^2 w_t}$$

where $C_{y,b}$ is the composite image of band b for year y, N is the number of images to be combined in that year, $B_{b,t}$ is the image data of band b for image t, and q_t and w_t are the weights for scene t described above. The sum of the combined weight terms, W_y, which is the denominator in the above equation, is also retained to interpolate data-poor areas using data from multiple years in the next step:

$$W_y = \sum_{t=1}^{N} q_t^2 w_t$$

3.2.2.6 Interannual Interpolation

For some areas and years, no obstruction-free view of the earth surface is available near the target date; these pixels have low values for W_y, indicating that little data were used to construct the composite in that area in year y. In these situations, data gaps are filled as an average of data values from nearby years. First, W_y is sigmoidally transformed such that weights of one or higher are near one, small values tend toward zero, and values near 0.5 remain near 0.5:

$$\hat{W}_y = \frac{2}{1 + \exp(-4W_y^2)} - 1$$

Then, a weighted mean, $M_{y,b}$, is constructed for each band for each year from all other yearly observations of that band. The weight for each year equals the product of $\hat{W}_y$ and a distance weight $s_{y,\hat{y}}$ that depends on both the year of the mean being computed, y, and the year of the other composite images used in the weighted mean, $\hat{y}$. This distance weight is high when those years are the same or near one another and quickly decreases as those years become more distant. Using all years ensures that in those rare cases when multiple consecutive years lack data some value will be used to fill in the data gaps.

$$M_{y,b} = \frac{\sum_{\hat{y}}^{T} s_{y,\hat{y}} \hat{W}_{\hat{y}} C_{\hat{y},b}}{\sum_{\hat{y}}^{T} s_{y,\hat{y}} \hat{W}_{\hat{y}}}$$

where $\hat{y}$ iterates over all of the years in the dataset to sum the weighted mean, and T is the total number of years.

The final interpolated composite image $A_{y,b}$ is then constructed by combining the composites with the weighted mean such that areas with few observations are filled in with the weighted mean and areas with many observations retain the values of the original composites:

$$A_{y,b} = \hat{W}_y C_{y,b} + (1 - \hat{W}_y) M_{y,b}$$

3.2.2.7 Generation of Vegetation Indices

Vegetation indices are calculated from the final, gap-filled composites. In this study, we explored the utility of four indices for predicting disturbance events: tassel cap angle (TCA, Powell et al. 2010), the normalized difference vegetation index (NDVI, Tucker 1979), the normalized burn ratio (NBR, van Wagtendonk et al. 2004), and the normalized difference moisture index (NDMI).

$$
\begin{aligned}
TCA &= \frac{4}{\pi}\text{atan}(\frac{TC_{Green}}{TC_{Bright}}) \\
NDVI &= \frac{TM_4 - TM_3}{TM_4 + TM_3} \\
NBR &= \frac{TM_4 - TM_7}{TM_4 + TM_7} \\
NDMI &= \frac{TM_4 - TM_5}{TM_4 + TM_5}
\end{aligned}
$$

where TM_n is the n-th Thematic Mapper band; TM_3 is red, TM_4 is near-infrared, and TM_5 and TM_7 are two short-wave infrared bands.

All of the following processing steps are applied to each of the vegetation indices independently.

3.2.2.8 Patch Creation (Spatial Segmentation)

When treating pixels independently, in moderately disturbed or mixed forest stands disturbance detection algorithms can create a pixelized mosaic where some pixels are marked as disturbed and others are not. This arises due to natural heterogeneity in the forest, which itself is increased by natural disturbance events (Chapter 2) coupled with the amplification of small differences when computing ratio-based Vegetation indices, particularly in dark areas such as hill-slopes opposite the sun. However, it is desirable in these cases to label the entire forest stand as disturbed and not just some selected pixels. To do this, some type of spatial filtering is required to tie neighboring pixels together, either as part of the detection process or afterward in a post-processing stage. Simply being nearby, though, is insufficient, as pixels on the borders of landscape features would be unduly influenced by changes in neighboring features. As such, a method that identifies and preserves patches of similar land cover while simultaneously homogenizing pixels within those patches is required.

The two-dimensional total variation regularization (TVR) (Rudin et al., 1992; Goldstein and Osher, 2009) is a spatial denoising algorithm that removes local variation while preserving object edges. TVR is applied to each vegetation index to create same-valued regions, thereby simplifying the image into a set of internally homogenous patches. Our approach uses Goldstein's Split Bregman algorithm to find a solution (U) that minimizes:

$$
\min_{U \in \mathbb{R}^2} \left[\|V - U\|^2 + \alpha \left|\nabla U\right| \right]
$$

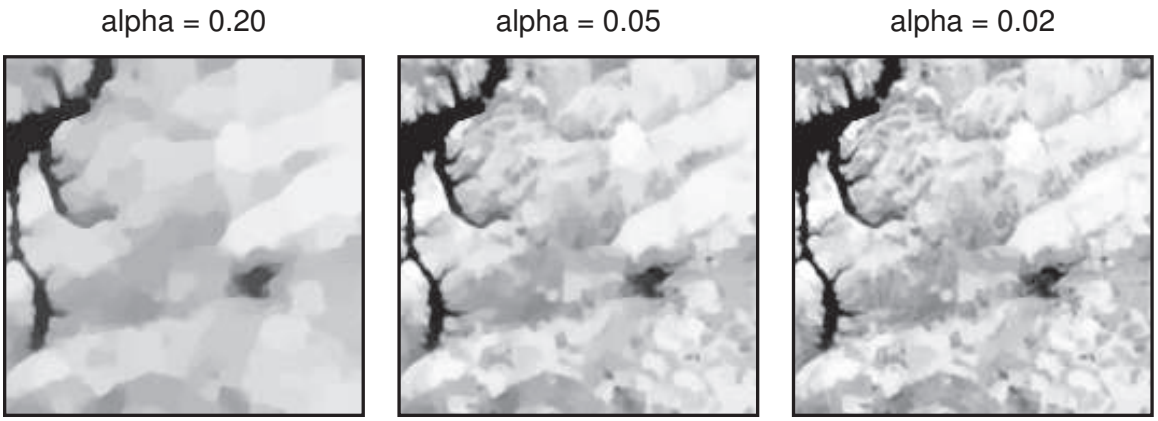

Figure 3.2: 2-dimensional TVR denoising at differing strengths subsumes more or less heterogeneity into patches.

were V is the original vegetation index image, and α controls the intensity of the denoising, and therefore the scale of the patches. For this study, a relatively low value of 0.05 was selected to make many small patches, and so preserves most image features while removing small-scale heterogeneity (Figure 3.2). Using larger values of α would create coarser patches that tie together larger sections of the landscape; future work that applies these methods using multiple values could be useful for detecting disturbance events at different spatial scales.

Segmentation of patches follows (Chapter 2). The Laplacian of Gaussian (LoG) algorithm (Marr and Hildreth, 1980) is applied to U, followed by a zero-crossing algorithm to identify boundaries of patches. Pixels on those boundaries are assigned to the neighboring patch with the most similar pixel value, and, finally, very small patches with total areas under 20 pixels into neighboring patches.

3.2.2.9 Breakpoint Identification (Temporal Segmentation)

Forest disturbances include events that are sudden, acute processes, such as landslides or fires, and events that are protracted over many years, such as pollution or insect outbreaks. General purpose detection algorithms should detect both kinds of events and also provide information on when, how long, and at what intensity these events occurred. The model of forest disturbance employed here conceives of forest change as a series of temporally non-overlapping periods of disturbance, regeneration, or stability. Each period is defined as the line segment between two break points in the vegetation index time series; the intensity of the change is defined as the slope of the segment.

This piecewise segmentation is found by treating the time-series as a noisy dataset and then searching for a piecewise linear approximation of that dataset. The fitting approach is to find a solution, u^*, to the total variation regularized functional with a $T \times T$ integration matrix, A, in the data fidelity term:

$$\min_u \left[\|f - Au\|^2 + \beta \left|\nabla_t u\right| \right]$$

where f is a time-series of vegetation index values, u is a proposed smoothed version of the time-series, β is a regularization parameter controlling the degree of smoothing (the analogue to α in the spatial denoising), and is chosen to be 0.03 in our detection algorithm, $\nabla_t u$ is the change is adjacent values of u, $\|\cdot\|$ is the L-2 norm and $|\cdot|$ is the L-1 norm, and where

$$A(i,j) = \begin{cases} 1 & , i >= j \\ 0 & , i < j \end{cases}$$

By comparing an integrated u to the data in the fidelity term, the minimization procedure finds a u^* that approximates the first derivative of f; integrating u^* generates an approximation of f. By using an L-1 norm in the regularization, the minimization is driven to a solution with a sparse $\nabla_t u$; that is, most between-year changes in u^* will be 0, with some terms being non-zero. By extension, that implies u^* will be composed of a series of same-valued runs with jumps in value between runs. And since u represents a derivative, this means that the integrated approximation will be composed of a series of line segments of same-valued slope, that is, a piecewise linear representation.

For efficient convergence, this functional is minimized using the fixed point lagged diffusivity method described in Vogel and Oman (1996):

```
D(i,j) := -1 , i==j or i==j-1
           0 , otherwise

A(i,j) :=  1 , i >= j
           0 , otherwise

u := initial guess
do:
        u1 := u
        E := β / (ϵ + |∇_t u|)
        L := D' diag(E)D
```

```
            u := (A'A + L)^-1 (A'f)
    until (u1 - u) > tol
```

D is a differentiation matrix and A is the integration matrix. Here, the initial guess of u is chosen to be f. In the Matlab implementation, u is updated using Matlab's backslash operator to avoid taking the inverse of the diffused Hessian term. The convergence loop is ended when the difference between consecutive steps is less than a selected error tolerance, tol; 1×10^{-5} is used here. Additionally, ϵ, a value used to prevent division by 0, is chosen as 1×10^{-6}; the algorithm is relatively insensitive to this value.

Although this algorithm provides an approximate piecewise linear function, the output has two drawbacks. First, solutions tend to be compressed in the sense that they have a smaller range than the original data. This is because small errors in the data fidelity term are made up for by reducing the size of jumps in the regularization term. Second, when β is small, as it must be when attempting to capture disturbance events that occur at fast scales, additional breakpoints can be introduced representing small changes in the slope. To combat both of these problems, a clean-up step is performed on the output to remove breakpoints with changes in slope less than 0.025, and then apply a piece-wise linear regression that, at each segment defined by the remaining breakpoints, simultaneously fits line segments to the data.

3.2.2.10 A note on TVR and choosing α and β

The value of the regularization term α or β controls the balance between the data fidelity and the smoothing term. Very low values cause the solution to simply equal the data, or its derivative in the linear fitting. Very high values cause the solutions to tend toward a single patch that equals the mean of the image in the spatial problem or a simple linear regression of the time-series. The best value, in the sense that it optimally solves the ill-posed inverse problem $\mathbf{Au} + \sigma = \mathbf{f}$ is to set α or β equal to the magnitude of the theoretical noise process that is 'contaminating' the data, σ. Empirically, as the signal to noise ratio approaches one, values of α should increase to around twice this magnitude, 2σ. Since vegetation indices are scaled between zero and one, this theoretical noise process will always have a magnitude of less than one, and for most imagery, closer to 0.01 to 0.10. However, it can never be directly known.

A different way to approach the regularization parameter is to think of it as a threshold for local heterogeneity. Assuming that heterogeneity on the landscape is built up by a series of different processes, then different choices of α and β represent different processes that are considered to be 'noise' versus 'signal'. Changing these parameters then becomes

a selective filter that determines the magnitude, and to some extent the scale, of processes that are included for that stage of analysis. Performing analyses over many values of α and β can allow for a hierarchical analysis of patterns at different spatial and temporal scales.

3.2.3 Accuracy Evaluation

From each tile, 16 evaluation pixels were chosen for evaluation, for a total of 192 evaluation pixels. Land cover change for those pixels was determined by a human operator using TimeSync (Cohen et al., 2010), which presents an operator with a Landsat time-series, and the option to select different views. It then allows the operator to select start and end dates of disturbed, regenerating, and stable periods, and then label those events with presumed causal agent, severity, and other attributes. While determining dates and attributes, the operator can incorporate other data sources such as Google Earth, Aerial Detection Survey data, and their own expert knowledge.

For each evaluation pixel and each vegetation index, the disturbance / regeneration trajectories identified by VeRDET are compared to the operator-defined trajectories from TimeSync by calculating the mean squared error (MSE) between the trajectories. By itself, MSE is not a meaningful measure of fit, so 1000 random trajectories are generated for comparison. Each random trajectory is assigned a random number of breakpoints, this number is resampled from a random TimeSync or VeRDET trajectory to approximate the distribution of real trajectories. Breakpoints are assigned to random years, and then the random trajectory is determined using piecewise linear regression. Finally, the MSE of the VeRDET trajectory is compared to the MSEs of the random trajectories. The number of random trajectories with higher MSE, converted to a percentile, is used as the dissimilarity score.

