2013

Continuous change detection and classification of land cover using all available Landsat data

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

CONTINUOUS CHANGE DETECTION AND CLASSIFICATION OF LAND COVER USING ALL AVAILABLE LANDSAT DATA

by

ZHE ZHU

B.E., Wuhan University, Wuhan, 2006

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2013

Approved by

First Reader [redacted]
Curtis E. Woodcock, Ph.D.
Professor of Earth and Environment

Second Reader [redacted]
Alan H. Strahler, Ph.D.
Professor of Earth and Environment

Third Reader [redacted]
Mark Friedl, Ph.D.
Professor of Earth and Environment

Acknowledgements

I would like to thank the faculty, staff, and graduate students of the Earth and Environment Department and the Center for Remote Sensing, Boston University for their continuous support and encouragement. The members of my committee have been extremely generous with their time and expertise. My deepest sense of gratitude goes to my primary advisor, Dr. Curtis E. Woodcock, who has, always been available to solve any problems, clarify any confusion, and to guide me through the whole process of my dissertation work. Without his continuous academic and financial support, this research would be simply impossible. Special thanks go to Dr. Mark Friedl for generously sharing his expertise in land cover classification and always provide me precious advices for my career development. Dr. Alan H. Strahler has advised me through the very beginning and always kept me on the right track. In particular, I would like to thank Dr. John Rogan and Thomas R. Loveland for serving as members of my committee and their extensive review of this manual and remarkable insights and valuable suggestions. I express my earnest acknowledgement to Dr. Ranga Myneni, Dr. Bruce Anderson, and Dr. Nathan Phillips for sharing with me their expertise on physical remote sensing, physical climatology, and ecological physiology. Thanks go to Crystal Schaaf for let me stay at your great lab and invited me to all kinds of great meals and parties. Laura Consalvo, Christian Cole, Christopher DeVits, Victorial Woodward, and Sayaka Yamaki have always been there to solve any administrative hassle. Don Johnson has ably succeeded in keeping the workstations up and running.

Thanks go to Dr. Pontus Olofsson, Dr. Jicheng Liu, Dr. Sung Bae Jeon, Dr. Qinchuan Xin, Dr. Qingling Zhang, Dr. Zhuoshen Wang, Dr. Feng Zhao, Dr. Xiaoyuan Yang, Dr. Tian Yao, Dr. Jian Sun, Dr. Miguel Roman, Dr. Yanmin Shuai, Dr. Jieping Li, Dr. Sangram Ganguly, Dr. Mitchell Schull, and Dr. Jun Wan for the numerous intellectual and technical conversations and their friendship during their stay here at Boston University.

I also wish to express my special thanks to Xiaoman Huang, Juanmin Chen, Haidi Chen, Qingsong Sun, JiHyun Kim, Zhan Li, Jian Bi, Liang Xu, Eli Melaas, Meghan Salmon, and all my other friends for their very friendly returns, association, help and joint programs due to which my stay with them turned out to be so rewarding and enjoyable.

My parents and my girlfriend have been fervent supporters. Without their continuous support and sacrifice, I would be not able to continue my advance education and to come to U.S. to finish my PH.D degree. In particular, I would like to thank Dr Xiuxiao Yuan at Wuhan University for his continuous support and encouragement. Finally, I also owe my genuine deference to my friends back in China for their friendship and encouragement which have long been making my life meaningful.

CONTINUOUS CHANGE DETECTION AND CLASSIFICATION OF LAND COVER USING ALL AVAILABLE LANDSAT DATA

(Order No.)

ZHE ZHU

Boston University Graduate School of Arts and Sciences, 2013

Major Professor: Curtis E. Woodcock, Professor of Earth and Environment

ABSTRACT

Land cover mapping and monitoring has been widely recognized as important for understanding global change and in particular, human contributions.

This research emphasizes the use of the time domain for mapping land cover and changes in land cover using satellite images. Unlike most prior methods that compare pairs or sets of images for identifying change, this research compares observations with model predictions. Moreover, instead of classifying satellite images directly, it uses coefficients from time series models as inputs for land cover mapping. The methods developed are capable of detecting many kinds of land cover change as they occur and providing land cover maps for any given time at high temporal frequency.

One key processing step of the satellite images is the elimination of "noisy" observations due to clouds, cloud shadows, and snow. I developed a new algorithm called Fmask that processes each Landsat scene individually using an object-based method. For a globally distributed set of reference data, the overall cloud detection accuracy is 96%. A second step further improves cloud detection by using temporal information.

The first application of the new methods based on time series analysis found change in forests in an area in Georgia and South Carolina. After the difference between observed and predicted reflectance exceeds a threshold three consecutive times a site is identified as forest disturbance. Accuracy assessment reveals that both the producers and users accuracies are higher than 95% in the spatial domain and approximately 94% in the temporal domain.

The second application of this new approach extends the algorithm to include identification of a wide variety of land cover changes as well as land cover mapping. In this approach, the entire archive of Landsat imagery is analyzed to produce a comprehensive land cover history of the Boston region. The results are accurate for detecting change, with producers accuracy of 98% and users accuracies of 86% in the spatial domain and temporal accuracy of 80%. Overall, this research demonstrates the great potential for use of time series analysis of satellite images to monitor land cover change.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACCA	Automated Cloud Cover Assessment
AVHRR	Advanced Very High Resolution Radiometer
B	Tasseled-Cap Brightness
BRDF	Bidirectional Reflectance Distribution Function
BT	Brightness Temperature
CCDC	Continuous Change Detection and Classification
CMFDA	Continuous Monitoring of Forest Disturbance Algorithm
CO2	Carbon Dioxide
DI	Disturbance Index
DN	Digital Number
ENVISAT	ENVIronmental SATellite
ETM+	Enhanced Thematic Mapper Plus
Fmask	Cloud detection algorithm
G	Tasseled-Cap Greenness
HOT	Haze Optimized Transformation
IFZ	Integrated Forest Z-score
ISCCP-FD	International Satellite Cloud Climatology Project – Flux Data
L1T	level-one terrain-corrected
LEDAPS	Ecosystem Disturbance Adaptive Processing System
MERIS	Medium Resolution Imaging Spectrometer
MODIS	Moderate Resolution Imaging Spectroradiometer
MR	Mean Reflectance
NBR	Normalized Burn Ratio
NCEP	National Centers for Environmental Prediction
NDSI	Normalized Difference Snow Index
NDVI	Normalized Difference Vegetation Index
NIR	Near Infrared
OLS	Ordinary Least Square
PCP	Potential Cloud Pixel
RIRLS	Robust Iteratively Reweighted Least Squares
RMSE	Root Mean Squared Error
RFC	Random Forest Classifier
SLC	Scan Line Corrector
SWIR	Short Wave Infrared

