UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

LANDSCAPE HETEROGENEITY AND SPATIO-TEMPORAL RESOLUTION CONSIDERATIONS FOR MAPPING LAND COVER CHANGES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

TRUNG VINH TRAN
Norman, Oklahoma
2013

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/7681

LANDSCAPE HETEROGENEITY AND SPATIO-TEMPORAL RESOLUTION CONSIDERATIONS FOR MAPPING LAND COVER CHANGES

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF GEOGRAPHY AND ENVIRONMENTAL SUSTAINABILITY

BY

_____
Dr. Jason P. Julian, Chair

_____
Dr. Kirsten M. de Beurs

_____
Dr. Sally Gros

_____
Dr. Charles G. Warnken

_____
Dr. Xiangming Xiao

© Copyright by TRUNG VINH TRAN 2013
All Rights Reserved.

I dedicate my dissertation work to my grandparents, Nguyễn Duy Hưng, Trần Thị Khanh, Phan Văn Ký, Trần Thị Sánh, for their love. A special feeling of gratitude to my loving parents, Trần Vĩnh Phước and Trần Thị Thu Hương, whose words of encouragement ring in my ears. My sister, Trần Thị Vân Anh, has never left my side.

I also dedicate this dissertation to my amazing wife, Trương Vũ Khánh Trâm. Her hours of work in loving our little son, Trần Vĩnh Nhật Nam, enabled the hours of research, contemplation, and writing necessary to complete this project.

ACKNOWLEDGEMENTS

I wish to thank my committee members who were more than generous with their expertise and precious time. I want to express my deep gratitude to Dr. Jason P. Julian and Dr. Kirsten M. de Beurs, who were my advisors, for their countless hours of reflecting, reading, encouraging, and most of all patience throughout the entire process. Thank you Dr. Sally Gros, Dr. Charles Warnken, and Dr. Xiangming Xiao for agreeing to serve on my committee. I would like to acknowledge and thank my Department of Geography and Environmental Sustainability for allowing me to conduct my research and for providing any assistance requested.

TABLE OF CONTENTS

Acknowledgements iv

List of Tables viii

List of Figures ix

Abstract xiv

Chapter 1. INTRODUCTION 1

1.1 Research objectives 6

1.2 Organization of dissertation 7

Chapter 2. LAND COVER HETEROGENEITY EFFECT ON SUBPIXEL CLASSIFICATION 8

2.1 Abstract 8

2.2 Introduction 9

2.3 Study area 11

2.4 Data 14

2.5 Methodology 15

2.5.1 Classification 15

2.5.2 Validation 17

2.5.3 Statistical analysis 18

2.6 Results and discussion 19

2.6.1 Land cover map and accuracy 19

2.6.2 Impact of heterogeneity on classification accuracies 26

2.7 Conclusion 30

Chapter 3. ACCOUNTING FOR LAND COVER HETEROGENEITY IN A SPATIAL TEMPORAL ADAPTIVE ALGORITHM FOR MAPPING REFLECTANCE CHANGE: A CASE STUDY IN SOUTHEAST OKLAHOMA 2000-2011..... 31
3.1 Abstract..... 31
3.2 Introduction..... 31
3.3 Study area..... 36
3.4 Data and pre-processing..... 36
3.4.1 Landsat data..... 36
3.4.2 MODIS data..... 37
3.4.3 Land cover data..... 40
3.4.4 Validation data..... 40
3.5 Methods..... 45
3.5.1 Spatial pattern analysis of NLCDs..... 45
3.5.2 Statistical test of tasseled cap indices..... 45
3.5.3 Data fusion..... 46
3.5.4 Validation..... 54
3.5.5 Mapping disturbances..... 56
3.6 Results..... 56
3.6.1 Land cover heterogeneity in the study area..... 56
3.6.2 Variability of tasseled cap indices due to different compositions of forest types..... 57
3.6.3 Method validation and comparison..... 59
3.6.4 Trends of area disturbed..... 64

3.6.5 Spatio-temporal distribution disturbances 65
3.7 Discussion 66
3.8 Conclusions 73
3.9 Acknowledgement 74
Chapter 4. MAPPING FIRE AND TIMBER HARVESTING DISTURBANCES USING HIGH TEMPORAL RESOLUTION TIME-SERIES 75
4.1 Abstract 75
4.2 Introduction 75
4.3 Study area 77
4.4 Data 79
4.5 Methods 82
4.5.1 Disturbance temporal characteristics 82
4.5.2 Disturbance type mapping and validation 84
4.6 Results 84
4.6.1 Temporal characteristics of disturbances 84
4.6.2 Map of disturbance types 88
4.7 Discussions 91
4.8 Conclusion 95
Chapter 5. CONCLUSIONS 97
Suggestions for future research 100
References 103

LIST OF TABLES

Table 2.1. Land cover classes used in the classifications. 17

Table 2.2. The cross tabulation matrix of the average sample for the two classifications: per-pixel (normal text) and subpixel (bold italicized text). Values in this matrix were calculated for the average sample, of which the class proportions were averages of the class proportions across all 700 validation samples. 22

Table 3.1. Landsat acquisition dates. 37

Table 4.1. Summary of the Mann-Whitney *U* test. This table shows that the temporal characteristics are significantly different between classes. 88

Table 4.2. Error matrix. 88

LIST OF FIGURES

Figure 2.1. The study area centered around Little Rock, Arkansas, USA. Base layer is the land cover map (2010; 30-meter resolution) generated by the per-pixel classification in this study. The four Omernik level III ecoregions are outlined in black lines. 13

Figure 2.2. Land cover percentages (2010; 30-meter resolution) generated by the subpixel classification. 21

Figure 2.3. Mean producer's and user's accuracies of per-pixel and subpixel classifications together with mean heterogeneity by classes. The means of producer's and user's accuracies were calculated as the averages of all producer's and user's accuracies taken from the individual cross tabulation matrices developed for all validation samplesThe asterisks indicate * $p<0.05$, ** $p<0.01$, and *** $p<0.001$ in Wilcoxon Signed-Rank test..... 25

Figure 2.4. Scatterplots of estimated versus referenced land cover proportions for the six classes. The solid line is the $y=x$ line. 26

Figure 2.5. Mean overall, producer's and user's accuracies of per-pixel (a) and subpixel (b) classifications by groups of heterogeneity. Error bar indicates +1 standard error (s.t.e). Numbers at the top of bars of a group of heterogeneity indicate the groups of heterogeneity from which it was statistically different ($p<0.05$) according to the Steel-Dwass test. P-CR / U-CR = producer's / user's accuracy of cropland, artificial surface (AR), barren (BA), grassland/shrub (GR), tree (TR), and water (WA). 29

Figure 3.1. The study area (185 km x 185 km; Landsat scene path 26, row 36) in southeast Oklahoma. The background shows land-cover according to the 2006 NLCD. The most dominant land-cover type is forest including deciduous, evergreen, and mixed forests (56%), followed by grassland, shrub and pasture hay (35%). Wetland occupies 1.3%, cultivated land 0.6%, non-vegetated surface (e.g. urban and barren land) 4.4%, and water 2.7%. 39

Figure 3.2. Field-trip routes and locations. The background shows land-cover according to the NLCD 2006. 41

Figure 3.3. Field-trip photos taken in different years, showing disturbance conditions and their initial dates of disturbance detected by the proposed MSTAARCH algorithm..... 43

Figure 3.4. Overview of the methodology with three primary steps (tasseled cap transformation, disturbance index calculation, and data fusion) and their corresponding outputs. (a) and (b) are Landsat and MODIS tasseled cap time-series; (c) and (d) are Landsat and MODIS DI profiles of the selected pixel in (a) and (b); note that in (c), a threshold of 2 is applied to the Landsat DI profile to convert it to a binary profile representing disturbance conditions; and (e) is the final fused binary DI time-series overlapped with the MODIS DI time-series of the selected pixel, for which the initial and final dates of a disturbance are 7/19/2004 and 9/29/2008, respectively..... 47

Figure 3.5. Ecoregional average (N=100 MODIS pixels per ecoregion) of MODIS pixel-based mean patch size and Shannon's diversity index calculated based on NLCD 2001. Error bars represent standard errors..... 57

Figure 3.6. Multiple disturbance index time-series of a MODIS disturbed evergreen pixel located at the red dot in NAIP photos at the top row. Each of these time-series was normalized by different reference sets: MODIS pure deciduous (blue), MODIS pure evergreen (green), and MODIS STAARCH (purple). Please see text for explanation of these reference sets. The red line represents the threshold of 2 used to determine disturbance condition, represented by the black line, of the analyzed pixel. Gray dots represent available Landsat images identifying the analyzed pixel as undisturbed whereas red dots represent available Landsat images identifying the analyzed pixel as disturbed. 58

Figure 3.7. Time-series of scene averages of tasseled cap indices for MODIS reference sets: pure deciduous (N=685), pure evergreen (N=274), and their combination (N=1,774). Red vertical lines represent the time steps used for the Mann-Whitney test. 60

Figure 3.8. Changes in area disturbed (30-meter resolution) detected by different methods for a typical evergreen forest heavily harvested since 2000. In this figure, the first column (NAIP) represents NAIP aerial photos in 2005 and 2008 and the second to the fifth columns represent the four binary DI time-series: Landsat DI, STAARCH DI, Threshold-2 DI, and MSTAARCH DI. Please see text for explanation on the binary time-series. 61

Figure 3.9. Producer's accuracies represented as color matrices over time and between methods of disturbance detection when validated against NAIP photos (A) and thematic burn severity dataset (B). Please see text for explanation on the binary time-series. Blank cells in the matrices represent no data. 62

Figure 3.10. Trends of yearly cumulative area disturbed derived from the four binary DI time-series: Landsat DI, STAARCH DI, Threshold-2 DI, and MSTAARCH DI. Please see text for explanation on these binary time-series. 64

Figure 3.11. The spatio-temporal distribution of disturbances over the period of 2000-2011 in southeast Oklahoma. The beginning dates of disturbances in southeast Oklahoma (B) are represented as a color scheme ranging from yellow (2000) to blue (2011). The three map insets (A, C, D) represents disturbances around the Eufaula Lake, west of the South Central Plains and middle of the Ouachita Mountains ecoregions, respectively. The MODIS DI time-series and MSTAARCH binary DI time-series of a forest-edge disturbance and a clear-cut disturbance are represented in chart E and F. In these charts, the blue lines represent the MODIS DI normalized by the combination reference set; the black line represent the MSTAARCH binary DI time-series; gray dots represents available Landsat images identifying the analyzed pixel as undisturbed whereas red dots represents available Landsat images identifying the analyzed pixel as disturbed. 67

Figure 4.1. Extent and context of the 185x185 km^2 study area located in southeast Oklahoma. The two map insets on the right show the cumulative (in 12 years from 2000 to 2011) disturbed (orange) and undisturbed (green) areas within (top inset) and outside (bottom inset) of a protected area in the Ouachita National Forest. Cumulative areas are created by spatially accumulate all disturbed/undisturbed pixels across all time steps..... 78

Figure 4.2. Example profiles of fire (A) and harvest (B) disturbances as well as other disturbances (C). Red solid lines represent the threshold of 2 used to determine disturbance conditions of a given pixel at a given time step. Solid blue lines represent MODIS DI profiles of the given pixel. Dotted blue lines represent disturbance conditions detected by MSTAARCH of the given pixel. Gray dots represent available Landsat images identifying the given pixel as undisturbed whereas red dots represent available Landsat images identifying the given pixel as disturbed. 86

Figure 4.3. Histograms of duration (top row) as well as beginning (middle row) and ending (bottom row) times of disturbances for fires (left column), harvest (center column), and other disturbances (right column). 87

Figure 4.4. Map of disturbance types on 8/29/2005. 90

Figure 4.5. Classification results for fire, harvest, and other-disturbance landscapes... 91

Figure 5.1. The Arkansas Red River Basin and its land cover distribution according to the National Land Cover Database 2006. 101

ABSTRACT

Land cover is changing dramatically worldwide from both anthropogenic and natural drivers. In the United States, the rates and types of land cover change have varied temporally due to government policy, environmental regulation, global and national economic conditions, and regional weather and climate variability. Land cover changes can cause environmental degradation that affects long-term sustainability of human societies. Therefore, balancing the human need and environmental degradation requires explicit knowledge about environmental changes over multiple scales and perspectives. Remote sensing has been used as an effective tool to assess land changes across broad scales with multiple resolutions. However, the extraction of information from remotely sensed images is still challenged by the complex interaction between land cover heterogeneity and spatial as well as temporal resolutions. This dissertation aims at exploring such interaction in data classification and data fusion to better extract useful information about land cover. To achieve such goal, this dissertation first analyzes the impact of land cover heterogeneity in per-pixel and subpixel classification. Furthermore, this study also analyzes and proposes a data fusion method to better detect forest disturbances with high spatial and temporal resolutions. Using a high spatio-temporal resolution map of forest disturbances, this study suggests the use of temporal characteristics of disturbances to identify disturbance types. This study uses the South-Central United States as a case study for all experiments.

CHAPTER 1. INTRODUCTION

Land cover is changing dramatically worldwide from both anthropogenic and natural drivers (Lambin et al. 2001; Meyer and Turner II 1996; Sterling et al. 2012). In the United States, the rates and types of land cover change have varied temporally due to government policy, environmental regulation, global and national economic conditions, and regional weather and climate variability (Sleeter et al. 2013). Most of those changes, however, have taken place in forests as a result of forest harvest, urbanization, and wildfire (Sleeter et al. 2013). These land cover changes can cause environmental degradation that affects long-term sustainability of human societies (Foley et al. 2005). Balancing the human need and environmental degradation requires explicit knowledge about environmental changes over multiple scales and perspectives.

Remote sensing has been used as an effective tool to assess land changes across broad scales with multiple resolutions (Turner et al. 2007). The twenty-first century has seen over a hundred of satellite platforms carrying Earth observation sensors launched in addition to many airborne and terrestrial sensors deployed (Boyd and Foody 2011). Remote sensing provides a systematic observation of the Earth's surface, which has no sampling bias problem traditionally seen in the use of survey data to investigate land cover change for a large area (Jensen 2005). In addition, as they are images remotely taken from the space (e.g. Landsat 7 altitude being 705 kilometers), remote sensing imagery data increase the opportunities to do land research for areas that are not accessible due to hazardous conditions (Kennedy et al. 2009). Moreover, with the capability of revisiting the same place (e.g. 16 days with Landsat sensor), remote sensors have produced archives of data that are useful to characterize the spatiotemporal

patterns of land cover types for a long-term period across large areas (Turner et al. 2007). For instance, time-series of Landsat images have been used to map forest disturbances in North American for the period of 1990 and 2000 (Masek et al. 2008) or to document changes in spatial patterns of land covers in southern Chile (Echeverría et al. 2012). However, users of remote sensing should consider limitations regarding the spatial and temporal resolutions of imagery data.

Spatial resolution of remotely sensed images refers to the size of an area (often square or rectangle) on the ground within which the intensity of reflected electromagnetic radiation of all land covers are measured by a sensor (Strahler et al. 1986; Woodcock and Strahler 1987). When the images are unprocessed (e.g. smoothing), their spatial resolution is often equivalent to their pixel size (Atkinson 2004). With a wide range of the currently available spatial resolutions ranging from meters (e.g. GEOEYE-1 multispectral sensor with a resolution of 1.65 meters) to kilometers (e.g. Advanced Very High Resolution Radiometer sensor with a resolution of roughly one kilometer), a selection of an appropriate resolution for a land change study is important because spatial resolution may impact the understanding of the spatial patterns, and thus processes, of a landscape under investigation (Turner et al. 1989). Indeed, a landscape may be homogeneous at one resolution but may become heterogeneous at another resolution (Walsh et al. 1999; Wu 2004).

In addition to the impact on spatial patterns, another impact of spatial resolution is on the accuracy of a classification performed to thematically map land cover types. Traditional image classification (or per-pixel classification) is a process of assigning categorical values (e.g. water, artificial surface, barren, forest, or grass) to pixels based

on their values of the intensity of reflected electromagnetic radiation. The accuracy of this assignment is driven by the spatial resolution. Technically, the coarser the resolution is the lower the accuracy (Aplin 2006). The reason is that as the resolution becomes coarser, there are more land cover types coexisting within pixels. As a result, the pixel values are representative of multiple land cover types rather than one. Consequently, the assignment of land cover types to pixels is compromised due to the uncertainty of what land cover type assigned to each pixel. Therefore, an alternative to this per-pixel approach is the subpixel approach that assigns fractions of land cover types to pixels (Foody 2006). This subpixel classification has been applied successfully for a variety of sensors including Landsat and MODIS (Moderate Resolution Imaging Spectroradiometer) (Shao and Lunetta 2011; Small 2001; Weng and Lu 2009). However, the accuracy of subpixel classification is still not perfect (Foody and Doan 2007). Indeed, the accuracy of subpixel classification is driven by the level of land cover heterogeneity, which is positively related to the intra-class variation, and thus, the classification accuracy (Foody and Doan 2007; Ngigi et al. 2009).

Beside spatial pattern, land change studies have also been analyzing changes of land cover. Recent studies in this area have gone beyond detecting changes between two time points to identify temporal trends and times of changes. In the latter case, while the temporal trends help to reveal the trajectories of changes, the times of changes help to answer questions regarding land cover compositions and configurations before and after the changes (Gillanders et al. 2008; Powell et al. 2010). To identify temporal trends and times of changes, a time-series of images needs to be used. The temporal resolution of

the time-series is important because it affects the results of trend analyses as well as time detections.

Temporal resolution of a time-series refers to the time interval between any two consecutive images in the time-series. Theoretically, temporal resolution of a time-series is equivalent to the revisit time of a sensor, which refers to the time period between repeat passes of the sensor over the same place on the Earth's surface. However, the temporal resolution is often longer than the revisit time due to cloud contamination or to satellite malfunction (e.g. the scan-line-corrector problem of Landsat 7) (Ju and Roy 2008; Turner et al. 2007). As the temporal resolution becomes longer, it reduces the opportunity to detect land cover change rate and timing (Lunetta et al. 2004). However, high temporal resolution data are often compromised by low spatial resolution (Townshend and Justice 1988). For instance, a time-series of MODIS has high temporal resolution (two days) but low spatial resolution (500 meters). As the spatial resolution becomes coarser, the time-series may not be able to effectively detect land cover changes of a heterogeneous landscape because there may be multiple changing processes (e.g. no change and clearcut harvest) occurring within the same pixels (Rindfuss et al. 2004; Townshend and Justice 1988). Therefore, since the last decade, there have been a few data fusion methods proposed to balance the advantages and disadvantages of high spatial resolution (e.g. Landsat) and high temporal resolution (e.g. MODIS) data by blending them together (Gao et al. 2006; Hilker et al. 2009; Zhu et al. 2010). For instance, the Spatial Temporal Adaptive Algorithm for mapping Reflectance Change (STAARCH) has been proposed since 2009 by (Hilker et al. 2009). The purpose of this algorithm is to blend Landsat and MODIS imagery together to

create a high spatial and temporal resolution time-series to detect forest disturbances. The algorithm requires a reference set representative of undisturbed forests against which other pixels are compared to determine their disturbance conditions (i.e. disturbed or undisturbed). However, testing and modifying this STAARCH algorithm and other algorithms for a variety of landscapes, especially those that are heterogeneous, are still in progress (eg. Emelyanova et al. 2013).

In general, both the spatial and temporal resolutions affect the analyses of land cover change. While the spatial resolution affects the accuracy of allocating land cover proportions to pixels during a classification process, the temporal resolution affects the accuracy of detecting the trend and time of changes based on a high temporal time-series of images. In either case, land cover heterogeneity plays an important role in diminishing those accuracies. It is because coarse spatial resolution coupled with high land cover heterogeneity results in high intra-class variation that may reduce the accuracies of both per-pixel and subpixel classification. Furthermore, the couple of coarse spatial resolution and high land cover heterogeneity also affects the performance of a data fusion method (e.g. the STAARCH) used to create the time-series for detecting the trend and time of changes. Therefore, the purpose of this dissertation is to explore and document the impact of land cover heterogeneity on both per-pixel and subpixel classification at a medium spatial resolution of Landsat (30 meters) as well as to test the STAARCH method and based on which to propose another data fusion method that accounts for land cover heterogeneity.

1.1 RESEARCH OBJECTIVES

The overall research goal of this dissertation is to document the impact of land cover heterogeneity on both per-pixel and subpixel classification at a medium spatial resolution of Landsat (30 meters) as well as to propose a data fusion method that accounts for land cover heterogeneity in the South-Central US. In order to pursue this study goal, three specific research objectives were formulated:

1. To explore the impact of land cover heterogeneity on image classification approaches including per-pixel and subpixel classifications with a case study area of 10,000 square kilometers centered at Little Rock (Arkansas);
2. To propose a data fusion method that is able to produce high spatial (30 meters) and temporal resolution (8 days) time-series for a disturbance detection that takes into account the heterogeneity of a forest landscape in southeast Oklahoma; and,
3. To determine whether the temporal characteristics of disturbances are sufficient to distinguish disturbance types (e.g. fire and harvest) across the study area of southeast Oklahoma within the study period from 7/19/2000 to 8/5/2011.

The two study areas (i.e. Little Rock and southeast Oklahoma) were selected because of their land cover heterogeneity. Little Rock was selected as a case study for the first study objective because there were a variety of land cover types (e.g. cropland, artificial surface, barren, tree, grassland/shrub, and water) in this study area. Southeast Oklahoma was selected as a typical study area to analyze forest disturbances for the second and third study objectives because it was dominated by forests (both deciduous

and evergreen). Southeast Oklahoma was also selected as a typical study area in other forest studies (Masek et al. 2008; Schleeweis et al. 2013).

1.2 ORGANIZATION OF DISSERTATION

This first chapter briefly introduces the effect of spatial and temporal resolutions on studies of land cover change and highlights research needs. It also gives an overview of the research objectives to be addressed in this dissertation. The pursuit of these objectives is presented in three self-contained research manuscripts. Chapter two investigates the impact of land cover heterogeneity on both per-pixel and subpixel classifications (objective one). Using statistical tests, this research documents the dependence of both classifications on land cover heterogeneity for each of the six land cover classes (cropland, artificial surface, barren land, tree, grassland/shrub, and water). This chapter will be submitted to the *International Journal of Remote Sensing*. Chapter three proposes a data fusion method based on an existing fusion algorithm (i.e., the STAARCH) to better monitor forest disturbances in a heterogeneous forest landscape (objective two). This chapter will be submitted to the *Remote Sensing of Environment* journal. Chapter four utilizes the disturbance map created in Chapter three to determine whether the temporal characteristics of disturbances are sufficient to classify disturbance types. This research classifies disturbance types by using a rule set of temporal characteristics of disturbances. Classification errors and potential solutions are also discussed in this chapter. This chapter will be submitted to the *Forest Ecology and Management* journal. Chapter five summarizes the conclusions of this dissertation and suggests potential directions for future research.

CHAPTER 2. LAND COVER HETEROGENEITY EFFECT ON SUBPIXEL CLASSIFICATION

2.1 ABSTRACT

Land cover maps play an important role in identifying the spatial and temporal patterns of land cover types to facilitate the understanding about the coupled human-environment system. With its capability of systematically capturing the Earth's surface over a large area, remote sensing is being used widely in land cover mapping. However, due to land cover heterogeneity, the accuracy of land cover maps created by classifying remotely sensed imagery data may not be high. Compared to the traditional per-pixel classification, the subpixel classification is expected to better mitigate the effect of land cover heterogeneity. However, in practice, the accuracy of subpixel classification is still not so high. Therefore, the objective of this chapter is to analyze the impact of land cover heterogeneity on a subpixel classification as well as a per-pixel classification for a highly heterogeneous region in central Arkansas, United States. The results demonstrate that both per-pixel and subpixel classifications successfully classify land cover types in the study area with high accuracy (81.87 percent for per-pixel and 82.28 percent for subpixel). Additionally, this study shows that land cover heterogeneity not only negatively affect the accuracy of the per-pixel classification but also the accuracy of the subpixel classification. Furthermore, this study points out that subpixel classification is not necessary always better than per-pixel classification. The use of subpixel classification is dependent on the degree of land cover heterogeneity of a study area and on the imagery data used because land cover may be heterogeneous at one resolution but may become homogeneous at another resolution.

2.2 INTRODUCTION

Land cover maps derived from remotely sensed images have been playing a key role in studies of the coupled human-environment system (Turner et al. 2007). Consequently, concern about the accuracy of these maps has grown. If accuracy refers to "the degree of 'correctness' of a map" (Foody 2002), accuracy assessment is a process of quantifying the degree to which the derived map agrees with reality or conforms to the 'truth' (Foody 2002). Currently, the confusion matrix is a key means for accuracy assessment because it quantifies not only the overall accuracy but also the errors of omission and commission associated with individual map class (Congalton 1994; Foody 2002). However, the confusion matrix provides no information about the spatial distribution of errors and thus, is inappropriate to validate the accuracy for sub-regions where local error rates may be much larger or smaller than the global measures performed on an entire dataset (McGwire and Fisher 2001).

Since the last decade, there has been a call to move beyond the confusion matrix to include the spatial pattern of classification errors when documenting the accuracy of land cover maps (Comber et al. 2012; Foody 2005; McGwire and Fisher 2001). Understanding the spatial variation of these errors helps scientists to identify whether regions of interest have sufficient accuracy (McGwire and Fisher 2001) or to pin point regions of low accuracy for further classification enhancement procedure (Foody 2005; Hubert-Moy et al. 2001). In order to reveal the spatial pattern of these errors, it is necessary to recognize their sources.

Besides the sensor's characteristics (e.g., instantaneous field of view, spectral and temporal response functions), the atmosphere (e.g., the scattering, absorption, and

emission that occur in the atmosphere between the radiation source and scene, and the scene and sensor), landscape characteristics (i.e. the spatial arrangement and properties of the covers) are also source of errors (Strahler et al. 1986). Elevation, for example, has been found to have an impact on classification accuracy in either positive or negative ways. In the former case, vegetated regions at higher elevation may have higher classification accuracy because phenology at higher elevation may be more homogeneous (Yu et al. 2008). In the latter case, change in elevation causes variation in brightness values between a horizontal surface and a sloped surface of the same cover class and therefore reduces classification accuracy (Fahsi et al. 2000).

Patch size and land cover heterogeneity (i.e. the number of classes found within a region) are also found to have an impact on classification accuracy. The larger the patch is the higher the accuracy and the higher the heterogeneity is the lower the accuracy (Lechner et al. 2009; Smith et al. 2003). For example, the probability of correct classification for a Landsat pixel will be greater than 0.5 if it is contained in a patch of 56x56 pixels or if the heterogeneity equals one (Smith et al. 2002). Depending on the image resolution and classification method used, the mechanism that patch size and heterogeneity influence the accuracy may be different. In per-pixel classification, when the image resolution is fine, heterogeneity results in high intra-class variation and thus, classification accuracy may be reduced (Aplin 2006). When the image resolution is coarse, heterogeneity results in mixed pixels representing areas comprising a mixture of more than one land cover type or areas at the boundaries of two or more land cover types and thus, classification accuracy is also reduced (Aplin 2006). For example, deforestation may be overestimated 50 percent if an one-kilometer-resolution imagery is

used (Skole and Compton 1993). However, the mixed-pixel problem may be resolved by applying subpixel classification to unmix the pixels to land cover proportions (Foody and Doan 2007; Lo and Choi 2004; Weng and Lu 2009). Although subpixel classification has potential in a variety of applications (Cross et al. 1991; Shao and Lunetta 2011; Weng et al. 2011), its accuracy is still affected by land cover heterogeneity with the mechanism similar to per-pixel classification applied to high resolution imagery. In other words, the increase of heterogeneity results in the increase of intra-class variation and thus, reduces subpixel classification accuracy (Foody and Doan 2007; Ngigi et al. 2009). Neither per-pixel nor subpixel classification is, therefore, appropriate for heterogeneous landscape (Mather 1999; Weng and Lu 2009).

Statistical analysis has been done to quantitatively examine the significance of the impact of land cover heterogeneity on per-pixel classification (Smith et al. 2003; Smith et al. 2002; van Oort et al. 2004) but not on subpixel classification. Therefore, the objective of this study is to systematically analyze the impact of land cover heterogeneity on subpixel classification as well as per-pixel classification for a highly heterogeneous region in central Arkansas, United States.

2.3 STUDY AREA

This study assesses a 10,000-square-kilometer area in central Arkansas (USA), centered around the capital of Little Rock (Figure 2.1). This area is selected for its heterogeneity in physiography and land cover. Little Rock is situated at the intersection of four Omernik Level III ecoregions (Omernik 1987). The Arkansas Valley ecoregion north of Little Rock is characterized by forested hills (31 percent forest in 2006) that bound large valleys covered in a mixture of agricultural activities (9 percent cropland).