This approach generates a score that takes into account the agreement in the position of breakpoints between VeRDET and the expert, and also the shape of the resulting trajectory. This means that omitted or additional breakpoints at subtle changes are weighted less than breakpoints at large changes. This is especially important because subtle changes are difficult to detect for both automated methods and human experts, so an evaluation method that is insensitive to the exact year it is detected is important.

In addition, for each year in each evaluation pixel, whether that pixel was undergoing disturbance, was stable, or was regenerating is determined from the slope of the piecewise linear function at that year. Negative slopes are labeled as undergoing disturbance, positive slopes as regenerating, and slopes near zero are labeled as stable. The threshold around zero was chosen empirically to minimize error. Labels represent trends occurring

between summers, and so do not correspond directly to calendar years. These labels are then compared to the disturbed, stable, and regenerating notations of each segment from the expert. Separate comparisons were performed for each vegetation index.

3.2.4 Changescape

Disturbances and regeneration are dynamic processes; however, a static spatial representation is most often desirable. To encode dynamic information in an easily understandable way, three independent axes of change are combined into a single image, which we term a *changescape*. The first axis, $Change$, represents the total amount of change in a given pixel, including both disturbances and regenerations, and is equal to:

$$Change = \sum_{y}^{T} |u_y^*|$$

where u_y^* is selected solution to the time-series simplification. $Change$ is high for disturbed regions and low for stable regions. $Change$ is mapped to the lightness of the changescape, with white indicating stable regions and dark/bright colors representing regions with substantial change.

The second axis, $Trend$ represents the overall direction of change during the time period of interest and is calculated as:

$$Trend = \sum_{y}^{T} u_y^* |u_y^*|$$

Although a simple linear regression provides similar information, it is sensitive to outlier values near end points, where data fidelity is often lowest. Instead, Trend sums up all changes, weighting sudden changes more heavily than slow changes. Trend is mapped to the hue of the changescape, with red values indicating disturbance, green values indicating regions with equal amounts of disturbance and regeneration (possibly none in stable regions), and blue values indicating regeneration.

Finally, $VegHi$, provides information on how vegetated a pixel was at some point during the time period of interest. It equals the 95th percentile value in the vegetation index time-series, to protect against noisy spikes. $VegHi$ is mapped to saturation such that areas that had low vegetation throughout the time period are gray and forested areas are the most colorful.

3.3 Results & Discussion

For disturbance segmentations derived from NBR and NDMI, VeRDET outperformed 80% of the random null models 89% percent of the time. For segmentations derived from NDVI, this only occurred 61% of the time, and for TCA a mere 42% (Figure 3.3). For segmentations derived from all vegetation indices, performance compared to null models was highly bimodal – either the segmentation performed very well, or the segmentation was significantly worse than random. This means that, although there are many cases where the automated method disagrees with the expert, VeRDET is not simply making a random guess, but is rather optimizing some criterion that is at odds with the criteria used by the expert.

Each segment was labeled by the expert as corresponding to a period of disturbance, stability, or regeneration and disturbance, regeneration, and stability determinations were made for each vegetation index in each year. The threshold around zero that determined whether a year was stable minimized error when set to ± 0.02. When comparing the disturbance, stability, or regeneration determination from each vegetation index with the expert's, a similar pattern in vegetation index effectiveness emerges (Table 3.1). Again, TCA performs worst, labeling the trend correctly in only 70% of the years. NDVI and NBR perform similarly overall, labeling the trend correctly in 83% of the year. However, NBR correctly labels 72% of disturbed years while doing less well in identifying years of regeneration. NDMI performs best, correctly identifying the trend in 85% of the years, and also has the highest percent correct of all methods for years that were disturbed and stable, and ties with NDVI for the highest percent correct in years that are regenerating.

Cohen et al. (2010) calculated confusion of disturbed, stable, and regenerating years for LandTrendr in the Pacific Northwest. Although an exact comparison cannot be made because of the different forest types and because their analysis focused on vertex years instead of all years, VeRDET's performance is similar to that reported for LandTrendr. VeRDET has more confusion between disturbed and regenerating years, though, suggesting that it may have difficulty pinpointing the exact year that a disturbance ends and the regeneration trajectory begins. This is to be expected in the complex forests of the eastern United States, given that disturbances are rarely as severe as the clear-cuts common in Pacific Northwest.

The high performance of NDMI may be explained in part by its sensitivity to biomass (Jin and Sader, 2005), which may enable it to detect actual disturbances and to track regenerating biomass instead of just quick regreening from forbs. (Cohen et al., 2010)

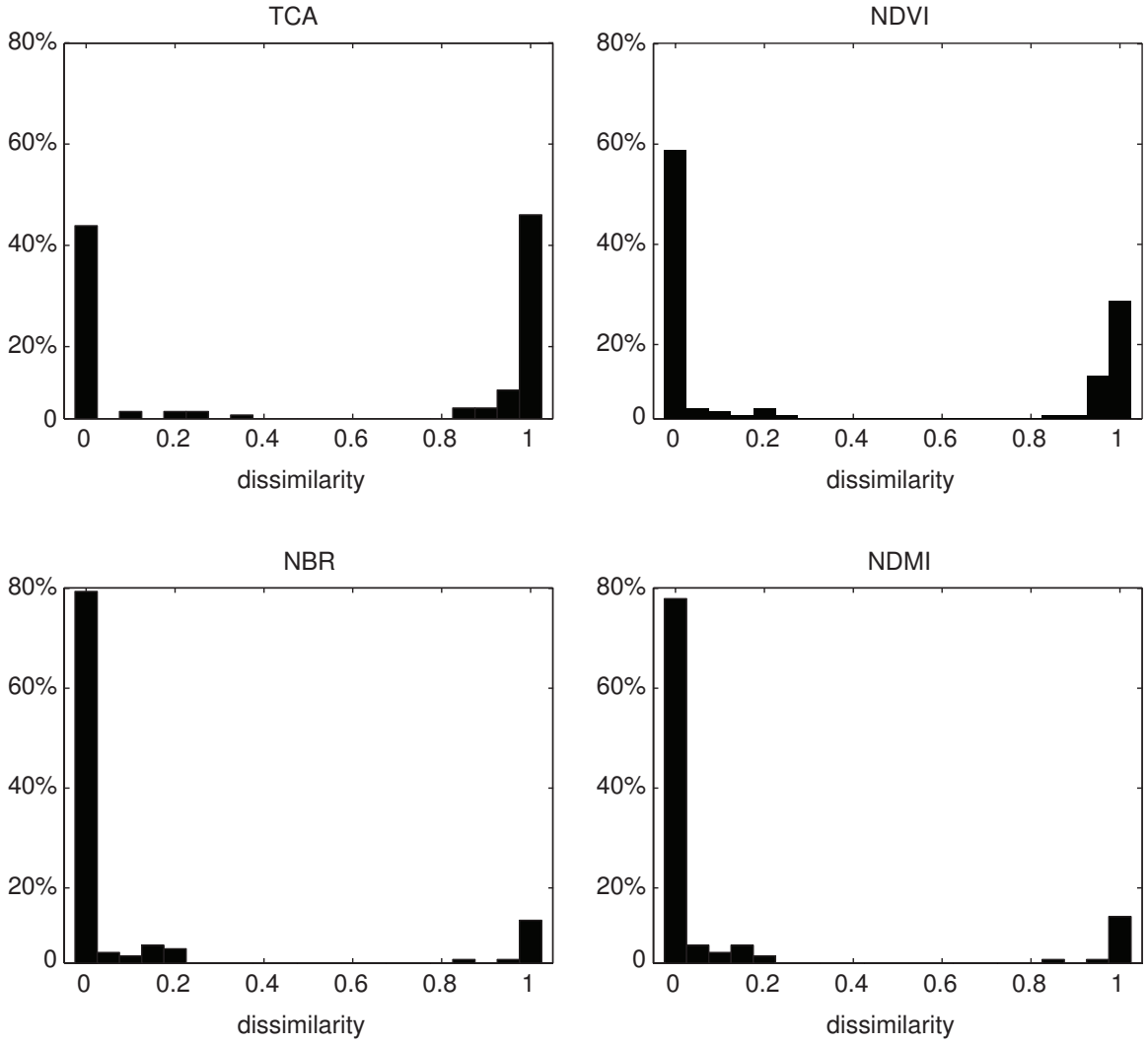

Figure 3.3: Dissimilarity scores for the 132 forested evaluation pixels, calculated by comparing the mean squared error between the VeRDET segmentation derived from each of four vegetation indices (TCA, NDVI, NBR, NDMI) and the expert segmentation, to the mean squared error between 1000 random segmentations and the expert segmentation and calculating its rank percentile. Scores near 0 are significantly more similar to the expert segmentation than random, whereas scores near 1 are significantly less similar to the expert segmentation than random.

Table 3.1: Agreement between the VeRDET segmentation for each of the four vegetation indices (TCA, NDVI, NBR, NDMI) and the expert segmentation on whether in each year the sample was being disturbed, was stable, or was regenerating.

	Expert Labeled as:		
	Disturbed	**Stable**	**Regen**
TCA classification:			
Disturbed	212	645	13
Stable	171	2655	35
Regen	33	310	38
% Correct	**51%**	**74%**	**44%**
NDVI classification:			
Disturbed	245	366	17
Stable	150	3103	19
Regen	21	141	50
% Correct	**59%**	**86%**	**58%**
NBR classification:			
Disturbed	301	377	19
Stable	75	3068	25
Regen	40	165	42
% Correct	**72%**	**85%**	**49%**
NDMI classification:			
Disturbed	321	358	12
Stable	63	3129	24
Regen	32	123	50
% Correct	**77%**	**87%**	**58%**

also found that Tassel-Cap Wetness, an index that is informationally similar to NDMI and is believed to provide information on vegetation structure, performed better than other indices. Although the use of NDMI or Wetness for vegetation studies is relatively uncommon, with NDVI or Tassel-Cap Greenness being preferred, these indices are promising indicators of disturbance and should be included in future forest disturbance studies.

Visually, time-series segmentations from VeRDET and the expert coincide well in forested patches with significant disturbance events (Figure 3.4). Many incorrectly identified breakpoints occur a year before or after the expert identified years. Additionally, VeRDET segmentations will often add short periods of stability whereas the expert often chooses one year as a breakpoint. This may be due to VeRDET having insufficient imagery during those years and thus the weighted averaging is interpreted as stability, whereas the

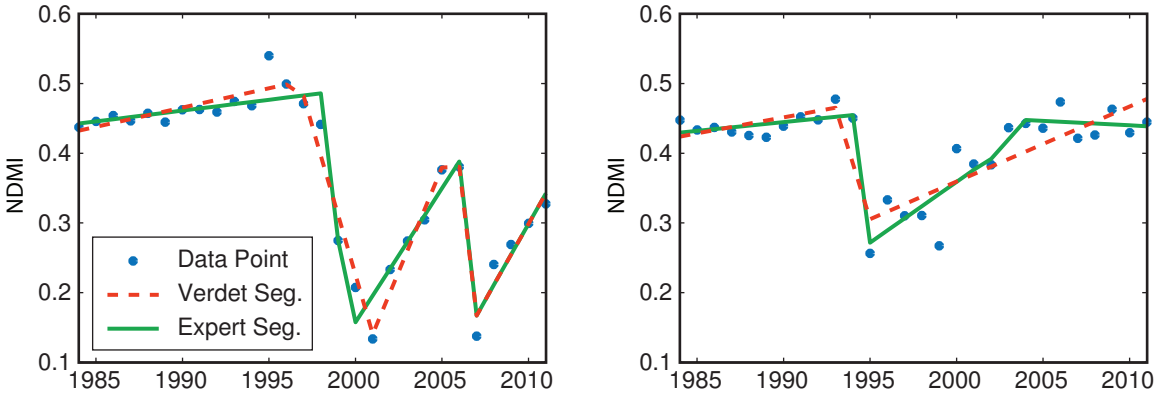

Figure 3.4: Segmented NDMI time-series (*dots*) for two different evaluation pixels from VeRDET (*dashed lines*) and the expert (*solid*).

expert had access to additional data sources and imagery dates and can therefore make an informed decision.