TM	Thematic Mapper
TOA	Top Of Atmosphere
USGS	U. S. Geological Survey
W	Tasseled-Cap Wetness
WRS	Worldwide Reference System

Chapter 1

1. Introduction

1.1. Research statement

Mapping and monitoring land cover has been widely recognized as an important scientific goal. Land cover influences the energy balance, carbon budget, and hydrological cycle by means of its different physical characteristics like albedo, emissivity, roughness, photosynthetic capacity, and transpiration. Land cover change can be natural or anthropogenic, but with human activities increasing, the Earth surface has been modified significantly in recent years by various kinds of land cover change such as deforestation, agriculture expansion and intensification, urban growth, and wetland loss (Jensen et al., 1995; Coppin & Bauer, 1996; Woodcock, et al., 2001; Seto et al., 2002; Galford et al., 2008). Satellite data has long been used to assess Earth surface because of repeated synoptic collection of consistent measurements.

The 40+ years of data in the Landsat archive is one of the most valuable dataset available for understanding the changes of the Earth surface. The opening of the Landsat archive in 2008 (Woodcock et al., 2008) has led to a boom in the use of Landsat data. The temporal domain of Landsat data has been found to have great potential for many applications, especially for forest change detection (Hostert et al., 2003; Kennedy et al., 2007; Goodwin, et al., 2008; Vogelmann et al., 2009; Kennedy et al., 2010; Hilker et al., 2009; Huang et al., 2010). To fully use the temporal domain of Landsat data, screening of clouds and their shadows in the data is a necessary first step (Simpson & Stitt, 1998; Irish, 2000; Arvidson et al., 2001). Clouds cover ~66% of the Earth surface (Zhang et al., 2004) and the presence of clouds and their shadows complicates the use of data in the optical domain from earth observation satellites. While Assessment of Cloud Cover Algorithm (ACCA) provides estimates of the cloud cover percentage for images in the Landsat archive, there is no operational product that can provide maps that show the locations of clouds and cloud shadows in each image. Therefore, in this research I developed a two-step algorithm that can identify clouds and their shadows in images from different parts of the word with high accuracy.

Landsat data has been extensively used for assessment of forest change (Collins & Woodcock, 1996; Hayes & Sader, 2001; Woodcock et al., 2001; Hostert et al., 2003; Healey et al., 2005; Healey et al., 2006; Kennedy et al., 2007; Goodwin, et al., 2008; Masek et al., 2008; Vogelmann et al., 2009; García-Haro et al., 2010; Huang et al., 2010; Kennedy et al., 2010). Most of these methods are based on a pair or set of images that capture the forest change over time intervals as long as five or ten years. Even for the most recent forest change detection algorithms (Huang et al., 2010; Kennedy et al., 2010), the best they can do is provide annual or biannual forest change maps. However, they are essentially retrospective and do not provide information in a timely fashion for applications like encroachment on protected area or illegal logging. To better use the temporal domain of Landsat for monitoring changes as they are occurring, one possibility is to use as many Landsat observations as possible. Therefore, in this research I used all available Landsat data to capture forest changes within a few weeks after their occurrence.

Forest change is only one kind of land cover change and there are many other kinds of land cover changes that are very important. Also, it would be beneficial to know the land cover class before and after the changes. However, most of the current change detection algorithms are focused on one kind of land cover (Coiner, 1980; Coppin & Bauer, 1994; Jensen, et al., 1995; Cohen et al., 1998; Seto et al., 2002, Masek et al., 2008). Post-classification comparison can provide the land cover information before and after change happens, but its accuracy is usually too low for land cover change because frequently the magnitude of the error in classification is much larger than the amount of land cover change (Fuller, 2003; Friedl et al., 2010). To monitor other kinds of land cover change and provide the land cover information before and after change occurs, I extended the forest change algorithm to include identification of a wide variety of land cover changes as well as land cover mapping based on time series analysis of Landsat data.

1.2. Structure of this dissertation

1.2.1. Cloud, cloud shadow, and snow masking for Landsat TM & ETM+ images

The major effort of this research is to build a cloud, cloud shadow, and snow detection algorithm for all Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) images. This work is based on a two-step algorithm. The first step uses a newly developed algorithm called "Fmask" that process each Landsat scene individually using an object-based method (Zhu & Woodcock, 2012). The second step further improves the results by using the additional dimension of temporal context.

1.2.2. Continuous monitoring of forest disturbance using all available Landsat data

In this research, I developed an algorithm called CMFDA (Continuous Monitoring of Forest Disturbance Algorithm) that can provide accurate forest disturbance maps at high spatial resolution (30m) and high temporal frequency (a few weeks) using all available Landsat data. This algorithm uses two years of Landsat data to estimates a time series model for surface reflectance that is used to predict the next year's data. The time series model is relatively simple and includes components that capture seasonality. CMFDA finds forest disturbance by differencing model predictions and clear observations and defines a change when the difference between observed and predicted exceeds the thresholds three consecutive times (Zhu et al., 2012). The goal of this approach is to find forest change as it is occurring.