The Mississippi Alluvial Plains to the east is a relatively flat ecoregion historically covered by forested wetlands and several large grasslands, but is now agriculturally dominated (54 percent cropland). South of Little Rock lays the South Central Plains ecoregion, composed of rolling forested plains (54 percent forest) with many small patches of urban (14 percent), agriculture (0.4 percent), and barren (0.3 percent) lands. The Ouachita Mountains ecoregion to the west is mostly forested (68 percent), with steep slopes along east-west trending ridges. Commercial logging is the major land use in these latter two ecoregions. More details of land cover composition within this study area, as well as temporal changes and their drivers, can be found in a study by Jawarneh and Julian (2012). In summary, land cover is most heterogeneous in the Arkansas Valley, followed by the South Coastal Plains, the Mississippi Alluvial Plain, and lastly the Ouachita Mountains.

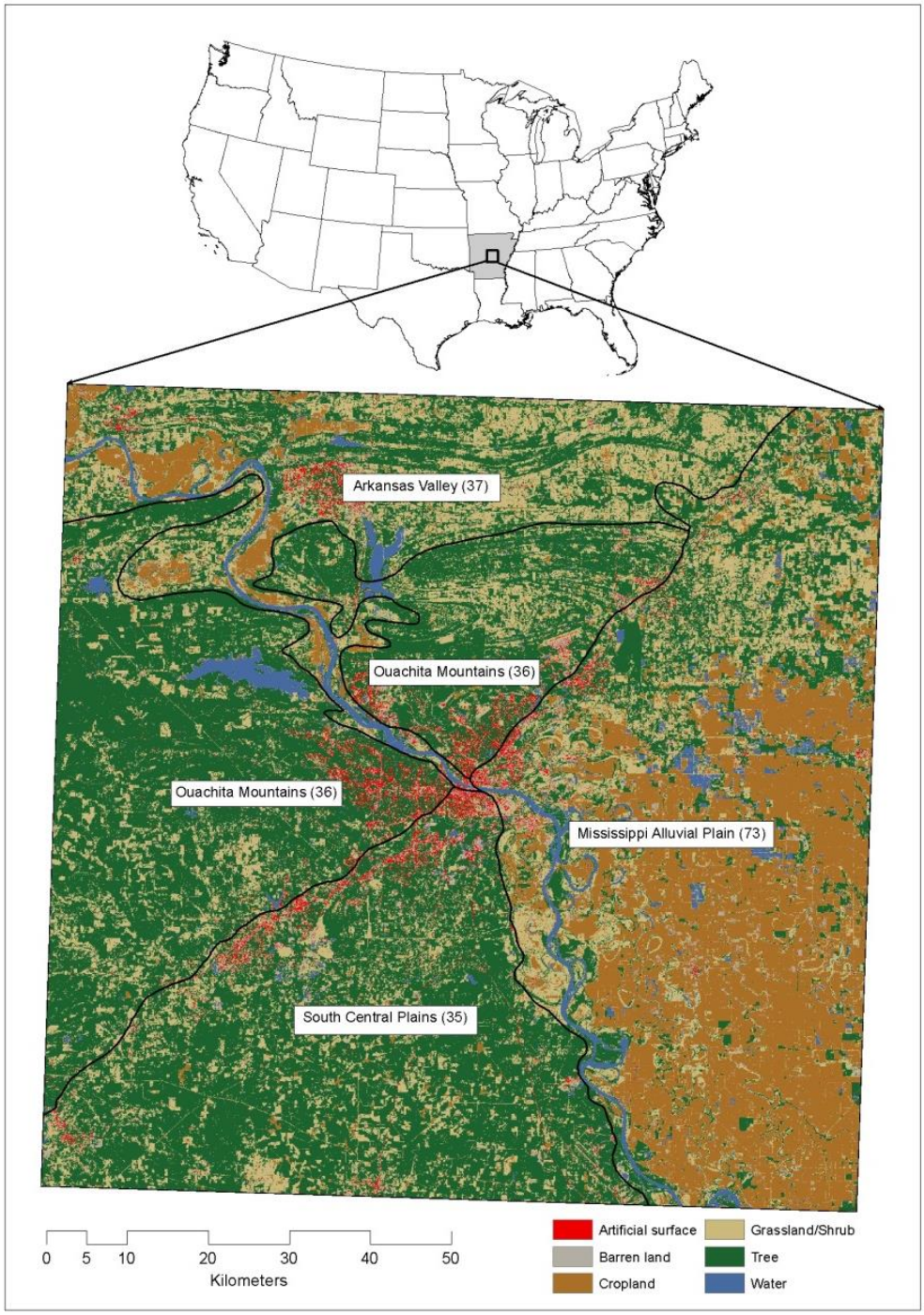


Figure 2.1. The study area centered around Little Rock, Arkansas, USA. Base layer is the land cover map (2010; 30-meter resolution) generated by the per-pixel classification in this study. The four Omernik level III ecoregions are outlined in black lines.

2.4 DATA

To analyze the impact of land cover heterogeneity on per-pixel classification and subpixel classification, Landsat imagery data were used as a primary input for the classification of land cover types in the study area. Land cover types were classified by a per-pixel classification using the supervised Maximum Likelihood Classification (MLC; Richards and Jia 1999) and by a subpixel classification using the supervised Fuzzy Maximum Likelihood Classification (FMLC; Wang 1990). Training and validation data for these classifications were developed based on the National Land Cover Database 2006 (NLCD 2006; Fry et al. 2011) and the National Agriculture Imagery Program (NAIP) aerial photos (U.S. Department of Agriculture 2010).

Landsat 5 TM images (path 24, row 36) at level 1T with six bands (excluding thermal band) were downloaded from USGS Earth Explorer. To reduce confusion between cropland and other land covers (e.g. barren and grassland/shrub), this study used images acquired on 12 April 2010, 19 September 2010, and 6 November 2010. These images were cloud-free and closest to the acquisition date of the NAIP photos. The Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) (ENVI 2009) was used to convert these Landsat images from top-of-atmosphere radiance to surface reflectance. Ultimately, all images were layer-stacked and subset to the predefined 100x100 kilometer study area.

The training and validation data were developed using the stratified random sampling technique. There were seven land cover strata delineated based on the NLCD 2006 downloaded from the web interface of the Multi-Resolution Land Characteristics Consortium (www.mrlc.gov) (Fry et al. 2011). This NLCD 2006 after downloaded was

reprojected to the Universal Transverse Mercator (UTM) zone 15N projection and clipped to the extent of the study area. Although the NLCD 2006 had 29 classes covering water, forest, shrub, herbaceous, and wetland (Fry et al. 2011), they were grouped to seven classes: water (NLCD class 11), artificial surface (21, 22, 23, 24), barren land (31), tree (41, 42, 43), grassland/shrub (52, 71, 81), cropland (82), and wetland (90, 95). The areas delineated by these classes then became land cover strata for the selection of training and validation samples.

As soon as the land cover strata were delineated, training and validation samples were randomly selected within these strata. For each sample, its land cover proportions were determined by visually interpreting the 2009 NAIP aerial photos obtained from the Arkansas' Official GIS Platform GeoStor. These one-meter-resolution natural-color photos were acquired during the agricultural growing seasons and were administered by the U.S. Department of Agriculture's Farm Service Agency through the Aerial Photography Field Office in Salt Lake City. NAIP photos were available as orthorectified photos with a reported horizontal accuracy of six meters (Adkins 2009; U.S. Department of Agriculture 2010).

2.5 METHODOLOGY

2.5.1 Classification

Training samples were suggested to be homogeneous (i.e. samples composing of only one land cover type) to make sure that the histogram of radiance data of a class was unimodal to facilitate the calculation of statistical measures (e.g. mean and covariance matrix) (Campbell and Wynne 2011). However, in a heterogeneous landscape, selecting homogeneous samples could be problematic because it was

difficult to have enough homogeneous samples located in all parts of the study area. Moreover, as the landscape became more heterogeneous with more mixed pixels, the radiance data extracted from homogeneous samples for a class might not be representative of other mixed pixels for the same class. It was because pixel radiance was a function of radiance from all classes within the pixel and its neighboring pixels (Wang 1990). Therefore, heterogeneous samples (i.e. samples composing of more than one land cover types) were used for subpixel classification (Foody 1999; Wang 1990). When used for per-pixel classification, these heterogeneous samples were hardened using the dominant rule. This study used 225 training samples per sampling class (i.e. a total of 1,575 samples). The size of training samples was one Landsat pixel. For each sample, its land cover proportions were estimated by visually interpreting the NAIP photos at the scale of 1:2,000.

As soon as the training samples were developed, six land cover types defined in Table 2.1 were identified from the Landsat imagery data by a per-pixel classification using the supervised Maximum Likelihood Classification (MLC; Richards and Jia 1999) and by a subpixel classification using the supervised Fuzzy Maximum Likelihood Classification (FMLC; Wang 1990). While the output of the Maximum Likelihood Classification was a categorical map with six classes color coded, the output of the Fuzzy Maximum Likelihood Classification was a fractional map that was actually a set of six proportional images, each of which represented proportions of a land cover type.

Table 2.1. Land cover classes used in the classifications.

Class	Definition
Cropland (CR)	Areas used for the production of crops such as corn, soybeans, vegetables, tobacco, and cotton. This class also includes fallow cropland.
Artificial surface (AR)	Construction materials such as asphalt, concrete, and rooftops.
Barren (BA)	Areas of bedrock, bare soil, quarries, and any accumulation of earthen material.
Tree (TR)	All trees over 5 m, including low-density trees in urban areas.
Grassland/Shrub (GR)	Areas with > 80% coverage of gramanoid or herbaceous vegetation; or areas with > 20% coverage of shrubs less than 5 m high
Water (WA)	Areas of open water with < 25% coverage of any other class.

2.5.2 Validation

This study used 100 validation samples per sampling class (i.e. a total of 700 samples). While the size of the training samples was one Landsat pixel, the size of validation samples was a block of 3x3 pixels to reduce the effect of misregistration between the referenced (i.e. NAIP photos) and the Landsat data (Powell et al. 2007; Song 2005). For each sample, its reference land cover proportions were estimated by visually interpreting the NAIP photos at the scale of 1:2,000. The cross tabulation matrix (Pontius and Cheuk 2006) proposed as a customization of the conventional confusion matrix (Congalton 1991) for a subpixel classification was used to validate the performances of the two classifications. An individual cross tabulation matrix was developed for each of the validation samples. For a given validation sample, while entries on the diagonal of the matrix were the overlaps of the reference proportions and the estimated proportions derived from a classification, entries off the diagonal were the non-overlaps of the reference proportions and the estimated proportions for every class.

The total of the diagonal entries was the overall accuracy whereas the ratios between each diagonal entry and its column total and row were the producer's and user's accuracies, respectively, of a given class for the given validation sample. To calculate the overall accuracy as well as the producer's and user's accuracies for the entire study area, a cross tabulation matrix was calculated for an *average sample*. The average sample was the sample whose reference proportion of each class was the average of the reference proportions of that class across all validation samples. Similarly, the estimated proportion of each class of the average sample was the average of the estimated proportions of that class across all validation samples. In this study, the cross tabulation matrix for the average sample provided information about the overall accuracy and a general understanding about the misclassifications between classes. Furthermore, the individual matrices provided information about the producer's and user's accuracies for validation samples and thus facilitated the statistical analysis of the impact of heterogeneity on the classification accuracies of the two classifications. Beside the cross tabulation matrices, multiple scatter plots were used to investigate the agreements between the referenced and estimated proportions for each of the land cover classes.

2.5.3 Statistical analysis

Statistical analysis was conducted to test the impact of heterogeneity on the classification accuracies of the two classifications. Samples used in this statistical analysis were the 700 validation samples. The overall accuracy as well as the producer's and user's accuracies of the validation samples were obtained from the individual cross tabulation matrices discussed above.

Two statistical tests were performed. First, the Wilcoxon Signed-Rank test was conducted to test whether the differences in the producer's and user's accuracies between the two classifications were significant. The null hypothesis for this test was that in the same area and using the same samples, there was no significant difference in the accuracies between per-pixel and subpixel classifications. Second, the Steel-Dwass test was conducted to test the impact of heterogeneity on the overall accuracy as well as the producer's and user's accuracies of both classifications. The null hypothesis of this test was that the classification accuracies of either the per-pixel or subpixel classification were not significantly different between groups of land cover heterogeneity. Land cover heterogeneity referred to the number of land cover types coexisting within samples. For instance, a sample occupied by two land cover types had its land cover heterogeneity of two. Samples having the same heterogeneity ranging from one to six were grouped into six groups (from one to six) of land cover heterogeneity. However, because there were too few (less than five) samples in group five and group six, these groups were excluded from the statistical analysis.

2.6 RESULTS AND DISCUSSION

2.6.1 Land cover map and accuracy

The results of per-pixel and subpixel classifications were presented as a categorical map and a fractional map in Figure 2.1 and 2.2, respectively. Visually, the overall distributions of land cover types obtained from these two maps were similar. The two maps demonstrated that while cropland was dominant in the Mississippi Alluvial Plain ecoregion, forest was dominant in both the Ouachita Mountains and South Central Plains ecoregions. Cropland was also found in the Arkansas Valley along

the Arkansas River. Beside cropland, the Arkansas Valley ecoregion was a mixture of forest and grassland/shrub. Urban areas characterized by artificial surfaces were mostly found at the intersection of the four ecoregions and along Interstate 30 running southwest-northeast. This land cover pattern was also identified and analyzed in another study (Jawarneh and Julian 2012). In this study chapter, it was found that both per-pixel and subpixel classification worked well on this particular study area. Typically, the overall accuracies of the two classifications were higher than 80 percent (81.87 percent for per-pixel and 82.28 percent for subpixel; Table 2). Although these overall accuracies were slightly lower than the common threshold of 85 percent (Thomlinson et al. 1999), they were high and acceptable because: (1) they were in a range of the accuracies published during the last decade (Wilkinson 2005), (2) they exceeded the level of 75 percent suggested by Goodchild *et al*. (1994), and (3) they were higher than the regional accuracy of NLCD 2001 – level I (79 percent for region 7; Wickham et al. 2010).

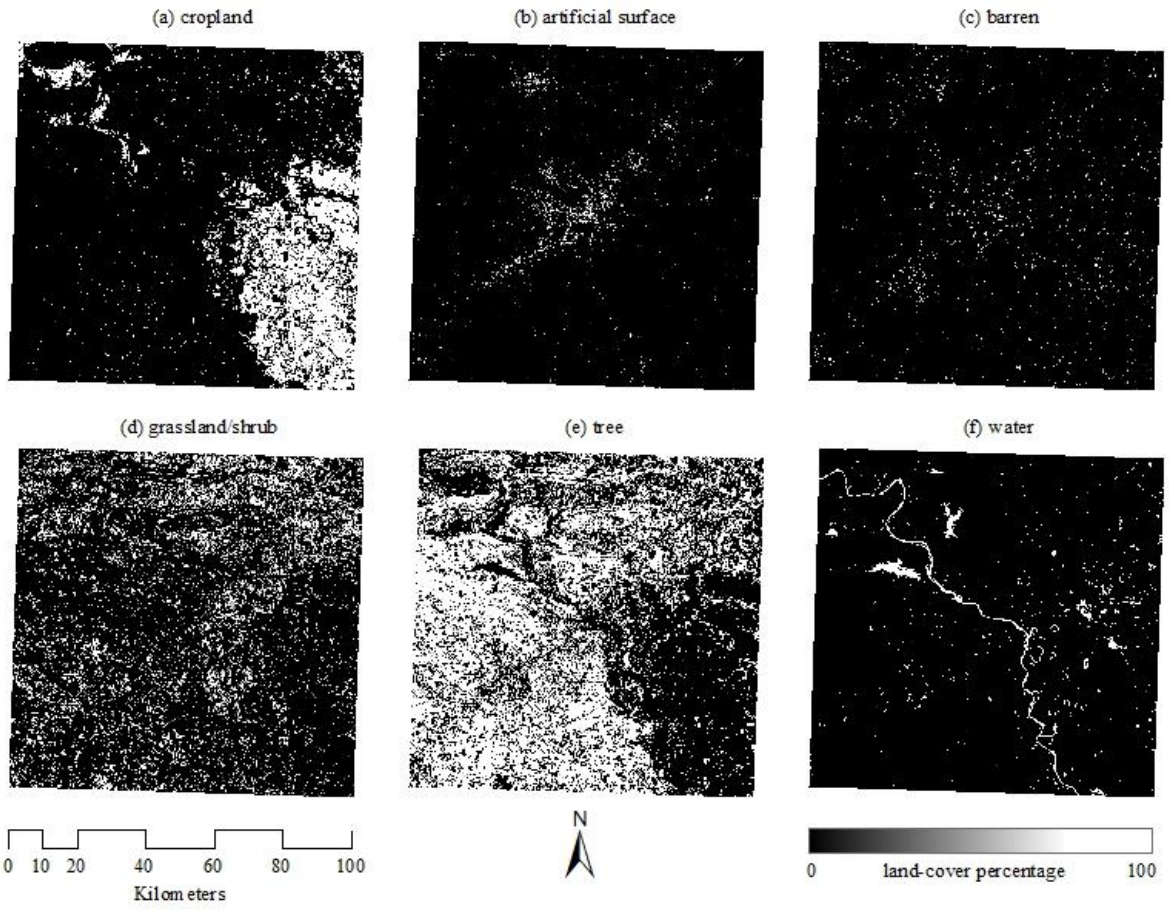


Figure 2.2. Land cover percentages (2010; 30-meter resolution) generated by the subpixel classification.

Among all classes, cropland was found to have highest producer's and user's accuracies (Figure 2.3). This supported the advantage of using multi-date Landsat data to identify cropland. However, the misclassification between fallow cropland and barren was still not completely avoided. Additionally, barren and artificial surface were easily confused with each other (Table 2.2). In fact, in the case of per-pixel classification, 27 percent of artificial surface was misclassified as barren and 9 percent of barren was misclassified as artificial surface. In the case of subpixel classification, 21 percent of artificial surface was misclassified as barren and 14 percent of barren was misclassified as artificial surface.

Table 2.2. The cross tabulation matrix of the average sample for the two classifications: per-pixel (normal text) and subpixel (bold italicized text). Values in this matrix were calculated for the average sample, of which the class proportions were averages of the class proportions across all 700 validation samples.

		Referenced percentage						
		Cropland	Artificial	Barren	Grassland / Shrub	Tree	Water	**Total**
Estimated percentage	Cropland	12.65	0.05	0.78	0.91	0.44	0.18	15.00
		12.40	***0.04***	***0.73***	***0.88***	***0.38***	***0.08***	***14.51***
	Artificial	0.01	4.87	0.78	0.62	0.37	0.04	6.68
		0.06	***5.87***	***1.15***	***1.09***	***0.51***	***0.07***	***8.75***
	Barren	0.38	2.33	6.01	1.17	0.61	0.31	10.81
		0.50	***1.80***	***5.78***	***0.90***	***0.58***	***0.33***	***9.88***
	Grassland / Shrub	1.61	0.65	0.37	17.93	1.58	0.19	22.33
		1.67	***0.25***	***0.22***	***17.56***	***1.35***	***0.21***	***21.24***
	Tree	0.09	0.59	0.25	2.94	29.11	0.24	33.22
		0.11	***0.52***	***0.29***	***3.13***	***29.29***	***0.21***	***33.54***
	Water	0.04	0.00	0.13	0.28	0.21	11.30	11.95
		0.04	***0.01***	***0.15***	***0.28***	***0.22***	***11.37***	***12.07***
	Total	14.78	8.48	8.32	23.84	32.32	12.26	100.00
		14.78	***8.48***	***8.32***	***23.84***	***32.32***	***12.26***	***100.00***
	Overall accuracy (%)	81.87		Kappa (%)	76.99			
		82.28			***77.54***			

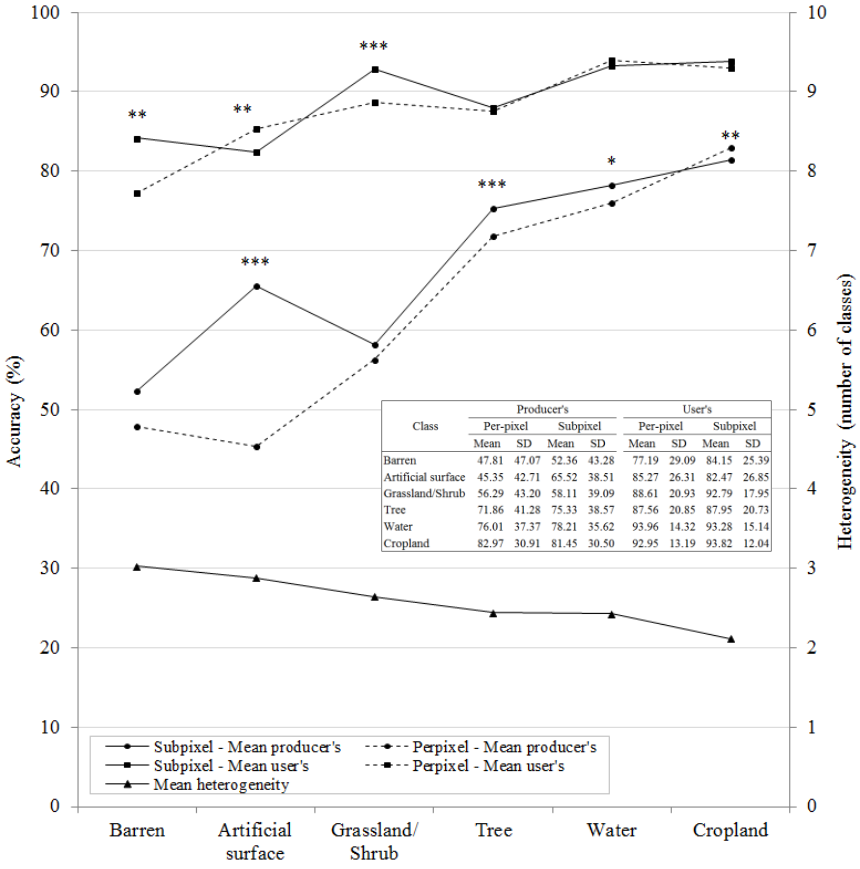
An exploration of these classification errors pointed out that the misclassifications between artificial surface and barren were mostly found for rooftops and parking lots in the Central Business District. These misclassifications were, therefore, due to two reasons. First, this study did not discriminate high-albedo (e.g., rooftop) from low-albedo (e.g., asphalt, roads) in the artificial surface class. Also, this study considered sand, dry soil, and gravel all belong to the barren class. Second, high-albedo artificial surfaces have been found to be spectrally confused with sand and dry soil (Hu and Weng 2011; Weng et al. 2009; Wu and Murray 2003) whereas low-albedo artificial surfaces have been found to be confused with gravel (Stefanov et al. 2001).

Additionally, artificial surface was also misclassified with grassland/shrub (nearly eight percent for per-pixel and three percent for subpixel) and tree (seven percent for per-pixel and six percent for subpixel). The reason was due to the spectral similarity in the near-infrared band between high-albedo artificial surface and vegetation (Wu and Murray 2003). Furthermore, barren was also found to be commonly misclassified with cropland, grassland/shrub, and tree. These misclassifications took place along class boundaries (e.g. narrow and/or linear farm paths, forest paths, or dirt shoulders along tree-lined roads). These misclassifications implied that not only was per-pixel classification not effective in identifying narrow and/or linear features (Lechner et al. 2009) but subpixel classification was not very improved as well. Furthermore, the misclassifications between tree and grassland/shrub were commonly found in low-intensity residence, forest regrown areas, and sparse forests. These misclassifications were attributed to the incapability of the classifications to identify narrow and/or linear features. Besides, these misclassifications were also due to the spectral confusion between them (Lu and Weng 2004).

Among all classes, barren and artificial surface were the two classes having lowest producer's and user's accuracies, except the producer's accuracy of subpixel classification for artificial surface (Figure 2.3). This was an expected result because on one hand, compared to other classes, pixels of barren and artificial surface were more heterogeneous (Figure 2.3). On the other hand, the accuracy of per-pixel classification was expected to be reduced as pixels became more heterogeneous (Smith et al. 2003; Smith et al. 2002).

Multiple Wilcoxon Signed-Rank tests were conducted to test the differences between the producer's accuracies of per-pixel and subpixel classification as well as the differences between the user's accuracies of per-pixel and subpixel classifications for each land cover class (Figure 2.3). The results demonstrated that the producer's accuracy of subpixel classification was significantly higher ($p<0.05$) than that of per-pixel classification for artificial surface, tree, and water classes (Figure 2.3). Notably, while the difference between the producer's accuracy of subpixel classification and producer's accuracy of per-pixel classification for most classes was small (less than five percent), such difference for artificial surface was over 20 percent for which the producer's accuracy of subpixel classification was higher (Figure 2.3). This finding was expected given that subpixel classification has been claimed to be more advantageous than per-pixel classification because subpixel classification relaxed the assumption that pixels were homogenous (Foody 2006; Foody and Cox 1994). In addition, many studies, especially urban studies characterizing artificial surface by using remote sensing data, have been successful in using subpixel classification to develop fractional maps of land cover types (Nichol et al. 2010; Shao and Lunetta 2011; Weng 2012). However, the conclusion that subpixel classification was more advantageous than per-pixel classification for artificial surface in this particular study area could be questionable because the user's accuracy of the subpixel classification for this class was significantly ($p<0.01$) lower than that of the per-pixel classification although the difference of the user's accuracy (nearly three percent) was only one-to-seventh of the difference (about 20 percent) of the producer's accuracy (Figure 2.3). Furthermore, for other classes, the conclusion that either the subpixel or the per-pixel classification was

more advantageous could not be drawn because not both of the producer's and user's accuracies of either the subpixel or the per-pixel classification were significantly different (Figure 2.3).



Class	Producer's				User's			
	Per-pixel		Subpixel		Per-pixel		Subpixel	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Barren	47.81	47.07	52.36	43.28	77.19	29.09	84.15	25.39
Artificial surface	45.35	42.71	65.52	38.51	85.27	26.31	82.47	26.85
Grassland/Shrub	56.29	43.20	58.11	39.09	88.61	20.93	92.79	17.95
Tree	71.86	41.28	75.33	38.57	87.56	20.85	87.95	20.73
Water	76.01	37.37	78.21	35.62	93.96	14.32	93.28	15.14
Cropland	82.97	30.91	81.45	30.50	92.95	13.19	93.82	12.04

Figure 2.3. Mean producer's and user's accuracies of per-pixel and subpixel classifications together with mean heterogeneity by classes. The means of producer's and user's accuracies were calculated as the averages of all producer's and user's accuracies taken from the individual cross tabulation matrices developed for all validation samplesThe asterisks indicate * p<0.05, ** p<0.01, and *** p<0.001 in Wilcoxon Signed-Rank test.

To further assess the output of subpixel classification, multiple scatter plots representing the class agreements between the referenced and estimated proportions were used (Figure 2.4). In these plots, the high diffusion of land cover proportions around the $y=x$ line of most classes, except water, implied that there were high intra-class variations of proportions and thus surface reflectance of these classes. These intra-

class variations were expected to affect the classification accuracies of the subpixel classification as mentioned by Foody and Doan (2007).

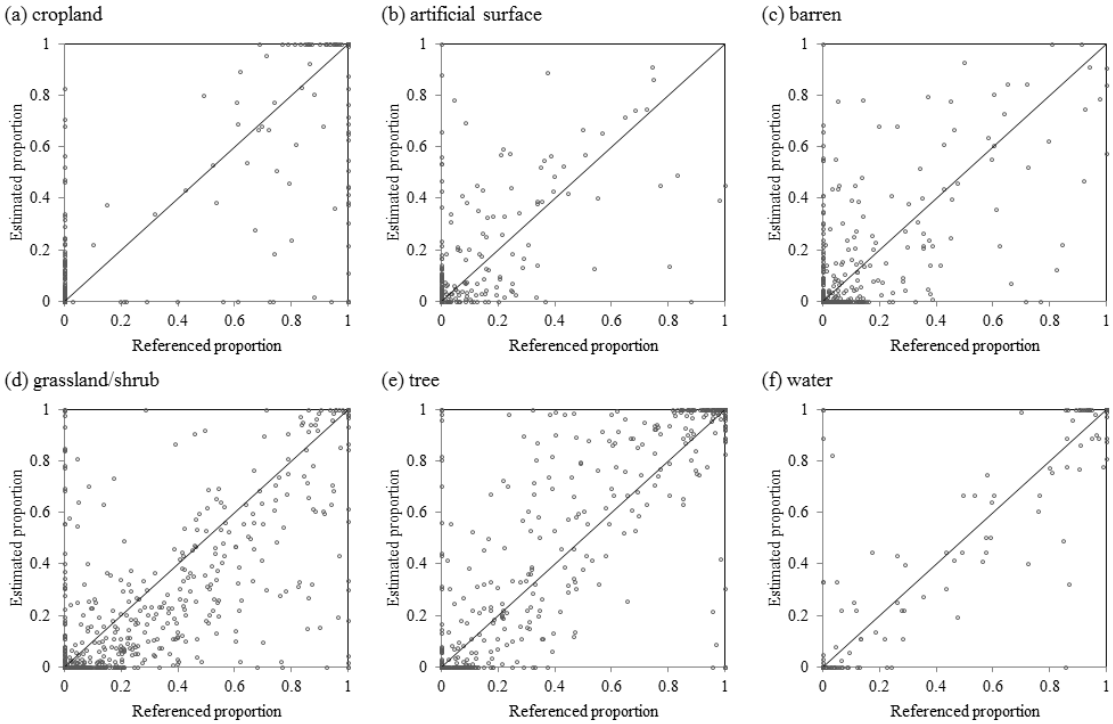


Figure 2.4. Scatterplots of estimated versus referenced land cover proportions for the six classes. The solid line is the *y=x* line.