The different ecoregions in the study region exhibited striking differences in the levels and spatial patterns of disturbance (Figure 3.5). The Central Appalachians (in the northwest of the map) have a mixture of stable forests punctuated with large regions of severe disturbance, primarily from coal mines. The Ridge and Valley region in the north-center of the map is stratified with forests on hilltops and agriculture in valleys. The agricultural areas have large amounts of interannual change, but no real trend; this is due to cropping and to human activity that periodically cuts fields to prevent them from succeeding to forest. The significant amount of regeneration in this area is likely due to old fields being allowed to transition to forest as agriculture and ranching become less economically important. In addition, this area has substantial human development, indicated by the very geometrically regular pattern of clear cuts and regrowth. The Blue Ridge ecoregion in the southeast-center is mostly stable forest thanks to its patchwork of protected areas and lack of accessibility. Disturbed sections in the northern part have strongly negative trends, and are mostly pine beetle die off that has then burned. Disturbed sections in the southern part are mostly caused by a mixture of pine beetle and hemlock woolly adelgid. Other tree and shrub species in these forests have mostly filled in the disturbed areas, resulting in a mixed disturbance / regeneration trend in the map. Finally, the Piedmont (southeast corner) is characterized by human development and rotating timber harvesting. Nearly all areas are affected by human activity.

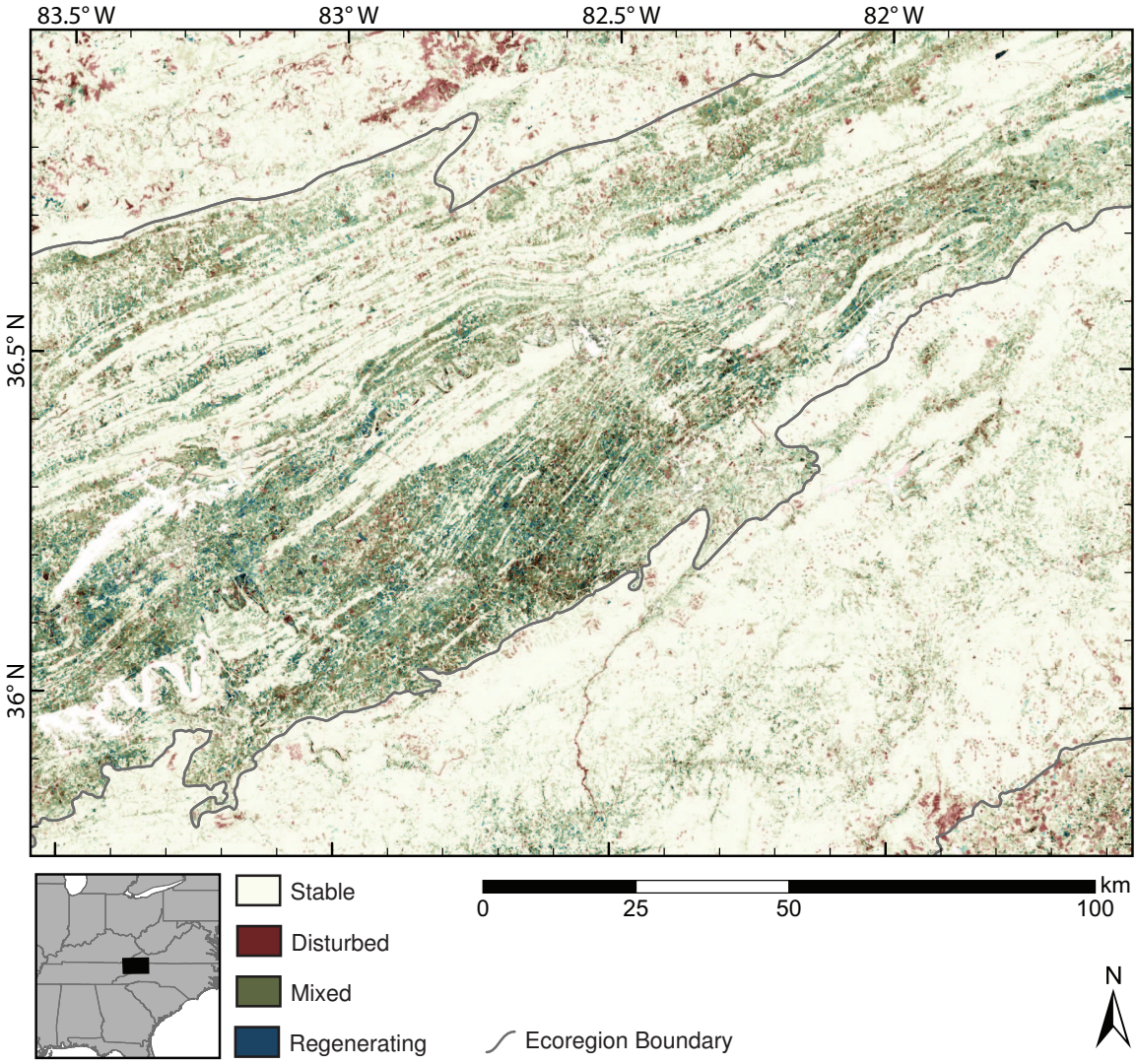

Figure 3.5: Changes over 28 years. Darker areas have higher $CHANGE$ than lighter areas, which are stable. Red areas experienced primarily disturbance (negative $TREND$), blue areas experienced primarily regeneration presumably from a previous disturbance (positive $TREND$), and green areas experienced a mix of disturbance and regeneration ($TREND$ near zero).

3.4 Conclusions

VeRDET welds multiple views from multiple projections, Landsat path/rows, and dates into a spatial representation of forest disturbance and regeneration. The resulting maps are notable for their visual lack of linear artifacts and abrupt changes near edges of individual Landsat images and VeRDET's native tile system. Nor do maps have similar artifacts or changes around removed cloud and cloud-shadow objects, strongly suggesting that these common contaminants are not sources of error in VeRDET. Additionally, using vegetation indices based on shortwave infrared bands, VeRDET is able to segment time-series well in 89% of the samples, though there is still room for improvement, as well as uncertainty around why VeRDET performs significantly worse than random in approximately 10% percent of cases. In a comparison of disturbed, stable, and regenerating years, the best performing vegetation index used by VeRDET labels those years correctly in 85% of years, which is comparable to the 86% accuracy found in LandTrendr (Cohen et al., 2010), even though VeRDET was operating in the more complex environment of eastern forests.

Such results enable the mapping of forest disturbance in the eastern US within a margin of error that is practical for forest monitoring at the stand level, and, when aggregated over larger spatial scales, for establishing historical baselines for use in regional planning or as inputs into process-based atmospheric carbon models. Overall, VeRDET is a promising tool to answer spatially and temporally explicit questions about vegetation change at a range of scales that can open up new avenues of inquiry in forest monitoring and habitat assessment.

Chapter 4

Algorithmic Attribution of Forest Disturbance Agents in Eastern Forests

Abstract

Millions of hectares of forest are disturbed each year by agents such as fire, storms, invasive insects, and disease, altering the forest carbon sink and changing forest nutrient cycles. The magnitude and direction of these changes, though, can be different for different disturbance agents. For example, trees that burn in severe fire rapidly release stored carbon into the atmosphere whereas standing deadwood from insect attacks degrade slowly while atmospheric carbon is fixed as regenerating vegetation. Usable estimates of these processes require accurate and reliable estimates of the extent and frequency of different disturbance agents. Here, a new method is presented that classifies disturbance events identified using time-series analysis of remotely sensed imagery using only the imagery itself. The method was chosen from a suite of neural network classifiers to explore the use of textural information in imagery to predict disturbance agent. The best classifier used a combination of three vegetation indices, NDVI, NBR, and NDMI, the patch-variance calculated from these indices, the three tassel-cap bands, and the local entropy over the tassel-cap bands calculated within a 9×9 pixel window. Classifiers were trained using data from the US Forest Service Aerial Detection Surveys and differentiate between fires, southern pine beetle, gypsy moth, hemlock woolly adelgid, beech bark disease, anthracnose, and storm events. In addition, the classifier returns a value of 'uncertain' when it is unable to make a clear determination. Overall classification accuracy is 81%, with a maximum accuracy of 90% for gypsy moth. The classifier performs poorly with storm events (3% accuracy), which it most often confuses for fire or beech bark disease. Reliabilities are similar to accuracies. The classifier is then applied to imagery spanning most of the eastern United States and the percent of forest cover affected by each disturbance agent each year is reported by ecoregion.

4.1 Introduction

Disturbance events are important ecological processes that strongly affect the composition, structure, and function of forest ecosystems. Forest disturbances cause direct carbon emissions and dramatically change energy flow, nutrient cycles, water balances and forest age structure (Running, 2008). They can maintain a given forest community type (Gilliam and Platt, 1999) or cause dramatic shifts in forest composition (Spaulding and Rieske, 2010; Fralish and McArdle, 2009).

An estimated 152 million ha of forests in the United States are affected each decade by various disturbances such as land-use change, fire, harvesting and insect attacks (Pan et al.,

2011), however, existing US land management inventories were not designed to monitor disturbance. Many studies of insect disturbance in forests of the eastern United States, for instance, seek to predict the spread of insects (Fitzpatrick et al., 2012; Costanza et al., 2012) or model future climate impacts on forest/insect dynamics (Paradis et al., 2008; Hicke et al., 2012). Although these studies require the amount of forest area affected as intermediate values in the models, this figure is not reported. As such, there is a lack of estimates of how much forest is affected by a given disturbance agent through time at regional scales. This information is integral to accurate models of carbon dynamics (Zhang et al., 2012), current and future climate change impacts (Ayres and Lombardero, 2000; Dale et al., 2001), and quantification of the synergistic impacts of multiple invasions.

Recently, the North American Forest Dynamics (NAFD) project sought to remedy this by quantifying forest disturbance through time using satellite imagery (Masek et al., 2013). Though this is one of the first studies to estimate annual rates of forest disturbance across the conterminous US, it uses biennial data, does not generate regional estimates, and provides no attribution to causal agent of disturbance. They note that future satellite-based studies of forest disturbance at regional and national scales should use methods with uninterrupted spatial extents at fine to moderate scale and annual temporal resolution for the best estimations.

The current study is a first attempt to detect disturbance events and then attribute these events to one of a variety of disturbance agents over the forests in the eastern United States. These forests are species rich and represent a substantial amount of the forest biomass in the United States, including much of the hardwood region (Woodall et al., 2006). In addition, because of the wide variety of species, they are also affected by a wide variety of natural disturbance agents, which create a challenging problem for discrimination using remote sensing.

As described in Chapter 3, disturbances increase heterogeneity in the forest canopy in eastern forests through selective defoliation or mortality. Spatially based indices of image texture were suggested as a path forward for discrimination between agents. Here, we assess two such measures. The first is patch-variance, which was described in Chapter 3. The second is local entropy, another measure of deviation from a spatial average that has been used in image processing applications for decades (Huang, 1975; Russ, 2007), but has seen little use in the remote sensing literature.

In this paper, a neural network classifier is developed to identify the causal agent of seven natural disturbances from remotely sensed imagery: southern pine beetle, gypsy moth, hemlock woolly adelgid, beech bark disease, anthracnose, fire, and severe weather.

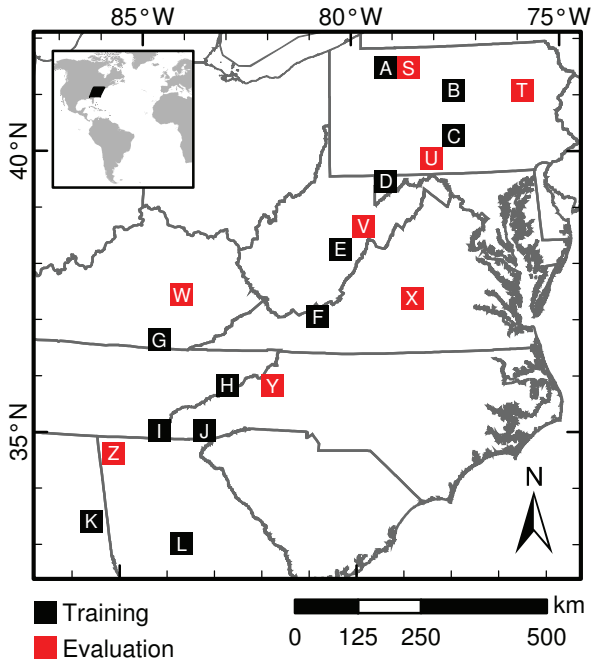

Figure 4.1: Locations of regions used in training (*black*) and evaluation (*red*). Letters correspond with rows in Table 4.1

.

Disturbed areas are identified using VeRDET (Chapter 3). and the agent classification is similarly derived from a yearly time-series of satellite imagery. Multiple indices, including texture-based indices, are explored and assessed for usefulness in the classifier. Finally, a first estimate of disturbance agent by ecoregion and by year is produced for the eastern United States.