1.2.3. Continuous change detection and classification of land cover using all available Landsat data

I developed an algorithm called CCDC (Continuous Change Detection and Classification) that extends the CMFDA to a longer time period and can find more kinds of land cover change (Zhu & Woodcock, in preparation). It is capable of detecting many kinds of land cover change continuously and providing land cover maps for any given time. The time series model is updated as new observations become available. It uses a more complex time series model that has components of seasonality, trend, and "break" for land cover change. The time series models estimate both surface reflectance and brightness temperature. To find many kinds of land cover change, the CCDC algorithm uses a data-driven threshold derived from all seven Landsat bands. Land cover classification is done after change detection. Coefficients from the time series models and the Root Mean Square Error (RMSE) from estimated time series models are used as classification inputs for the Random Forest Classifier (RFC).

In conclusion, the overall goal of this research is to fully explore the temporal domain of Landsat archive and use this important dimension to find many kinds of land cover changes as they are occurring and provide land cover maps at the same time. The effort in screening cloud, cloud shadow, and snow in Chapter 2 made applications that use lots of images possible. Chapter 3 is the first test of the time series model for detecting changes, as it only applies for forest disturbance and find forest disturbance in one year. Chapters 4 extended the number of detected land cover classes (a total of 16 different land cover classes) and longer change detection time period (almost 30 years).

Chapter 2

2. Cloud, cloud shadow, and snow masking for Landsat TM & ETM+ images

2.1. Introduction

The long history of Landsat data is one of the most valuable datasets available for studying land cover change and human influences on the land surface (Coiner, 1980; Coppin & Bauer, 1994; Cohen et al., 1998; Seto et al., 2002), especially since the first Thematic Mapper (TM) sensor was launched in 1982, which provided higher spatial resolution and more spectral bands. However, many of the Landsat images are inevitably covered by cloud, especially in the tropics (Asner, 2001). The presence of clouds and their shadows complicates the use of data in the optical domain from earth observation satellites. The brightening effect of the clouds and the darkening effect of cloud shadows influence many kinds of data analyses, causing problems for many remote sensing activities, including inaccurate atmospheric correction, biased estimation of Normalized Difference Vegetation Index (NDVI) values, mistakes in land cover classification, and false detection of land cover change. Therefore, clouds and cloud shadows are significant sources of noise in the Landsat data, and their detection is an initial step in most analyses (Simpson & Stitt, 1998; Irish, 2000; Arvidson et al., 2001). Generally, clouds can be divided into two categories: thick opaque clouds and thin semitransparent clouds. The thick opaque clouds are relatively easier to identify because of their high reflectance in the visible bands. The identification of thin semitransparent clouds is difficult as their signal includes both clouds and the surface underneath (Gao & Kaufman, 1995; Gao et al., 1998; Gao et al., 2002).

Due to the high spectral variability of clouds, cloud shadows, and the Earth's surface, automated accurate separation of clouds and cloud shadows from normally illuminated surface conditions is difficult. Intuitively, it seems that clouds and cloud shadows are easily separable from clear-sky measurements, as clouds are generally white, bright, and cold compared to the Earth's surface, while cloud shadows are usually dark. Nevertheless, there are clouds that are not white, bright, or cold and cloud shadows even brighter than the average surface reflectance. Part of the difficulty arises from the wide range of reflectance and temperature observed on the surface (Irish, 2000). One common approach is to screen clouds and cloud shadows manually. However, this approach is time consuming and will limit efforts to mine the Landsat archive to study the history of the Earth's surface.

Over the years, a number of methods were developed for cloud identification. However, most of them are designed for moderate spatial resolution sensors such as Advanced Very High Resolution Radiometer (AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS). These sensors are usually equipped with more than one thermal band, or with water vapor/CO2 absorption bands, both of which are useful for thin semitransparent cloud detection (Derrien, et al., 1993, Ackerman, et al., 1997, Saunders & Kriebel, 1998). For high spatial resolution sensors like Landsat, with only one thermal band and 6 optical bands placed in atmospheric windows, accurate cloud identification is difficult. And, cloud shadow identification is even more difficult.

Clouds cast shadows on any type of land cover. When cloud shadows fall on urban, snow, ice, or bright rocks, they can be very bright compared to the average surface reflectance. Moreover, when the cloud is semitransparent, the darkening effect of the cloud shadow can be subtle, making the cloud shadow hard to detect. Therefore, how to detect clouds, cloud shadows, and especially thin clouds and their shadows in Landsat images is still an important issue in the remote sensing community, particularly as we try to use increasingly automated methods to analyze large volumes of data.