2.6.2 Impact of heterogeneity on classification accuracies

Among the six classes, barren pixels were most heterogeneous with an average heterogeneity of 3.02 (Figure 2.3). In contrast, cropland pixels were least heterogeneous with an average heterogeneity of 2.1. Given that barren and artificial surface were mostly found in urban areas while grassland/shrub and cropland were found in rural areas, the trend of heterogeneity decreasing from barren to cropland implied that the urban landscape was much more heterogeneous than the rural landscape in this study area (Figure 2.1, 2.3). This trend of heterogeneity reflected the reversed trend of the producer's and user's accuracies of the per-pixel classification. In other words, as land

cover was less heterogeneous, the accuracies of per-pixel classification were higher. This result was similar to the findings from other scholars for per-pixel classification (Smith et al. 2003; Smith et al. 2002; van Oort et al. 2004). Interestingly, we found that this rule was also applicable to the subpixel classification, except artificial surface. For this class, although the accuracies of per-pixel classification were low, the accuracies of subpixel classification were much higher (e.g. 20 percent higher for the producer's accuracy; Figure 2.3). This could be the reason for many urban studies using subpixel classification to characterize the urban landscape dominated by artificial surfaces (e.g. Berezowski et al. 2012; Small 2001).

The impact of heterogeneity on classification accuracies was further analyzed by testing the differences in mean rank of the overall accuracy as well as the producer's and user's accuracies between groups of heterogeneity using the Steel-Dwass non-parametric test (Figure 2.5). The result demonstrated that while the accuracies of pixels having one class (i.e. heterogeneity being one) were the highest, these accuracies were reduced significantly when the pixels were occupied by more than one class. Besides, in most cases, the accuracies were not significantly different between the second group and groups of heterogeneity greater than two. These findings implied that land cover heterogeneity negatively influenced the accuracies of both the per-pixel and subpixel classifications. Subsequently, when there was more than one class coexisting in pixels, the classification accuracies were reduced. However, the accuracies of pixels having two or more classes were not significantly different.

In summary, with Landsat data, both per-pixel and subpixel classifications well performed on the study area centered at Little Rock with high accuracy. However,

apparent misclassifications were still identified between classes. One of the reasons of the misclassifications was due to the small size and linear shape of land cover features within pixels. Another reason was attributed to the heterogeneity of land cover. In concert with other studies, this study found that as the land cover became more heterogeneous, the classification accuracy was reduced significantly. The comparisons of classification accuracies between per-pixel and subpixel classifications pointed out that for this particular study area, it was not clear to conclude which classification was more advantageous. It was because while the producer's accuracies of the subpixel classification for some classes were higher than those of the per-pixel classification, the user's accuracies of the subpixel classification for those classes were not higher than those of the per-pixel classification. Artificial surface, however, was the most potential class that demonstrated the advantage of the subpixel classification over the per-pixel classification. It was because the producer's accuracy of the subpixel classification for artificial surface was twenty percent higher than that of the per-pixel classification. However, the user's accuracy of the subpixel classification for artificial surface was three percent lower than that of the per-pixel classification. As a result, the conclusion that the subpixel classification was better than per-pixel classification for artificial surface could be questionable. It was mentioned in other studies that neither per-pixel nor subpixel classifications would be good for a heterogeneous landscape (Mather 1999; Weng and Lu 2009). The reason was because as the landscape became heterogeneous, the intra-class variation of surface reflectance increased and thus reduced the classification accuracies of both per-pixel and subpixel classifications. Another reason particularly for this study could be due to the resolution of Landsat,

which was not coarse enough to weaken the performance of the per-pixel classification so that the contrast between the per-pixel and subpixel classifications would have been more apparent.

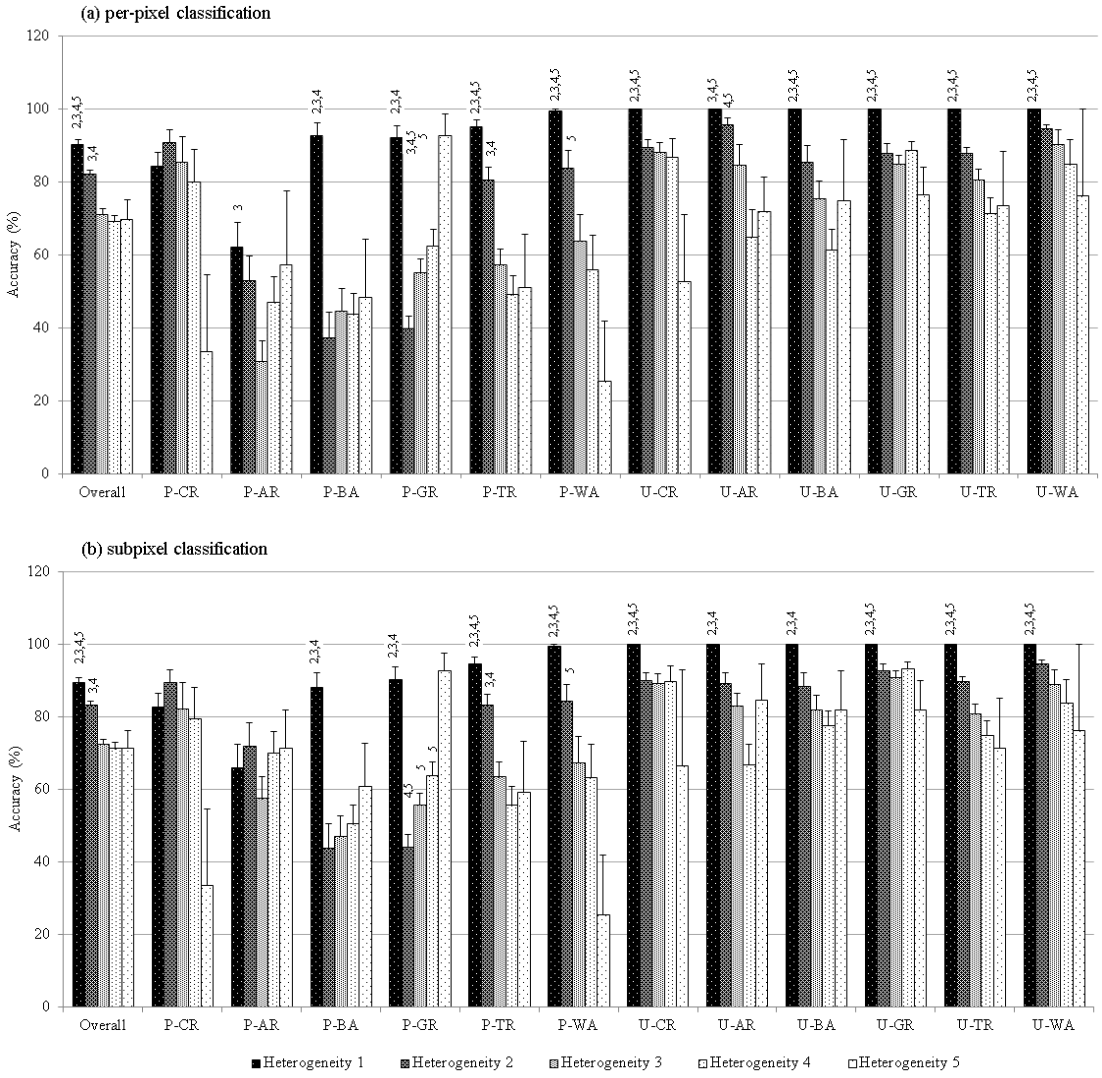


Figure 2.5. Mean overall, producer's and user's accuracies of per-pixel (a) and subpixel (b) classifications by groups of heterogeneity. Error bar indicates +1 standard error (s.t.e). Numbers at the top of bars of a group of heterogeneity indicate the groups of heterogeneity from which it was statistically different (p<0.05) according to the Steel-Dwass test. P-CR / U-CR = producer's / user's accuracy of cropland, artificial surface (AR), barren (BA), grassland/shrub (GR), tree (TR), and water (WA).

2.7 CONCLUSION

In this study, Landsat data was used to classify land cover types using a per-pixel classification and a subpixel classification. Although the classification accuracies were high (81.87 percent for the per-pixel and 82.28 percent for subpixel classifications) apparent misclassifications still existed. Land cover heterogeneity was found to have negative influence on the classification accuracies. Indeed, the classification accuracies decreased significantly when pixels were occupied by more than one class. In addition, the comparisons between the accuracies of the per-pixel classification and those of the subpixel classification for each class demonstrated that for this particular study area, it was not clear to conclude which classification was more advantageous. It could be because the land cover heterogeneity of this study area was high and thus resulted in the high intra-class variation that reduced the classification accuracies of both per-pixel and subpixel classifications. Another reason could be related to the resolution of Landsat data. With a resolution of 30 meters, Landsat data might not be coarse enough to produce enough mixed pixels for the subpixel classification to weaken the per-pixel classification. Findings from this study had several meanings. First, it systematically confirmed that the accuracies of both per-pixel and subpixel classifications were negatively influenced by land cover heterogeneity. Second, the use of per-pixel classification or subpixel classification was dependent on the study area and on the imagery data used because a study area could be heterogeneous at one resolution but could become homogeneous at another resolution.

CHAPTER 3. ACCOUNTING FOR LAND COVER HETEROGENEITY IN A SPATIAL TEMPORAL ADAPTIVE ALGORITHM FOR MAPPING REFLECTANCE CHANGE: A CASE STUDY IN SOUTHEAST OKLAHOMA 2000-2011

3.1 ABSTRACT

Monitoring forest disturbances using remote sensing data with high spatial and temporal resolutions is an important requirement for revealing the relationship between forest disturbances and the dynamics of terrestrial carbon stock as well as ecosystem dynamics. Landsat and MODIS are the two most common imagery data source for monitoring forest disturbances. Therefore, to have imagery data with high spatial and temporal resolutions, it is necessary to perform a data fusion procedure to blend a Landsat imagery time-series with a MODIS time-series. The Spatial Temporal Adaptive Algorithm for mapping Reflectance Change is one of the popular fusion methods. However, it does not explicitly account for land cover heterogeneity. In this study I not only test the STAARCH but also propose the MSTAARCH framework, which is a modification of the STAARCH, to account for land cover heterogeneity. The results show that the MSTAARCH produces a fusion imagery with much higher accuracy than the STAARCH.

3.2 INTRODUCTION

Forest disturbance has received significant attentions from scientists because of its impact on the source/sink dynamics of the aboveground terrestrial carbon (Williams et al. 2012) as well as ecosystem dynamics (Turner 2010). On the global scale, Hansen *et al*. (2010) reported that between 2000 and 2005 the global rate of gross forest cover

loss was 0.6 percent per year, which made up an area of 1,011,100 square kilometers, about 3.1 percent of the estimated forest area in 2000. Of the countries with more than a million square kilometers of forest cover, the United States had the highest rate of gross forest cover loss (1.2 percent per year) mostly due to fire in the western states and logging in the southeastern states, west coast, and in the upper Midwest (Hansen et al. 2010; Masek et al. 2008). However, the loss of US national forest cover has been surpassed by the gain from agricultural abandonment, fire suppression, and logging reduction (King et al. 2007). This may be the reason that the conterminous US forest net carbon uptake from the atmosphere increased 20 times in the last two decades (Williams et al. 2012).

At the regional and global scales, studies related to time since disturbance have been conducted based on a time-series of remotely sensed imagery data, such as Landsat or MODIS (Moderate Resolution Imaging Spectroradiometer) (Frolking et al. 2009). For instance, using a time-series of annual Landsat TM and ETM+ images, Masek *et al*. (2008) developed the wall-to-wall maps of extent and rate of stand-clearing disturbances (clear-cut harvest and fire) for the conterminous US and Canada between 1990 and 2000. Huang *et al*. (2010) proposed the Vegetation Change Tracker (VCT) procedure to map forest disturbances using a time-series of biennial Landsat images. The primary advantage of Landsat data was its spatial resolution (30 meters), which was good enough to differentiate features that might have different disturbance regimes (Townshend and Justice 1988; Turner 2010).

However, Landsat time-series might not be useful in the detection of time since disturbance because of its 16-day temporal resolution, which could be markedly

extended to years due to cloud contamination (Ju and Roy 2008). This has become a major concern in geographic areas that undergo rapid regeneration, such as those with high temperature and high levels of precipitation (Lunetta et al. 2004; Misson et al. 2005) or those where only partial harvesting occurred (Asner et al. 2004). It was found that the coarser the temporal resolution (e.g. greater than two years) was, the lower the accuracy of disturbance detection was (Lunetta et al. 2004; Masek et al. 2008). To overcome the limitation of temporal resolution, some studies tested the use of MODIS data because of its high temporal resolution (e.g. 8 days) (Mildrexler et al. 2009; Pape and Franklin 2008). However, due to its moderate spatial resolution (about 500 meters) MODIS might not be reliable to locate much land conversion and logging activity (Hansen et al. 2008; Pape and Franklin 2008).

Solutions have been devised to use both Landsat and MODIS data in detection of forest disturbances. One approach was to blend Landsat and MODIS data together to create a high spatial (30 meters) and temporal (8 days) data (Gao et al. 2006; Hilker et al. 2009). Gao *et al*. (2006) introduced the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) to predict surface reflectance of 30-meter pixels at a point in time (T_2) based on known reflectance of a reference Landsat imagery at another point in time (T_1) and known reflectance of reference MODIS images at both (T_1 and T_2) points in time using a spatial moving window technique. However, this approach suffered from a couple of disadvantages that could make it inappropriate for disturbance detection of a heterogeneous area.

One of the disadvantages was that STARFM assumed that MODIS pixels were relatively homogeneous and thus, their surface reflectance was proportional to surface

reflectance of corresponding Landsat pixels (Zhu et al. 2010). As such, changes in surface reflectance of MODIS pixels could be used to predict surface reflectance of corresponding Landsat pixels. This assumption of pixel homogeneity might not hold in such forest landscapes that often had mixtures of needle-leaf-evergreen (e.g. pine) and broad-leaf-deciduous (e.g. hickory) forests, which were different in spectral sensitivity (Asner 1998). The second disadvantage of STARFM was its choice of reference images (i.e. Landsat image at T_1 and MODIS images at T_1 and T_2) (Hilker et al. 2009). Originally, STARFM chose reference images independently from disturbances. As a result, reference images before disturbances might be used to predict surface reflectance values after disturbances. This method of prediction might not be effective if MODIS pixels were a mixture of forest types and disturbances happened to the types of which Landsat pixels were not representative.

In order to solve the first limitation of STARFM, Zhu et al. (2010) proposed the Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM) by introducing the spectral unmixing theory into the algorithm. A comparison between STARFM and ESTARFM, however, showed that ESTARFM was not always better than STARFM but rather dependent on the level of heterogeneity of a study area (Emelyanova et al. 2013).

Beside STARFM, the algorithm Spatial Temporal Adaptive Algorithm for mapping Reflectance Change (STAARCH) was introduced by Hilker et al. (2009). Unlike STARFM predicting surface reflectance, STAARCH predicts disturbance conditions (undisturbed or disturbed). To do so, STAARCH first determines disturbance conditions of Landsat and MODIS pixels independently and then perform a rule-based

fusion based on the disturbance conditions of Landsat and MODIS pixels to develop a binary fused time-series of disturbances. Disturbance conditions of Landsat and MODIS pixels are identified by a disturbance index (DI; Healey et al. 2005) measuring the spectral disturbance from the pixels to reference sets representative of unchanged and mature, and thus undisturbed forests.

Although STAARCH fusion framework was successfully used to detect forest disturbances for a boreal forest dominated by evergreen trees in west central Alberta, Canada (Hilker et al. 2009), its capability was limited due to the use of one single reference set to detect disturbance conditions of MODIS pixels. STAARCH, therefore, did not account for the fact that at the resolution of 500 meters, surface reflectance of MODIS pure pixels (i.e. pixels occupied by only one forest type) might be significantly different from surface reflectance of MODIS mixed pixels (i.e. pixels occupied by more than one forest types and/or other land cover types such as water, rangeland, and artificial surface). As a result, STAARCH might overestimate disturbance duration of broad-leaf-deciduous forests or underestimate disturbance duration of needle-leaf-evergreen forests because broad-leaf-deciduous forests have had higher spectral sensitivity, especially in infrared bands (Asner 1998).

The goal of this dissertation chapter was to propose an effective fusion framework to detect forest disturbances in southeast Oklahoma where (1) forests were highly mixed between deciduous and evergreen (Fry et al. 2011) and (2) forests have been disturbed due to both natural (e.g. fire) and anthropogenic (e.g. logging) forces (Harper and Johnson 2012; Ouachita Ecoregional Assessment Team 2003). To do so, I

modified STAARCH by proposing the MSTAARCH (modified STAARCH) framework to better develop reference sets for MODIS data.

3.3 STUDY AREA

The study area (approximately 185 x 185 km^2) covering primarily southeast Oklahoma and part of Arkansas and Texas is of a Landsat scene path 26 and row 36 (Figure 3.1). Most of this area (96 percent) is situated across three Level III ecoregions (Omernik, 1987): Arkansas Valley, Ouachita Mountains, and South Central Plains. The Arkansas Valley ecoregion is characterized by a mosaic of land cover types across plains, hills, floodplains, terraces and scattered mountains, with most of its forests occurring on steep slopes (Oklahoma Forestry Services 2010). Pasture and hay are found on gently sloping uplands while croplands mostly occur in bottomlands. The South Central Plains are dominated by agriculture (Figure 3.1), but have many large areas of medium-tall to tall mixed forests. Compared to the other two ecoregions, forests dominate the Ouachita Mountains (75 percent), which are characterized by steep slopes along east-west trending ridges. Over the entire study area, forests cover 56 percent of the land, of which 59 percent is deciduous, 30 percent evergreen, and 11 percent mixed forests (Figure 2.1; Fry et al. 2011). Forests in the study area are primarily used for logging and recreation (Oklahoma Forestry Services 2010).

3.4 DATA AND PRE-PROCESSING

3.4.1 Landsat data

Twenty-one cloud-free (cloud cover $\leq$ 5 percent) Landsat images (path 26, row 36) acquired in summer months (June-September; Table 3.1) were derived from the Landsat Climate Data Record (CDR) Surface Reflectance images downloaded from the

US Geological Survey Earth Explorer portal (http://earthexplorer.usgs.gov). These CDR images, projected to the Universal Transverse Mercator (UTM) zone 15N, included 17 bands, six of which represented surface reflectance of the six Landsat reflectance bands (excluding the thermal infrared) and two of which represented flags of clouds and cloud shadows (Masek et al. 2006). Surface reflectance data of the CDR images were produced by the 6S approach and the cloud and cloud shadow flags were produced by the Automatic Cloud-cover Assessment (ACCA) algorithm (Department of the Interior 2013; Masek et al. 2006). In this study, the cloud and cloud shadow flags were used as masks to remove cloud and cloud shadow pixels from the six reflectance bands. These cloud and cloud shadow pixels were then filled by using a temporally linear interpolation method suggested in (Huang et al. 2010). A Landsat time-series (i.e. a temporal sequence of Landsat images) was then created by layer stacking all twenty-one Landsat images in a temporal sequence manner.

Table 3.1. Landsat acquisition dates.

Sensor	Date of acquisition	Sensor	Date of acquisition
Landsat 5 TM	2000-07-19	Landsat 5 TM	2004-07-14
Landsat 7 ETM+	2000-08-12	Landsat 5 TM	2004-08-31
Landsat 5 TM	2000-08-20	Landsat 5 TM	2005-06-15
Landsat 7 ETM+	2000-08-28	Landsat 5 TM	2005-08-18
Landsat 7 ETM+	2000-09-29	Landsat 5 TM	2005-09-19
Landsat 5 TM	2001-07-22	Landsat 5 TM	2008-09-27
Landsat 7 ETM+	2002-06-15	Landsat 5 TM	2009-06-26
Landsat 7 ETM+	2002-08-18	Landsat 5 TM	2010-07-31
Landsat 5 TM	2002-09-11	Landsat 5 TM	2011-07-02
Landsat 5 TM	2002-09-27	Landsat 5 TM	2011-08-03
Landsat 5 TM	2003-07-28		

3.4.2 MODIS data

Another study showed that the Nadir BRDF-Adjusted Reflectance (NBAR) 16-day MODIS composites (MCD43A4) provided the most accurate synthetic imagery

when fused with Landsat imagery (Walker et al. 2012). Therefore, NBAR composites were used in this study. A total of 604 NBAR composites (tile h10v05; 8-day temporal resolution; 500-meter spatial resolution) between 5/24/2000 and 9/30/2011 were obtained from the Land Processes Distributed Active Archive Center (LP DAAC; https://lpdaac.usgs.gov). These composites were presented as eight-day composites based on 16-days of data (Lucht et al. 2000). The temporal extent of these composites was seven time steps before and seven time steps after the first and last images, respectively, in the Landsat time-series. Using additional MODIS composites before and after the study period was necessary for the Savitzky-Golay algorithm (Savitzky and Golay 1964) used to smooth the time-series of MODIS disturbance index (detail in section 3.4.3). All of these composites were reprojected to the UTM zone 15N and clipped to the extent of the Landsat time-series. These composites were then layer stacked by date to form a MODIS time-series.

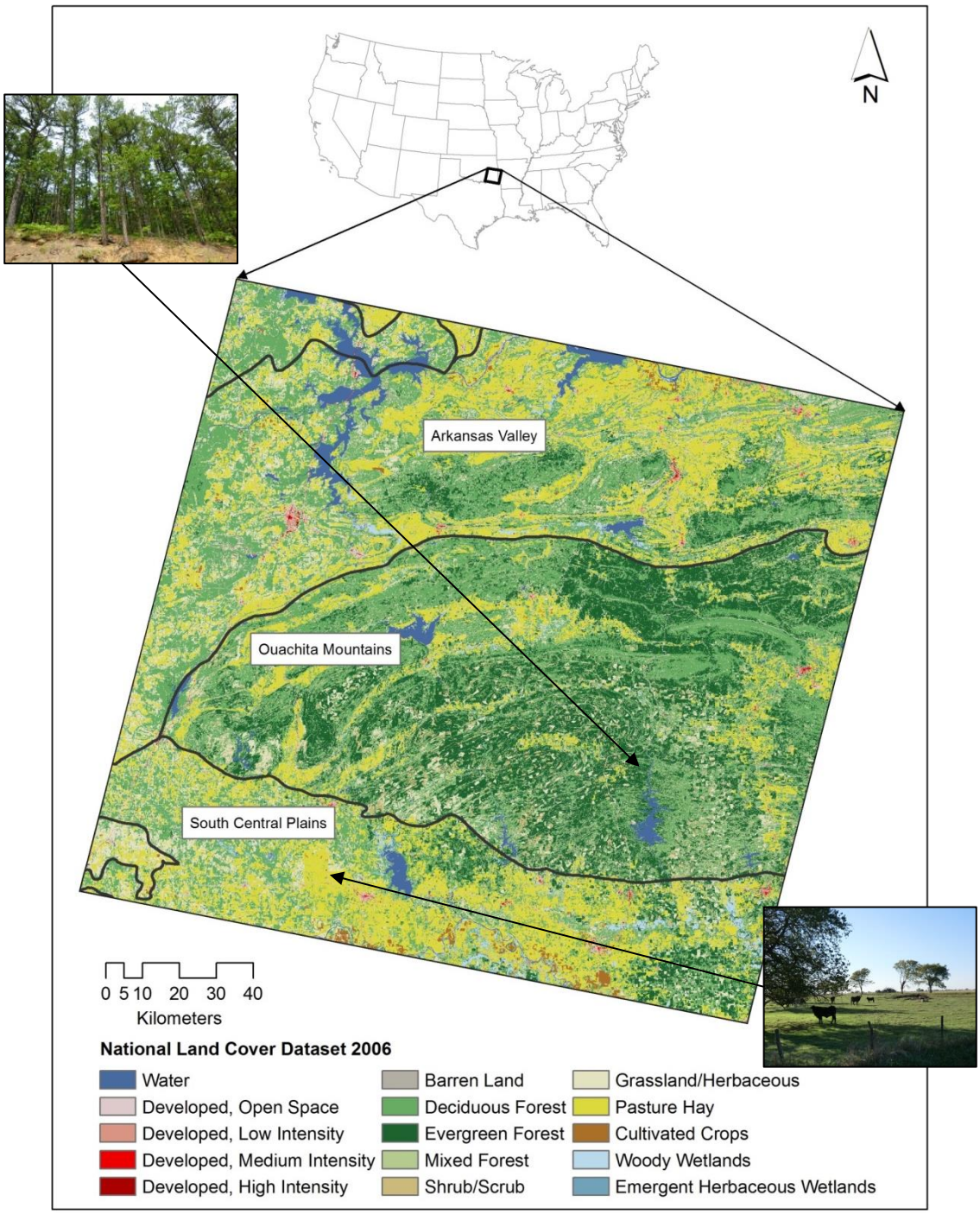


Figure 3.1. The study area (185 km x 185 km; Landsat scene path 26, row 36) in southeast Oklahoma. The background shows land-cover according to the 2006 NLCD. The most dominant land-cover type is forest including deciduous, evergreen, and mixed forests (56%), followed by grassland, shrub and pasture hay (35%). Wetland occupies 1.3%, cultivated land 0.6%, non-vegetated surface (e.g. urban and barren land) 4.4%, and water 2.7%.

3.4.3 Land cover data

The National Land Cover Databases (NLCDs) 2001 and 2006 (Fry et al. 2011; Homer et al. 2004) were used in this study for two purposes: to identify input forest pixels for the disturbance detection and to develop reference sets for the disturbance detection. The reference sets included pixels spectrally defined as unchanged and mature forests against which all other pixels were spectrally compared for the determination of their disturbance conditions (i.e. disturbed or undisturbed; section 3.4.3). Both NLCDs were downloaded from the web interface of the Multi-Resolution Land Characteristics Consortium (www.mrlc.gov). NLCDs after downloaded were reprojected to the UTM zone 15N and clipped to the extent of the Landsat time-series. Although these NLCDs had 29 classes covering water, forest, shrub, herbaceous, and wetland (Homer et al. 2004), only forest classes (deciduous 41 and evergreen 42) were used for this study.

3.4.4 Validation data

Field-trip photos

Georeferenced field-trip photos were used to validate locations of disturbances. Beside the existing photos taken in June and July 2011 (Julian 2011), I conducted a couple field trips to southeast Oklahoma in October 2012 and June 2013 to gain knowledge about the dynamics and spatial patterns of disturbed and undisturbed forests. During the field trips, I used a Casio EX-H20G GPS enabled camera to photograph disturbance conditions resulting from forest harvests, urban development, and road clearing. Because forest was dominant in the Ouachita Mountains ecoregion, I mainly explored this ecoregion during the field trips. The explored area was generally enclosed

by the Talimena Scenic Drive in the north, Highway 259 in the east, and Highway 3 in the south (Figure 3.2). Besides those roads, I also explored highway 144 and a couple of forest paths.