4.2 Methods

4.2.1 Classifier Construction

Twenty 45×45 km regions with at least one widespread disturbance agent, as identified by USDA Forest Service Aerial Detection Survey (ADS) polygons, were selected to provide training and evaluation data for the neural network classifiers (Figure 4.1). These twenty regions were split into twelve training regions and eight evaluation regions with an intent to distribute regions in each set throughout the study area and to ensure that polygons representing different disturbance agents were well distributed among the sets. Eight labeled disturbance agents were included in both the training and evaluation sets: fire, southern pine beetle (SPB), gypsy moth (AGM), hemlock woolly adelgid (HWA), beech

Table 4.1: Number of pixels used in training and evaluation by region and disturbance agent.

Training:	Fire	SPB	AGM	HWA	BBD	Anthr.	Weather	Unknown
A		53			5449	234	20	120
B	393		9617					
C			2893					
D			2511					
E				43	9539			232
F		54		2				
G	221	284						
H	185	787						
I	50	570		73				
J	5	39		40				
K	7	282					6	
L	14	469						
Evaluation:	Fire	SPB	AGM	HWA	BBD	Anthr.	Weather	Unknown
S		14			3410	160	464	217
T			1756	56				
U			1497					
V				83	54945			166
W		128						
X	207	7						44
Y		4685		7				
Z	32	186					6	

bark disease (BBD), anthracnose (Anthr.), weather, and disturbed areas for which cause was unknown (Table 4.1).

Inputs to the suite of neural networks were derived from declouded and composited Landsat TM imagery from 1995 to 2011 using SPARCS (Hughes and Hayes, 2014). Images were then analyzed to create patches and find disturbances using VeRDET (Chapter 3). Inputs include five years of observations centered around the disturbance date noted in the ADS dataset for: four vegetation indices, tassel-cap angle (TCA) (Powell et al., 2010), the normalized difference vegetation index (NDVI) (Tucker, 1979), the normalized burn ratio (NBR) (van Wagtendonk et al., 2004), and the normalized difference moisture index (NDMI); the patch-variance calculated from each of those vegetation indices for the same five years; the three tassel-cap (Crist, 1985) bands for those five years; the local entropy calculated within a 9x9 pixel window (ignoring patch boundaries) for each of the three tassel-cap bands for those five years; and the disturbance slope for each of the vegetation

indices during the year in the ADS record. Local entropy, $H_{i,j}$ is simply the Shannon entropy calculated over the vegetation index values, $V_{i,j}$, of the 81 pixels within the 9x9 pixel window (Russ, 2007):

$$H_{i,j} = \sum_{\hat{i}=i-4}^{i+4} \sum_{\hat{j}=j-4}^{j+4} -V_{\hat{i},\hat{j}} \ln V_{\hat{i},\hat{j}}$$

For training, to ensure that pixels within ADS polygons represented disturbed vegetation and are not simply lumped in with nearby disturbed areas (Meddens et al., 2012; McConnel et al., 2000), two filters were applied. The first selects vegetated pixels as those with a TCA greater than 0.4 at least twice between 1984 and 2011. The second selects non-stable pixels by identifying those with a range of NDMI of at least 0.1 over the time period. These are conservative filters meant to include shrublands and weak disturbances. In addition, to reduce the effects of spatial autocorrelation, pixels were thinned by selecting only 2 pixels from every 4×4 pixel region in a structured, repeating pattern (Figure 4.2). All remaining pixels in the twelve training regions were aggregated by disturbance agent type. Because some agents are much more common on the landscape, and because neural network classifiers are sensitive to the distribution of training samples, each disturbance agent was limited to a maximum of 2000 training samples, selecting samples randomly during each training run for those disturbance agents with more than 2000 samples.

Neural networks are supervised classification algorithms that accept a set of training data, consisting of a set of inputs labeled with a desired class, and then learn a non-linear function mapping the inputs to a set of outputs, where each class corresponds to one output. The training data is partitioned into learning data and validation data. The samples in the learning data are presented to the network algorithm, and it is this data to which the non-linear mapping is fit. The validation data is used to test the mapping while learning is occurring to prevent over fitting. Here, 20% of the training data was used for validation and was randomly selected. In a crisp classification, such as the one used here, the output that has the highest value for a given set of inputs is considered the class of the inputs. When that value is high, the classifier is quite certain in the classification; the values of other outputs are typically low when this occurs. When the highest output is low, the classifier is typically uncertain; here, this information is used to provide an additional 'uncertain' classification outcome.

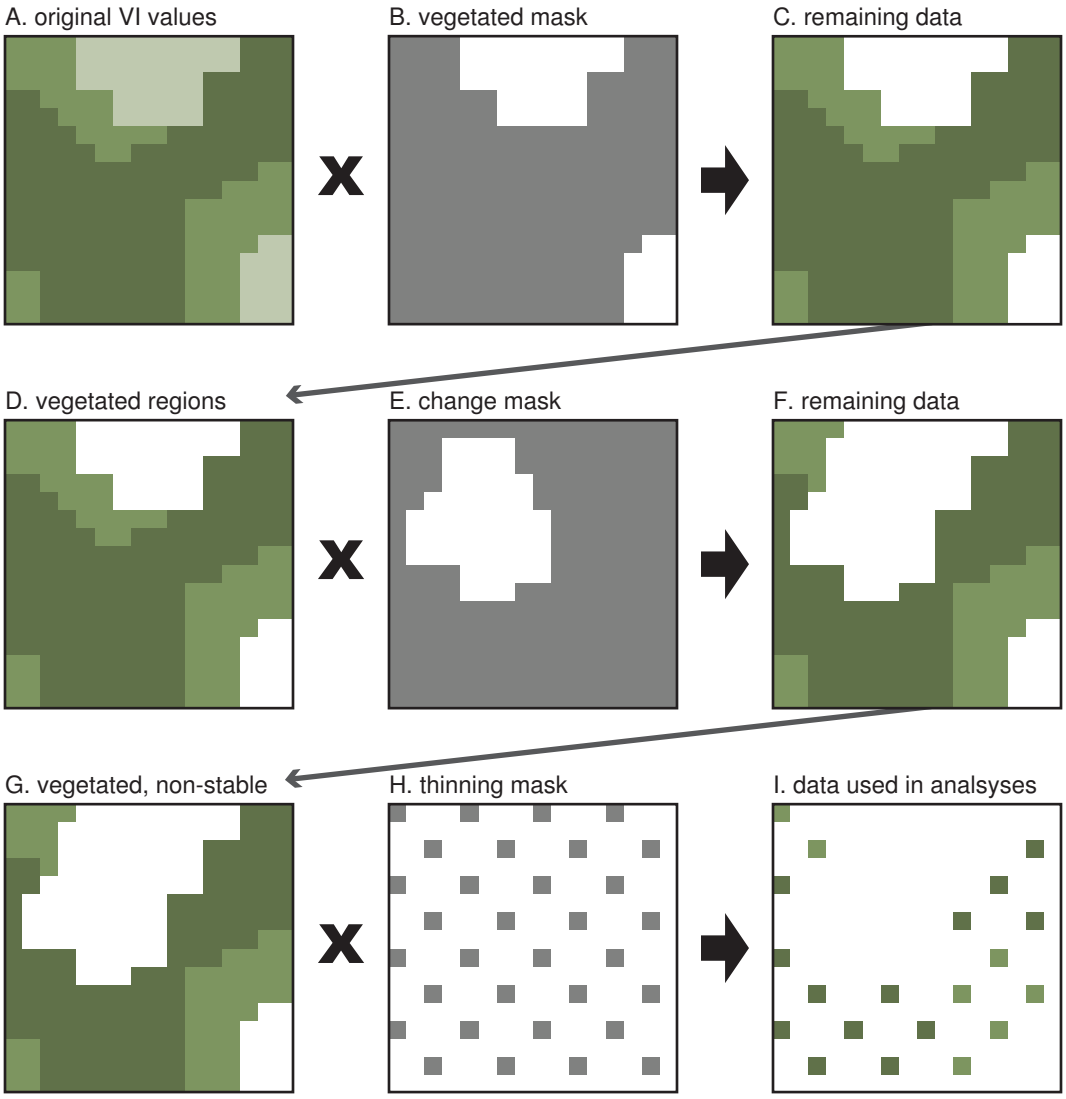

Figure 4.2: Data within ADS polygons is filtered and thinned to ensure that only vegetated, disturbed pixels are used in training and evaluation, and to reduce the effects of autocorrelation. Original vegetation index values (*A*) are filtered using a mask of vegetated regions (*B*) such that non-vegetated areas (*B*, *white*) are omitted. The remaining data (*C*, *D*) is then filtered to remove areas that do not change (*E*, *white*). Finally, the vegetated, non-stable regions (*F*, *G*) are thinned using a repeating, structured mask (*H*) that ensures neighboring pixels are not used in analyses (*I*).

4.2.2 Classifier Assessment

To assess the role of network size and complexity on the disturbance agent classification problem, networks with 10, 20, 40, and 80 hidden nodes were tested. As the number of hidden nodes increases, the more complex the function the neural network can approximate, which should increase classification accuracy in complex problems. However, with increasing free parameters, networks with more hidden nodes can overfit the training data and result in poor general classifiers. The sizes examined in this study were selected by examining performance over a range of network sizes in a small set of exploratory networks.

A total of 5120 neural networks were trained to test the effects of the different inputs and network size in all combinations (1024 different network structures) in five replications. Replications reduce the likelihood of the neural network converging on a local minimum and mitigate the effect of sample selection. Training employed a feedforward network with Levenberg-Marquardt backpropagation (Marquardt, 1963). A validation subset of 20% of training samples, which were not presented to the learning algorithm, was used to determine when to stop training. If performance on the validation set did not increase after 6 epochs, the default value for Matlab's training function, training was stopped to prevent overfitting. Training would also have ended after 2 hours or after 1000 epochs, but these halting criteria were never reached.

Networks were evaluated using pixels within the ADS polygons in the eight evaluation regions. As with the testing samples, the evaluation pixels were similarly filtered to exclude non-forest pixels, stable pixels, and neighboring pixels. For evaluation, disturbance agents were again limited to 2000 samples. Total accuracy and kappa scores (Cohen, 1960) were calculated for each network and compared using a multi-way ANOVA. Statistical significance between network sizes was determined using Tukey's Honestly Significant Difference (HSD) (Kramer, 1956). Accuracy and confusion results are reported over both samples taken only from the evaluation regions as well as samples from both the evaluation and training regions which were not used during training and were at least 60 m away from samples that were used during training. No pixel used in training was used at any point during evaluation.

4.2.3 Estimation of Natural Forest Disturbance by Agent

The best performing neural network classifier was selected to classify forested, disturbed pixels within the entire study region between the years 1986 and 2009. Forested pixels

were defined as those with a TCA greater than 0.65 in at least two years between 1984 and 2011; this threshold omits agricultural lands while keeping small forest stands. Disturbed pixels are those for which VeRDET returned a slope smaller than -0.02 using the NDMI vegetation index. Additionally, forested pixels with an NDMI-derived disturbance slope greater than 0.02 were defined as regenerating, those between -0.02 and 0.02 were defined as stable. These thresholds were found to optimize the fit between VeRDET output and an expert operator's determination (Chapter 3).

Neural network classifiers provide a set of continuous memberships in each class. In a typical crisp classification, a sample is assigned to the class with the highest membership. The membership values themselves, however, provide information on how certain the classifier is; when the maximum value is high, the classifier is very certain and is also likely to be correct, when it is low the classifier is less certain and less likely to be correct. To quantify classifier certainty, maximum membership values were binned and the percent of correct samples within each bin calculated. Then, the maximum membership value that corresponds to 50% correct was found. Pixels with maximum membership values below this threshold have even odds of being correct or incorrect and are instead labeled as 'uncertain'.

Finally, the forested area affected by each disturbance agent in each year was aggregated by ecoregion (Omernik, 1987).

4.3 Results & Discussion

4.3.1 Classifier Evaluation

Over the evaluation regions, each vegetation index, the inclusion of the tassel-cap bands, as well as the inclusion of patch-variance for the vegetation indices and local entropy for the tassel-cap bands were statistically significant (Table 4.2). The inclusion of the disturbance slope from VeRDET was not statistically significant, likely because that information could be readily and linearly approximated by the learning algorithm from the five-year window of each vegetation index.

Surprisingly, on average, both TCA and NDVI decreased classifier effectiveness of samples in the evaluation region, as measured by changes in kappa between networks that include them and those that do not (Table 4.3). This disappears when comparing against all regions, and suggests that the increases to accuracy seen during training when including these indices was often simply from overfitting the network to the training data and not the result of learning general patterns. The negative trend is largely driven by networks

Table 4.2: Multi-way ANOVA over the kappa scores from classifiers with different combinations of inputs calculated from classifying all pixels within ADS polygons from the 8 evaluation regions.