Historically, screening of clouds in Landsat data has been performed by the Automated Cloud Cover Assessment (ACCA) system (Irish, 2000; Irish 2006). By applying a number of spectral filters, and depending heavily on the thermal infrared band, ACCA generally works well for estimating the overall percentage of clouds in each Landsat scene, which was its original purpose. However, it does not provide sufficiently precise locations and boundaries of clouds and their shadows to be useful for automated analyses of time series of Landsat images. Additionally, ACCA fails to identify warm cirrus clouds and falsely identifies snow/ice in high latitude areas as clouds (Irish, 2000; Irish 2006). Wang et al. (1999) proposed the use of two multi-temporal Landsat TM images to find clouds and their shadows by image differencing. This method can successfully provide an accurate cloud and cloud shadow mask, but it is highly dependent on the input images. Since the Landsat sensors are not always turned on, it can be months between successive acquisitions. Also, it is possible that the next Landsat observation is cloudy in the same location as the previous Landsat image, which would further limit the utility of the proposed algorithm. As cloud and snow/ice are very hard to distinguish from each other in high latitude areas, Choi and Bindschadler (2004) suggested a method for detecting clouds over ice sheets by using a shadow matching technique and an automatic Normalized Difference Snow Index (NDSI) threshold. This method matches the possible cloud and cloud shadow edges iteratively to find the optimal NDSI threshold for cloud detection. It works well over ice sheets but it is time consuming and only works for the surface of ice sheets. The Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) atmosphere correction tool also generates an internal cloud mask (Vermote & Saleous, 2007). It uses two passes. There are four tests in the first pass and a thermal test in the second pass which is similar to ACCA, except that the second pass generates a cloud mask while the second pass of ACCA only provides the percentage of cloud cover. This algorithm needs other ancillary data like the surface temperature provided from National Centers for Environmental Prediction (NCEP) to help generate a coarse resolution surface temperature reference layer for cloud detection. This algorithm has already been used extensively for atmospheric correction of Landsat images and has shown to be a better method for cloud detection in low and middle latitudes compared to ACCA. However, it may not work well when the clouds cover a large percentage of the image (large amount of leakage were observed) or in sun glint and turbid water conditions (Vermote, Landsat Science Team Meeting, 2010). Hégarat-Mascle and André (2009) developed an approach that uses only two bands, Green and Short Wave Infrared (SWIR), to generate a "clear-sky line" and use the distance from the tested points to this line to detect cloud pixels. This method was originally used by Zhang et al. (2002) to correct for haze in Landsat imagery. It has been shown to be accurate for retrieving clouds over vegetated areas, but it fails when the surface reflectance is bright, as is the case for rocks, snow, ice, sand, etc (Zhang et al., 2002). By implementing a cloud-mask algorithm originally developed for the MODIS Land bands on Landsat data, Oreopoulos et al. (2011) proposed an algorithm that performs on par with the ACCA algorithm without using the thermal band.

Detecting cloud shadow is more difficult than detecting cloud. Previously, cloud shadow identification was based on spectral tests. Though it works sometimes, most of the time it will inevitably include other dark surfaces that have similar spectral signatures (like topographic shadows or wetlands) and exclude cloud shadows that are not dark enough (Ackerman et al., 1998, Hutchison et al., 2009). Recently, geometry-based cloud shadow detection has been shown to be feasible and more accurate. Currently, there are three kinds of geometry-based cloud shadow detection methods in the literature: object matching, lapse rate, and scattering differencing. The object matching algorithm detects cloud shadow by matching cloud shadows with cloud objects (Berendes et al., 1992, Simpson & Stitt, 1998, Simpson et al., 2000, Hégarat-Mascle & André, 2009). The lapse rate method used a constant lapse rate to estimate cloud top height by brightness temperature and use the cloud pixels to cast shadows (Vermote & Saleous, 2007). This latter method works well for thick clouds but is not accurate when the clouds are semitransparent, as the brightness temperature will be a mixture of thin cloud and the surface, making cloud height estimation problematic. As cloud shadow scattering is stronger in the short wavelengths (especially the Blue band), Luo et al. (2008) proposed to use this physical characteristic, scattering differences between the short wavelength and Near Infrared (NIR) or SWIR, combined with the geometry, to produce cloud shadow masks. This new method works well over vegetated area, but is less accurate when the cloud shadow falls on bright surfaces or the cloud shadow comes from a very thin cloud.

2.2. Methodology

The input data are Top of Atmosphere (TOA) reflectances for Bands 1, 2, 3, 4, 5, 7 and Band 6 Brightness Temperature (BT) (Table 2.1). For Landsat level-one terrain-corrected (L1T) images, Digital Number (DN) values are converted to TOA reflectances and BT (Celsius degree) with the LEDAPS atmosphere correction tool (Masek et al., 2006; Vermote & Saleous, 2007). Then, rules based on cloud and cloud shadow physical properties are used to extract a potential cloud layer and a potential cloud shadow layer. Finally, the segmented potential cloud layer and the geometric relationships are used to match the potential cloud shadow layer, leading to the production of the final cloud and cloud shadow mask. If the Landsat scene has snow, Fmask will also produce a snow mask in addition to the cloud and cloud shadow mask.

Next, a multitemporal method for automatically identifying clouds, cloud shadows, and snow is provided. With the long history Landsat archive data, we can build up a cloud, cloud shadow free dataset, and using this to help us in screening them out. In order to build up this dataset, we firstly use a single-image based method (Fmask) to exclude most of the "noisy" pixels and then use the robust linear least square fitting method to get the predicted cloud, cloud shadow, and snow free dataset. With this dataset, we can improve the previous single-image based mask results and use it in future dataset analysis.

Table 2.1. Landsat TM/ETM+ spectral bands

TM bands (μm)	ETM+ bands (μm)
Band 1 (0.45-0.52)	Band 1 (0.45-0.515)
Band 2 (0.52-0.60)	Band 2 (0.525-0.605)
Band 3 (0.63-0.69)	Band 3 (0.63-0.69)
Band 4 (0.76-0.90)	Band 4 (0.75-0.90)
Band 5 (1.55-1.75)	Band 5 (1.55-1.75)
Band 6 (10.40-12.50)	Band 6 (10.40-12.50)
Band 7 (2.08-2.35)	Band 7 (2.09-2.35)

2.2.1. Layers of potential clouds, cloud shadows, and snow

2.2.1.1. Potential cloud layer – Pass one

The Fmask algorithm first combines several spectral tests to identify the Potential Cloud Pixels (PCPs) – the pixels that may possibly be cloudy and may sometimes be clear pixels. Otherwise, the pixels are considered to be absolutely clear-sky pixels. This first pass includes a number of spectral tests as follows:

$$\textit{Basic Test} = \textit{Band } 7 > 0.03 \textit{ AND } BT < 27 \textit{ AND } NDSI < 0.8 \textit{ AND } NDVI < 0.8 \qquad (2.1)$$

Where,

$NDSI = (Band\ 2 - Band\ 5) / (Band\ 2 + Band\ 5)$

$NDVI = (Band\ 4 - Band\ 3) / (Band\ 4 + Band\ 3)$

This "Basic Test" is one of the most fundamental tests for cloud identification. Due to the bright and cold nature of clouds, all kinds of clouds should have Band 7 TOA reflectance larger than 0.03 (heritage from LEDAPS internal cloud masking algorithm) and BT less than 27 °C (heritage from ACCA). The NDSI and NDVI values of clouds are usually around zero because of their "white" character in optical spectral bands. For certain cloud types, such as very thin clouds over highly vegetated area or icy clouds, the NDVI and NDSI values can be larger, but both of them cannot be higher than 0.8. ACCA also uses NDSI threshold of 0.8 to separate clouds from snow pixels in the first pass. Therefore, Fmask uses NDSI and NDVI thresholds of 0.8 to separate PCPs from some of the vegetated or snow covered areas.