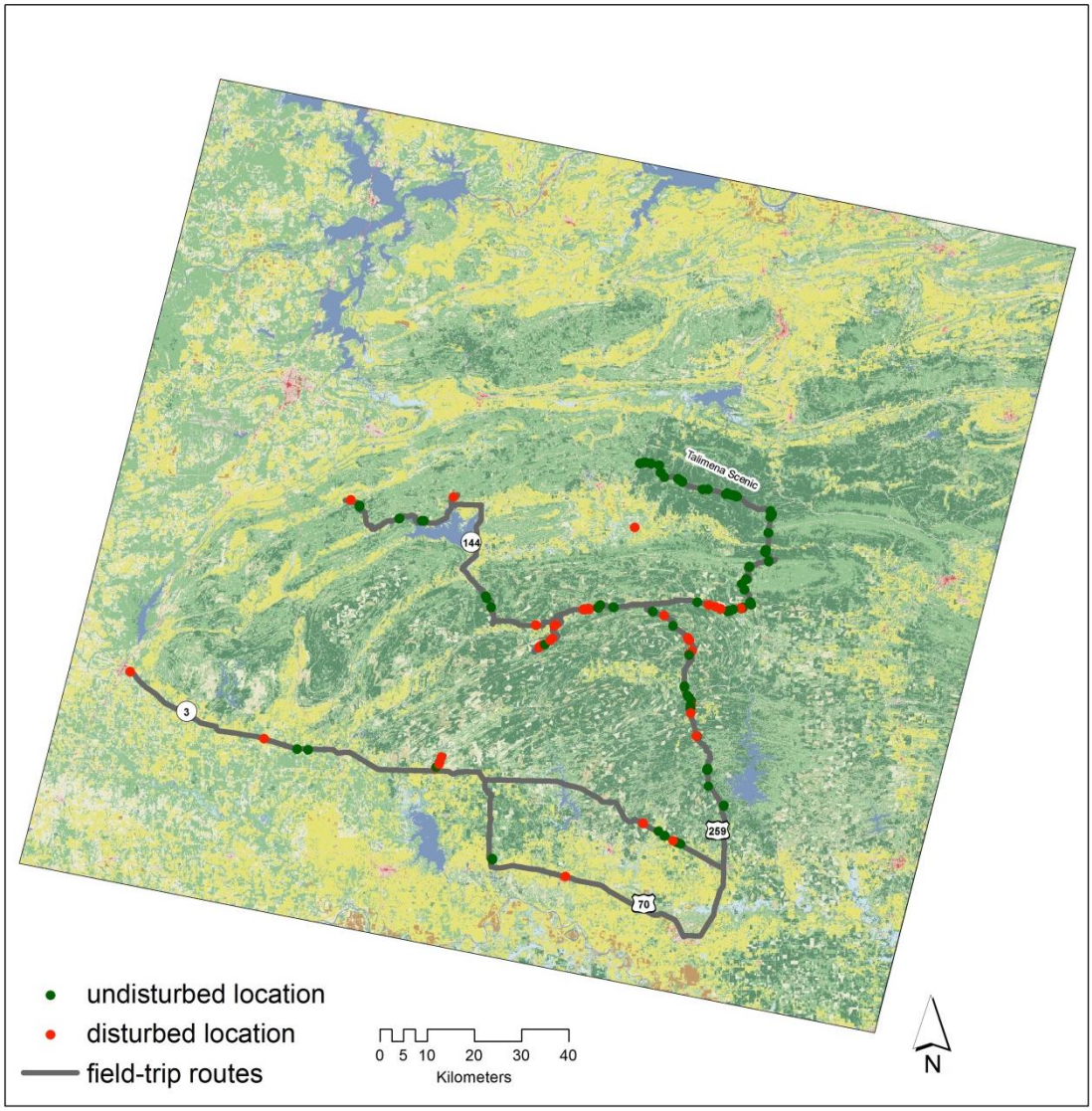


Figure 3.2. Field-trip routes and locations. The background shows land cover types according to the NLCD 2006.

The Talimena Scenic Drive going east-west has been a National Scenic Byway in southeastern Oklahoma. Along this drive was dominated by deciduous trees. Unlike the Talimena Scenic Drive, Highway 259 going north-south and cutting across multiple

mountains has had both deciduous and evergreen forests, some of which were clear-cut areas. Highway 3 was dominated by evergreen forests. I found this route interesting because it had many new clear-cut areas. In fact, Google Earth™ showed that along Highway 3, more evergreen forests have recently been cut compared to 2003. Along Highway 144 was dominant with evergreen. These evergreen areas are harvested or replaced by agricultural fields. Driving into the forests through forest trails, I found that there have been urban and/or agricultural expansions going on, proven by areas cleared for artificial surfaces.

Besides the knowledge of the dynamics and spatial patterns of disturbances, the other achievement of these field trips was 206 georeferenced photos used for validation. Photos were taken at least 30 meters apart so that they would not fall into the same Landsat pixel. In addition, because the locations of these photos were used for validation, I estimated the distances from the camera positions to the sites captured by the camera. These distances were then used to offset the photo locations so that they represented the disturbance conditions in the photos. Disturbance conditions of forests in the photos were determined based on the NLCD definitions of shrub and grassland/herbaceous. Therefore, areas that were occupied by trees less than five meters tall or by more than 80 percent of gramanoid or herbaceous vegetation were considered as disturbed (Figure 3.3).



(a) Closed-canopy deciduous forest, detected as persistently undisturbed (photo taken on 6/22/2013)



(b) Open-canopy evergreen forest, detected as persistently undisturbed (taken on 6/22/2013)



(c) Disturbed due to harvesting, detected on 7/4/2011 (taken on 7/13/2011)



(d) Disturbed due to harvesting, detected on 7/4/2011 (taken on 10/7/2012)



(e) Regeneration of an evergreen forest from a harvest detected on 6/26/2009 (taken on 6/22/2013)

Figure 3.3. Field-trip photos taken in different years, showing disturbance conditions and their initial dates of disturbance detected by the proposed MSTAARCH algorithm.

Aerial photos

Time-stamped NAIP photos were used to validate changes in spatial extent of areas disturbed. These photos in county-mosaic format were ordered through the U.S. Department of Agriculture Geospatial Data Gateway (http://datagateway.nrcs.usda.gov/) for the sixteen counties in Oklahoma in six years (2003, 2004, 2005, 2006, 2008, and 2010). These one-meter-resolution natural-color photos were acquired during the agricultural growing seasons in the continental U.S (U.S. Department of Agriculture 2010). The actual dates of acquisition of these orthorectified NAIP photos were derived from their metadata shapefiles representing digital-ortho-quarter-quad-tiles (DOQQs) or flying paths. To create the NAIP validation dataset, I randomly digitized a total of 350 clear-cut areas (i.e. NAIP patches) from NAIP photos. For each NAIP patch, the dates that it was disturbed were recorded by comparing its location with the NAIP metadata shapefiles. These vector-based NAIP patches were converted to raster-based (30 meters) NAIP patches for the validation.

Thematic burn severity dataset

The thematic burn severity dataset was used to validate the spatial extents of areas disturbed due to fire at a yearly time scale. This 30-meter raster-based thematic burn severity dataset was obtained from the Monitoring Trends in Burn Severity Geospatial Database (http://www.mtbs.gov). The dataset represented large (over two square kilometers in East and four square kilometers in West) fires in the conterminous US (Schwind 2007). Each pixel in this dataset was classified as (1) unburned to very low, (2) low, (3) moderate, (4) high, (5) increased greenness, or (6) non-processing mask (Schwind 2007). In this study, only pixels of class 3 (moderate) and class 4 (high)

were labeled as burnt pixels and used for the validation because they captured most severe fires.

3.5 METHODS

3.5.1 Spatial pattern analysis of NLCDs

Because of the effect of land cover heterogeneity on the disturbance detection, especially when it was performed based on the moderate-resolution MODIS imagery, land cover heterogeneity in southeast Oklahoma was investigated. The investigation was conducted based on both the NLCD 2001 and 2006. FRAGSTATS software (McGarigal et al. 2012) was used to calculate the landscape-level Shannon's Diversity Index (SHDI) for the entire region with all fifteen NLCD land cover types. This diversity index has been widely used to measure the heterogeneity (the higher Shannon's index, the more heterogeneous) of a landscape through its proportions of all land cover types. Additionally, FRAGSTATS was also used to calculate the class-level number of patch (NP) index for the deciduous and evergreen classes within 300 randomly selected MODIS pixels using both NLCD 2001 and 2006 to better understand the degree of heterogeneity of MODIS pixels in the study area. The smaller the number of patches within MODIS pixels was, the less heterogeneous they were.

3.5.2 Statistical test of tasseled cap indices

To better understand the variability of tasseled cap indices due to different compositions of forest types within MODIS pixels, the Mann-Whitney test was performed for ten MODIS tasseled cap images (Brightness, Greenness, Wetness) randomly selected from the MODIS tasseled cap time-series. This non-parametric test was used to relax the assumption of data normality. The significance of the test was

assessed at α=0.05. For each tasseled cap image, the test was perform for three groups of pixels (pure deciduous, pure evergreen, and combination) with the hypothesis that the tasseled cap indices of these groups of pixels were not significantly different. The developments of the tasseled cap time-series and the three testing groups were presented in section 3.5.3 below.

3.5.3 Data fusion

In this study, the MSTAARCH fusion framework was developed as a modification of the original STAARCH. The modification focused on the selection of reference sets used to develop MODIS time-series of disturbance index. Similarly to STAARCH, the MSTAARCH included three main steps: tasseled cap transformation, disturbance index calculation, and data fusion (Figure 3.4; Hilker et al. 2009).

Tasseled cap transformation

The purpose of this step was to transform the Landsat time-series and MODIS time-series to multiple time-series of tasseled cap indices (Brightness, Greenness, and Wetness; Kauth and Thomas 1976). Separate transformation coefficients were used for Landsat TM and Landsat ETM+ (Crist 1985) and for MODIS (Lobser and Cohen 2007). These transformed time-series were then used as inputs to the calculation of disturbance index.

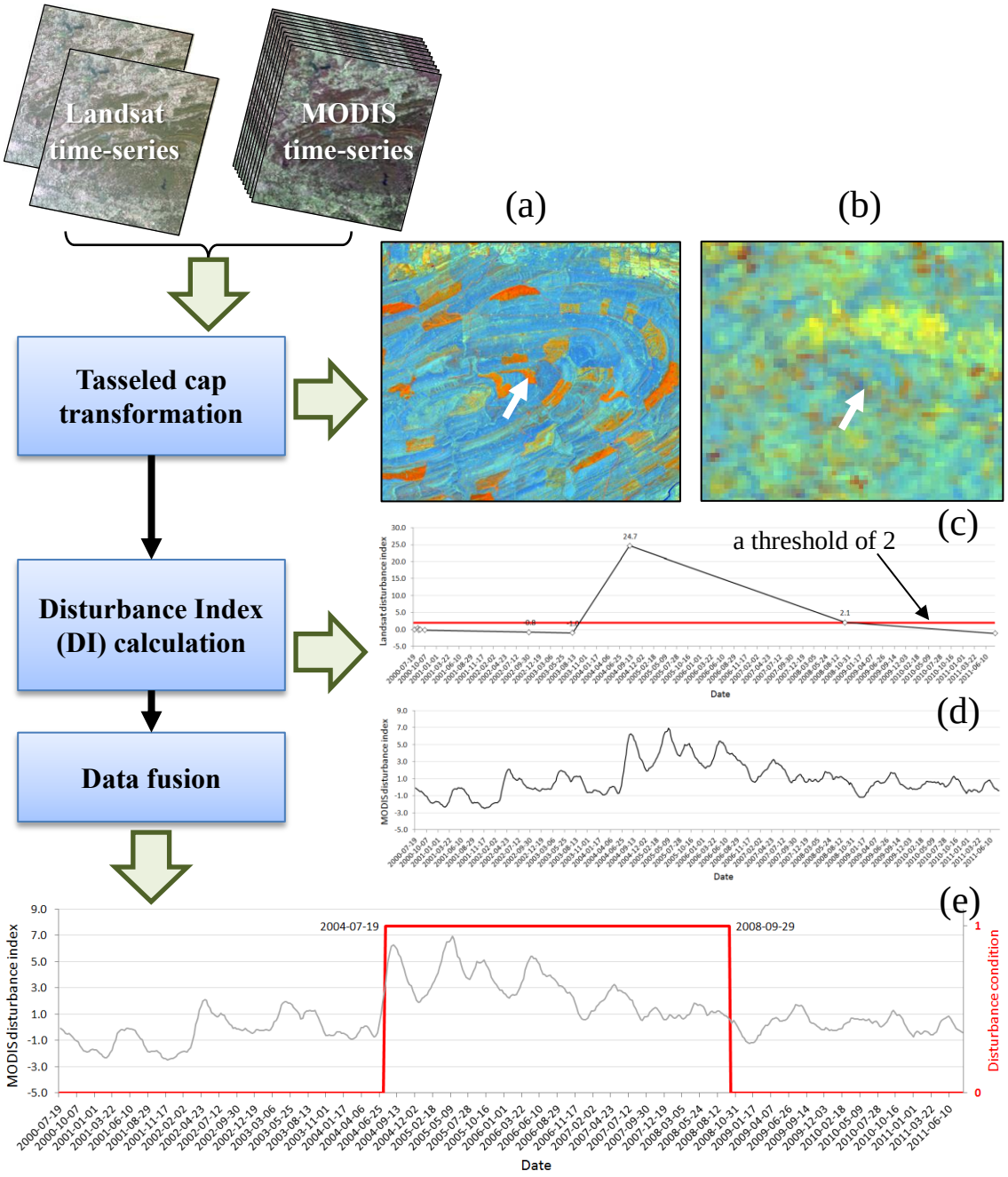


Figure 3.4. Overview of the methodology with three primary steps (tasseled cap transformation, disturbance index calculation, and data fusion) and their corresponding outputs. (a) and (b) are Landsat and MODIS tasseled cap time-series; (c) and (d) are Landsat and MODIS DI profiles of the selected pixel in (a) and (b); note that in (c), a threshold of 2 is applied to the Landsat DI profile to convert it to a binary profile representing disturbance conditions; and (e) is the final fused binary DI time-series overlapped with the MODIS DI time-series of the selected pixel, for which the initial and final dates of a disturbance are 7/19/2004 and 9/29/2008, respectively.

Development of reference sets and calculation of disturbance index

The purpose of this step was to create a Landsat time-series of disturbance index (i.e. Landsat DI time-series) and a MODIS time-series of disturbance index (i.e. MODIS DI time-series). These time-series were created by calculating disturbance index (DI; Healey et al. 2005) for every pixel on the scene-by-scene basis using equation (1).

$$DI = Brightness_n - (Greenness_n + Wetness_n) \quad (1)$$

DI was developed in the form of equation (1) to identify stand-replacing disturbances which often resulted in large contrast between the brightness and the total of greenness and wetness indices (Healey et al. 2005). For each scene (either Landsat or MODIS), DI was calculated for every pixel as a linear function of its three tasseled cap indices normalized by the scene mean and standard deviation of the tasseled cap indices of a reference set composing of pixels representing unchanged and mature forest (Masek et al. 2008). This normalization was expected to reduce the effects of solar geometry and vegetation phenologic variability on the change of tasseled cap indices so that this change would be strongly associated with disturbances (Healey et al. 2005; Masek et al. 2008). Because of the normalization, DI actually represented a spectral distance of a given pixel to an undisturbed forest represented by the reference set. The larger the distance was the more likely the given pixel became disturbed. A threshold suggested to be 2 could be applied across time for every pixel to determine its temporal disturbance conditions (i.e. when it was disturbed with DI $\geq$ the threshold or undisturbed; Healey et al. 2005; Masek et al. 2008).

Similarly to STAARCH, in MSTAARCH, DI was calculated separately for Landsat and MODIS time-series and based on separate approaches of developing reference sets. For the Landsat time-series, multiple reference sets were developed, each of which represented a forest type. Because there were primarily two forest types (deciduous and evergreen) in the study area, two Landsat reference sets were developed to normalize deciduous and evergreen pixels identified by the NLCD 2001. The Landsat deciduous reference set was created by extracting pixels labeled as deciduous on both NLCD 2001 and NLCD 2006. Similarly, the Landsat evergreen reference set was created by extracting pixels labeled as evergreen on both NLCD 2001 and NLCD 2006. These Landsat reference sets were used to normalize the Landsat tasseled cap time-series to create a Landsat DI time-series.

For the MODIS time-series, STAARCH used only one reference set, named here as STAARCH reference set. This reference set represented all MODIS pixels assumed to be unchanged without accounting for their variability in forest compositions. However, because of the heterogeneity of the study area, most MODIS pixels often composed of multiple forest types and/or other land cover types that could change from time to time with different frequency and intensity, and thus without careful selection, MODIS pixels could hardly become a good reference set of undisturbed forest. In addition, STAARCH reference set could include MODIS pixels (especially those in the South Central Plains) that composed of a little amount of forest but almost a full amount of rangeland. This situation would result in the STAARCH reference set being similar to a disturbed pixel having very high brightness and low greenness and wetness. Consequently, STAARCH reference set might not be able to

effectively detect disturbances. Therefore, I argued that using one reference set to normalize MODIS time-series would be inappropriate for a heterogeneous landscape.

In a heterogeneous landscape, MODIS pixels would be mixtures of forest types and/or other land cover types (e.g. water, grass, artificial surface). This mixed-pixel problem has been a challenge of not only land cover classification (Foody 2006) but also data fusion (Gao et al. 2006). For instance, normalizing a disturbed evergreen pixel (i.e. an evergreen pixel partly covered by evergreen forest) to a reference undisturbed evergreen pixel (i.e. an evergreen pixel fully covered by evergreen forest) would produce a DI higher (Figure 3.6, green line) than if the disturbed pixel was normalized by a reference pixel partly covered by evergreen forest (Figure 3.6, purple line) or by a reference deciduous undisturbed pixel (Figure 3.6, blue line). This was because of the higher brightness and lower wetness of the reference pixel partly covered by evergreen forest or of the reference undisturbed deciduous pixel compared to the reference evergreen pixel. Consequently, the disturbance condition, and thus the beginning and ending dates of a disturbance of the disturbed pixel might not be effectively detected by applying a threshold (e.g. 2; Figure 3.6, red line) to the pixel's DI if the reference pixel partly covered by evergreen forest or if the reference undisturbed deciduous pixel was used in the normalization.

In another instance, normalizing an undisturbed deciduous pixel (i.e. a deciduous pixel fully covered by deciduous forest) to a reference undisturbed evergreen pixel would produce a DI higher than if the undisturbed deciduous pixel was normalized by a reference undisturbed deciduous pixel. Consequently, the undisturbed deciduous pixel might be detected as disturbed even though in fact, it was not.

Therefore, I suggested that it was necessary to (1) use multiple reference sets representative of forest types and their combinations and (2) use pixels fully covered rather than partly covered by forests for the reference sets. In MSTAARCH, I proposed to have three reference sets: pure deciduous, pure evergreen, and a combination of deciduous and evergreen. The pure deciduous reference set composed of unchanged pixels fully covered by deciduous forest. It was used to normalize pixels fully covered by deciduous forests at the beginning of the study period. Similarly, the pure evergreen reference set was used to normalize pixels fully covered by evergreen forests at the beginning of the study period. The combination class included pixels that were fully covered by a combination of both deciduous and evergreen forests and was used to normalize any pixels that were not fully covered by either deciduous or evergreen forests at the beginning of the study period.

MODIS pixels to be normalized were identified by first aggregating deciduous and evergreen pixels extracted from the NLCD 2001 to 500 meters. Those 500-meter pixels having 100 percent of deciduous were normalized by the pure deciduous reference set. Similarly, those 500-meter pixels having 100 percent of evergreen were normalized by the pure evergreen reference set. All other pixels were normalized by the combination reference set. The three MODIS reference sets were created by first aggregating the Landsat deciduous and evergreen reference sets to 500 meters. Among these 500-meter pixels, those having 100 percent of deciduous were used as the MODIS pure deciduous reference set. Similarly, those having 100 percent of evergreen were used as the MODIS pure evergreen reference set. Among the rest of these 500-meter pixels, those having a sum of deciduous and evergreen proportions greater than or equal

to 99 percent were used as the MODIS combination reference set. The proportion of 99 percent was used instead of 100 percent because in the latter case, there were too few pixels. All these three reference sets were then used to normalize the MODIS tasseled cap time-series to create a MSTAARCH MODIS DI time-series.

In addition to the three proposed reference sets, the STAARCH reference set was also created by merging all three reference sets (pure deciduous, pure evergreen, and combination) together. This STAARCH reference set was used to create a STAARCH MODIS DI time-series for comparison purpose.

All MODIS DI time-series after created was run for the Savitzky-Golay filter (Savitzky and Golay 1964) to reduce noises caused by cloud contamination, atmospheric variability, and bi-directional effects (Chen et al. 2004; Hird and McDermid 2009). This Savitzky-Golay filter used a polynomial function to average data within a moving temporal window. The filter had two important parameters including the window size and the degree of the polynomial function. The appropriate window size was recommended between 9 and 15 and the degree between 2 and 4 (Chen et al. 2004). In this study, I used the window size of 15 and the degree of 2 to maximize the smoothing capability of the filter. After being filtered, the MODIS DI time-series was temporally subset to represent the period of 7/19/2000 and 8/5/2011, which was closest to the Landsat period. The MODIS DI time-series was then resized to 30 meters using the cubic convolution resampling method.

Image fusion

This image fusion steps composed of two procedures. The first procedure was to apply a MODIS DI threshold to a MODIS DI time-series to produce a MODIS binary

DI time series. In the original STAARCH framework, the DI threshold was not constant but rather varied from pixel to pixel because this STAARCH DI threshold was a linear function of a pixel's DI_{min} and DI_{max}, which were the minimum and maximum, respectively, of the pixel's DI across time. This threshold was applied to the STAARCH MODIS DI time-series to create a STAARCH MODIS binary DI time-series. Because the use of DI_{min} and DI_{max} as a threshold could be affected by DI outliers caused by noises, in MSTAARCH, I used a more common threshold of 2 (Healey et al. 2005; Hilker et al. 2009). To understand the effect of this threshold on the final result, I applied this threshold 2 to the STAARCH MODIS binary DI time-series to produce a Threshold-2 MODIS binary DI time-series for comparison purpose. The threshold 2 was also applied to the MSTAARCH MODIS DI time-series to produce a MSTAARCH MODIS binary DI time-series. These three binary DI time-series were then used for the fusion procedure.

In each case of the binary DI time-series, a fused binary DI time-series was initially created with null images. This time-series had a spatial resolution of 30 meters and temporal resolution of eight days from 7/19/2000 to 8/5/2011. Every Landsat image from the binary Landsat DI time-series was copied to replace a null image that was temporally closest to the date of the Landsat image. As a result, there were multiple sub-periods whose beginning and ending images had binary values taken from the Landsat binary DI time-series. Values for images within the sub-periods were determined as following.

For every sub-period, pixels identified as disturbed (or undisturbed) by both the beginning and ending images of the sub-period was considered as disturbed (or

undisturbed) for the entire sub-period and thus, their entire sub-period temporal profile was assigned with a value of one (or zero if undisturbed). In another case, pixels identified as undisturbed by the beginning image but disturbed by the ending image of the sub-period were considered as changed pixels. Their beginning time of disturbance was the date that they were first detected as disturbed by the binary MODIS DI time-series within the sub-period. Similarly, the ending time of disturbance of pixels identified as disturbed by the beginning image but undisturbed by the ending image of a sub-period was the date that they were first detected as undisturbed by the binary MODIS DI time-series within the sub-period. At the end of the image fusion step, three binary DI time-series were created: STAARCH binary DI time-series, Threshold-2 binary DI time-series, and MSTAARCH binary DI time-series corresponding to the three MODIS binary DI time-series created above.

3.5.4 Validation

The validations were performed on a pixel-by-pixel basis for (1) the Landsat binary DI time series, (2) the STAARCH binary DI time-series, (3) the Threshold-2 binary DI time series, and (4) the MSTAARCH fused binary DI time-series using field-trip photos, NAIP aerial photos, and the thematic burn severity dataset.

Field-trip photos were used to validate locations of areas disturbed and undisturbed, which were detected by the last image in each of the four binary DI time-series to be validated. The disturbance conditions of pixels were compared with those of the photos at the same locations. Pixels having the same disturbance conditions as the photos were considered as agreed pixels. The ratio between the number of agreed pixels and the total number of photos was the overall accuracy.

Digitized NAIP patches were used to validate the locations and extents of areas disturbed for every of the 50 NAIP dates. To do this, the raster-based NAIP patches were used to validate images in each of the four binary DI time-series on the dates that were closest to the NAIP dates. For instance, if a NAIP patch was observed on 6/7/2003, it would be used to validate the image on 6/10/2003. In the case of the Landsat binary DI time-series, because there was no Landsat image in 2006 and 2007, these years were excluded from the validation of Landsat binary DI time-series. Producer's accuracies were calculated for each NAIP date and for each binary DI time-series. These accuracies were also summarized by year and presented in a color matrix, in which each cell represented the accuracy of a given time-series in a given year. Summary of the accuracies in this manner allowed for a synoptic assessment of the four time-series over time and between the methods used to create them.

The annual thematic burn severity dataset was used to validate the locations and extents of areas disturbed due to fire at a yearly time scale. To do this, each of the four time-series to be validated was accumulated by year to create a yearly time-series. Consequently, each image of these yearly time-series represented all disturbances happening between January 1st and December 31st of a given year (except the year 2000 from 7/19/2000 to 12/31/2000 and the year 2011 from 1/1/2011 to 8/5/2011). The data accumulation process was done by copying all disturbed pixels from all dates within a year into one image representative of that year. These yearly time-series, which would include twelve images for twelve years (from 2000 to 2011) were then validated against the thematic burn severity dataset. In this validation, the producer's accuracy was

calculated as a ratio between the number of pixels in a yearly time-series matched with the burnt pixels and the total of burnt pixels.

3.5.5 Mapping disturbances

To understand the spatio-temporal distribution of disturbances in southeast Oklahoma, the beginning time of disturbances were mapped. Because there could be multiple disturbances happing at the same location within the twelve-year period, for each pixel, its beginning time of the disturbance having longest duration was mapped. It was assumed that disturbances having longer duration were more severe.

3.6 RESULTS

3.6.1 Land cover heterogeneity in the study area

Land cover in southeast Oklahoma was heterogeneous (Shannon's Index over 1.8) in both 2001 and 2006. This heterogeneity resulted in highly heterogeneous MODIS pixels (average of Shannon's Index over 0.8; Figure 3.5) all over the landscape of southeast Oklahoma spreading across the Arkansas Valley, Ouachita Mountains, and South Central Plains ecoregions. As a result, forest patches within MODIS pixels was, on average, as small as 28 percent of the size of a MODIS pixel (Figure 3.5).

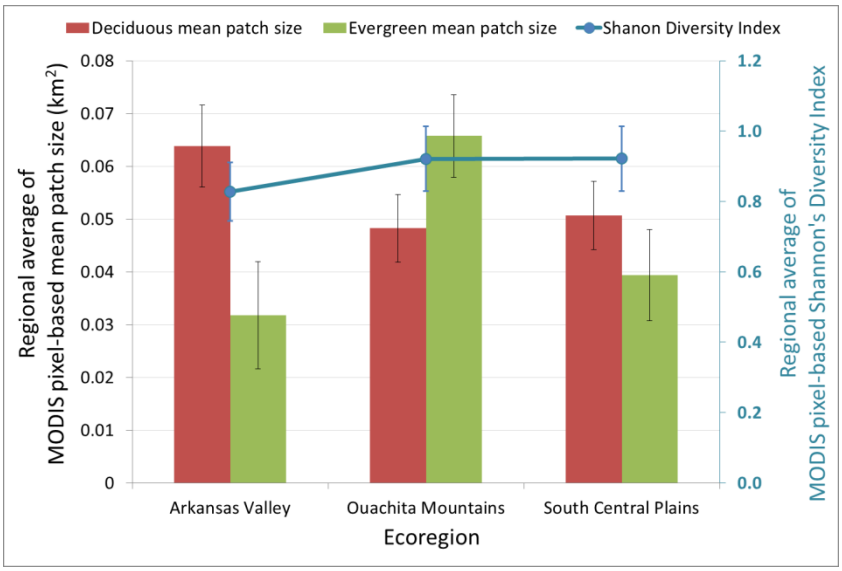


Figure 3.5. Ecoregional average (N=100 MODIS pixels per ecoregion) of MODIS pixel-based mean patch size and Shannon's diversity index calculated based on NLCD 2001. Error bars represent standard errors.

3.6.2 Variability of tasseled cap indices due to different compositions of forest types

The Mann-Whitney test showed that there were significant differences (p<0.05) in MODIS tasseled cap indices between the three MODIS reference sets (pure deciduous, pure evergreen, and combination) proposed by this study (Figure 3.7). The brightness of pure deciduous reference set was significantly higher than that of pure evergreen reference set because reflectance values of deciduous forests were higher than those of evergreen forests in infrared bands. Unlike brightness, the difference in greenness between the two reference sets was seasonally dependent. Typically, greenness of pure deciduous reference set was higher than that of pure evergreen reference set during leaf-on (e.g. summer) season but became lower during leaf-off (e.g. winter) season. In contrast with brightness, wetness of pure deciduous reference set was lower than that of pure evergreen reference set (Figure 3.6) typically because pure

deciduous forests were more sensitive to the infrared bands. The tasseled cap indices of the combination reference set were generally between those of the evergreen and deciduous reference sets.