Source	Sum Sq	df	Mean Sq	F	p
TCA	0.035	1	0.035	14.57	< 0.0001
NDVI	0.25	1	0.25	102.86	< 0.0001
NBR	0.173	1	0.173	71.39	< 0.0001
NDMI	0.642	1	0.642	264.38	< 0.0001
Tassel-Cap	1.645	1	1.645	677.75	< 0.0001
T-C Entropy	1.465	1	1.465	603.69	< 0.0001
PatchVar	0.141	1	0.141	58.02	< 0.0001
Disturb Sev	0.01	1	0.01	4.13	0.043
Nodes	0.061	3	0.02	8.38	< 0.0001
Error	2.457	1012	0.003		
Total	6.879	1023			

Table 4.3: Effect size of input parameters on classification accuracy as measured by the difference in kappa calculated over samples from only the evaluation regions and from all regions.

	TCA	NDVI	NBR	NDMI	Tassel-Cap	TC Entropy	Patch Var.	Disturb. Severity
Evaluation	-0.012	-0.031	0.026	0.050	0.080	0.076	0.023	-0.006
All	0.028	0.024	0.054	0.054	0.083	0.048	0.024	-

with 10 and 20 hidden nodes, which can be more easily overwhelmed when attempting to summarize large numbers of inputs at the hidden layer. However, despite the negative average trend, the performance of many networks did increase when TCA or NDVI was included. Still, in the top ten performing networks, two included neither of these greeness detecting indices and none included both, implying that the two indices are redundant.

Increasing the number of hidden nodes increased the effectiveness of networks when comparing over both the evaluation regions (Figure 4.3A) and over all regions (Figure 4.3B). A network size of 40 hidden nodes was optimal and was sufficient to generalize the patterns in the data without overfitting. Increasing from 40 to 80 hidden nodes did not increase accuracy in either case. Indeed, in the top ten performing networks, three had 80 hidden nodes and the rest, 40. The range of performance within the set of networks that share the same number of hidden nodes, however, was wide. For example, in samples from all regions, kappa ranged from 0.21 to 0.68 in networks with 40 nodes.

The best performing network has 40 hidden nodes, uses the NDVI, NBR, and NDMI vegetation indices, patch-variances calculated over those indices, the three tassel-cap bands, and the local entropy calculated from the those bands. It does not use TCA or

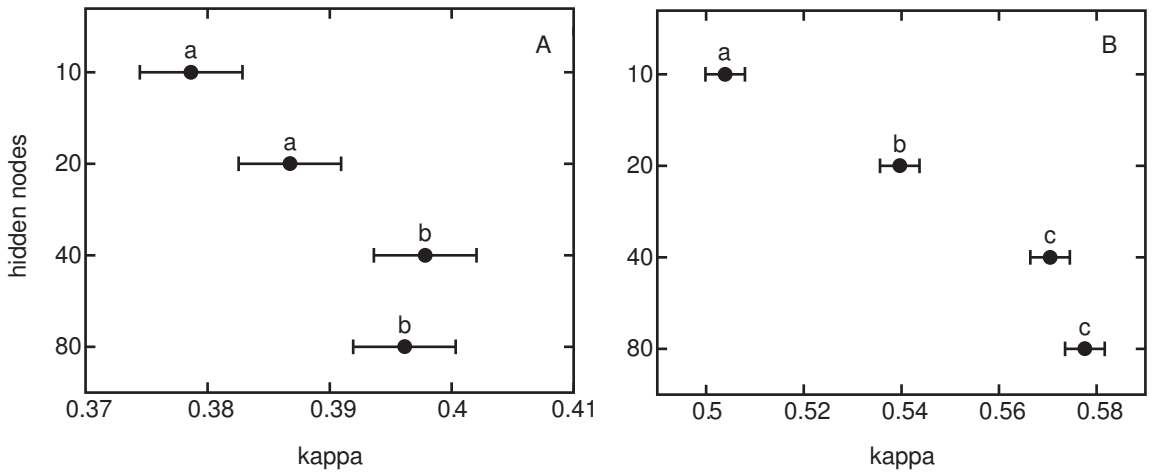

Figure 4.3: Average kappa of networks with different numbers of hidden nodes for samples from evaluation regions (*A*) and from all regions (*B*). Levels that share the same letter within each plot are not significantly different by Tukey's HSD.

disturbance severity from VeRDET. In samples from the evaluation regions only, this network has a total classification accuracy of 66% with a kappa of 0.54. In evaluation samples from all regions, its accuracy is 76% with a kappa of 0.68.

In neural network classifiers, each sample is assigned membership values in each class; the maximum membership value corresponds to the class the network believes the sample belongs. To quantify classifier certainty, these maximum membership values were binned into 11 equally spaced groups. The percent of correct samples within each bin is given in Figure 4.4; the shape of the curve is typical for other classifiers examined. Samples with a maximum membership value below 0.38 have a 50% probability of being incorrectly identified. Instead of mapping these pixels as the classifier's best guess, the algorithm instead maps pixels below this cutoff as 'uncertain'. When calculating error ignoring those samples that the classifier admits it does not know, classifier accuracy increases from 66% to 71% and kappa increases from 0.54 to 0.60 in samples from the eight evaluation regions. In samples from all regions, classifier accuracy increases from 76% to 81%, and kappa increases from 0.68 to 0.75.

The classifier performs best at identifying regions with gypsy moth, with an omission error of 10% and a commission error of 16% (Table 4.4). Southern pine beetle, beech bark disease, and fire are also identified well; the most confusion among these comes from samples labeled as fire in the ADS dataset that are classified as southern pine beetle. The classifier correctly identifies less than half of the hemlock woolly adelgid and anthracnose affected samples, and similarly misattributes disturbances to each of these agents about a third of the time. Finally, the classifier performs poorly on disturbances caused by weather

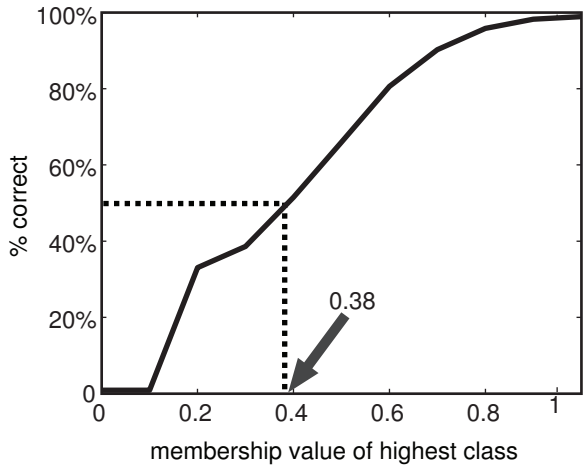

Figure 4.4: Percent of samples that are correct given the membership value of the highest membership class. Samples with a maximum membership class less than 0.38 are 50% likely to be wrong, and are labeled as 'uncertain' in the classifier.

events, such as storm damage and frost. The likely arises due to having too few examples of weather damage during training. Those disturbances that are labeled as 'unknown' in the ADS dataset are correctly classified about half of the time; the other half of the time, the classifier labels those samples with a known disturbance, suggesting that many unknown disturbances in the ADS dataset may be caused by known agents and are simply unidentified.

The confusion between southern pine beetle and fires can be explained by two processes. Areas affected by southern pine beetle contain patches of dead trees that are conducive to fire (Schowalter et al., 1981); since these two disturbances co-occur spatially, this confusion is natural. Figure 4.5 shows a burned area identified in the ADS dataset that was previously affected by southern pine beetle. The classifier identified the area as a spatially intermixed disturbance of fire and southern pine beetle, most likely with the heavily burned regions identified as fire. The classifier was uncertain in some regions on the borders between the intensive fire and southern pine beetle outbreak; those areas are likely intermediate or mixed at a spatial scale below the patch sizes used in some inputs. This spatial heterogeneity, however, is not captured in the ADS dataset (McConnel et al., 2000), which means that at least some of the 'error' arises from the classifier's increased spatial precision.

Table 4.4: Agreement matrix between ADS-labeled pixels (columns) and pixels classified using the best neural network classifier (rows) over sample pixels taken from all regions for southern pine beetle (SPB), gypsy moth (AGM), hemlock woolly adelgid (HWA), beech bark disease (BBD), anthracnose, fire, weather, and samples labeled 'unknown' in the ADS database. Samples were limited to 4000 pixels per ADS-labeled class. No evaluation pixels were used during training.

	Labeled as:								
	SPB	AGM	HWA	BBD	Anthr.	Fire	Weath.	Unkwn.	
Classed as:									
SPB	3280	156	44	310	17	147	59	66	80%
AGM	148	3593	18	239	110	48	91	68	84%
HWA	18	5	116	29	0	9	0	9	64%
BBD	166	112	44	2883	37	19	150	176	80%
Anthracnose	24	28	1	47	219	2	0	11	66%
Fire	356	83	30	258	7	893	162	33	48%
Weather	0	5	1	16	1	2	14	1	37%
Unknown	8	18	21	218	4	2	20	417	57%
	82%	90%	42%	71%	55%	80%	3%	53%	

4.3.2 Natural Disturbance by Agent in the Eastern US

The highly developed Southeastern Plains and Piedmont ecoregions experienced the most disturbance, largely driven by southern pine beetle outbreaks (Table 4.5). The Blue Ridge ecoregion, with its low development and large swaths of protected areas, experienced the least disturbance, followed by the Central Appalachians. These observations correspond with Pan et al. (2011), who found young forests (appox. 30 years) in the Piedmont, moderately aged forests in the southern Appalachians (appox. 70 years) and older forests in the northern Appalachians. Inverting these average forest ages provides disturbance rates for the Piedmont of approximately 3.3%, for the southern Appalachians of approximately 1.4%, and less than 1% for the northern Appalachians. These rates are similar to the overall rates of disturbance reported in Powell et al. (2014), but are approximately half those detected in this study. However, because VeRDET is capable of detecting subtle disturbances that do not cause mortality, these estimates are expected to be higher than those in other studies.

The percent forest impacted by southern pine beetles is highest in the Southeastern Plains and Piedmont, with peaks in the early 1990s, and again in the late 1990's to early 2000's (Figure 4.6A). Nowak et al. (2008) confirms this detected outbreak, and notes it impacted over 400,000 ha of forest with an estimated economic loss of $1 billion. Another disturbance event from southern pine beetle during 2006 to 2008 spans both the Southeastern Plains and the Piedmont. These dates for outbreaks are consistent with the

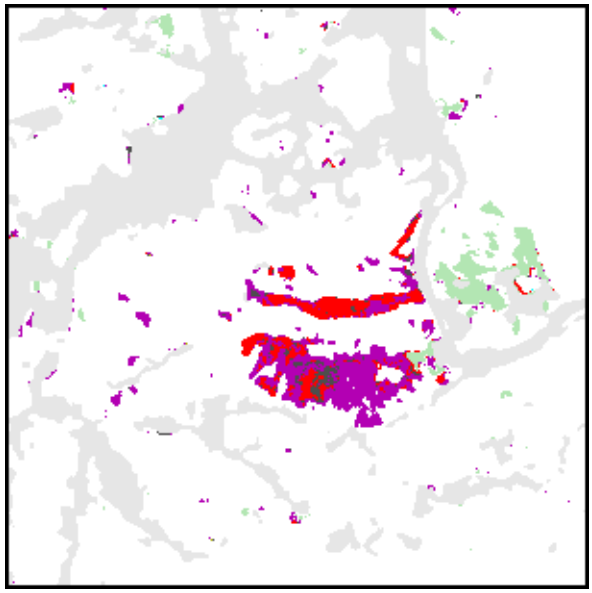

Figure 4.5: Classifier results for a known fire in 2002 preceded by a southern pine beetle outbreak outside of Newport, Tennessee. In the main disturbed area, pixels are are classified as a mix of fire (*red*), southern pine beetle (*blue*, and as uncertain (*dark gray*. Non-forest (*light gray*), stable (*white*), and regenerating (*green*) areas are also shown.

outbreak frequency of 7-10 years in the Piedmont and Southeastern Plains reported in Price et al. (1998). The peaks for all three SPB events are roughly the same size, though the latest event is worse, possibly because of synergistic effects with widespread drought in those years.

Fires in the Ridge and Valley and Central Appalachians range from 0.2% to 1.5% of forested land each year (Figure 4.6B). In the Central Appalachians, the peaks of fire activity in 1989 and 1999 correspond very well to the peaks reported by Lafon et al. (2005), as does the decline in activity in the 1990s. Additionally, in the Ridge and Valley, peaks correlated well with southern pine beetle activity (Schowalter et al., 1981). In both, peaks in the late 1990s and around 2008 correspond with severe droughts that affected the entire southeast. As such, these fires were likely the result of standing dead pine left in areas disturbed by drought and southern pine beetle (Xi et al., 2009).