Mean Visible = (Band 1 + band 2 + band 3) / 3

Whiteness Test = $\sum_{i=1}^{3} | (Band\ i - Mean\ Visible) \ / \ Mean\ Visible| \ < \ 0.7$ (2.2)

This "Whiteness" index was originally proposed by Gomez-Chova et al. (2007). As clouds always appear white due to their "flat" reflectance in the visible bands, they used the sum of the absolute difference between the visible bands and the overall brightness to capture this cloud property. This index works well in ENVIronmental SATellite (ENVISAT) Medium Resolution Imaging Spectrometer (MERIS) multispectral image as it has many narrow visible bands. However, it is not that useful for Landsat sensor which only has three visible bands. By dividing the difference by the average value of the visible bands, the new "Whiteness" index works well for Landsat imagery and 0.7 (sensitivity analysis of the global cloud reference dataset) appears to be an optimal threshold for excluding clear-sky pixels that exhibit high variability in the visible bands. All the sensitivity analyses in this paper are based on a set of 142 Landsat reference images. To find the optimal threshold for "Whiteness", we let the "Whiteness" threshold vary from 0.5 to 0.9 (at 0.1 intervals) and chose the one with the highest average cloud overall accuracy (Figure 2.1). The above "Whiteness" index is used to exclude pixels that are not "white" enough to be clouds. Note that this "Whiteness Test" may also include some pixels of bare soil, sand, and snow/ice as they may also have "flat" reflectance in the visible bands.

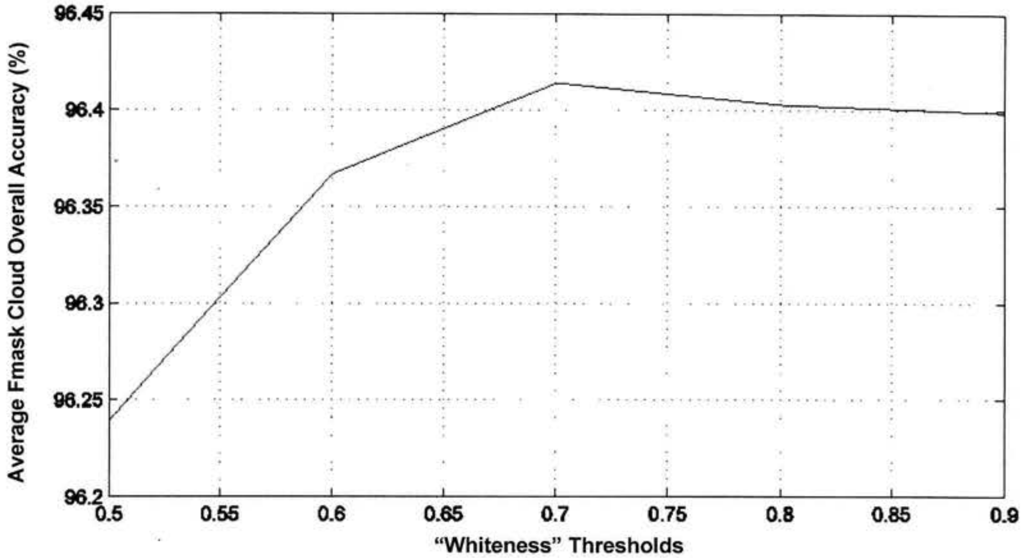


Figure 2.1. An example of choosing an optimal threshold for "Whiteness" based on sensitivity analysis. Note that a threshold of 0.7 shows the highest average Fmask cloud overall accuracy.

$$HOT\ Test = Band\ 1 - 0.5 \bullet Band\ 3 - 0.08 > 0 \qquad (2.3)$$

This Haze Optimized Transformation (HOT) for Landsat data was firstly proposed by Zhang et al (2002). It is based on the idea that the visible bands for most land surfaces under clear-sky conditions are highly correlated, but the spectral response to haze and thin cloud is different between the blue and red wavelengths. The "HOT" in Zhang et al.

(2002) is built empirically from regression of DN values from clear-sky pixels. If we use TOA reflectance as inputs for regression, Equation 2.3 is retrieved for most of the Landsat images. The results are especially helpful for separating haze and thin cloud from clear-sky pixels. Similar test has also been used in the LEDAPS internal cloud masking algorithm. All kinds of clouds (thin and thick) and possibly thick aerosols will be identified by this test. Note that this "HOT Test" may also include some bright pixels like rocks, turbid water, or snow/ice surface due to their large TOA reflectance in the visible bands.

$$B4/B5\ Test = Band\ 4\ /\ Band\ 5 > 0.75 \tag{2.4}$$

This spectral test is similar to a test in ACCA (Irish, R., 2000) in which a Band 4 and Band 5 ratio larger than 1 is used to exclude bright rock and desert due to the fact that they tend to exhibit higher reflectance in Band 5 than in Band 4, whereas the reverse is true for clouds. However, this threshold may also exclude some thin clouds. Therefore, we reduced this threshold to 0.75 (sensitivity analysis of the global cloud reference dataset) in Fmask to include all possible cloud pixels. This test may also include other noncloud pixels, but the main focus of this test is separating most of bright rocks from clouds.