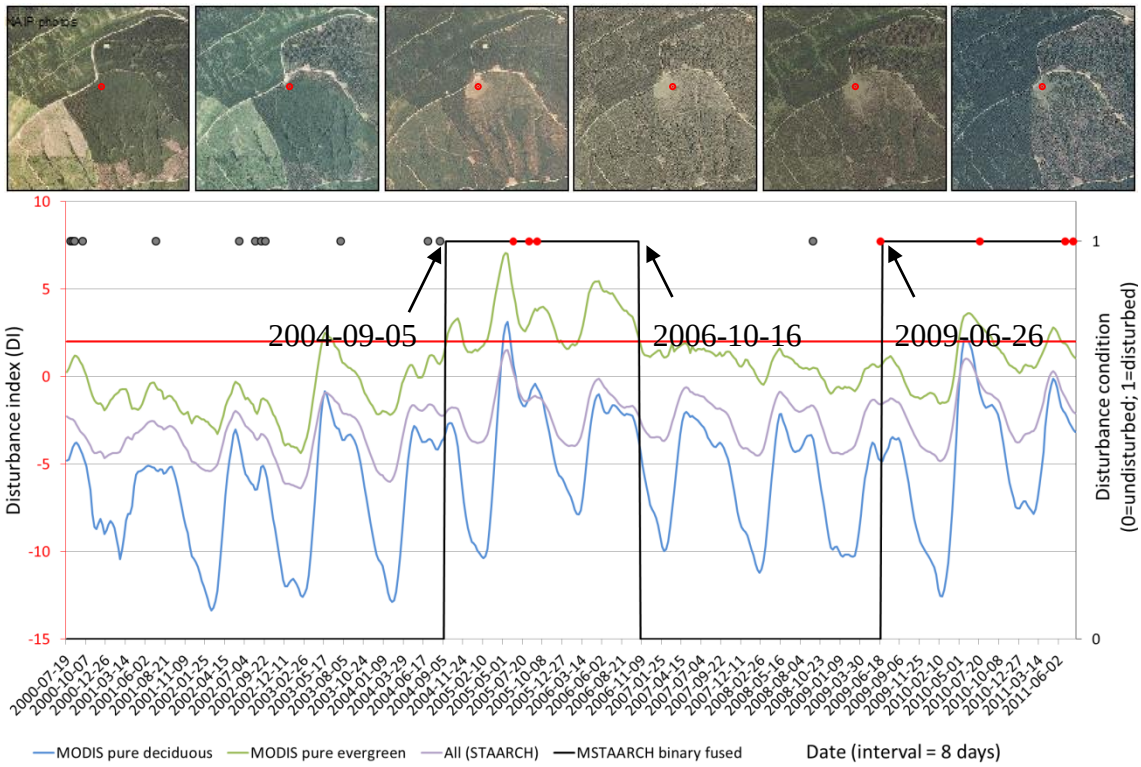


Figure 3.6. Multiple disturbance index time-series of a MODIS disturbed evergreen pixel located at the red dot in NAIP photos at the top row. Each of these time-series was normalized by different reference sets: MODIS pure deciduous (blue), MODIS pure evergreen (green), and MODIS STAARCH (purple). Please see text for explanation of these reference sets. The red line represents the threshold of 2 used to determine disturbance condition, represented by the black line, of the analyzed pixel. Gray dots represent available Landsat images identifying the analyzed pixel as undisturbed whereas red dots represent available Landsat images identifying the analyzed pixel as disturbed.

The tasseled cap indices of the STAARCH reference set were also presented in Figure 3.7 for comparison purpose. In general, the tasseled cap indices of the STAARCH reference set were in between those of the combination and pure deciduous

reference sets. Indeed, while brightness and wetness of STAARCH reference set were as high as those of the pure deciduous reference set, greenness of STAARCH reference set was similar to that of combination reference set.

3.6.3 Method validation and comparison

Field-trip photos were first used to validate locations of disturbances on the last date (8/3/2011) of the Landsat binary DI time-series. Note that the last Landsat binary DI image in this time-series was used as the last binary DI image in all other binary DI time-series produced by the three fusion methods. Of the 206 photos, 198 of them (141 undisturbed and 57 disturbed) were correctly detected. The overall accuracy, therefore, was 96 percent. To further investigate the capability of the Landsat binary DI time-series and the other three fused binary DI time-series NAIP photos and the burn severity dataset were used.

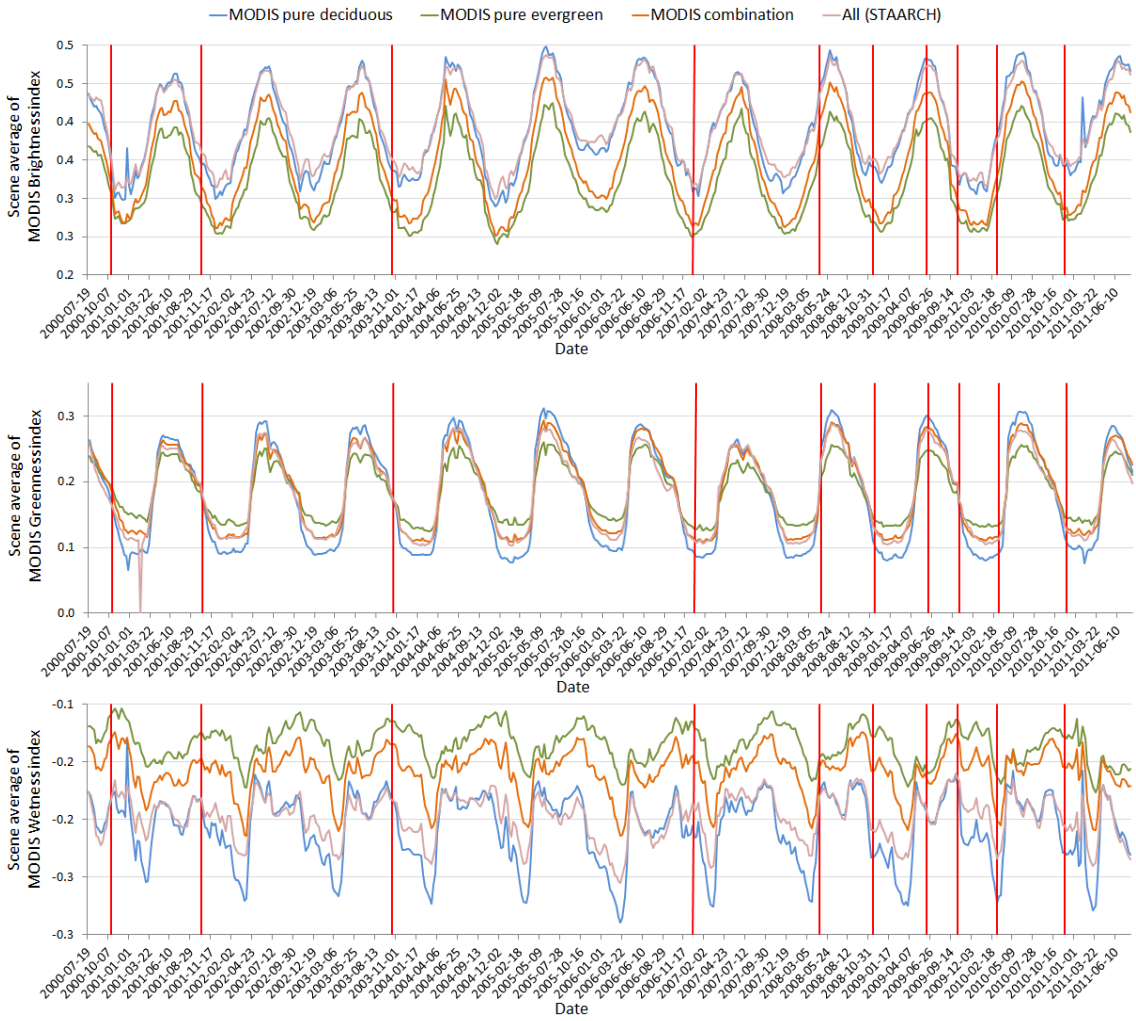


Figure 3.7. Time-series of scene averages of tasseled cap indices for MODIS reference sets: pure deciduous (N=685), pure evergreen (N=274), and their combination (N=1,774). Red vertical lines represent the time steps used for the Mann-Whitney test.

The four binary DI time-series of disturbances were first compared visually against NAIP photos and against each other (Figure 3.8). While the Landsat binary DI time-series (Figure 3.8 b,g) was created by using only the Landsat time-series, the other three fused binary time series were created by using both the Landsat time-series and MODIS time-series based on a fusion algorithm. Recall that the Threshold-2 binary DI time-series (Figure 3.8 d,i) was created by modifying the STAARCH (Figure 3.8 c,h) with the use of a MODIS DI threshold 2. This threshold was expected to reduce the

impact of noises causing extremes in MODIS DI and therefore, to be improved from the original STAARCH. The MSTAARCH binary DI time-series (Figure 3.7 e,j) was created by modifying the STAARCH with not only the use of the MODIS DI threshold 2 but also the use of multiple MODIS reference sets. Because of these modifications, MSTAARCH was expected to be significantly improved from STAARCH.

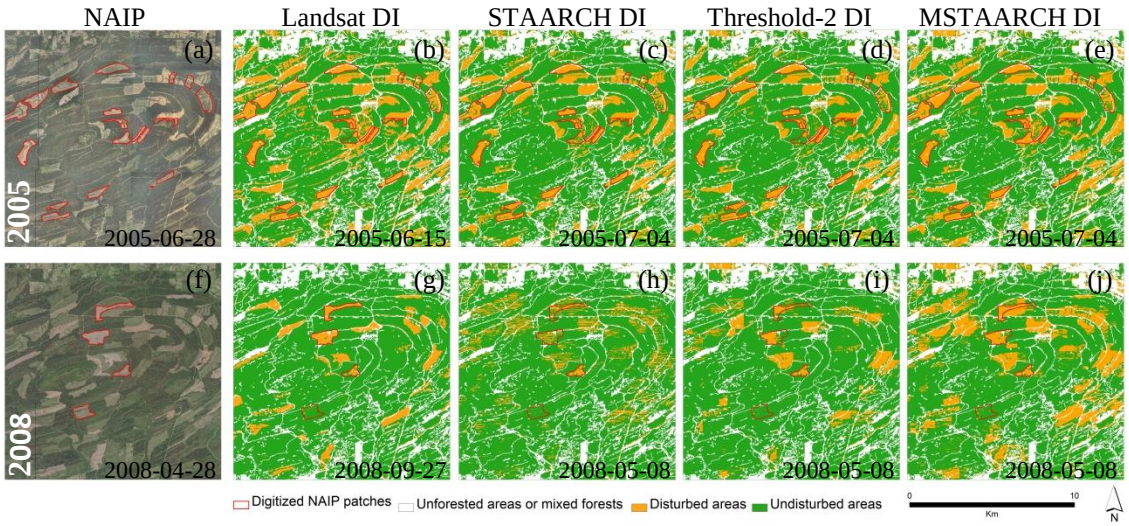


Figure 3.8. Changes in area disturbed (30-meter resolution) detected by different methods for a typical evergreen forest heavily harvested since 2000. In this figure, the first column (NAIP) represents NAIP aerial photos in 2005 and 2008 and the second to the fifth columns represent the four binary DI time-series: Landsat DI, STAARCH DI, Threshold-2 DI, and MSTAARCH DI. Please see text for explanation on the binary time-series.

In Figure 3.8, NAIP aerial photos (Figure 3.8 a,f) were presented as a base line for comparison. The four binary DI time-series were presented on two dates in 2005 and 2008. In general, with its high spatial resolution, the Landsat binary DI time-series well captured the locations and extents of disturbances (Figure 3.8 b). However, its moderate temporal resolution avoided it from being a perfect time-series capturing disturbances over time (Figure 3.8 g). For instance, the Landsat binary DI image on 9/27/2008 was not capable of capturing disturbed patches found five months before (i.e. on 4/28/2008).

It was because forests in these patches, which could have become disturbed well before 4/28/2008, grew to a level that was not able to be spectrally detected by a Landsat sensor.

A fusion of Landsat time-series with such a high temporal resolution data as the MODIS time-series was expected to solve this problem. However, it was found that both the STAARCH binary DI time-series and the Threshold-2 binary DI time-series were much worse (Figure 3.8 h,i). In contrast, the MSTAARCH binary DI time-series was much better (Figure 3.8 j). It was able to not only detect disturbances on the date close to the Landsat date (Figure 3.8 e) but also on the date that was further from the Landsat date (Figure 3.8 j).

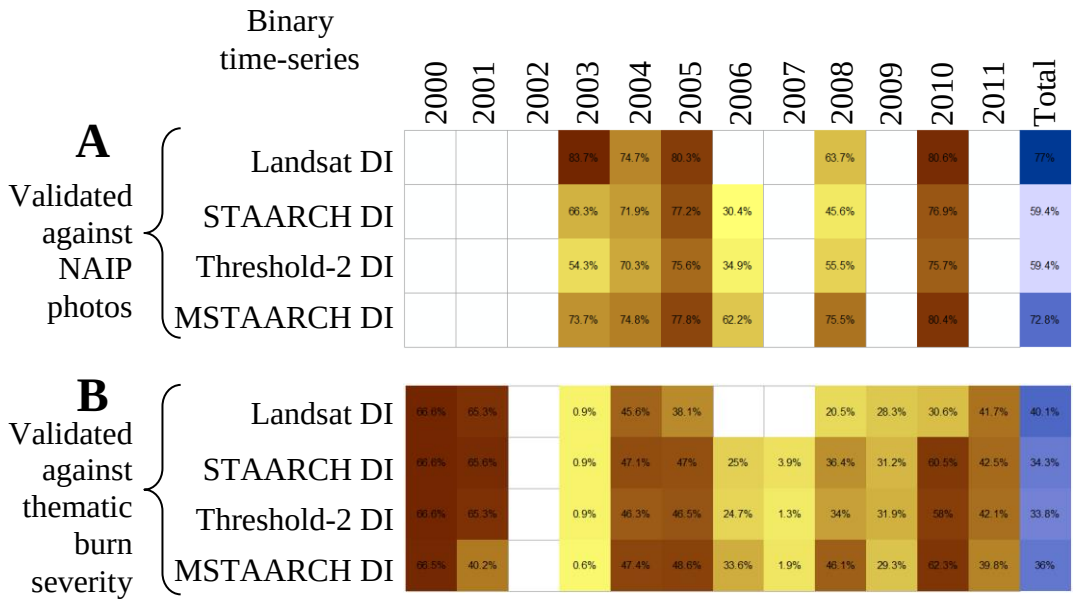


Figure 3.9. Producer's accuracies represented as color matrices over time and between methods of disturbance detection when validated against NAIP photos (A) and thematic burn severity dataset (B). Please see text for explanation on the binary time-series. Blank cells in the matrices represent no data.

Using NAIP to validate the four time-series revealed that the Landsat binary DI time-series had highest total producer's accuracy (77 percent), followed by the MSTAARCH binary DI time-series (4.2 percent lower than that of the Landsat; Figure 3.9 A). The producer's accuracies of the STAARCH and Threshold-2 binary DI time-series were lowest (18 percent lower than that of the Landsat). This same pattern of producer's accuracy among the four time-series was found for the yearly producer's accuracy as well. For example, in 2003, the producer's accuracy of the Landsat binary DI time-series was highest (83.7 percent), followed by the MSTAARCH binary DI time-series (73.7 percent) and the STAARCH binary DI time series (66.3 percent). It was also found that among the six years of NAIP availability, the years of 2005 and 2010 had higher producer's accuracies than other years for all four binary DI time-series. Conversely, the producer's accuracies of all four binary DI time-series were low in 2006 and 2008.

Specifically to fire, all four binary DI time-series did not demonstrate high producer's accuracies when they were validated against the thematic burn severity dataset. Among the four binary time-series, the Landsat binary DI time-series had the highest (40 percent) total producer's accuracy (Figure 3.8 B). For all time-series, the produce's accuracy was highest (66 percent) in 2000 while lowest (0.9 percent) in 2003. In 2001, compared to other time-series, the MSTAARCH binary DI time-series had the lowest producer's accuracy. In the years that did not have Landsat images (2006 and 2007) the producer's accuracies were low for all three fused time-series (i.e. STAARCH, Threshold-2, and MSTAARCH binary DI time-series), but those accuracies in 2006 were much higher than those in 2007.

3.6.4 Trends of area disturbed

The trends of yearly cumulative area disturbed derived from the four binary DI time-series were presented in Figure 3.10. In general, the cumulative area disturbed in southeast Oklahoma dropped in 2001 but then gradually increased from 2002 to 2005. In 2006 and 2007, there was again a large drop of cumulative area disturbed. Beginning in 2008, this cumulative area disturbed increased gradually.

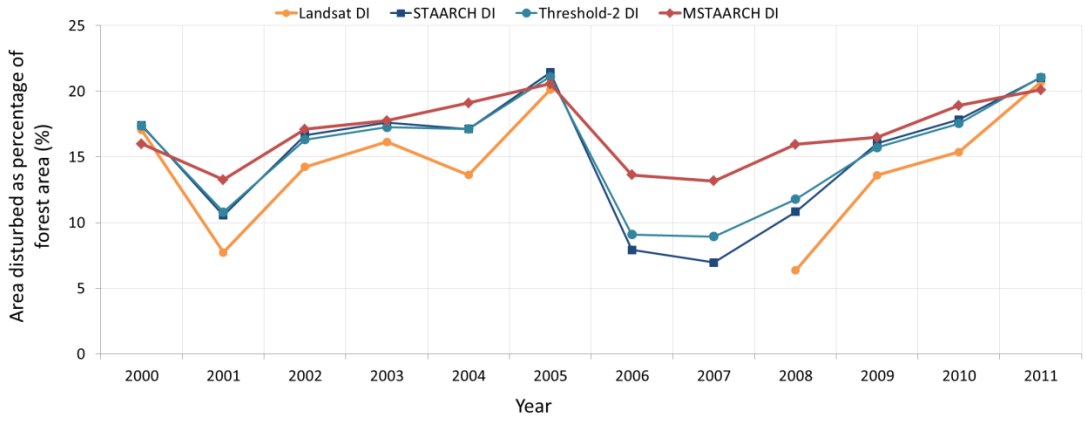


Figure 3.10. Trends of yearly cumulative area disturbed derived from the four binary DI time-series: Landsat DI, STAARCH DI, Threshold-2 DI, and MSTAARCH DI. Please see text for explanation on these binary time-series.

It was shown that the yearly cumulative areas disturbed detected by the Landsat binary DI time-series were lowest whereas those detected by the MSTAARCH binary DI time-series was highest. This was because compared to the fused binary DI time-series, the temporal resolution of the Landsat binary DI time-series was the lowest (a maximum of five images found in 2000; Table 1). Consequently, when accumulated, the areas disturbed derived from the Landsat binary DI time-series was not as much as the one derived from the fused binary time-series. Furthermore, because of its moderate temporal resolution, the Landsat binary DI time-series was not able to account for disturbed patches being undisturbed before the next Landsat image in the time-series.

Consequently, the trend line of the Landsat binary DI time-series was more jagged than the one of the MSTAARCH binary DI time-series. The moderate temporal resolution of the Landsat binary DI time-series also resulted in data gaps in 2006 and 2007. Compared to the other two fused time-series, the areas disturbed obtained from the MSTAARCH binary DI time-series was higher (Figure 3.10). Additionally, its trend line was also smoother, typically between 2005 and 2008, when there were no Landsat images.

3.6.5 Spatio-temporal distribution disturbances

Over the twelve-year study period, a majority of disturbances occurred in the evergreen forest in middle of the Ouachita Mountains and in the evergreen forest southeast of the South Central Plains (Figure 3.11 A). These disturbances resulted in large and homogeneous disturbed patches having regular shape (i.e. shape similar to a square; Figure 3.11 D). These disturbed patches tended to be extended from year to year. At the same time, small, elongated and heterogeneous disturbed patches were found in the deciduous forest in the Arkansas Valley and western of the South Central Plains where cultivated land and rangeland were dominated (Figure 3.11 A,C). These edge disturbed patches were primarily found at forest edges along roads, tree lines on rangeland, or riparian forests (Figure 3.11 C). In 2000, a large fire in northeast of the Arkansas Valley caused a largely disturbed patch northeast of the ecoregion. Besides, many oil sites were also found as small disturbed patches in the Arkansas Valley. Compared to disturbed patches at forest edges, disturbed patches in evergreen forests exhibited a homogeneous temporal pattern with one or two disturbed periods (Figure

3.11 F). In contrast, disturbed patches at forest edges exhibited a more heterogeneous temporal pattern with multiple disturbed periods (Figure 3.11 E).

3.7 DISCUSSION

This study found that the forest landscape in southeast Oklahoma has been heterogeneous with small forest patch size. This heterogeneity has resulted in the challenge of the mixed-pixel problem at the MODIS resolution (500 meters) when applying the STAARCH framework to fuse Landsat and MODIS time-series together to detect forest disturbances. Indeed, the core component of the STAARCH fusion framework was the calculation of disturbance index (DI; Healey et al. 2005) used to determine disturbance conditions of a given MODIS pixel. Accordingly, disturbance conditions of the MODIS pixel was determined by its spectral distance to a MODIS reference set composing of undisturbed pixels. In STAARCH, these reference undisturbed pixels were not separated by forest types as well as by forest compositions.

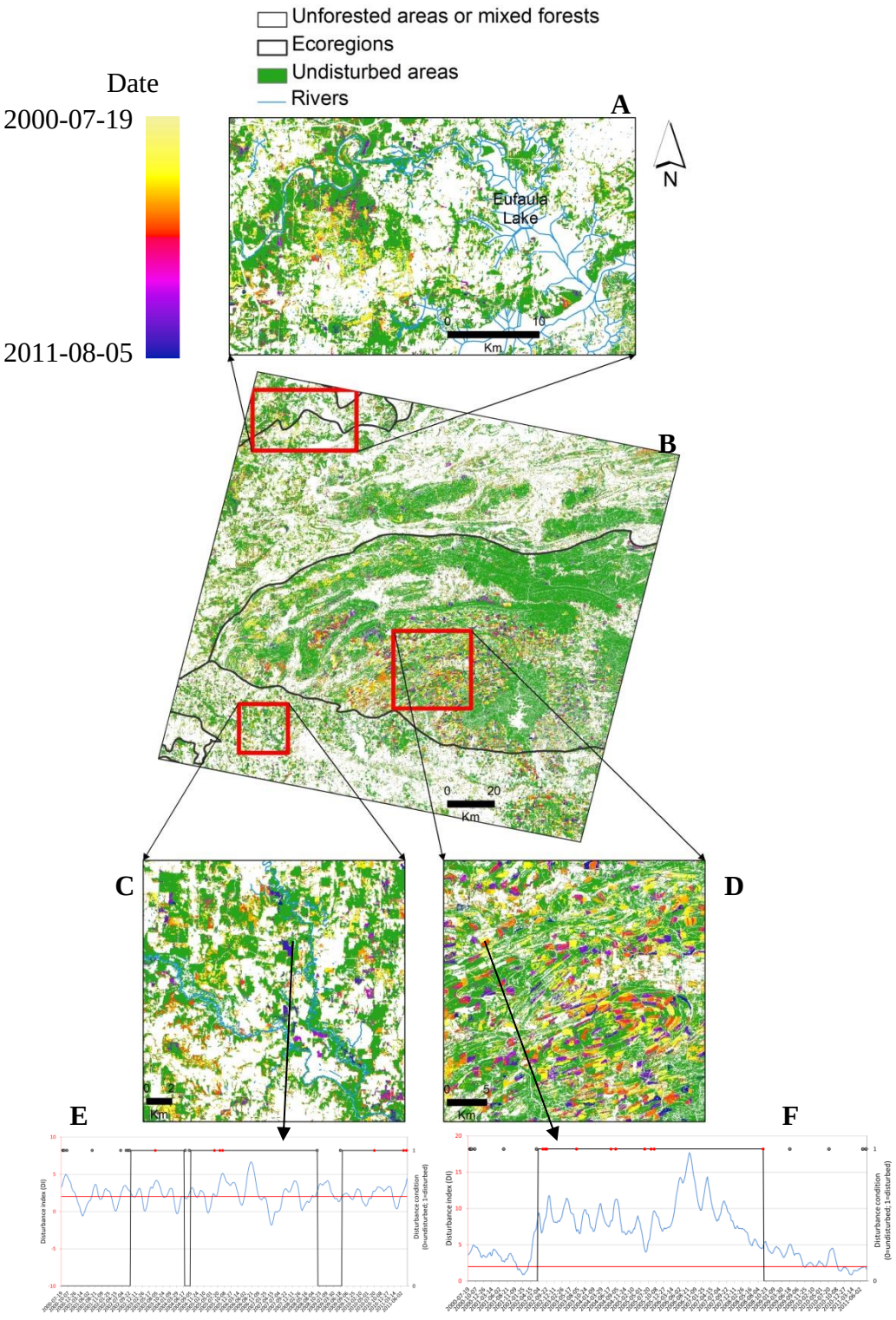


Figure 3.11. The spatio-temporal distribution of disturbances over the period of 2000-2011 in southeast Oklahoma. The beginning dates of disturbances in southeast Oklahoma (B) are represented as a color scheme ranging from yellow (2000) to blue (2011). The three map insets (A, C, D) represents disturbances around the Eufaula

Lake, west of the South Central Plains and middle of the Ouachita Mountains ecoregions, respectively. The MODIS DI time-series and MSTAARCH binary DI time-series of a forest-edge disturbance and a clear-cut disturbance are represented in chart E and F. In these charts, the blue lines represent the MODIS DI normalized by the combination reference set; the black line represent the MSTAARCH binary DI time-series; gray dots represents available Landsat images identifying the analyzed pixel as undisturbed whereas red dots represents available Landsat images identifying the analyzed pixel as disturbed.

However, I found that the spectral variability of different forest types was significantly different (e.g. tasseled cap indices of the MODIS pure deciduous reference set being significantly higher than those of the MODIS pure evergreen reference set). Consequently, the spectral distance of the MODIS pixel to different reference sets of various forest types was expected to be different. For instance, spectral distance of a disturbed evergreen pixel to the pure evergreen reference set was higher than to the pure deciduous reference set. In addition, the spectral distance was also affected by forest composition. Indeed, it was found by this study that the tasseled cap indices of the MODIS combination reference set, which composed of MODIS pixels having both deciduous and evergreen forests, were significantly higher and lower than those of the MODIS pure evergreen and MODIS pure deciduous reference sets, respectively. This mixed-pixel problem was mentioned by Healey *et al*. in their paper (2005).

To go over this mixed-pixel problem, I proposed the MSTAARCH framework to use multiple reference sets. I suggested the number of reference sets to be (2^n-1) with n to be the number of forest types. Among these reference sets, there were n pure reference sets composing of pixels covered by 100 percent of a given forest type. The other sets were combinations of forest types. In addition, pixels of these combination

reference sets were covered by 100 percent of forests even though they were combinations of multiple forest types.

The use of this approach was successfully tested for southeast Oklahoma to produce the MSTAARCH fused binary DI time-series having its accuracy of 13 percent higher than the original STAARCH approach using only one MODIS reference set. The high accuracy of the MSTAARCH fused binary DI time-series implied that the disturbance index (DI) and its common threshold of 2 first proposed by Healey *et al*. to detect disturbances from Landsat data could be used for MODIS data with high accuracy as well. However, this usage needed to account for land cover heterogeneity especially at the MODIS resolution.

It was apparent from Figure 3.6 that the capability of MSTAARCH might be reduced when detecting small disturbed patches. The reason was that when a disturbed patch was small and was surrounded by undisturbed forest, the MODIS pixel covering this disturbed patch was more likely to represent the undisturbed forest around the patch. As a result, its disturbance index was not high enough to be detected as disturbed, especially when the patch was in a regeneration process. In this case, the use of appropriate reference set was important to mitigate the problem. Therefore, compared to STAARCH, MSTAARCH was still able to detect part of the disturbed period of the patch in Figure 3.6. The problem may be profound if there were not enough Landsat images to help adjust the detection from MODIS.