Disturbance from gypsy moth peaks with two outbreaks estimated in the North Central Appalachian ecoregion in the early 90's, and then two again in the north central ecoregion after 2005 (Figure 4.6C). Additionally, in the Southeastern Plains, autocorrelation analysis (Maronna et al., 2006) of these time-series finds an average distance between peaks of 4.75 years in the Southeastern Plains and 4.5 years in the North Central Appalachians. As reported by Johnson et al. (2006), defoliation from gypsy moth is cyclic with a frequency

Table 4.5: Mean percent of forested area affected by each identified disturbance agent, plus the area undergoing regeneration, per year by ecoregion. Stable regions are those that are neither disturbed nor regenerating. Since many disturbance events last multiple years, percents do not represent new disturbance events, but simply the area affected.

Ecoregion	SPB	AGM	HWA	BBD	Anthr	Fire	Weath	Unkwn	Uncrt	Stable	Regen
Piedmont	4.59	0.12	0.02	0.22	0.02	0.72	0.18	0.01	0.57	84.47	9.09
N Central Apps.	0.66	0.13	0.03	0.45	0.01	0.97	0.03	0.04	0.54	94.31	2.85
SE Plains	5.02	0.31	0.06	0.41	0.08	0.93	0.16	0.02	0.74	81.05	11.23
Blue Ridge	1.01	0.02	0.01	0.18	0.00	0.33	0.04	0.01	0.29	96.05	2.05
Ridge & Valley	1.92	0.17	0.03	0.41	0.03	0.75	0.07	0.02	0.63	91.65	4.30
SW Apps.	2.27	0.10	0.02	0.22	0.02	0.50	0.09	0.01	0.45	91.83	4.48
Central Apps.	0.95	0.10	0.02	0.36	0.02	0.55	0.04	0.01	0.43	95.21	2.31
W Alleg. Plat.	1.16	0.07	0.02	0.39	0.02	0.45	0.05	0.01	0.51	94.70	2.63

of approximately 4 to 5 years, precisely what is seen in the satellite record and detected by these methods. The intervening years in the North Central Appalachians with little gypsy moth activity may be due to substantial control efforts (Campbell and Schlarbaum, 1994).

For hemlock woolly adelgid, the classifier detects outbreaks in the Central Appalachians in the late 1980s, and then again in the Blue Ridge in the late 1990s and in 2005 (Figure 4.6D). All of these outbreaks are reflected in the Ridge and Valley ecoregion as well. These dates correspond to the spread of hemlock woolly adelgid southward over that time period (Krapfl et al., 2011), and the comparatively high levels in the Ridge and Valley during the 1990s suggest that migration could have occurred along those highlands. However, while the relative patterns may be correct, the classifier underestimates the absolute amount of hemlock woolly adelgid in the region. Using numbers from 2009, the amount may be off by an order of magnitude (Potter and Conkling, 2010). Detecting eastern hemlock dieback is difficult for two reasons. First, hemlocks often grow in mixed forests with a rhododendron understory, which causes remotely-sensed imagery to remain green and structurally complex. Second, hemlocks typically grows in hill-shaded valleys that cause normalized difference vegetation indices to be less sensitive to small changes.

4.4 Conclusions

The results presented here represent the first yearly, regional-scale estimates of forest disturbance partitioned by disturbance agent. With this first attempt, we find good correspondence with previously described patterns of disturbance in southern pine beetle, gypsy moth, and fire, including direct observational results of their predicted periodicities

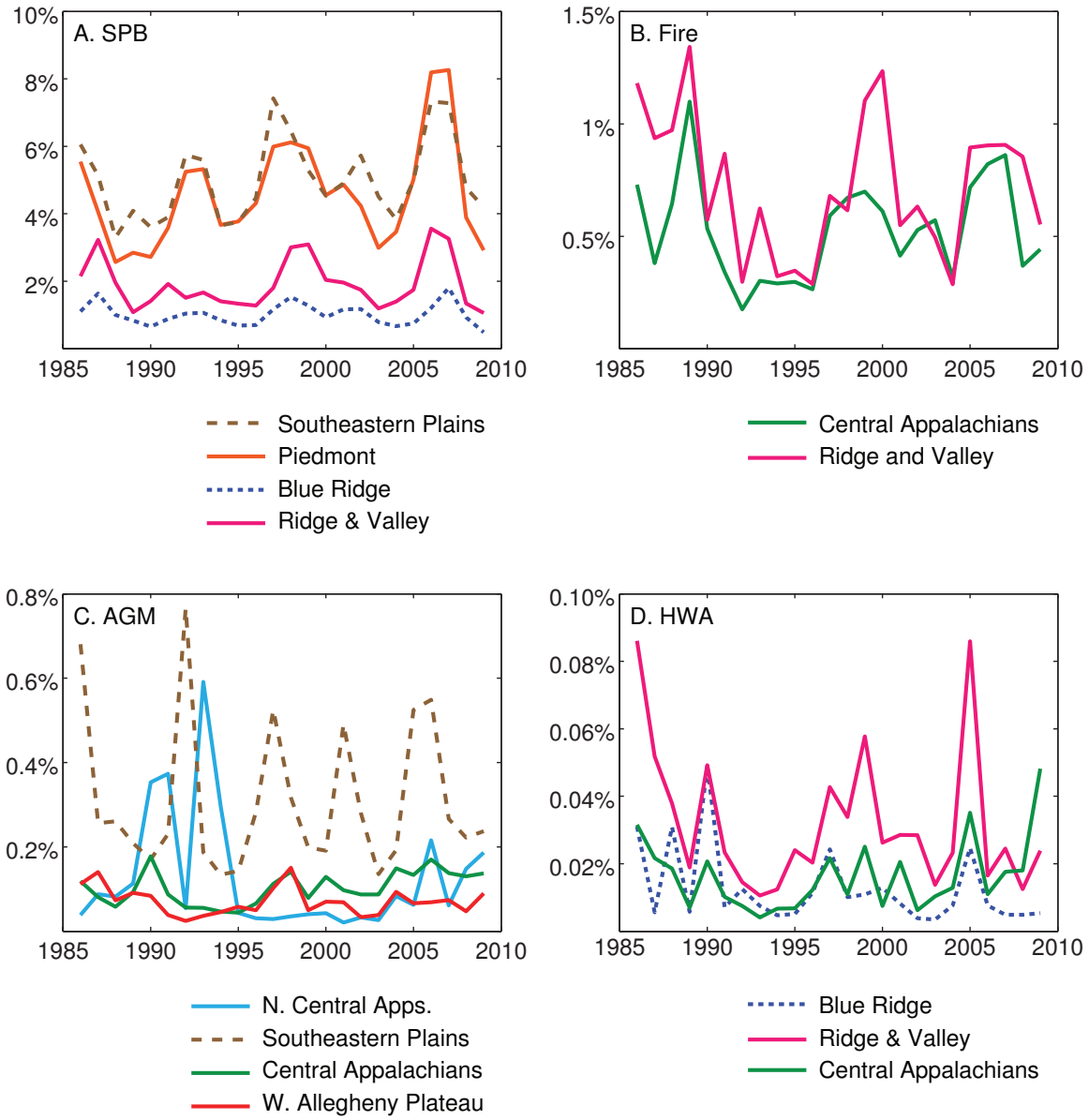

Figure 4.6: Percent of forest area affected by southern pine beetle (*A*), fire (*B*), gypsy moth (*C*), and hemlock woolly adelgid (*D*) for selected ecoregions by year.

over entire ecoregions. Such estimates are vital for forest monitoring efforts and for predictive models of impacts by ecoregion in order to preserve the aesthetic and economic uses of forest ecosystems. The role of disturbance on the forest carbon sink is a major source of uncertainty in atmospheric carbon models (Pan et al., 1998), and more accurate estimates of turnover are necessary. Estimates that differentiate between different causal agents, which transfer carbon from living forests into functionally different pools such as standing deadwood or down, charred wood, are especially important. Additionally, modeling the periodic dynamics of forest insect outbreaks can help provide more accurate yearly representations of forest carbon flux in models that explicitly model disturbances (Weng et al., 2012).

With an overall classifier accuracy of 81%, combined with the 85% accuracy in VeRDET, these methods are suitable for reasonable and useful estimates of forest disturbance over moderate spatial scales, such as aggregates of ecoregions or even quarter-degree grid cells. However, these methods are not reliable enough for accurate mapping of disturbance agent at the native 30 m resolution. As such, interpretation should be limited to better understanding broad spatial and temporal patterns of forest change and the processes that drive them. Future iterations of these algorithms will improve accuracy and reliability. To do so, additional comparative data is needed to assess the accuracy of beech bark disease and anthracnose and additional spatially-explicit data on hemlock woolly adelgid and weather-related disturbance events is needed to increase classifier accuracy for these agents. Additionally, the inclusion of anthropogenic causal agents in future classifiers is an essential next step toward the accurate attribution of forest disturbances.

Finally, an important goal was to assess the use of spatial information in automated classification of forest disturbance agents. The use of texture-based indices were found to be useful for discriminating disturbance agents. Both patch-variance and local entropy were included in the best classifiers of disturbance agent. From this, it is clear that different disturbance agents impact forest canopies in different ways, a finding that reinforces the need to discriminate between disturbance agents in ecologically-informed models.

Conclusions

The methods described in the preceding chapters significantly advance the science of the remote sensing of forest disturbance. These automated algorithms systematically quantify vegetation change over a large area at fine spatial and temporal resolutions with minimal operator involvement. Additionally, though the methods are exclusively applied here to the problem of disturbance in complex eastern forests, much of the approach can be adapted to detecting generic land cover change, such as shifts from shrub to grassland, progressive wetland inundation, or desertification. To do so, the vegetation indices used here would simply need to be replaced with new indices that vary over the desired transition.

The cloud and cloud-shadow detection algorithm described in Chapter 1 is a top, perhaps best, method for single-date Landsat TM/ETM+ imagery currently available. However, with the decommissioning of Landsat 5 in 2011 and the scan-line failure in Landsat 7, the future of Landsat imagery will be from Landsat 8, which has different band specifications from the earlier TM and EMT+ sensors. The relatively simple task of training an additional network over Landsat 8 imagery must be completed before SPARCS is viable for widespread adoption. The spatial post-processing procedures would remain the same. Additionally, cloud and cloud-shadow detection would likely benefit from the information in multi-temporal image stacks, though this is a more research-intensive undertaking.

The method detailed in Chapter 4 is a first attempt at disturbance agent classification and produces accuracies of 70-80%. However, while this level of explained variation is exceptional and sufficient for average incidence rates summed over relatively large areas, such as quarter-degree cells for climate models, 20-30% error rates are nevertheless too high for fine-scale management and monitoring decisions. Spatially and temporally reliable estimates will require additional information for both rare and difficult-to-detect disturbance agents. Such a task is straightforward in the sense of simply needing time, effort, and dedicated researchers to uncover and synthesize the data required. In addition, because a great many possible inputs to the neural network could provide meaningful

information about disturbance agents, including summaries of different aspects of pattern and texture or environmental gradients, testing all possible combinations becomes computationally infeasible, even with supercomputers. A different approach to training the network that algorithmically weeds out non-contributing inputs by evolving a subset of weights to zero, called L-1 regularization or weight-decay, is a promising solution to this feature selection problem.

Finally, the disturbance detection method described in Chapter 3 advances the field toward a truly patch-based approach for disturbance detection while retaining features that allow for the simultaneous detection of both swift and slow disturbance events. Future work that builds from the method presented should focus on merging the 2-dimensional spatial total variation regularized denoising with the 1-dimensional temporal total variation regularized differentiation operator. Doing so will enable the algorithm to account for the spatio-temporal covariance structure in the imagery stacks and define a spatio-temporal manifold of disturbance and regeneration events. Additionally, combining this approach with a variable scale parameter (α) will lead to methods that are able to deconvolve multiple disturbance processes from the patterns those processes leave on the landscape. Such methods will require new data mining visualization approaches to extract meaning from the data.