$$\begin{aligned} Water\ Test = {} & (NDVI < 0.01\ AND\ Band\ 4 < 0.11) \\ & OR\ (NDVI < 0.1\ AND\ Band\ 4 < 0.05) \end{aligned} \tag{2.5}$$

This "Water Test" divides all pixels into two categories – water and land pixels. The thick clouds will be identified as land pixels whether they are over land or water (thick clouds block all information for land and water separation), while the thin clouds over water may still be able to be identified as water pixels. NIR band reflectance is a good indicator for water identification, as water is generally dark in this band while land is usually bright. Additionally, NDVI values are especially useful for separating water pixels from land pixels, as land NDVI values are generally above 0.10 and water NDVI values are usually less than 0.10 (Vermote & Saleous, 2007). Most of the water pixels are identified by NDVI less than 0.1 and Band 4 less than 0.05. Some of the water pixels may have relatively large Band 4 reflectance because of influence of thin clouds or turbid conditions, and they will be identified by using the test NDVI less than 0.01 and Band 4 less than 0.11. The "Water Test" is mainly used for separating cloud probability calculation over water and land in pass two.

By applying the spectral tests above, Fmask will identify PCP as follows:

$$PCP = Basic\ Test\ (true)\ AND\ Whiteness\ Test\ (true)\ AND\ HOT\ Test\ (true)\ AND\ B4/B5\ Test\ (true) \tag{2.6}$$

If the PCPs are more than 99.9% of the scene, they will be used for the final cloud mask directly, as there are not enough clear-sky pixels (approximately 50,000 pixels) for statistic analyses in the second pass. If the PCPs are less than 99.9% of the scene, the PCPs and the absolute clear-sky pixels will be sent to the second pass. As the algorithm tends to include all possible cloudy pixels (it overestimates cloud fraction) in the first pass, Fmask requires a relatively small percentage (0.1%) of the scene to be absolutely clear to allow the second pass to work.

2.2.1.2. Potential cloud layer – Pass two

After identification of all PCPs, the rest of the pixels (absolute clear-sky pixels) can be used for computing cloud probability for all pixels in the image. As the temperature distributions and the range of reflectances for land and water can be quite variable in space and time, Fmask computes cloud probability separately for water and land. The water and land pixels are classified by the "Water Test" applied in pass one.

The cloud probability for water (*wCloud_Prob*) is a combination of temperature probability (*wTemperature_Prob*) and brightness probability (*Brightness_Prob*) computed as follows:

Temperature probability for water:

$$Clear\text{-}sky\ Water = Water\ Test\ (true)\ AND\ Band\ 7 < 0.03 \tag{2.7}$$

$$T_{water} = 82.5\ percentile\ of\ Clear\text{-}sky\ Water\ pixels'\ BT \tag{2.8}$$

$$wTemperature\_Prob = (T_{water} - BT) / 4 \tag{2.9}$$

The difference between the estimated clear-sky water temperature (T_{water}) and the pixels' BT are normalized by 4 °C to compute the temperature probability for water (Equation 2.9). The clear-sky water pixels are identified with a "Water Test" and a low Band 7 reflectance threshold (Equation 2.7). T_{water} is estimated with the upper level (82.5 percentile) of clear-sky water temperature, in purpose of exclude other atmospheric influences that are usually making water temperature colder. A constant of 4 °C is used for re-scale the temperature probability because a pixel would have a high probability of being a cloud pixel if its BT is 4 °C colder than the surface temperature (Vermote & Saleous, 2007). As temperature is one of the most important dimensions in cloud detection, the temperature probability can be higher than one if the BT of the pixel is more than 4 °C colder than the estimated clear-sky water temperature.

Brightness probability:

Water is generally dark, especially in Band 5 reflectance. The existence of clouds over water can increase Band 5 reflectance greatly. Fmask uses the normalized Band 5 reflectance to calculate the brightness probability for cloud detection over water. Usually Band 5 reflectance of water is less than 0.05, except for some turbid or shallow water pixels that may have higher reflectance. The brightest water may have Band 5 reflectance as high as 0.11. Fmask calculates the normalized brightness probability with Equation 2.10.

$$Brightness\_Prob = \min(Band\ 5,\ 0.11) / 0.11 \qquad (2.10)$$

Cloud probability for water:

The cloud probability for water pixels is computed by combining both the temperature probability and the brightness probability (Equation2. 11). The temperature probability may contribute more than the brightness probability for some very cold pixels because of its wider probability range.

$$wCloud\_Prob = wTemperature\_Prob \bullet Brightness\_Prob \qquad (2.11)$$

As BT and Band 5 reflectance for clear-sky water pixels are very homogenous, Fmask uses a fixed threshold to retrieve clouds over water. A water pixel is identified as a cloud pixel if *wCloud_Prob* is larger than 0.5. This fixed threshold works well for detecting clouds over water. By combining temperature and brightness probabilities, bright water pixels (like shallow or turbid water pixels) or cold water pixels (higher elevation water) will be easily excluded from cloud pixels because if one of the probabilities is close to zero, no matter how large the other probability is, the cloud probability for water will still be close to zero.

The cloud probability for land (*lCloud_Prob*) is a combination of temperature probability (*lTemperature_Prob*) and variability probability (*Variability_Prob*) computed as follows:

Temperature probability for land:

$$\textit{Clear-sky Land} = \textit{PCP (false) AND Water Test (false)} \tag{2.12}$$

$$(T_{low}, T_{high}) = \textit{(17.5, 82.5) percentile of Clear-sky Land pixels' BT} \tag{2.13}$$

If clear-sky land pixels cover less than 0.1% (minimum necessary pixels for statistic analysis) of the total observations in the scene, Fmask will use the clear-sky pixels (from both land and water) for calculating temperature probability instead of only using the clear-sky land pixels. T_{low} and T_{high} calculated from Equation 2.13 provide the temperature interval for clear-sky land pixels. The 17.5% and 82.5% thresholds were derived from a sensitivity analysis of the global cloud reference masks. As land temperatures can differ greatly, Fmask uses the upper and lower level of the clear-sky land temperature to normalize the temperature probability for land (Equation 2.14). Normally, if the pixel's BT is 4 °C colder than T_{low}, the pixel has a high probability of being a cloud. On the other hand, if the pixel's BT is 4 degree warmer than T_{high}, the pixel is most likely clear. Because temperature is one of the most important dimensions in cloud detection, the temperature probability for land can be larger than one if the BT of the pixel is more than 4 °C colder than T_{low}.