Validating disturbance events at eight-day interval has been difficult due to the lack of high temporal resolution validation data (Hilker et al. 2009; Thomas et al. 2011). In this study, I utilized the available time-stamped NAIP aerial photos to validate the

fused binary time-series for specific dates. The accuracy results of this validation, therefore, were useful to not only evaluate the extent but also the time of disturbances. The total producer's accuracies provided by this validation were highest (> 75 percent) in 2005 and 2010 for all of the binary DI time-series. This could be because in these years, there were more of new disturbed patches having high contrast in reflectance between the area disturbed inside the patches and the area undisturbed outside the patches. This contrast reduced the uncertainty of the disturbance detection. In addition, the high accuracies of the three fused binary time-series were also attributed to the availability of Landsat images in 2004 and 2009. In 2006 and 2008, however, those total producer's accuracies were lowest, especially for the STAARCH and Threshold-2 binary DI time-series. This was, perhaps, because of the regeneration of disturbed patches, which resulted in low contrast between area disturbed and area undisturbed. Furthermore, the low accuracies of the three fused binary DI time-series in 2008 were also attributed to the unavailability of Landsat images in 2006 and 2007.

The NAIP total producer's accuracy of the Landsat binary DI time-series (77 percent) was within the range of previous studies relating to the detection of disturbances from Landsat (Masek et al. 2008; Thomas et al. 2011). The accuracy found by Hilker *et al*. (2009) when detecting disturbances from Landsat using the original STAARCH was between 87 percent and 89 percent. This high accuracy was perhaps due to the fact that Hilker *et al*. validated each Landsat image against a yearly cumulative disturbances data. Consequently, this validation protocol ignored the mismatches in the times of disturbances (i.e. the beginning and ending times of disturbances) due to data accumulation. Indeed, when exploring their results against

Landsat images by dates, they found that there were disturbed patches not captured by STAARCH (Hilker et al. 2009).

The use of thematic burn severity dataset for validation was quite problematic. There were a couple reasons to the problem. First, given that burnt pixels were often scattered with small patches and that there could be misregistrations between the time-series and the thematic burn severity dataset due to its projection transformation, thematic disagreements between pixels from the time-series and those from the thematic burn severity dataset were expected. Second, because the burn conditions of pixels in the thematic burn severity dataset were subjectively interpreted, there could be interpretation errors that accidently interpreted no-burnt (or low burnt) pixels as burnt (or high burnt) pixels. Even though the total producer's accuracy of the Landsat binary DI time-series (40.1 percent) in this case was not high, it was slightly higher than the findings from a relevant study (from 31.9 percent to 36.9 percent; He et al. 2011). In fact, in this study, there were years (2000, 2001, and 2010) when the producer's accuracies of the time-series went over 60 percent.

The accuracy results presented in Figure 3.9 and the trends of area disturbed presented in Figure 3.10 further confirmed the advantage of the MSTAARCH fusion framework. Compared to the STAARCH and Threshold-2 binary DI time-series, the MSTAARCH binary DI time-series had higher accuracies. In addition, compared to the Landsat binary DI time-series, the MSTAARCH binary DI time-series provided data of area disturbed for a long time period without interruption due to data lacking. Furthermore, the high temporal resolution of the MSTAARCH binary DI time-series allowed it to capture disturbances not at one specific time point but rather from the

beginning to the end of the disturbances. Subsequently, the trend of area disturbed derived from the MSTAARCH binary DI time-series was smoother and more realistic. Accordingly, both of the drops in area disturbed in 2001 and 2006 was due to the economic recessions (Hodges et al. 2011; Masek et al. 2013). The drop in 2006, however, might also be due partly to the error in disturbance detection (Figure 3.6) as a result of lacking Landsat imagery. In addition, while large disturbed patches, which was often detected as being disturbed one or two times during the study period, found in the Ouachita Mountains and southeast of the South Central Plains were due to harvesting, edge disturbed patches, which was often detected as being disturbed more than two times during the study period, found in the Arkansas Valley and western South Central Plains were due to natural disturbance agents. For instance, trees near forest edges, especially those facing south or southwest, were more prone to damage due to high wind speed and shallow soil (Harper et al. 2005). Additionally, trees in riparian forests might be killed by floods, bank erosion, or avulsions (Moore and Richardson 2012). The fact that these edge disturbed patches were often detected as being disturbed multiple times during the study period could be due to canopy shadow. For instance, forest trails should have been detected as being disturbed persistently during the entire study period. However, due to the shadows of trees along these trails, the detected disturbance conditions of these trails changed over time as shadows would be detected as undisturbed due to its low brightness.

Land cover heterogeneity has been a problem of the STAARCH fusion framework. It was because on one hand, land cover heterogeneity results in mixed MODIS pixels which might not be spectrally representative of a reference set of

undisturbed pixels. On the other hand, the STAARCH fusion framework did not propose a procedure to select MODIS pixels for a reference set. Rather, STAARCH uses all forest pixels as a reference set. This resulted in low accuracy in detecting disturbances for a heterogeneous landscape. Therefore, in this study, I proposed the MSTAARCH framework with details on developing multiple MODIS reference sets. The number of reference sets was suggested to be (2^n-1) to include both pure reference sets and combination reference sets. The MSTAARCH proved to be much more improved than the original STAARCH in detecting disturbance in southeast Oklahoma.

3.8 CONCLUSIONS

In this study, a Landsat time-series and a MODIS time-series were fused using the proposed fusion framework, MSTAARCH, which was a modification of the original STAARCH proposed by Hilker *et al*. (2009). The primary advantage of MSTAARCH was its capability of accounting for land cover heterogeneity at the MODIS resolution by using multiple MODIS reference sets to detect disturbances. Through this study I suggested the number of reference sets to be (2^n-1) where n was the number of forest types whose tasseled cap indices were significantly different from each other. Among these (2^n-1) reference sets, n of them were pure reference sets and the rest of them were combinations of the pure reference sets. This framework was successfully used for southeast Oklahoma to detect its forest disturbances between 2000 and 2011. In general, the accuracy of MSTAARCH was 77.8 percent and was 13 percent higher than the accuracy of the original STAARCH. In addition, with the advantage of high temporal resolution inherited from the MODIS time-series, the MSTAARCH binary DI time-series, unlike the Landsat binary DI time-series, did not have data gap. Consequently,

the MSTAARCH binary DI time-series would be more useful than the Landsat time-series in temporal analyses regarding disturbances at multiple temporal resolutions ranging from week to year. The success of MSTAARCH encouraged the next step of classifying disturbances into different disturbance types, such as fire and harvest, given that they have had different impacts on wildlife habitat as well as biodiversity.

3.9 ACKNOWLEDGEMENT

I want to thank Paul de Beurs for letting me use his IDL tool of MODIS Index Batch.

CHAPTER 4. MAPPING FIRE AND TIMBER HARVESTING DISTURBANCES USING HIGH TEMPORAL RESOLUTION TIME-SERIES

4.1 ABSTRACT

Mapping disturbance types at high spatial and temporal resolution is necessary to help forest managers and planners in developing informative forest management plans to balance the economic and ecological benefits of a forest ecosystem. Fire and clearcut timber harvesting have been the two most common disturbance agents occurring in the forests of the United States. This study uses the temporal characteristics (i.e. beginning time, ending time, and duration) of disturbances to classify them into fire, harvest, and other. Given that the temporal characteristics of fires, clearcuts, and other disturbance types were relatively distinct, the time-series maps of disturbance types were created with high overall accuracies (75.5 percent).

4.2 INTRODUCTION

Forest disturbances are discrete events causing abrupt changes in the structure of forest, and include fire, drought, disease, insect defoliation, land use change, and timber harvesting, with clearcut being the most common technique of harvesting (Turner 2010). Between 2000 and 2005, global forests lost 3.1 percent (over one million square kilometer) due mainly to the disturbances of fire, clearcut, and land use change (Hansen et al. 2010). During this same period, both temperate and boreal forests in the US lost 6.0 percent (120,000 square kilometers), with most of the losses attributed to fire in the western United States and timber harvesting clearcuts in the eastern US (Hansen et al. 2010; Masek et al. 2008). Fire has been the main cause of forest losses in western US

because of the increase in forest areas together with increase in droughts of the region (Westerling et al. 2006). Conversely, forests in eastern US have been used for commercial timber production in several decades (Drummond and Loveland 2010).

Satellite images have been widely used to monitor forest disturbances due to their capability of providing information about disturbances over a large area and during a long time period. Landsat time-series, for instance, have been often used because the images are available at no-cost for large areas (185x185 square kilometers), and at high spatial resolution (30 meters) (Main-Knorn et al. 2013; Masek et al. 2008; Vogelmann et al. 2009; Woodcock et al. 2008). The high spatial resolution of Landsat makes it a reliable data source to detect locations and extents of disturbances (Hansen et al. 2010; Masek et al. 2011). However, its moderate temporal resolution (16 days), which can be markedly extended to years due to cloud contamination, prevents Landsat from accurately detecting temporal characteristics of disturbances such as beginning time, ending time, duration, and frequency (Frolking et al. 2009; Ju and Roy 2008; Reynolds-Hogland and Mitchell 2007; Turner 2010). These temporal characteristics are important for characterizing disturbance types and impacts (Gillanders et al. 2008; Turner and Dale 1998).

Recently, image processing algorithms have been proposed to blend Landsat time-series with MODIS (Moderate Resolution Imaging Spectroradiometer) time-series (temporal resolution is daily, with 8-day and 16-day composites) to improve the capability of Landsat in detecting land cover changes as well as forest disturbances (Gao et al. 2006; Hilker et al. 2009). Tests and modifications of these algorithms have also been proposed and recommended (Chapter 3 of this study; Emelyanova et al. 2013;

Walker et al. 2012; Zhu et al. 2010). Despite the progress of automating the detection of forest disturbances using the blended high spatial (30 meters) and temporal (8 days) resolution time-series, little work has been done with respect to the application of these high spatio-temporal resolution time-series to classify disturbance types; most classification studies still use Landsat time-series. Therefore, there has been a need of an approach that is able to classify disturbance types based on the high spatio-temporal resolution time-series. Given that the MSTAARCH binary DI time-series can provide information about the beginning time, ending time, and duration of all disturbances within the study period over the entire study area, this study aims at determining whether the temporal characteristics (i.e. the beginning time, ending time, and duration) of disturbances are sufficient to classify disturbance types across the study area of southeast Oklahoma within the study period from 7/19/2000 to 8/5/2011.

4.3 STUDY AREA

The study area (185x185 square kilometers) comprises a Landsat scene (path 26 and row 36) that covers most of southeast Oklahoma and parts of western Arkansas and northern Texas (Figure 4.1). The area contains six Omernik level III ecoregions (Omernik 1987): the Arkansas Valley, Ouachita Mountains, South Central Plains, Cross Timbers, Central Irregular Plains, and East Central Texas Plains. Deciduous and evergreen forests cover most of the study area at 16,098 square kilometers (National Landcover Dataset; Homer et al. 2004). The area receives an average annual precipitation of 1,219 millimeters and an average annual temperature of 16.4 degree Celsius (Oklahoma Climatological Survey 2013).

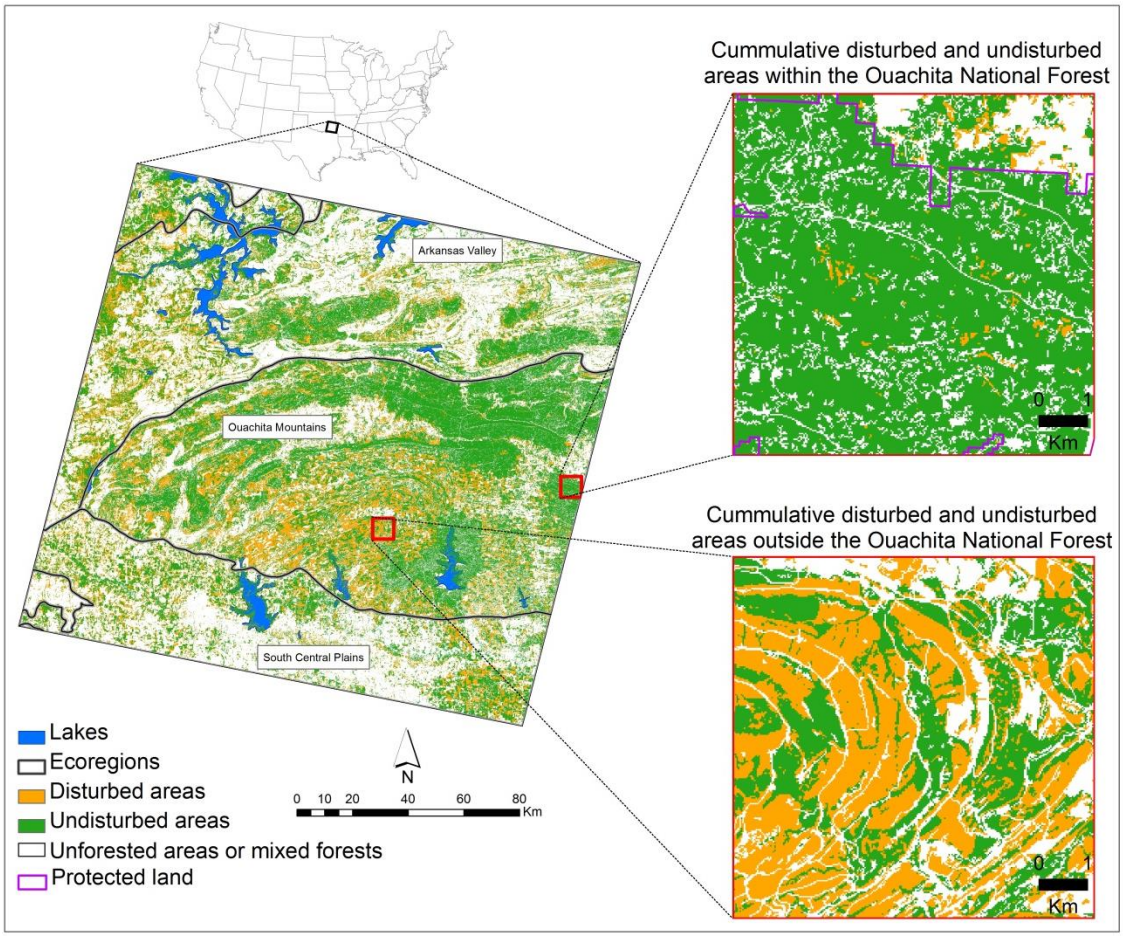


Figure 4.1. Extent and context of the 185x185 km^2 study area located in southeast Oklahoma. The two map insets on the right show the cumulative (in 12 years from 2000 to 2011) disturbed (orange) and undisturbed (green) areas within (top inset) and outside (bottom inset) of a protected area in the Ouachita National Forest. Cumulative areas are created by spatially accumulate all disturbed/undisturbed pixels across all time steps.

The dominant evergreen species in southeast Oklahoma have been shortleaf pine (*Pinus echinata*), but loblolly pine (*Pinus taeda*) has become more common with the widespread introduction of plantation forests (Harper 2010; Oklahoma Forestry Services 2010). Deciduous forests have been composed mostly of oak (*Quercus alba*, *Quercus stellata*, *Quercus rubra*) and hickory (*Carya texana*, *Carya tomentosa*) (Harper 2010; Oklahoma Forestry Services 2010; Rice and Penfound 1959). Forests in southeast Oklahoma has long been disturbed due to harvesting for commercial timber production

(Harper and Johnson 2012; Johnson 2011). Trees have been harvested on a 28-32 year rotation (Ouachita Ecoregional Assessment Team 2003). As of 2009, nearly 100 percent (46,519 thousand cubic feet) of Oklahoma roundwood products were from evergreen forests and 82 percent (29,743 thousand cubic feet) from deciduous forests in southeast Oklahoma (Johnson 2011).

It has been found that forests in the east Oklahoma region including both the northeast and southeast Oklahoma regions have been disturbed not only due to timber harvesting but also due to natural disturbance agents (i.e. disease, insect defoliation, and weather events), fire, and land use change (Harper and Johnson 2012). Of these disturbance agents, weather events especially ice storms have resulted in the highest (48 percent) area of forest land disturbed annually; fire 36.5 percent; land use change 10.2 percent, and other natural disturbances 5.3 percent. Because more than 73 percent of forests in the east Oklahoma region have situated in southeast Oklahoma (Harper and Johnson 2012), this study has assumed that the area of forest land disturbed by disturbance types in southeast Oklahoma have followed the pattern of area of forest land disturbed by disturbance type in the east Oklahoma region. Subsequently, beside timber harvesting, weather events and fires have been the most common disturbance agents in southeast Oklahoma. Additionally, forests in southeast Oklahoma could be disturbed due to tourism and/or oil and gas industries (Boyd 2002; Oil and Gas Conservation Division 2011; Oklahoma Forestry Services 2010).

4.4 DATA

The main input of this study was the MSTAARCH binary DI time-series developed in Chapter 3. This time-series was able to detect disturbances with an overall

accuracy of 96 percent when comparing with field-trip photos and a total producer's accuracy of 72.8 percent when comparing against the National Agriculture Imagery Program (NAIP) aerial photos (Chapter 3). The time-series was used in this study chapter to classify disturbances into three classes: fire, harvest, and other based on their temporal characteristics.

The fire class included both wildfires and prescribed-fires. The harvest class included all timber harvesting types (i.e. clearcut harvest, partial harvest, and commercial thinning) as defined by the US Department of Agriculture (US Department of Agriculture Forest Service 2013a). This was because as inherited from the original STAARCH, MSTAARCH detected disturbances resulting in removals of almost all forest canopy without discriminating tree qualities used to define timber harvesting types (Healey et al. 2005; Hilker et al. 2009). The other class represented all other possible disturbances including disease, insect defoliation, weather events, land use change, and forest-edge disturbances (Chapter 3; Harper and Johnson 2012). Forest-edge disturbances were mainly found at forest edges, such as those along roads, tree lines on rangeland, rivers, and lake shores.

A supervised decision tree classification was used in this study to classify disturbances into fire, harvest, and other. Therefore, it was necessary to develop reference disturbances including training and validation disturbances for each class. For fire, the reference disturbances (including training and validation disturbances) were identified based on a thematic burn severity dataset obtained from the Monitoring Trends in Burn Severity Geospatial Database (http://www.mtbs.gov). This yearly dataset available from 2000 to 2011 represented large (over two square kilometers in

the East and four square kilometers in the West) fires in the conterminous U.S (Schwind 2007). Because only moderate and high severity fires were considered in this study, other classes (unburned to very low, low, increased greenness, and non-processing mask classes) were masked out from the thematic burn severity dataset. For the harvest class, a shapefile of time-stamped clearcut patches digitized from NAIP aerial photos was used. Details about this digitizing process were presented in Chapter 3.

For the other class, because it was unknown about locations of weather events and land use changes, these disturbances were not accounted during the process of identifying reference disturbances. To identify reference disturbances for disease and insect defoliation, the Insect and Disease Detection Survey data (IDS) was used. This data was downloaded from the IDS explorer (US Department of Agriculture Forest Service 2013b) as a state-wide geodatabase about forest health of Oklahoma. Locations and extents of tree mortalities due to disease or insect defoliation were represented as polygons associated with their damage type and survey year from 2002 to 2012.

Additionally, because forest-edge disturbances could be found along roads, tree lines on rangeland, rivers, and lake shores, these features were initially considered to inform the selection of reference disturbances taking place at forest edges. However, because from the field trips, I am only certain that forests along Highway 70 (Figure 3.1) were either not disturbed or disturbed neither due to fire nor to harvest, this highway was used to inform the selection of reference disturbances taking place at forest edges. The 2013 TIGER road shapefile (US Census Bureau 2013) was used to extract this Highway 70 for the selection of reference disturbances.

4.5 METHODS

4.5.1 Disturbance temporal characteristics

The objective of this study was to classify disturbances into three classes (fire, harvest, and other) based on their temporal characteristics. Therefore, it was necessary to analyze the temporal characteristics of the three disturbance types before classification. Temporal characteristics of a disturbance included its beginning time, ending time, and duration derived from the MSTAARCH binary DI time-series (Figure 4.2). The beginning time was the date a given pixel switched from undisturbed to disturbed whereas the ending time was the date the pixel switched from disturbed back to undisturbed. Duration was the difference of the beginning and ending times.

The analysis of temporal characteristics of disturbances was conducted using their reference data. Fire reference data were selected based on yearly fire strata. A fire stratum of a given year included all fire pixels taken from the thematic burn severity dataset for that year. For instance, the fire stratum of 2000 included all fire pixels taken from the 2000-thematic burn severity dataset. For each of these fire strata, ten percent of the fire pixels were randomly selected as reference fire pixels. The temporal characteristics of these reference fire pixels were then extracted from the MSTAARCH binary DI time-series and used as a reference for fire disturbances. A total of 3,363 reference fire disturbances were identified through this selection process.

To create reference harvest disturbances, the time-stamped NAIP patches were first grouped by year and then used as yearly strata for the selection of reference clearcut pixels. Using the same selection process as selecting reference fire disturbances, a total of 4,114 reference clearcut disturbances were identified. These

reference disturbances included those starting before and ending after the study period. Although it was uncertain about the durations of these reference disturbances, they were included to classify disturbances, especially those in 2000 and 2011, either starting before or ending after the study period.

Reference data for the class of other type included reference data of disease and insect defoliation as well as reference data of forest-edge disturbances. The reference data of disease and insect defoliation was developed from the Insect and Disease Detection Survey data (IDS). Polygons representing the extents of yearly tree mortalities were used as yearly strata to select reference pixels. Because the number of reference pixels in this case was too small (20 pixels), all of them were used as reference pixels for the identification of reference disturbances. Additionally, the reference data of forest-edge disturbances were created by first creating a 300-meter buffer around Highway 70 as a stratum to select reference pixels. Using this buffer as a stratum, twenty percent of all disturbed pixels in this stratum were randomly selected for the identification of reference disturbances. Using twenty percent in this case was to make sure that the number of reference disturbances in this class was close to the number of reference disturbances in the other two classes. A total of 2,743 reference disturbances was identified for the class of other disturbance types. Again, these disturbances included those starting before and ending after the study period.

The three temporal characteristics of disturbances were then statistically tested between classes to determine whether they were significantly different. The Mann-Whitney *U* test was used because (1) the beginning and ending times of disturbances were ordinal data and (2) duration of disturbances was not normally distributed. The

hypothesis for this test was that there was no difference in each of the temporal characteristics between any two classes.

4.5.2 Disturbance type mapping and validation

To map disturbance types, a classification was performed for every disturbance of all pixels across the study area based on a rule set constructed by a classification tree. In this study, the CHAID (CHi-squared Automatic Interaction Detection; Kass 1980) classification tree method available in the SPSS® software package was used to build the rule set. CHAID was used to automatically partition the independent variables (i.e. beginning time, ending time, and duration) into mutually exclusive groups that best described the dependent variable (i.e. classification type).

I used 70 percent of the reference dataset as training data for each disturbance type. Thus, the number of training disturbance pixels for the fire class was 2,354; 2,879 for clearcut; and 1,920 for other disturbance types. The other 30 percent of the reference data were used to validate the disturbance type maps. To validate the disturbance type maps, this study used the error matrix (Congalton 1991). The overall accuracy, kappa, producer's and user's accuracies were calculated based on the agreements and disagreements between the reference and classified disturbances.

4.6 RESULTS

4.6.1 Temporal characteristics of disturbances

Temporal characteristics of fire, clearcut, and other disturbance types were found to be visually and statistically different (Figures 4.2 and 4.3). It was found that durations of disturbances as a result of fires were often shorter than those of harvestings, although there were exceptions when fires were large enough to destroy a

wide area (e.g. 12 square kilometers) or when fires were followed by clearcuts (Figure 4.2 A). Although fires and harvestings often occurred at the same pixels one or two times during the study period, other disturbances could happen more than two times and thus could result in shorter durations (Figure 4.2 C). These findings were confirmed by the temporal distribution presented in Figure 4.3.

This study found that there were 75 percent of fires ended in 200 eight-day time steps (about four years) and 18 percent of fires ended between 240-300 eight-day time steps (about five to five-and-a-half years) (Figure 4.3 A). Besides, there were 99.7 percent of harvest disturbances ended in 340 eight-day time steps (about seven-and-a-half years; Figure 4.3 B). Other disturbances, such as forest-edge disturbances, often occurred several times at the same pixels during the study period. Thus, most (78.5 percent) of their durations were short and less than 100 eight-day time steps (about two years; Figure 4.4 C). Furthermore, there were 34 percent of all three types of disturbances happened in July (Figure 4.3 D, E, F). Fires, however, were also found frequently in August (Figure 4.3 D). While not many fires occurred in winter months, most (91 percent) harvest disturbances occurred between March and July (Figure 4.3 E). Other disturbances, in contrast, were found regularly between June and October (Figure 4.3 F).

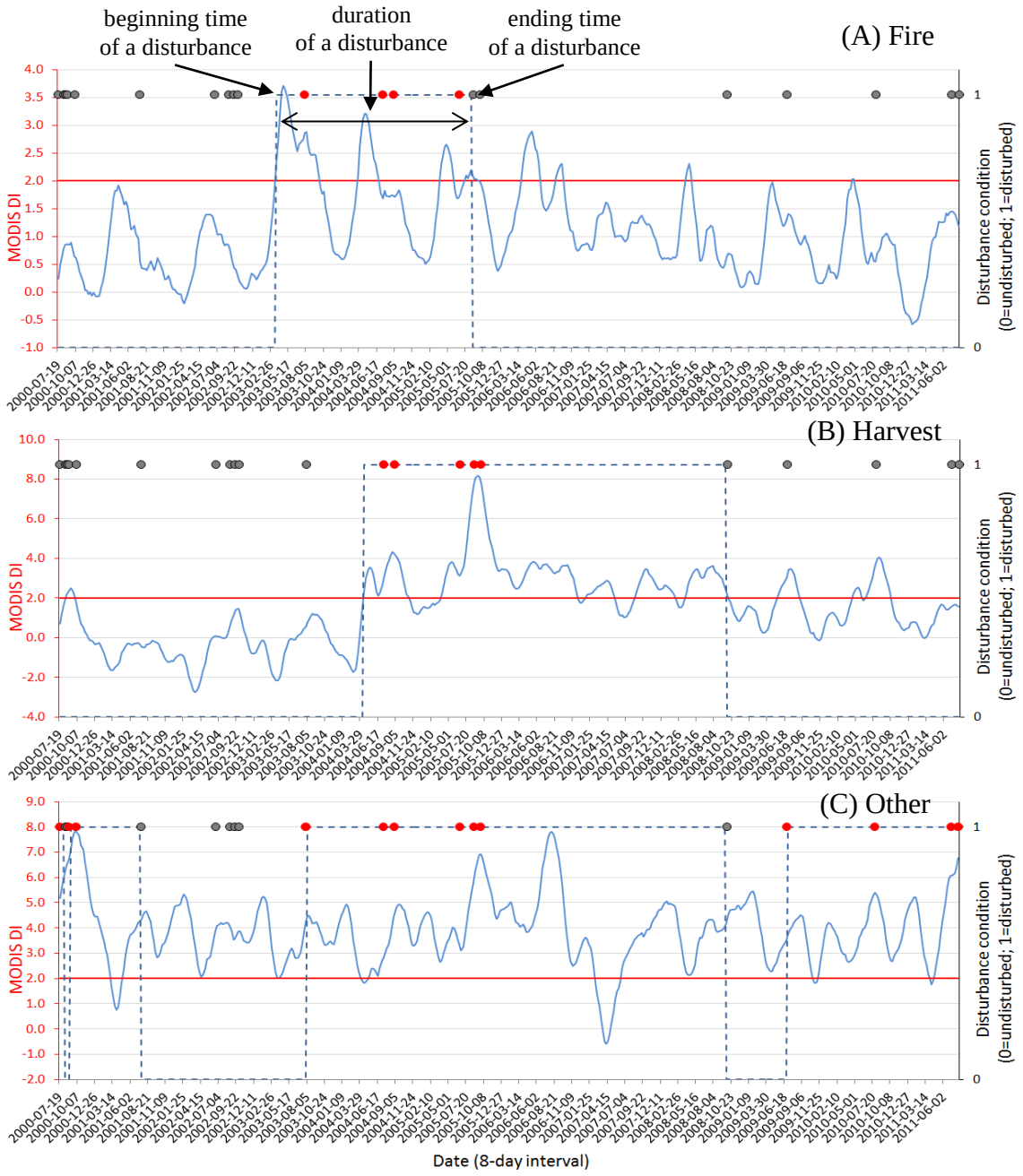


Figure 4.2. Example profiles of fire (A) and harvest (B) disturbances as well as other disturbances (C). Red solid lines represent the threshold of 2 used to determine disturbance conditions of a given pixel at a given time step. Solid blue lines represent MODIS DI profiles of the given pixel. Dotted blue lines represent disturbance conditions detected by MSTAARCH of the given pixel. Gray dots represent available Landsat images identifying the given pixel as undisturbed whereas red dots represent available Landsat images identifying the given pixel as disturbed.