Bibliography

Aber, J. D., Ollinger, S. V., Driscoll, C. T., Likens, G. E., Holmes, R. T., Freuder, R. J., and Goodale, C. L. (2002). Inorganic nitrogen losses from a forested ecosystem in response to physical, chemical, biotic, and climatic perturbations. *Ecosystems*, 5:648–658. [pg 41]

Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., and Gumley, L. E. (1998). Discriminating clear sky from clouds with MODIS. *Journal of Geophysical Research*, 103(24):32141–32157. [pg 9]

Ayres, M. P. and Lombardero, M. J. (2000). Assessing the consequences of global change for forest disturbance from herbivores and pathogens. *Science of The Total Environment*, 262(3):263–286. [pg 64]

Beale, M. H., Hagan, M. T., and Demuth, H. B. (2013). *Neural Network Toolbox User's Guide R2013b*. The MathWorks, Inc, Natick, MA. [pg 15]

Berendes, T., Sengupta, S., Welch, R., Wielicki, B., and Navar, M. (1992). Cumulus cloud base height estimation from high spatial resolution Landsat data: a Hough transform approach. *IEEE Transactions on Geoscience and Remote Sensing*, 30(3):430–443. [pg 9]

Bernal, S., Hedin, L. O., Likens, G. E., Gerber, S., and Buso, D. C. (2012). Complex response of the forest nitrogen cycle to climate change. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9):3406–11. [pg 41]

Bianchini, E., Pimenta, J. A., and dos Santos, F. A. M. (2001). Spatial and temporal variation in the canopy cover in a tropical semi-deciduous forest. *Brazilian Archives of Biology and Technology*, 44(3):269–276. [pg 38]

Bleby, T. M., McElrone, A. J., and Jackson, R. B. (2010). Water uptake and hydraulic redistribution across large woody root systems to 20 m depth. *Plant, Cell & Environment*, 33(12):2132–48. [pg 41]

Bradley, B. A. (2013). Remote detection of invasive plants: a review of spectral, textural and phenological approaches. *Biological Invasions*, 16(7):1411–1425. [pg 39]

Campbell, F. T. and Schlarbaum, S. E. (1994). *Fading Forests: North American Trees and the Threat of Exotic Pests*. Natural Resources Defense Council, Inc. [pg 76]

Chander, G., Markham, B. L., and Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, 113(5):893–903. [pp 13, 44]

Chavez, P. S. (1988). An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24(3):459–479. [pp 44, 46]

Chavez, P. S. (1996). Image-based atmospheric corrections - revisited and improved. *Photogrammetric Engineering & Remote Sensing*, 62(9):1025–1036. [pp 13, 44]

Choi, H. (2004). Cloud detection in Landsat imagery of ice sheets using shadow matching technique and automatic normalized difference snow index threshold value decision. *Remote Sensing of Environment*, 91(2):237–242. [pg 9, 9]

Cihlar, J. and Howarth, J. (1994). Detection and removal of cloud contamination from AVHRR images. *IEEE Transactions on Geoscience and Remote Sensing*, 32(3):583–589. [pg 9]

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46. [pg 69]

Cohen, W. B., Yang, Z., and Kennedy, R. E. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSyncTools for calibration and validation. *Remote Sensing of Environment*, 114(12):2911–2924. [pp 54, 56, 56, 61]

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., and Lambin, E. (2004). Digital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing*, 25(9):1565–1596. [pg 32]

Corder, G. W. and Foreman, D. I. (2009). *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*. [pg 36]

Costanza, J. K., Hulcr, J., Koch, F. H., Earnhardt, T., McKerrow, A. J., Dunn, R. R., and Collazo, J. A. (2012). Simulating the effects of the southern pine beetle on regional dynamics 60 years into the future. *Ecological Modelling*, 244:93–103. [pg 64]

Crist, E. P. (1985). A TM Tasseled Cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, 17(3):301–306. [pp 12, 14, 66]

Dale, V. H., Joyce, L. A., Mcnulty, S., Neilson, R. P., Ayres, M. P., Flannigan, M. D., Hanson, P. J., Irland, L. C., Lugo, A. E., Peterson, C. J., Simberloff, D., Swanson, F. J., Stocks, B. J., and Michael Wotton, B. (2001). Climate Change and Forest Disturbances. *BioScience*, 51(9):723. [pp 42, 64]

de Jong, R., Verbesselt, J., Schaepman, M. E., and Bruin, S. (2012). Trend changes in global greening and browning: contribution of short-term trends to longer-term change. *Global Change Biology*, 18(2):642–655. [pg 43]

Derrien, M., Farki, B., and Harang, L. (1993). Automatic cloud detection applied to NOAA-11/AVHRR imagery. *Remote Sensing of Environment*, 46(3):246–267. [pg 9]

du Buf, J., Kardan, M., and Spann, M. (1990). Texture feature performance for image segmentation. *Pattern Recognition*, 23(3-4):291–309. [pg 39]

EPA (2005). *Mountaintop Mining / Valley Fills in Appalachia Final Programmatic Environmental Impact Statement*. Number October. Philadelphia, PA. [pg 3]

Eyer, F. H. (1980). *Forest Cover Types of the United States and Canada.* Society of American Foresters, Washington DC. [pg 42]

Fitzpatrick, M. C., Preisser, E. L., Porter, A., Elkinton, J., and Ellison, A. M. (2012). Modeling range dynamics in heterogeneous landscapes: invasion of the hemlock woolly adelgid in eastern North America. *Ecological Applications*, 22(2):472–486. [pg 64]

Fralish, J. S. and McArdle, T. G. (2009). Forest dynamics across three century-length disturbance regimes in the Illinois Ozark hills. *The American Midland Naturalist*, 162(2):418–449. [pg 63]

Frolking, S., Palace, M. W., Clark, D. B., Chambers, J. Q., Shugart, H. H., and Hurtt, G. C. (2009). Forest disturbance and recovery: A general review in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy structure. *Journal of Geophysical Research*, 114(G2):1–27. [pp 31, 32, 38]

Gao, B.-C. and Kaufman, Y. J. (1995). Selection of the 1.375-m MODIS Channel for Remote Sensing of Cirrus Clouds and Stratospheric Aerosols from Space. *Journal of the Atmospheric Sciences*, 52(23):4231–4237. [pg 9]

García-Haro, F. J., Gilabert, M. A., and Meliá, J. (2001). Monitoring fire-affected areas using Thematic Mapper data. *International Journal of Remote Sensing*, 22(4):533–549. [pg 42]

Garren, K. H. (1943). Effects of fire on vegetation of the southeastern United States. *Botanical Review*, 9(9):617–654. [pg 31]

Ge, S., Carruthers, R., Gong, P., and Herrera, A. (2006). Texture analysis for mapping Tamarix parviflora using aerial photographs along the Cache Creek, California. *Environmental monitoring and assessment*, 114(1-3):65–83. [pg 39]

Gilliam, F. S. and Platt, W. J. (1999). Effects of long-term fire exclusion on tree species composition and stand structure in an old-growth *Pinus palustris* (Longleaf pine) forest. *Plant Ecology*, 140(1):15–26. [pg 63]

Glenn, E. P., Huete, A. R., Nagler, P. L., and Nelson, S. G. (2008). Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: what vegetation indices can and cannot tell us about the landscape. *Sensors*, 8(4):2136–2160. [pg 39]

Goldstein, T. and Osher, S. (2009). The split Bregman method for L1-regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343. [pp 12, 14, 33, 50]

Goodwin, N. R., Collett, L. J., Denham, R. J., Flood, N., and Tindall, D. (2013). Cloud and cloud shadow screening across Queensland, Australia: An automated method for Landsat TM/ETM+ time series. *Remote Sensing of Environment*, 134:50–65. [pp 9, 22]

Goodwin, N. R., Coops, N. C., Wulder, M. A., Gillanders, S., Schroeder, T. A., and Nelson, T. (2008). Estimation of insect infestation dynamics using a temporal sequence of Landsat data. *Remote Sensing of Environment*, 112(9):3680–3689. [pg 32]

Goward, S. N., Masek, J. G., Cohen, W., Moisen, G., Collatz, G. J., Healey, S., Houghton, R. A., Huang, C., Kennedy, R., Law, B., Powell, S., Turner, D., and Wulder, M. A. (2008). Forest disturbance and North American carbon flux. *Eos, Transactions of the American Geophysical Union*, 89(11):105. [pg 41]

Hagolle, O., Huc, M., Pascual, D., and Dedieu, G. (2010). A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENS, LANDSAT and SENTINEL-2 images. *Remote Sensing of Environment*, 114(8):1747–1755. [pg 9]

Hall, D., Foster, J., Verbyla, D., Klein, A., and Benson, C. (1998). Assessment of snow-cover mapping accuracy in a variety of vegetation-cover densities in central Alaska. *Remote Sensing of Environment*, 66(2):129–137. [pg 12]

Hansen, M. C. and Loveland, T. R. (2012). A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment*, 122:66–74. [pg 32]

Haykin, S. (2008). *Neural Networks and Learning Machines*. Prentice Hall, Saddle River, NJ, 3rd edition. [pp 10, 11]

Hicke, J. A., Allen, C. D., Desai, A. R., Dietze, M. C., Hall, R. J., Ted Hogg, E. H., Kashian, D. M., Moore, D., Raffa, K. F., Sturrock, R. N., and Vogelmann, J. (2012). Effects of biotic

disturbances on forest carbon cycling in the United States and Canada. *Global Change Biology*, 18(1):7–34. [pp 32, 64]

Hollingsworth, B. V., Chen, L., Reichenbach, S. E., and Irish, R. R. (1996). Automated cloud cover assessment for Landsat TM images. In Descour, M. R. and Mooney, J. M., editors, *SPIE's 1996 International Symposium on Optical Science, Engineering, and Instrumentation*, pages 170–179. International Society for Optics and Photonics. [pg 9]

Holmgren, P. and Thuresson, T. (1998). Satellite remote sensing for forestry planning – A review. *Scandinavian Journal of Forest Research*, 13(1-4):90–110. [pg 32]

Huang, C., Thomas, N., Goward, S. N., Masek, J. G., Zhu, Z., Townshend, J. R. G., and Vogelmann, J. E. (2010). Automated masking of cloud and cloud shadow for forest change analysis using Landsat images. *International Journal of Remote Sensing*, 31(20):5449–5464. [pp 9, 9, 9, 25, 42]

Huang, T. S. (1975). *Picture processing and digital filtering*. Springer-Verlang, New York. [pg 64]

Hughes, M. and Hayes, D. (2014). Automated detection of cloud and cloud shadow in single-date Landsat imagery using neural networks and spatial post-processing. *Remote Sensing*, 6(6):4907–4926. [pp 8, 35, 48, 48, 66]

Javier Traver, V. and Bernardino, A. (2010). A review of log-polar imaging for visual perception in robotics. *Robotics and Autonomous Systems*, 58(4):378–398. [pg 28]

Jin, S. and Sader, S. A. (2005). Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote Sensing of Environment*, 94(3):364–372. [pg 56]

Johnson, D. M., Liebhold, A. M., and Bjørnstad, O. N. (2006). Geographical variation in the periodicity of gypsy moth outbreaks. *Ecography*, 29(3):367–374. [pg 75]

Johnson, E. W. and Ross, J. (2008). Quantifying error in aerial survey data. *Australian Forestry*, 71(3):216–222. [pg 39]

Ju, J. and Roy, D. P. (2008). The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally. *Remote Sensing of Environment*, 112(3):1196–1211. [pg 8]

Kennedy, R. E., Cohen, W. B., and Schroeder, T. A. (2007). Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sensing of Environment*, 110(3):370–386. [pg 9]

Kennedy, R. E., Yang, Z., and Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendrTemporal segmentation algorithms. *Remote Sensing of Environment*, 114(12):2897–2910. [pp 5, 43, 43]

Kneeshaw, D. D. and Prévost, M. (2007). Natural canopy gap disturbances and their role in maintaining mixed-species forests of central Quebec, Canada. *Canadian Journal of Forest Research*, 37(9):1534–1544. [pg 38]

Kramer, C. Y. (1956). Extension of multiple range tests to group means with unequal numbers of replications. *Biometrics*, 12(3):307–310. [pp 18, 69]

Krapfl, K. J., Holzmueller, E. J., and Jenkins, M. A. (2011). Early impacts of hemlock woolly adelgid in *Tsuga canadensis* forest communities of the southern Appalachian mountains. *The Journal of the Torrey Botanical Society*, 138(1):93–106. [pg 76]

Lafon, C., Hoss, J., and Grissino-Mayer, H. (2005). The contemporary fire regime of the central Appalachian mountains and its relation to climate. *Physical Geography*, 26(2):126–146. [pp 3, 75]

Lundquist, J. E. and Beatty, J. S. (2002). A method for characterizing and mimicking forest canopy gaps caused by different disturbances. *Forest Science*, 48(3):582–594. [pg 38]

Lunetta, R. S., Knight, J. F., Ediriwickrema, J., Lyon, J. G., and Worthy, L. D. (2006). Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sensing of Environment*, 105(2):142–154. [pg 42]

Luo, Y., Trishchenko, A., and Khlopenkov, K. (2008). Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven MODIS land bands over Canada and North America. *Remote Sensing of Environment*, 112(12):4167–4185. [pp 9, 15]