$lTemperature\_Prob = (T_{high} + 4 - BT) / (T_{high} + 4 - (T_{low} - 4))$ (2.14)

Variability probability:

Due to the large variability of reflectance for land pixels, the brightness probability does not work well over land for cloud detection. However, as the cloud spectral reflectances in the optical bands are very consistent, Fmask uses the probability of the spectral variability to identify clouds over land. The NDVI, NDSI, and "Whiteness" values are used to capture the spectral variability in NIR/Visible, SWIR/Visible, and within the Visible. Fmask uses 1 minus the largest value among the three indices to represent the spectral variability. The NDVI and NDSI based spectral variability may not be accurate when dealing with saturated pixels. In this case, a modified NDVI and NDSI are used in Equation 2.15. The NDSI and NDVI values are modified as follows: if a pixel is saturated in Band 2 and has Band 5 larger than Band 2, Fmask gives a zero value for this pixel's NDSI; the same rule is applied for the modified NDVI, that is, if a pixel is saturated in Band 3 and has Band 4 larger than Band 3, Fmask gives a zero value for this pixel's NDVI. This is because compared to NIR and SWIR bands, the Landsat visible bands are easily saturated for bright pixels. Theoretically, for bright cloud pixels, all the optical bands TOA reflectance will be close to 1, making the NDSI and NDVI values close to 0. However, if visible bands become saturated at a small value, for example 0.5 (Dozier, 1989), while NIR and SWIR bands do not (close to 1), it would make the absolute values of NDSI and NDVI much larger than 0, making probability of spectral variability lower for cloud pixels.

Variability_Prob = 1 – max (abs (modified NDVI), abs (modified NDSI),

$AND\ Whiteness)$ (2.15)

Cloud probability for land:

The cloud probability for land pixels is computed by combining both temperature probability and variability probability as follows. The temperature probability may contribute more than the variability probability for some very cold pixels because of its wider probability range.

$$lCloud\_Prob = lTemperature\_Prob \bullet Variability\_Prob \tag{2.16}$$

The threshold for defining cloud over land is consisted by the upper level (82.5 percentile) of clear-sky land pixels' probability plus a constant of 0.2 (based on sensitivity analysis) shown in Equation 2.17. Fmask identifies a pixel as cloud if the land pixel's *lCloud_Prob* is larger than this scene-based threshold.

$$Land\_threshold = 82.5\ percentile\ of\ lCloud\_Prob\ (Clear\text{-}sky\ Land\ pixels) + 0.2 \tag{2.17}$$

Therefore, by combining the cloud probability and the previously identified PCPs, Fmask generates the potential cloud layer in Equation 2.18. Due to the possibility of omitting clouds in PCPs, Fmask finds missed cloud pixels if the *lCloud_Prob* is extremely large (more than 99%) over land or BT is extremely cold (35 °C colder than T_{low}) .

$$\begin{aligned} &Potential\ Cloud\ Layer = (PCP\ (true)\ AND\ Water\ Test\ (true)\ AND\ wCloud\_Prob > 0.5) \\ &OR\ (PCP\ (true)\ AND\ Water\ Test\ (false)\ AND\ lCloud\_Prob > Land\_threshold) \\ &OR\ (lCloud\_Prob > 0.99\ AND\ Water\ Test\ (false))\ OR\ (BT < T_{low} - 35) \end{aligned} \tag{2.18}$$

Finally, Fmask will spatially improve the cloud mask by using the rule that sets a pixel to cloud if five or more pixels in its 3-by-3 neighborhood are cloud pixels; otherwise, the pixel stays clear.

2.2.1.3 Potential cloud shadow layer

Because beam solar radiation is blocked by clouds, the cloud shadows are mainly illuminated by scattered light. As the atmospheric scattering is stronger at shorter wavelengths (for example visible bands), the diffusive radiation in the shadows will be relatively smaller at longer wavelengths (for example NIR and SWIR bands), making the shadowed pixels darker than their surroundings (Luo et al., 2008). Moreover, as NIR reflectance is usually high (including vegetation, snow, ice, and rock), the darkening effect of cloud shadows is most obvious in this Band. Therefore, a morphological transformation called flood-fill is performed for Band 4 reflectance (NIR band) that brings the intensity values of dark areas that are surrounded by lighter areas up to the same intensity level as the surrounding pixels (Soille, 1999). In field of morphology, the gray-scaled image is viewed as a "digital elevation model". Therefore, all cloud shadows are located at places with regional minima due to their relatively darker Band 4 reflectance compared to their surroundings. The flood-fill transformation is defined as the reconstruction by erosion of the input digital elevation model using a marker image which is set to the maximum height of the digital elevation model except along its borders and at the bottom of natural depressions where it inherits the values of the input digital elevation model (Soille et al., 2003). In this case, the difference between the filled

Band 4 reflectance and the original Band 4 reflectance will include the darkening effect of the cloud shadows. If the cloud shadow is located at the edge of the scene, the flood-fill transformation will not be able to identify it. Therefore, Fmask fills the edge of the scene with the lower level (17.5 percentile) of the clear-sky land Band 4 reflectance to catch all potential cloud shadows.

$$\textit{Potential Cloud Shadow Layer} = \textit{Flood-fill Band 4} - \textit{Original Band 4} > 0.02 \tag{2.19}$$