Of the twelve months, August was the most common time all types of disturbances ended (Figure 4.3 G, H, I). September was also the time when areas disturbed due to harvesting became undisturbed (Figure 4.3 H). The Mann-Whitney *U* tests were performed for each temporal characteristic between classes and showed that these characteristics were significantly different between classes (Table 4.2).

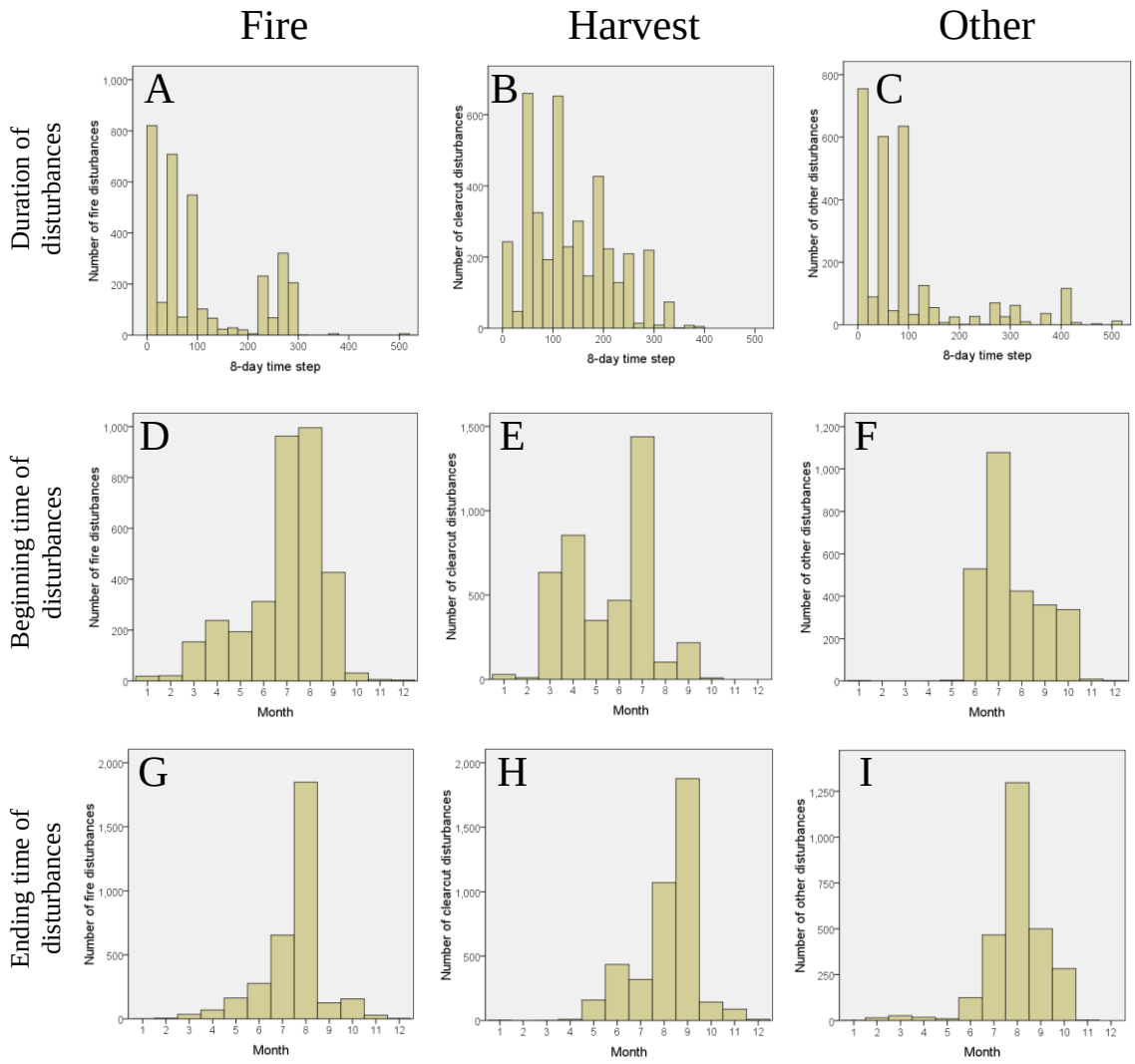


Figure 4.3. Histograms of duration (top row) as well as beginning (middle row) and ending (bottom row) times of disturbances for fires (left column), harvest (center column), and other disturbances (right column).

Table 4.1. Summary of the Mann-Whitney *U* test. This table shows that the temporal characteristics are significantly different between classes.

Temporal character-istics	Descriptive statistics *Mean (standard deviation)*			Mann-Whitney U test results		
	Fire n = 3,364	Harvest n = 4,114	Other n = 2,743	Fire vs. Harvest	Harvest vs. Other	Fire vs. Other
Beginning time	6.9 (1.7)	5.6 (1.8)	7.6 (1.3)	$U = 3.9 \times 10^{-6}$ Z = -33.4 p<0.001	$U = 2.4 \times 10^{-6}$ Z = -42.0 p<0.001	$U = 4.0 \times 10^{-6}$ Z = -9.4 p<0.001
Ending time	7.5 (1.3)	8.2 (1.3)	8.0 (1.2)	$U = 4.4 \times 10^{-6}$ Z = -27.9 p<0.001	$U = 4.9 \times 10^{-6}$ Z = -9.7 p<0.001	$U = 3.5 \times 10^{-6}$ Z = -17.2 p<0.001
Duration	104.1 (99.4)	134.6 (81.3)	94.5 (109.1)	$U = 4.8 \times 10^{-6}$ Z = -22.3 p<0.001	$U = 3.3 \times 10^{-6}$ Z = -29.7 p<0.001	$U = 4.4 \times 10^{-6}$ Z = -2.7 p=0.006

4.6.2 Map of disturbance types

The map of disturbance types was created by classifying the temporal characteristics of disturbances (Figure 4.4). The overall accuracy of the classification was high (75.5 percent, kappa = 62.7 percent; Table 4.1). Among the three classes, harvest disturbances were best classified (producer's accuracy = 85.9 percent; user's accuracy = 81.4 percent). However, harvest disturbances were frequently misclassified with fires. The classes of fire and other were often misclassified with each other and resulted in their low producer's accuracies (66.7 percent for fire and 70.8 percent for other; Table 4.1).

Table 4.2. Error matrix.

	Reference			User's accuracy
	Fire	Harvest	Other	
Fire	1,570	239	287	74.9%
Harvest	291	2,472	273	81.4%
Other	493	168	1,360	67.3%
Producer's accuracy	66.7%	85.9%	70.8%	
Overall accuracy		75.5%		
Kappa		62.7%		

A visual analysis between the classification result and NAIP aerial photos was also conducted to better understand the classification performance (Figure 4.5). In Figure 4.5, three typical landscapes of fire (Figure 4.5 A, D), clearcut harvest (Figure 4.5 B, E), and riparian forest (Figure 4.5 C, F) representing the three classes, respectively, were presented. Figure 4.5 D showed the misclassification between fire and harvesting for a fire on 8/5/2006 (black arrow in Figure 4.5 D). Figure 4.5 E demonstrated that although there were misclassifications between harvest disturbances and other disturbances (black arrow in Figure 4.5 E), harvest disturbances were still well identified by the classification (blue arrow in Figure 4.5 E). Additionally, the classification was also useful in detecting forest edges of clearcut patches (yellow arrow in Figure 4.5 E). Figure 4.5 F pointed out that the classification well identified disturbances along riparian forests and correctly considered them as other disturbances.

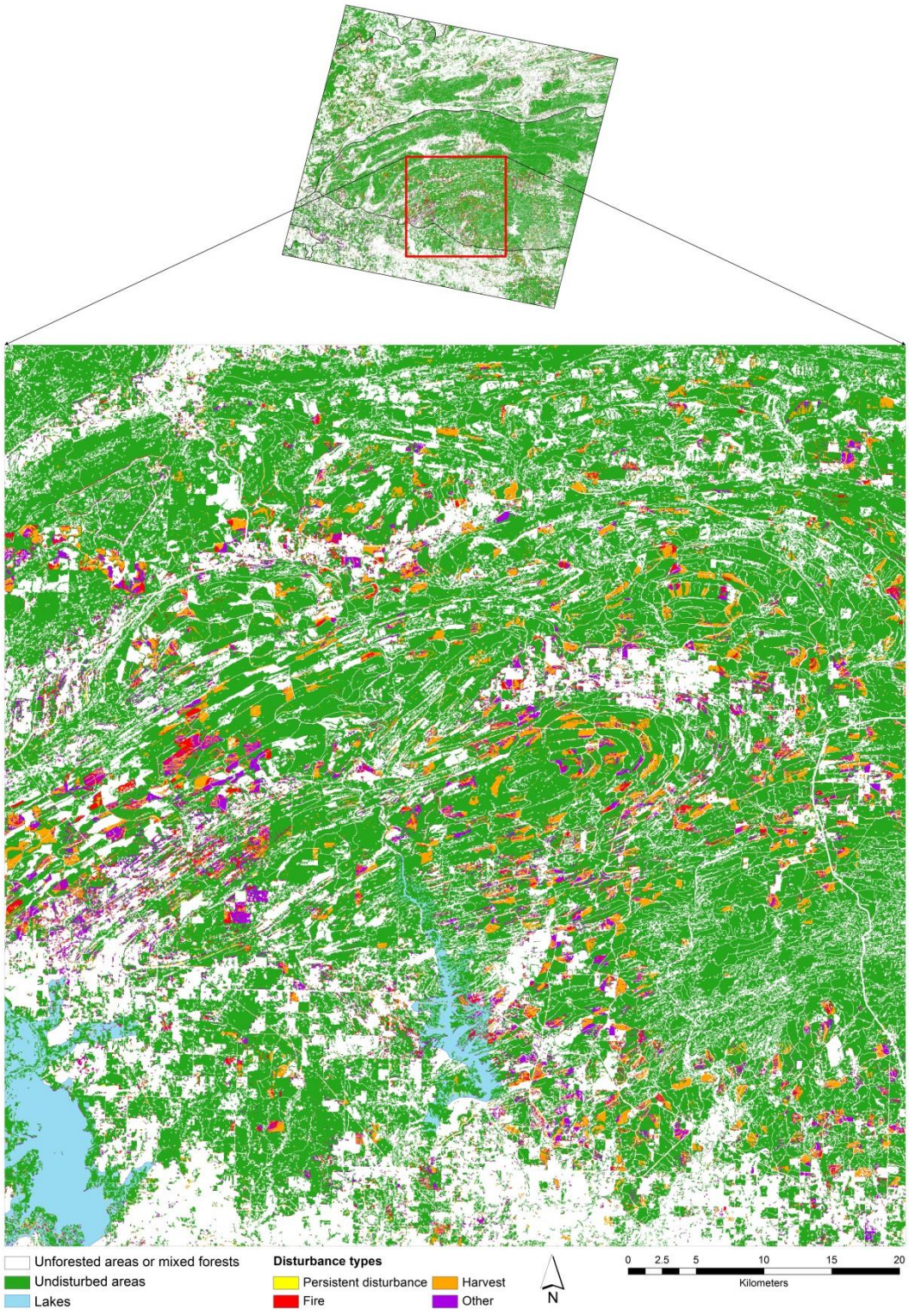


Figure 4.4. Map of disturbance types on 8/29/2005.

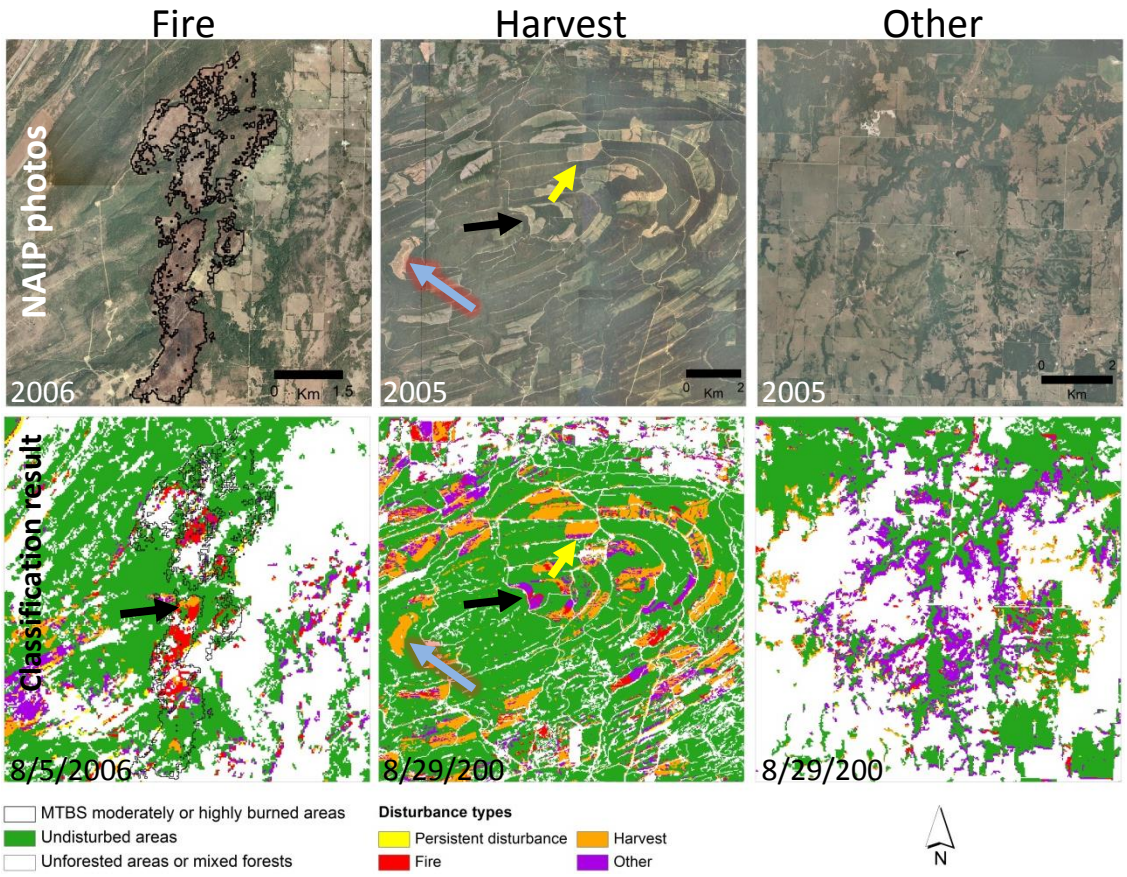


Figure 4.5. Classification results for fire, harvest, and other-disturbance landscapes. The first row represents NAIP aerial photos showing a 2006-fire, 2005-harvest, and 2005-other disturbances. The second row represents the classification results. Black arrow points out a misclassification of a clearcut harvest; blue arrow for correct classification of a clearcut harvest; yellow arrow for a correct classification of a clearcut harvest and its edge. The MTBS moderately or highly burned areas are taken from the thematic burn severity dataset.

4.7 DISCUSSIONS

This study aimed at classifying disturbance types based on their temporal characteristics, extracted from the MSTAARCH binary DI time-series developed in Chapter 3. The high accuracies of the classification (Table 4.1) demonstrated that with high temporal resolution (i.e. 8 days) it was possible to distinguish disturbance types (i.e. fire, harvest, and other disturbances) using their temporal characteristics. That the

temporal resolution was high was important and required in this study because at high temporal resolution, variations in temporal characteristics between disturbance types were better detected and used to classify disturbance types.

The histogram in Figure 4.3 showed that there were very small amounts of disturbances in winter months (from November to February; Figure 4.3 D, E, F). One of the reasons was because of the artifact of the MSTAARCH method used to detect disturbances. Indeed, the MSTAARCH as inherited from STAARCH (Hilker et al. 2009) was not able to detect disturbances during winter months when all leaves were off across the entire study area. When this happened, the reference sets against which other pixels were compared in order to determine their disturbance conditions would have the same disturbance index as the index of disturbed pixels. This issue was expected to be prominent in deciduous forests where fires and other disturbances often occurred because they were more sensitive to seasonal effect than evergreen forests. In evergreen forests this issue could also happen although evergreen trees were not supposed to be as sensitive to seasonal effect. The reason was that there were cases where evergreen trees were mixed with deciduous trees in areas delineated as evergreen forests. Because evergreen forests where most clearcuts occurred were less sensitive to seasonal effect, harvest disturbances were still well captured in March (Figure 4.3 E). I also found that many disturbances occurred in summer months especially in July and August (Figure 4.3 D, F). This was partly because these months often had lowest precipitation (Oklahoma Climatological Survey 2012) leading to droughts and thus, resulting in mortalities of deciduous forests. Furthermore, droughts in these months could also

result in high chance of fires, insects and diseases in deciduous forests (Zhang et al. 2013).

Because the temporal characteristics of fire, harvest, and other disturbances were significantly different from each other, the classifications were successfully conducted to produce maps of disturbance types with high accuracies (Table 4.1, Figure 4.5). In other words, this study found that fire, harvest, and forest-edge disturbances could be successfully recognized using their temporal characteristics. Indeed, the overall accuracy of the classification was 75.5 percent, which satisfied the accuracy threshold proposed by (Goodchild et al. 1994). However, this accuracy could have been improved if ancillary data had been used. For instance, other studies showed that clearcut disturbances were larger and more irregularly shaped than fire disturbances (Cohen and Goward 2004; Gluck and Rempel 1996). In addition, clearcut patches were also suggested to be close to roads providing access to logging areas (Market Segment Specialization Program 2013; Potapov et al. 2008). It was also found that locations of fires could be a function of slope, aspect, and valley orientation (McRae et al. 2001).

Although the accuracies were high (Table 4.1), there were apparent classification errors in Figure 4.4 and 4.5. First, part of these classification errors was expected to be from a portion of the 39 percent of disturbances starting before or ending after the study period. Classification of these disturbances was expected to be confusing because their temporal characteristics were uncertain. Second, in this study area, beside fires and harvests, other disturbance types often included forest-edge disturbances mainly found along roads, tree lines on rangeland, rivers, and lake shores (Chapter 3). These disturbances could be real disturbances due to high wind speed or shallow soil

often found at the edge of forests or due to floods, bank erosion, or avulsions along riparian forests (Harper et al. 2005; Moore and Richardson 2012). However, other disturbance types could also be false changes resulting from tree shadows or fluctuation of water level in wetland areas (Schroeder et al. 2011). Therefore, it was expected that other disturbance types would be misclassified as either fires or harvests and thus, resulted in classification errors. Although other disturbance types caused omission and commission errors of fire and harvest classes, those other disturbance types mainly representing forest-edge disturbances were well detected along forest edges such as those along riparian forests in the South Central Plains ecoregion (Figure 4.5 F). This finding suggested that the classification accuracy could have been improved if the classification had been applied to multiple subregions, such as ecoregions, that had been relatively homogeneous of a disturbance type to reduce possible confusions between disturbance types during a classification.

The success of using the temporal characteristics of disturbances to classify fires, harvests, and other disturbance types indicates that the MSTAARCH binary DI time-series is useful to detect not only the spatial properties (location and extent) and temporal properties (beginning time, ending time, and duration) of disturbances but also type of disturbances. This success recommends that a fused time-series can be used beside a Landsat time-series in monitoring and modeling forest disturbances. In fact, the advantage of high temporal resolution of the fused time-series is helpful in determining the right temporal extent and grain of an ecological process (Reynolds-Hogland and Mitchell 2007). Additionally, the advantage of high temporal resolution of a time-series of disturbance types may be useful for forest models, especially processed based models

(Fontes et al. 2010). Although the accuracy is high, the use of temporal characteristics to classify disturbance types is still challenged by the uncertainly in determining the beginning time and ending time of disturbances starting before or ending after the study period. The challenge also exists due to the overlaps of temporal characteristics between classes. However, these challenges can be mitigated by using ancillary data and/or by conducting the multiple classifications for multiple subregions relatively homogeneous in disturbance type.

4.8 CONCLUSION

In this study, the differences in temporal characteristics (i.e. beginning time, ending time, and duration) of fire, harvest, and other disturbance types were investigated. It was found that their temporal characteristics were significantly different and thus could be used as inputs of a classification to discriminate disturbance types based on the time-series of disturbances developed in the previous chapter. The classification result demonstrated that the temporal characteristics of disturbances could be used to classify disturbances into fires, harvests, and other types with high accuracy.

Forest disturbances have caused forest landscapes to be fragmented into multiple smaller undisturbed patches disconnected by those disturbed. Subsequently, disturbances have resulted in habitat degradation as well as biodiversity decline and have increased potential of insect outbreak as well as fire (Gillanders et al. 2008; Nepstad et al. 1999; Sharitz et al. 1992). Fire and clearcut, however, have had different impacts on both the carbon dynamics and wildlife habitat as well as biodiversity (Taylor et al. 2013; Williams et al. 2012). Because this study pointed out that disturbance types could be mapped with high spatial and temporal resolutions, future studies would be

suggested to identify the spatial and temporal patterns of disturbance types and to quantify the environmental consequences of those patterns.

CHAPTER 5. CONCLUSIONS

Land cover and land use are changing due to both natural and anthropogenic forces. Changes in both land cover and land use cause changes in global and regional climate, losses of biodiversity, degradations of water and air quality, and increases of infectious disease. Remote sensing plays a critical role in long-term monitoring of land cover over a large area but our strategies for handling remote sensing data are not yet fully developed. Typically, although there are many land cover classification methods currently available, their accuracy is still not perfect due to the impact of land cover heterogeneity. The higher the heterogeneity is, the less accuracy the classification is. Land cover heterogeneity may also diminish the accuracy of a data fusion algorithm used to create a high spatio-temporal resolution time-series of images to monitor land cover change. It is because the data fusion requires a selection of reference pixels representative of unchanged land cover types. However, due to land cover heterogeneity, multiple land cover types having different changing regimes may coexist within pixels leading to the uncertainty in the selection of unchanged pixels.

Prior to this research, a systematically quantitative analysis of the impact of land cover heterogeneity on both per-pixel and subpixel classification for a large area consisting of multiple land cover types had yet to be done. In addition, there was no study testing and modifying the Spatial Temporal Adaptive Algorithm for mapping Reflectance Change (STAARCH) fusion method for a heterogeneous forest landscape where both deciduous and evergreen forests could be intermixed. The overall goal of this dissertation, therefore, included a documentation of the impact of land cover heterogeneity on both per-pixel and subpixel classifications at a medium spatial

resolution of Landsat (30 meters) as well as a proposal of a data fusion method that accounts for land cover heterogeneity in the South-Central US. Three research objectives were formulated to address this goal:

- To explore the impact of land cover heterogeneity on image classification approaches including per-pixel and subpixel classifications with a case study area of 10,000 square kilometers centered at Little Rock (Arkansas);
- To propose a data fusion method that is able to produce high spatial (30 meters) and temporal resolution (8 days) time-series for a disturbance detection that takes into account the heterogeneity of a forest landscape in southeast Oklahoma; and
- To determine whether the temporal characteristics of disturbances are sufficient to distinguish disturbance types (e.g. fire and harvest) across the study area of southeast Oklahoma within the study period from 7/19/2000 to 8/5/2011.

This PhD research has successfully accomplished each of these three objectives. The first objective is presented in the second chapter of this dissertation. In this chapter, both per-pixel and subpixel classifications were performed for a 10,000 square kilometer study area centered at Little Rock (Arkansas) to create land cover maps of cropland, artificial surface, barren, tree, grassland/shrub, and water. While the output land cover map of the per-pixel classification was a categorical map, the output land cover map of the subpixel classification was a fractional map. The accuracies of these land cover maps were both high (81.87 percent and 82.28 percent for per-pixel and subpixel, respectively). Further statistical comparisons between the producer's accuracies of per-pixel and the producer's accuracies of subpixel classifications as well as the comparisons between the user's accuracies of per-pixel and the user's accuracies

of subpixel classifications demonstrated that it was not clear to conclude whether per-pixel or subpixel classification was better for the study area. Moreover, the analysis of producer's and user's accuracies for each classification by degrees of heterogeneity revealed that the accuracies decreased significantly when the degree of heterogeneity increased from one to two. However, the differences in accuracies between the degrees of three, four, and five were not significant.

The concern about land cover heterogeneity learnt from Chapter 2 was carried over to the third chapter when solving the second objective. In this third chapter, a test of the STAARCH fusion method was conducted for a forest landscape in southeast Oklahoma. Because this STAARCH method used only one reference set to determine the disturbance conditions of MODIS pixels, the method was not successful to fuse Landsat and MODIS time-series for a detection of forest disturbances in southeast Oklahoma. Therefore, another framework of selecting reference sets for STAARCH was introduced in this chapter. As a result, a new fusion method (MSTAARCH) was proposed. In this MSTAARCH method, instead of using one reference set, it was suggested to use multiple reference sets to represent undisturbed pure and undisturbed mixed pixels.

With a time-series of disturbance maps created in Chapter 3 through the MSTAARCH fusion method, this fourth chapter solved the third objective by analyzing the temporal characteristics of a variety of disturbance types including fire, harvest, and other types. The analysis demonstrated that the temporal characteristics were significantly different and thus could be used to create a time-series of disturbance type maps. Disturbance type maps were created by classifying disturbances based on their

temporal characteristics using a classification tree method. The high classification accuracy (75.5 percent) indicated that the temporal characteristics were sufficient for a classification of disturbance types. However, solutions, such as using ancillary data to inform the classification, could be applied to even improve the classification accuracy.

In conclusion, this dissertation demonstrates that when the spatial resolution is coarse, classification accuracy is easy to be affected by land cover heterogeneity. Furthermore, land cover heterogeneity may also diminish the accuracy of a data fusion method if it does not separate pure pixels and mixed pixels during the fusion process. Moreover, using a high temporal resolution time-series of disturbances, a classification of disturbance types can be performed well with high accuracy using their temporal characteristics derived from the time-series.

SUGGESTIONS FOR FUTURE RESEARCH

While this dissertation has a number of research contributions, some specific research issues require further study as well as a large number of potential research applications exist. The research issues and applications most relevant are suggested below.

Subpixel classification has been claimed to be more advantageous than per-pixel classification because subpixel classification is able to solve the mixed-pixel problem (DeFries et al. 2000; Foody 2006). However, the analysis in the second chapter about the impact of heterogeneity on per-pixel and subpixel classifications demonstrated that it was not clear to conclude which classification was better for the 10,000-square-kilometer study area around Little Rock (Arkansas). This could be because of two reasons. First, the study area was too heterogeneous and thus resulted in a reduction of

classification accuracy for both per-pixel and subpixel classifications. Second, the Landsat resolution was perhaps fine enough to weaken the mixed-pixels problem and thus the contrast between per-pixel and subpixel classifications were not very high. A follow-up question was that whether per-pixel or subpixel classification would be better to characterize land cover changes for a larger area, such as the Arkansas Red River Basin.

The Arkansas Red River Basin occupies an area of about 583,660 square kilometers going across the Arkansas, Colorado, Kansas, Louisiana, Missouri, New Mexico, Oklahoma, and Texas states (Figure 5.1). It goes across different land cover conditions including the one specifically to the Rocky Mountain environment in the West to the one specifically to the flood plain of the Mississippi River in the East. This basin is expected to include a large variety of land cover types. In addition, creating historical maps of land cover types for this basin using Landsat data is necessary given that its land cover is changing due to the westward movement of population.

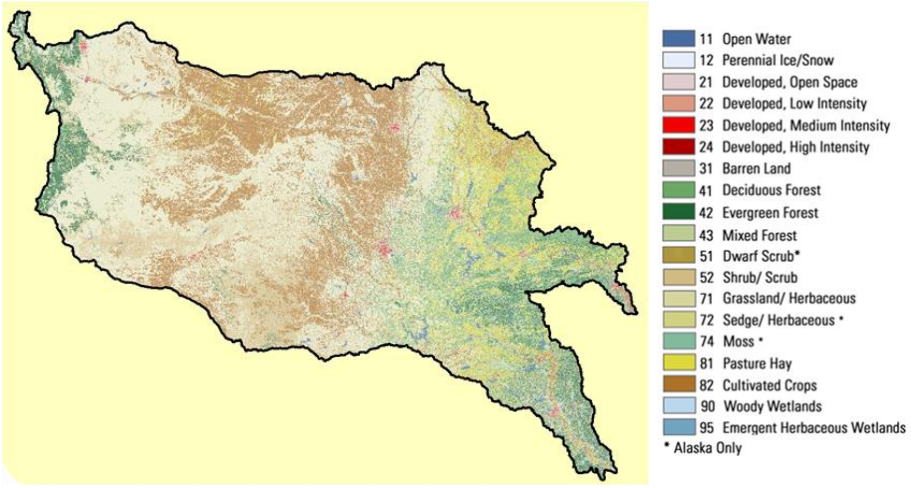


Figure 5.1. The Arkansas Red River Basin and its land cover distribution according to the National Land Cover Database 2006.