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the fifth Berkeley symposium on Math, Statistics and Probability*, 1(233):281–297. [pg 13]

Maronna, R. A., Martin, R. D., and Yohai, V. J. (2006). Robust Statistics. *Statistics*, 108:888–889. [pg 75]

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441. [pg 69]

Marr, D. and Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)*, 207:187–217. [pp 33, 51]

Martinuzzi, S., Gould, W., and González, O. (2007). Creating cloud-free Landsat ETM+ data sets in tropical landscapes: cloud and cloud-shadow removal. *General Technical Report IITF-GTR-32*, pages 1–18. [pg 9, 9]

Masek, J. G., Goward, S. N., Kennedy, R. E., Cohen, W. B., Moisen, G. G., Schleeweis, K., and Huang, C. (2013). United States forest disturbance trends observed using Landsat time series. *Ecosystems*, 16(6):1087–1104. [pg 64]

Masek, J. G., Huang, C., Wolfe, R., Cohen, W., Hall, F., Kutler, J., and Nelson, P. (2008). North American forest disturbance mapped from a decadal Landsat record. *Remote Sensing of Environment*, 112(6):2914–2926. [pg 42]

Mathworks (2013). *Statistics Toolbox User's Guide R2013b*. The MathWorks, Inc, Natick, MA. [pg 36, 36]

McConnel, T. J., Johnson, E. W., and Burns, B. (2000). *A Guide to Conducting Aerial sketchmapping Surveys FHTET 00-01*. USDA Forest Service, Fort Collins Colorodo. [pp 39, 67, 73]

McGarigal, K., Tagil, S., and Cushman, S. A. (2009). Surface metrics: an alternative to patch metrics for the quantification of landscape structure. *Landscape Ecology*, 24(3):433–450. [pg 32]

Meddens, A. J. H., Hicke, J. A., and Ferguson, C. A. (2012). Spatiotemporal patterns of observed bark beetle-caused tree mortality in British Columbia and the western United States. *Ecological Applications*, 22(7):1876–1891. [pp 39, 67]

Moran, M., Jackson, R., Slater, P., and Teillet, P. (1992). Evaluation of simplified procedures for retrieval of land surface reflectance factors from satellite sensor output. *Remote Sensing of Environment*, 41(2-3):169–184. [pg 13]

Nowak, J., Asaro, C., Klepzig, K., and Billings, R. (2008). The southern pine beetle prevention initiative: Working for healthier forests. *Journal of Forestry*, 106(5):261–267. [pg 74]

Omernik, J. (1987). Ecoregions of the conterminous United States. Map (scale 1:7,500,000). *Annals of the Association of American Geographers*, 77:118–125. [pg 70]

Oreopoulos, L., Wilson, M. J., and Várnai, T. (2011). Implementation on Landsat Data of a Simple Cloud-Mask Algorithm Developed for MODIS Land Bands. *IEEE Geoscience and Remote Sensing Letters*, 8:597–601. [pg 9, 9]

Orwig, D. A., Thompson, J. R., Povak, N., Manner, M., Niebyl, D., and Foster, D. R. (2012). A foundation tree at the precipice: *Tsuga canadensis* health after the arrival of *Adelges tsugae* in central New England. *Ecosphere*, 3:1–16. [pg 31]

Oumar, Z. and Mutanga, O. (2011). The potential of remote sensing technology for the detection and mapping of *Thaumastocoris peregrinus* in plantation forests. *Southern Forests: a Journal of Forest Science*, 73:23–31. [pg 32]

Pan, Y., Chen, J. M., Birdsey, R., McCullough, K., He, L., and Deng, F. (2011). Age structure and disturbance legacy of North American forests. *Biogeosciences*, 8(3):715–732. [pp 63, 74]

Pan, Y., Melillo, J. M., McGuire, A. D., Kicklighter, D. W., Pitelka, L. F., Hibbard, K., Pierce, L. L., Running, S. W., Ojima, D. S., Parton, W. J., and Schimel, D. S. (1998). Modeled responses of terrestrial ecosystems to elevated atmospheric CO2 : a comparison of simulations by the biogeochemistry models of the Vegetation/Ecosystem Modeling and Analysis Project (VEMAP). *Oecologia*, 114(3):389–404. [pg 78]

Paradis, A., Elkinton, J., Hayhoe, K., and Buonaccorsi, J. (2008). Role of winter temperature and climate change on the survival and future range expansion of the hemlock woolly adelgid (*Adelges tsugae*) in eastern North America. *Mitigation and Adaptation Strategies for Global Change*, 13(5-6):541–554. [pg 64]

Potter, K. M. and Conkling, B. L. (2010). Forest health monitoring: national status, trends, and analysis 2010. Technical report, U.S. Department of Agriculture Forest Service, Southern Research Station, Asheville, NC. [pg 76]

Powell, S. L., Cohen, W. B., Healey, S. P., Kennedy, R. E., Moisen, G. G., Pierce, K. B., and Ohmann, J. L. (2010). Quantification of live aboveground forest biomass dynamics

with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sensing of Environment*, 114:1053–1068. [pp 35, 49, 66]

Powell, S. L., Cohen, W. B., Kennedy, R. E., Healey, S. P., and Huang, C. (2014). Observation of trends in biomass loss as a result of disturbance in the conterminous U.S.: 19862004. *Ecosystems*, 17(1):142–157. [pp 3, 74]

Price, T. S., Doggett, C., Pye, J. M., and Smith, B. (1998). *A History of Southern Pine Beetle Outbreaks In The Southeastern United States*. http://www.barkbeetles.org/spb. [pg 75]

Pugh, E. and Gordon, E. (2013). A conceptual model of water yield effects from beetle-induced tree death in snow-dominated lodgepole pine forests. *Hydrological Processes*, 27(14):2048–2060. [pg 41]

Rudin, L., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60:259–268. [pp 11, 12, 14, 33, 50]

Running, S. W. (2008). Ecosystem disturbance, carbon, and climate. *Science*, 321(5889):652–653. [pg 63]

Russ, J. C. (2007). *The Image Processing Handbook*. Taylor and Francis, Boca Raton, USA, 5th editio edition. [pp 64, 67]

Saunders, R. and Kriebel, K. (1988). An improved method for detecting clear sky and cloudy radiances from AVHRR data. *International Journal of Remote Sensing*, 9(1):123–150. [pg 9]

Scaramuzza, P. L., Bouchard, M. A., and Dwyer, J. L. (2012). Development of the Landsat Data Continuity Mission cloud-cover assessment algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, 50(4):1140–1154. [pp 10, 12]

Schimmel, J. and Granstrom, A. (1996). Fire severity and vegetation response in the boreal Swedish forest. *Ecology*, 77(5):1436–1450. [pg 31]

Schowalter, T. D., Coulson, R. N., and Crossley, D. A. (1981). Role of southern pine beetle and fire in maintenance of structure and function of the southeastern coniferous forest. *Environmental Entomology*, 10(6):821–825. [pp 73, 75]

Shands, W. E. and Healy, R. G. (1977). *The Lands Nobody Wanted*. Conservation Foundation, Washington DC. [pg 3]

Simpson, J. and Stitt, J. (1998). A procedure for the detection and removal of cloud shadow from AVHRR data over land. *IEEE Transactions on Geoscience and Remote Sensing*, 36(3):880–890. [pg 9, 9]

Spaulding, H. L. and Rieske, L. K. (2010). The aftermath of an invasion: structure and composition of central Appalachian hemlock forests following establishment of the hemlock woolly adelgid, *Adelges tsugae*. *Biological Invasions*, 12(9):3135–3143. [pg 63]

Strother, D. H. (1857). The Virginian Canaan. *Harper's New Monthly Magazine*, 8:31–63. Republished in: Gohdes C. (1967). Hunting in the Old South: Original Narratives of the Hunters. Louisiana State University Press, Kingsport TN. 200 pp 31-63. [pg 1]

Townsend, P. A., Singh, A., Foster, J. R., Rehberg, N. J., Kingdon, C. C., Eshleman, K. N., and Seagle, S. W. (2012). A general Landsat model to predict canopy defoliation in broadleaf deciduous forests. *Remote Sensing of Environment*, 119:255–265. [pp 31, 32]

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2):127–150. [pp 35, 49, 66]

Turner, M. G. (2005). Landscape ecology: what is the state of the science? *Annual Review of Ecology, Evolution, and Systematics*, 36(1):319–344. [pg 38]

Turner, M. G., Gardner, R. H., and O'Neill, R. V. (2001). *Landscape Ecology in Theory and Practice*. Springer Science+Business Media, New York, NY. [pg 33]

van Wagtendonk, J. W., Root, R. R., and Key, C. H. (2004). Comparison of AVIRIS and Landsat ETM+ detection capabilities for burn severity. *Remote Sensing of Environment*, 92(3):397–408. [pp 35, 49, 66]

Vogel, C. R. and Oman, M. E. (1996). Iterative methods for total variation denoising. *SIAM Journal on Scientific Computing*, 17(1):227–238. [pg 52]

Vogelmann, J. E., Tolk, B., and Zhu, Z. (2009). Monitoring forest changes in the southwestern United States using multitemporal Landsat data. *Remote Sensing of Environment*, 113(8):1739–1748. [pg 42]

Weed, A. S., Ayers, M. P., and Hicke, J. A. (2013). Consequences of climate change for biotic disturbances in North American forests. *Ecological Monographs*, 83(4):441–470. [pg 42]

Weng, E., Luo, Y., Wang, W., Wang, H., Hayes, D. J., McGuire, a. D., Hastings, A., and Schimel, D. S. (2012). Ecosystem carbon storage capacity as affected by disturbance

regimes: A general theoretical model. *Journal of Geophysical Research*, 117(G3):1–15. [pg 78]

White, P. S., Collings, B. S., and Wein, G. R. (2011). Natural disturbances and early successional habitats. In Greenberg, C. H., Collins, B. S., and Thompson, F. R., editors, *Sustaining Young Forest Communities*, pages 27–40. Springer. [pg 41]

Whittaker, R. H. (1956). Vegetation of the Great Smoky Mountains. *Ecological Monographs*, 26(1):1–80. [pg 3]

Williams, C. A., Collatz, G. J., Masek, J., Huang, C., and Goward, S. N. (2013). Impacts of disturbance history on forest carbon stocks and fluxes: Merging satellite disturbance mapping with forest inventory data in a carbon cycle model framework. *Remote Sensing of Environment*, in press. [pg 41]

Wilson, E. H. and Sader, S. A. (2002). Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sensing of Environment*, 80(3):385–396. [pg 42]

Woodall, C. W., Perry, C. H., and Miles, P. D. (2006). The relative density of forests in the United States. *Forest Ecology and Management*, 226:368–372. [pg 64]

Xi, W., Waldron, J. D., Lafon, C. W., Cairs, D. M., Birt, A. G., Tchakerian, M. D., and Klepzig, K. D. (2009). Modeling Long-term effects of altered fire regimes following southern pine beetle outbreaks (North Carolina). *Ecological Restoration*, 27(1):24–26. [pg 75]

Zar, J. H. (2010). *Biostatistical Analysis*. Pearson, 4th edition. [pg 17]

Zhang, F., Chen, J. M., Pan, Y., Birdsey, R. A., Shen, S., Ju, W., and He, L. (2012). Attributing carbon changes in conterminous U.S. forests to disturbance and non-disturbance factors from 1901 to 2010. *Journal of Geophysical Research: Biogeosciences*, 117(G2). [pg 64]

Zhu, Z. and Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118:83–94. [pp 9, 9, 10]

Vita

Michael Joseph (Joe) Hughes was born in Huntington, West Virginia, on November 19, 1980. He grew up wandering the forests around his home in nearby Ona, and graduated from Cabell Midland High School in 1999. He then studied Integrated Science and Technology, with a concentration in Environmental Policy and Assessment, at Marshall University and received his Bachelors of Science in 2003. He proceeded to pursue a Masters of Science, also at Marshall University, concentrating in Geobiophysical Modeling and graduated in 2006 with a thesis on determining assemblages of periphyton, benthic macroinvertebrates, heavy metals, and water quality indicators on the heavily coal-mine impacted Stony River in Grant County, West Virginia. While writing his thesis, Joe worked for Vandalia Research, where he developed The Mystery of Lyle and Louise, six interdisciplinary educational kits in forensic science for high school students and undergraduate non-science majors, which are currently available through Fisher Scientific Education. He then worked at The Herald Dispatch as a web developer, database administrator, and systems programmer for editorial operations. In 2009, he accepted a SCALE-IT Fellowship at the University of Tennessee, Knoxville, to pursue computational ecology, from which he graduated with a doctorate in Ecology and Evolutionary Biology in August, 2014.