In the fourth chapter, although the classification of disturbance types was successful with high accuracy, apparent classification errors still existed. These errors were mostly due to the confusion between fire or harvest and other disturbance types. Further study should consider using spatial characteristics of disturbances, such as patch size, shape, and location together with temporal characteristics to better classify disturbance types. Moreover, a specifically well-suited application for the MSTAARCH method proposed in the third chapter and the disturbance classification in the fourth chapter is to investigate the spatial pattern (i.e. composition and configuration) of disturbance types in southeast Oklahoma. This investigation helps to understand the causes and consequences of each disturbance type and thus contributes to the Oklahoma forest resource assessment (Oklahoma Forestry Services 2010).

In summary, based on the findings of this research, a variety of future applications may be developed. One of them can be a development of historical maps of land cover types for the Arkansas Red River Basin, an area highly dynamic in terms of land cover. These historical maps may be created by either per-pixel or subpixel classifications. A test to determine which one of them is better may be conducted given that subpixel classification is claimed to be more advantageous than per-pixel classification for a heterogeneous landscape. Another application is the use of MSTAARCH and the disturbance classification proposed in the fourth chapters to investigate the spatial pattern of forest disturbance types in southeast Oklahoma. The result of this investigation helps to better assess forest resources in Oklahoma.

REFERENCES

Adkins, Z. (2009). NAIP 2008 Absolute Ground Control: From the Ground Up. In: U.S. Department of Agriculture

Aplin, P. (2006). On scales and dynamics in observing the environment. *International Journal of Remote Sensing, 27*, 2123-2140

Asner, G.P. (1998). Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment, 64*, 234-253

Asner, G.P., Keller, M., & Silva, J.N.M. (2004). Spatial and temporal dynamics of forest canopy gaps following selective logging in the eastern Amazon. *Global Change Biology, 10*, 765-783

Atkinson, P.M. (2004). Resolution Manipulation and Sub-Pixel Mapping. In S.M.d. Jong & F.D.v.d. Meer (Eds.), *Remote Sensing Image Analysis: Including the Spatial Domain* (pp. 51-70): Kluwer Academic Publishers

Berezowski, T., Chormański, J., Batelaan, O., Canters, F., & Van de Voorde, T. (2012). Impact of remotely sensed land-cover proportions on urban runoff prediction. *International Journal of Applied Earth Observation and Geoinformation, 16*, 54-65

Boyd, D.S., & Foody, G.M. (2011). An overview of recent remote sensing and GIS based research in ecological informatics. *Ecological Informatics, 6*, 25-36

Boyd, D.T. (2002). Map of Oklahoma Oil and Gas Fields. In. Norman, Oklahoma: Oklahoma Geological Survey

Campbell, J.B., & Wynne, R.H. (2011). *Introduction to Remote Sensing*: The Guilford Press

Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., & Eklundh, L. (2004). A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sensing of Environment, 91*, 332-344

Cohen, W.B., & Goward, S.N. (2004). Landsat's Role in Ecological Applications of Remote Sensing. *BioScience, 54*, 535-545

Comber, A., Fisher, P., Brunsdon, C., & Khmag, A. (2012). Spatial analysis of remote sensing image classification accuracy. *Remote Sensing of Environment, 127*, 237-246

Congalton, R.G. (1991). A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote Sensing of Environment, 37*, 35-46

Congalton, R.G. (1994). Accuracy assessment of remotely sensed data: future needs and directions. In, *Pecora 12 land information from space-based systems* (pp. 383-388): ASPRS

Crist, E.P. (1985). A TM Tasseled Cap Equivalent Transformation for Reflectance Factor Data. *Remote Sensing of Environment, 17*, 301-306

Cross, A.M., Settle, J.J., Drake, N.A., & Paivinen, R.T.M. (1991). Subpixel measurement of tropical forest cover using AVHRR data. *International Journal of Remote Sensing, 12*, 1119-1129

DeFries, R.S., Hansen, M.C., & Townshend, J.R.G. (2000). Global continuous fields of vegetation characteristics: a linear mixture model applied to multi-year 8 km AVHRR data. *International Journal of Remote Sensing, 21*, 1389-1414

Department of the Interior (2013). Product Guide: Landsat Climate Data Record (CDR) Surface Reflectance. In: U.S. Geological Survey

Drummond, M.A., & Loveland, T.R. (2010). Land-use pressure and a transition to forest-cover loss in the eastern United States. *BioScience, 60*, 286-298

Echeverría, C., Newton, A., Nahuelhual, L., Coomes, D., & Rey-Benayas, J.M. (2012). How landscapes change: Integration of spatial patterns and human processes in temperate landscapes of southern Chile. *Applied Geography, 32*, 822-831

Emelyanova, I.V., McVicar, T.R., Van Niel, T.G., Li, L.T., & van Dijk, A.I.J.M. (2013). Assessing the accuracy of blending Landsat–MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics: A framework for algorithm selection. *Remote Sensing of Environment, 133*, 193-209

ENVI (2009). Atmospheric Correction Module User's Guide. In

Fahsi, A., Tsegaye, T., Tadesse, W., & Coleman, T. (2000). Incorporation of digital elevation models with Landsat-TM data to improve land cover classification accuracy. *Forest Ecology and Management, 128*, 57-64

Foley, J.A., Defries, R., Asner, G.P., Barford, C., Bonan, G., & al., e. (2005). Global Consequences of Land Use. *Science, 309*, 570-574

Fontes, L., Bontemps, J.-D., Bugmann, H., Van Oijen, M., Gracia, C., Kramer, K., Lindner, M., Rötzer, T., & Skovsgaard, J.P. (2010). Models for supporting forest management in a changing environment. *Forest Systems, 3*, 8-29

Foody, G.M. (1999). The Continuum of Classification Fuzziness in Thematic Mapping. *Photogrammetric Engineering & Remote Sensing, 65*, 443-451

Foody, G.M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment, 80*, 185-201

Foody, G.M. (2005). Local characterization of thematic classification accuracy through spatially constrained confusion matrices. *International Journal of Remote Sensing, 26*, 1217-1228

Foody, G.M. (2006). Sub-Pixel Methods in Remote Sensing. In S.M.D. Jong & F.D.V.d. Meer (Eds.), *Remote Sensing Image Analysis: Including The Spatial Domain* (pp. 37-49): Springer Netherlands

Foody, G.M., & Cox, D.P. (1994). Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *International Journal of Remote Sensing, 15*, 619-631

Foody, G.M., & Doan, H.T.X. (2007). Variability in Soft Classification Prediction and its implications for Sub-pixel Scale Change Detection and Super Resolution Mapping. *Photogrammetric Engineering & Remote Sensing, 73*, 923-933

Frolking, S., Palace, M.W., Clark, D.B., Chambers, J.Q., Shugart, H.H., & 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*, G00E02

Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N., & Wickham, J. (2011). Completion of the 2006 National Land Cover Database for the Conterminous United States. *Photogrammetric Engineering & Remote Sensing, 77*, 858-864

Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing, 44*, 2207-2218

Gillanders, S.N., Coops, N.C., Wulder, M.A., Gergel, S.E., & Nelson, T. (2008). Multitemporal remote sensing of landscape dynamics and pattern change: describing natural and anthropogenic trends. *Progress in Physical Geography, 32*, 503-528

Gluck, M.J., & Rempel, R.S. (1996). Structural characteristics of post-wildfire and clearcut landscapes. *Global to Local: Ecological Land Classification* (pp. 435-450): Springer

Goodchild, M.F., Biging, G.S., Congalton, R.G., Langley, P.G., Chrisman, N.R., & Davis, F.W. (1994). *Final Report of the Accuracy Assessment Task Force*. , California Assembly Bill AB1580, Santa Barbara, University of California, National Center for Geographic Information and Analysis (NCGIA)

Hansen, M.C., Stehman, S.V., & Potapov, P.V. (2010). Quantification of global gross forest cover loss. *Proceedings of the National Academy of Sciences, 107*, 8650-8655

Hansen, M.C., Stehman, S.V., Potapov, P.V., Loveland, T.R., Townshend, J.R.G., DeFries, R.S., Pittman, K.W., Arunarwati, B., Stolle, F., & Steininger, M.K. (2008).

Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data. *Proceedings of the National Academy of Sciences, 105*, 9439-9444

Harper, K.A., Macdonald, S.E., Burton, P.J., Chen, J., Brosofske, K.D., Saunders, S.C., Euskirchen, E.S., Roberts, D.A.R., Jaiteh, M.S., & Esseen, P.A. (2005). Edge influence on forest structure and composition in fragmented landscapes. *Conservation Biology, 19*, 768-782

Harper, R.A. (2010). East Oklahoma, 2008: Forest Inventory and Analysis Factsheet. In: U.S. Department of Agriculture

Harper, R.A., & Johnson, T.G. (2012). Forest Resources of East Oklahoma, 2008. In R.B. SRS-187 (Ed.) (p. 112). Asheville, NC: Department of Agriculture Forest Service, Southern Research Station

He, L., Chen, J.M., Zhang, S., Gomez, G., Pan, Y., McCullough, K., Birdsey, R., & Masek, J.G. (2011). Normalized algorithm for mapping and dating forest disturbances and regrowth for the United States. *International Journal of Applied Earth Observation and Geoinformation, 13*, 236-245

Healey, S.P., Cohen, W.B., Zhiqiang, Y., & Krankina, O.N. (2005). Comparison of Tasseled Cap-based Landsat data structures for use in forest disturbance detection. *Remote Sensing of Environment, 97*, 301-310

Hilker, T., Wulder, M.A., Coops, N.C., Linke, J., McDermid, G., Masek, J.G., Gao, F., & White, J.C. (2009). A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sensing of Environment, 113*, 1613-1627

Hird, J.N., & McDermid, G.J. (2009). Noise reduction of NDVI time series: An empirical comparison of selected techniques. *Remote Sensing of Environment, 113*, 248-258

Hodges, D.G., Hartsell, A.J., Brandeis, C., Brandeis, T.J., & Bentley, J.W. (2011). Recession effects on the forests and forest products industries of the south. *Forest Products Journal, 61*, 614-624

Homer, C.C., Huang, L.Y., Wylie, B., & Coan, M. (2004). Development of a 2001 National Landcover Database for the United States. *Photogrammetric Engineering and Remote Sensing, 70*, 829-840

Hu, X., & Weng, Q. (2011). Estimating impervious surfaces from medium spatial resolution imagery: a comparison between fuzzy classification and LSMA. *International Journal of Remote Sensing, 32*, 5645-5663

Huang, C., Goward, S.N., Masek, J.G., Thomas, N., Zhu, Z., & Vogelmann, J.E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment, 114*, 183-198

Hubert-Moy, L., Cotonnec, A., Le Du, L., Chardin, A., & Perez, P. (2001). A Comparison of Parametric Classification Procedures of Remotely Sensed Data Applied on Different Landscape Units. *Remote Sensing of Environment, 75*, 174-187

Jawarneh, R.N., & Julian, J.P. (2012). Development of an Accurate Fine-resolution Land Cover Timeline: Little Rock, Arkansas, USA (1857 – 2006). *Applied Geography, 35*, 104-113

Jensen, J.R. (2005). *Introductory Digital Image Processing - A Remote Sensing Perspective*: Prentice Hall

Johnson, T.G. (2011). Oklahoma's Timber Industry - An Assessment of Timber Product Output and Use, 2009. In: Southern Research Station, Forest Service, United States Department of Agriculture

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

Julian, J.P. (2011). Field photographs, unpublished data

Kass, G.V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied statistics*, 119-127

Kauth, R.J., & Thomas, G.S. (1976). The Tasselled Cap: a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In, *Proceedings second ann. symp. machine processing of remotely sensed data* (p. 159). West Lafayette' Purdue University Lab. App. Remote Sensing

Kennedy, R.E., Townsend, P.A., Gross, J.E., Cohen, W.B., Bolstad, P., Wang, Y.Q., & Adams, P. (2009). Remote sensing change detection tools for natural resource managers: Understanding concepts and tradeoffs in the design of landscape monitoring projects. *Remote Sensing of Environment, 113*, 1382-1396

King, A.W., dilling, L., Zimmerman, G.P., Fairman, D.M., Houghton, R.A., Marland, G., Rose, A.Z., & Wilbanks, T.J. (2007). *The first state of the carbon cycle report (SOCCR): The North American carbon budget and implications for the global carbon cycle, synthesis and assessment product 2.2*: National Oceanic and Atmospheric Admistration, National Climate Data Center, Asheville, N.C.

Lambin, E.F., Turner II, B.L., Geist, H.J., Agbola, S.B., Angelsen, A., & al., e. (2001). The causes of land-cover change: moving beyond the myths. *Global Environmental Change, 11*, 261-269

Lechner, A.M., Stein, A., Jones, S.D., & Ferwerda, J.G. (2009). Remote sensing of small and linear features: Quantifying the effects of patch size and length, grid position and detectability on land cover mapping. *Remote Sensing of Environment, 113*, 2194-2204

Lo, C.P., & Choi, J. (2004). A hybrid approach to urban land use/cover mapping using Landsat 7 Enhanced Thematic Mapper Plus (ETM+) images. *International Journal of Remote Sensing, 25*, 2687-2700

Lobser, S.E., & Cohen, W.B. (2007). MODIS tasselled cap: land cover characteristics expressed through transformed MODIS data. *International Journal of Remote Sensing, 28*, 5079-5101

Lu, D., & Weng, Q. (2004). Spectral Mixture Analysis of the Urban Landscape in Indianapolis with Landsat ETM+ Imagery. *Photogrammetric Engineering & Remote Sensing, 70*, 1053-1062

Lucht, W., Schaaf, C.B., & Strahler, A.H. (2000). An algorithm for the retrieval of albedo from space using semiempirical BRDF models. *Geoscience and Remote Sensing, IEEE Transactions on, 38*, 977-998

Lunetta, R.S., Johnson, D.M., Lyon, J.G., & Crotwell, J. (2004). Impacts of imagery temporal frequency on land-cover change detection monitoring. *Remote Sensing of Environment, 89*, 444-454

Main-Knorn, M., Cohen, W.B., Kennedy, R.E., Grodzki, W., Pflugmacher, D., Griffiths, P., & Hostert, P. (2013). Monitoring coniferous forest biomass change using a Landsat trajectory-based approach. *Remote Sensing of Environment, 139*, 277-290

Market Segment Specialization Program (2013). Harwood Timber Industry Guidelines. In: National Timber Tax

Masek, J.G., Cohen, W.B., Leckie, D., Wulder, M.A., Vargas, R., de Jong, B., Healey, S., Law, B., Birdsey, R., & Houghton, R.A. (2011). Recent rates of forest harvest and conversion in North America. *Journal of Geophysical Research, 116*, G00K03

Masek, J.G., Goward, S.N., Kennedy, R.E., Cohen, W.B., Moisen, G.G., Schleeweis, K., & Huang, C. (2013). United States Forest Disturbance Trends Observed Using Landsat Time Series. *Ecosystems*, 1-18

Masek, J.G., Huang, C., Wolfe, R., Cohen, W., Hall, F., Kutler, J., & Nelson, P. (2008). North American forest disturbance mapped from a decadal Landsat record. *Remote Sensing of Environment, 112*, 2914-2926

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

Mather, P.M. (1999). Land cover classification revisited. In P.M. Atkinson & N.J. Tate (Eds.), *Advances in Remote Sensing and GIS* (pp. 7-16). New York: John Wiley & Sons

McGarigal, K., Cushman, S.A., & Ene, E. (2012). FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. In

McGwire, K.C., & Fisher, P. (2001). Spatially variable thematic accuracy: Beyond the confusion matrix. In C.T. Hunsaker, M.F. Goodchild, M.A. Friedl & T.J. Case (Eds.), *Spatial uncertainty in ecology: Implications for remote sensing and GIS applications* (pp. 308-329). New York: Springer-Verlag

McRae, D.J., Duchesne, L.C., Freedman, B., Lynham, T.J., & Woodley, S. (2001). Comparisons between wildfire and forest harvesting and their implications in forest management. *Environmental Reviews, 9*, 223-260

Meyer, W.B., & Turner II, B.L. (1996). Land-use/land-cover change: challenges for geographers. *GeoJournal, 39*, 237-240

Mildrexler, D.J., Zhao, M., & Running, S.W. (2009). Testing a MODIS global disturbance index across North America. *Remote Sensing of Environment, 113*, 2103-2117

Misson, L., Tang, J., Xu, M., McKay, M., & Goldstein, A. (2005). Influences of recovery from clear-cut, climate variability, and thinning on the carbon balance of a young ponderosa pine plantation. *Agricultural and Forest Meteorology, 130*, 207-222

Moore, R.D., & Richardson, J.S. (2012). Natural disturbance and forest management in riparian zones: comparison of effects at reach, catchment, and landscape scales. *Freshwater Science, 31*, 239-247

Nepstad, D.C., Verssimo, A., Alencar, A., Nobre, C., Lima, E., Lefebvre, P., Schlesinger, P., Potter, C., Moutinho, P., Mendoza, E., Cochrane, M., & Brooks, V. (1999). Large-scale impoverishment of Amazonian forests by logging and fire. *Nature, 398*, 505-508

Ngigi, T.G., Tateishi, R., & Gachari, M. (2009). Global mean values in linear spectral unmixing: double fallacy! *International Journal of Remote Sensing, 30*, 1109-1125

Nichol, J.E., Wong, M.S., Corlett, R., & Nochol, D.W. (2010). Assessing avian habitat fragmentation in urban areas of Hong Kong (Kowloon) at high spatial resolution using spectral unmixing. *Landscape and Urban Planning, 95*, 54-60

Oil and Gas Conservation Division (2011). Report on Oil and Natural Gas Activity Within the State of Oklahoma. In. Oklahoma City, Oklahoma: Technical Services Department, Oklahoma Corporation Commision

Oklahoma Climatological Survey (2012). The normal annual precipitation for a 30-year period: 1981-2010. In

Oklahoma Climatological Survey (2013). Temperature/Precipitation history - Annual, Southeast Oklahoma. In

Oklahoma Forestry Services (2010). The Oklahoma Forest Resource Assessment, 2010. In: Oklahoma Department of Agriculture, Food, and Forestry

Omernik, J. (1987). Ecoregions of the Conterminous United States. *Annals of the Association of American Geographers, 77*, 118-125

Ouachita Ecoregional Assessment Team (2003). Ouachita Mountains Ecoregional Assessment. In (p. 231)

Pape, A.D., & Franklin, S.E. (2008). MODIS-based change detection for Grizzly Bear habitat mapping in Alberta. *Photogrammetric Engineering and Remote Sensing, 74*, 973-985

Pontius, R.G., & Cheuk, M.L. (2006). A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions. *International Journal of Geographical Information Science, 20*, 1-30

Potapov, P., Hansen, M.C., Stehman, S.V., Loveland, T.R., & Pittman, K. (2008). Combining MODIS and Landsat imagery to estimate and map boreal forest cover loss. *Remote Sensing of Environment, 112*, 3708-3719

Powell, R., Roberts, D.A., Dennison, P.E., & Hess, L.L. (2007). Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. *Remote Sensing of Environment, 106*, 253-267

Powell, S.L., Cohen, W.B., Healey, S.P., Kennedy, R.E., Moisen, G.G., Pierce, K.B., & 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

Reynolds-Hogland, M.J., & Mitchell, M.S. (2007). Three axes of ecological studies. *Temporal Dimensions of Landscape Ecology* (pp. 174-194): Springer

Rice, E.L., & Penfound, W.T. (1959). The Upland Forests of Oklahoma. *Ecology, 40*, 593-608

Richards, J.A., & Jia, X. (1999). *Remote Sensing Digital Image Analysis: An Introduction*. Berlin, Germany: Springer-Verlag

Rindfuss, R.R., Walsh, S.J., Turner II, B.L., Fox, J., & Mishra, V. (2004). Developing a science of land change: Challenges and methodological issues. *PNAS, 101*, 13976-13981

Savitzky, A., & Golay, M.J.E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry, 36*, 1627-1639

Schleeweis, K., Goward, S.N., Huang, C., Masek, J.G., Moisen, G., Kennedy, R.E., & Thomas, N.E. (2013). Regional Dynamics of Forest Canopy Change and Underlying Causal Processes in the Contiguous US. *Journal of Geophysical Research: Biogeosciences*

Schroeder, T.A., Wulder, M.A., Healey, S.P., & Moisen, G.G. (2011). Mapping wildfire and clearcut harvest disturbances in boreal forests with Landsat time series data. *Remote Sensing of Environment, 115*, 1421-1433

Schwind, B. (2007). Monitoring Trends in Burn Severity: Report on the 2004 Mapped Fires. In

Shao, Y., & Lunetta, R.S. (2011). Sub-Pixel Mapping of Tree Canopy, Impervious Surfaces, and Cropland in the Laurentian Great Lakes Basin Using MODIS Time-Series Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 4*, 336-347

Sharitz, R.R., Boring, L.R., Van Lear, D.H., & Pinder Iii, J.E. (1992). Integrating ecological concepts with natural resource management of southern forests. *Ecological Applications*, 226-237

Skole, D., & Compton, T. (1993). Tropical Deforestation and Habitat Fragmentation in the Amazon: Satellite Data from 1978 to 1988. *Science, 260*, 1905-1910

Sleeter, B.M., Sohl, T.L., Loveland, T.R., Auch, R.F., Acevedo, W., Drummond, M.A., Sayler, K.L., & Stehman, S.V. (2013). Land-cover change in the conterminous United States from 1973 to 2000. *Global Environmental Change, 23*, 733-748

Small, C. (2001). Estimation of urban vegetation abundance by spectral mixture analysis. *International Journal of Remote Sensing, 22*, 1305-1334

Smith, J.H., Stehman, S.V., Wickham, J.D., & Yang, L.M. (2003). Effects of landscape characteristics on land-cover class accuracy. *Remote Sensing of Environment, 84*, 342-349

Smith, J.H., Wickham, J.D., Stehman, S.V., & Yang, L. (2002). Impacts of Patch Size and Land-Cover Heterogeneity on Thematic Image Classification Accuracy. *Photogrammetric Engineering & Remote Sensing, 68*, 65-70

Song, C. (2005). Spectral mixture analysis for subpixel vegetation fractions in the urban environment: How to incorporate endmember variability? *Remote Sensing of Environment, 95*, 248-263

Stefanov, W.L., Ramsey, M.S., & Christensen, P.R. (2001). Monitoring urban land cover change: An expert system approach to land cover classification of semiarid to arid urban centers. *Remote Sensing of Environment, 77*, 173-185

Sterling, S.M., Ducharne, A., & Polcher, J. (2012). The impact of global land-cover change on the terrestrial water cycle. *Nature Climate Change, 3*, 385-390

Strahler, A.H., Woodcock, C.E., & Smith, J.A. (1986). On the Nature of Models in Remote Sensing. *Remote Sensing of Environment, 20*, 121-139

Taylor, A.R., Hart, T., & Chen, H.Y.H. (2013). Tree community structural development in young boreal forests: A comparison of fire and harvesting disturbance. *Forest Ecology and Management, 310*, 19-26

Thomas, N.E., Huang, C., Goward, S.N., Powell, S., Rishmawi, K., Schleeweis, K., & Hinds, A. (2011). Validation of North American Forest Disturbance dynamics derived from Landsat time series stacks. *Remote Sensing of Environment, 115*, 19-32

Thomlinson, J.R., Bolstad, P.V., & Cohen, W.B. (1999). Coordinating methodologies for scaling landcover classifications from site-specific to global: Steps toward validating global map products. *Remote Sensing of Environment, 70*, 16-28

Townshend, J.R.G., & Justice, C.O. (1988). Selecting the spatial resolution of satellite sensors required for global monitoring of land transformations. *International Journal of Remote Sensing, 9*, 187-236

Turner, B.L., Lambin, E.F., & Reenberg, A. (2007). The emergence of land change science for global environmental change and sustainability. *Proceedings of the National Academy of Sciences, 104*, 20666-20671

Turner, M.G. (2010). Disturbance and landscape dynamics in a changing world. *Ecology, 91*, 2833-2849

Turner, M.G., & Dale, V.H. (1998). Comparing large, infrequent disturbances: what have we learned? *Ecosystems, 1*, 493-496

Turner, M.G., O'Neill, R.V., Gardner, R.H., & Milne, B.T. (1989). Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology, 3*, 153-162

U.S. Department of Agriculture (2010). National Agriculture Imagery Program (NAIP) Information Sheet. In

US Census Bureau (2013). TIGER/Line® Shapefiles. In

US Department of Agriculture Forest Service (2013a). The Forest Inventory and Analysis Database: Database Description and Users Manual Version 5.1.6 for Phase 2. In. US Department of Agriculture Forest Service

US Department of Agriculture Forest Service (2013b). Insect and Disease Detection Survey Data Explorer. In

van Oort, P.A.J., Bregt, A.K., de Bruin, S., de Wit, A.J.W., & Stein, A. (2004). Spatial variability in classification accuracy of agricultural crops in the Dutch national land-cover database. *International Journal of Geographical Information Science, 18*, 611-626

Vogelmann, J.E., Tolk, B., & Zhu, Z. (2009). Monitoring forest changes in the southwestern United States using multitemporal Landsat data. *Remote Sensing of Environment, 113*, 1739-1748

Walker, J.J., de Beurs, K.M., Wynne, R.H., & Gao, F. (2012). Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology. *Remote Sensing of Environment, 117*, 381-393

Walsh, S.J., Evans, T.P., Welsh, W.F., Entwisle, B., & Rindfuss, R.R. (1999). Scale-Dependent Relationships between Population and Environment in Northeastern Thailand. *Photogrammetric Engineering & Remote Sensing, 65*, 97-105

Wang, F. (1990). Fuzzy Supervised Classification of Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing, 28*, 194-201

Weng, Q. (2012). Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sensing of Environment, 117*, 34-49

Weng, Q., Hu, X., & Liu, H. (2009). Estimating impervious surfaces using linear spectral mixture analysis with multitemporal ASTER images. *International Journal of Remote Sensing, 30*, 4807-4830

Weng, Q., & Lu, D. (2009). Landscape as a continuum: an examination of the urban landscape structures and dynamics of Indianapolis City, 1991–2000, by using satellite images. *International Journal of Remote Sensing, 30*, 2547-2577

Weng, Q., Rajasekar, U., & Hu, X. (2011). Modeling Urban Heat Islands and Their Relationship With Impervious Surface and Vegetation Abundance by Using ASTER Images. *IEEE Transactions on Geoscience and Remote Sensing, 49*, 4080-4089

Westerling, A.L., Hidalgo, H.G., Cayan, D.R., & Swetnam, T.W. (2006). Warming and earlier spring increase western US forest wildfire activity. *Science, 313*, 940-943

Wickham, J.D., Stehman, S.V., Fry, J.A., Smith, J.H., & Homer, C.G. (2010). Thematic accuracy of the NLCD 2001 land cover for the conterminous United States. *Remote Sensing of Environment, 114*, 1286-1296

Wilkinson, G.G. (2005). Results and implications of a study of fifteen years of satellite image classification experiments. *IEEE Transactions on Geoscience and Remote Sensing, 43*, 433-440

Williams, C.A., Collatz, G.J., Masek, J., & Goward, S.N. (2012). Carbon consequences of forest disturbance and recovery across the conterminous United States. *Global Biogeochemical Cycles, 26*, GB1005

Woodcock, C.E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S.N., Helder, D., & Helmer, E. (2008). Free access to Landsat imagery. *Science (New York, NY), 320*, 1011

Woodcock, C.E., & Strahler, A.H. (1987). The Factor of Scale in Remote Sensing. *Remote Sensing of Environment, 21*, 311-332

Wu, C., & Murray, A.T. (2003). Estimating impervious surface distribution by spectral mixture analysis. *Remote Sensing of Environment, 84*, 493-505

Wu, J. (2004). Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecology, 19*, 125-138

Yu, Q., Gong, P., Tian, Y.Q., Pu, R., & Yang, J. (2008). Factors affecting spatial variation of classification uncertainty in an image object-based vegetation mapping. *Photogrammetric Engineering & Remote Sensing, 74*, 1007-1018

Zhang, Y., Peng, C., Li, W., Fang, X., Zhang, T., Zhu, Q., Chen, H., & Zhao, P. (2013). Monitoring and estimating drought-induced impacts on forest structure, growth, function, and ecosystem services using remote-sensing data: recent progress and future challenges. *Environmental Reviews, 21*, 103-115

Zhu, X., Chen, J., Gao, F., Chen, X., & Masek, J.G. (2010). An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment, 114*, 2610-2623