2.2.1.4. Potential snow layer

$$\begin{aligned}\textit{Potential Snow Layer} = {} & \textit{NDSI} > 0.15 \textit{ AND BT} < 3.8 \\ & \textit{AND Band 4} > 0.11 \textit{ AND Band 2} > 0.1\end{aligned} \tag{2.20}$$

Most of the spectral tests used here (BT less than 3.8, Band 4 more than 0.11, and Band 2 more than 0.1) are from the MODIS snow mapping algorithm (Hall et al., 2001). The only difference is the NDSI thresholds used. The MODIS snow algorithm uses NDSI larger than 0.4 as its threshold to identify pixels that are approximately 50% or greater covered by snow. We lower the NDSI threshold to 0.15 for Fmask to include pixels with snow coverage less than 50% and snow contaminated forest areas in which snow are partly blocked by the forests. At the same time, for all clear (snow and cloud free) land pixels in Landsat data, the NDSI values are always lower than 0.15. Therefore, with a NDSI threshold of 0.15, we can separate snow free and snow contaminated pixels accurately in Landsat data. This threshold has already been used for operational snow mapping in Meteosat Spinning Enhanced Visible Infra-Red Imager (SEVIRI) imagery (Wildt et al., 2007).

2.2.2. Object-based cloud and cloud shadow match

The basic idea of this cloud and cloud shadow matching approach is that by knowing the view angle of the satellite sensor, the solar zenith angle, the solar azimuth angle, and the relative height of the cloud, we can predict the cloud shadow location based on the geometric relationship between a cloud and its shadow. Because the first three factors are known, we can use them to calculate the projected direction of the cloud shadow. Along this direction, Fmask matches the cloud object with the potential shadow layer by using the idea that a cloud and its shadow have similar shape (Figure 2.2.). The original cloud object is excluded from the calculated shadow, as the pixels cannot be cloud and shadow at the same time. The match similarity for each cloud object is the ratio of the overlap area between the calculated shadow and the potential cloud or shadow layers to the calculated shadow area. To match the correct cloud shadow, iteration of the cloud height continues if similarity is increasing or not decreasing to 98% of the maximum similarity; otherwise, the iteration will stop. If similarity is larger than a given threshold, the cloud shadow is matched, otherwise, it is rejected. The similarity threshold can be any value from 0.2 to 0.5, which all provide similar cloud shadow results. A threshold of 0.3 is applied for Fmask as it keeps a balance between omission and commission errors of cloud shadows.

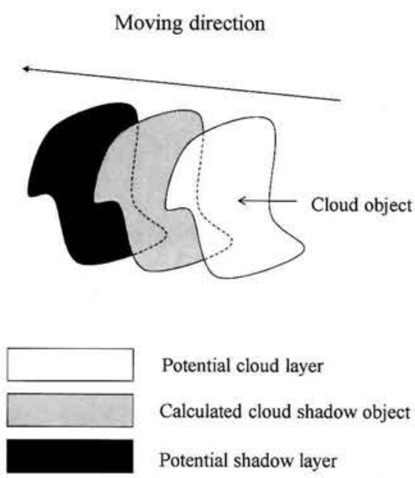


Figure 2.2. Illustration of cloud and cloud shadow matching based on similarity

The cloud objects are derived by segmentation of the potential cloud layer, that is, the potential cloud pixels adjacent to other potential cloud pixels (using 8-way connectedness) are identified as one cloud object. The shapes of cloud objects are not always the same as their cloud shadows (Figure 2.2. is an ideal case assuming cloud objects are flat), because some kinds of clouds having large vertical extents may cast cloud shadow that extend further than the flat cloud approximation allows. This can also occur with small vertical extent clouds at very low solar elevations angles. Therefore, Fmask treats each cloud as a 3D object with a base height retrieved by matching clouds and cloud shadows, and a top height estimated by a constant lapse rate and its corresponding base height.

As cloud base height can be any value from 200 m to 12,000 m, it would be time consuming and may cause false matches if we iterate cloud height across this entire range for each cloud object to find its cloud shadow. Fmask algorithm narrows the cloud base height range by using cloud object BT. For standard atmosphere conditions, the adiabatic lapse rate for dry air is -9.8 K km^{-1} and for wet air is -6.5 K km^{-1} (Hartmann, 1994). However, this is not always true for thin clouds, as their BT is influenced by the warmer ground surface underneath. In this case, Fmask uses a reduced wet adiabatic lapse rate of -1 K km^{-1} to capture thin cloud shadows. Therefore, we can predict the minimum and maximum cloud object base height range as:

$$H_{cloud\_base} = (max\ (0.2,\ (T_{low} - 4 - T_{cloud\_base})\ /\ 9.8),$$
$$min\ (12,\ (T_{high} + 4 - T_{cloud\_base})))\ km \qquad (2.21)$$

For each cloud object, T_{cloud_base} should have the highest BT due to the fact that the cloud base pixels are the lowest cloud pixels that the sensor can detect. Nevertheless, for both thick and thin clouds, the warmest cloud pixels located at the edge do not represent the actual BT of the cloud base due to influences from the neighboring warm ground. Therefore, it is necessary to use pixels far enough from the edge of the cloud to represent the cloud base BT and adjust the edge pixels that are warmer than this value. For the purpose of simplify cloud base BT calculation, Fmask assumes each cloud object is round and 8 cloud edge pixels are influenced by the neighboring warm surface. If the calculated radius of cloud object is less than 8 pixels, Fmask uses the minimum BT of the cloud object as its cloud base BT. Therefore, we can calculate the cloud base BT (Equation 2.22) and adjust the influenced cloud BT with this value (Equation 2.23) as follows:

$$If\ \ R \geq 8 \qquad T_{cloud\_base} = 100(R\text{-}8)^2/R^2\ percentile\ of\ cloud\ object\ BT$$
$$Else \qquad T_{cloud\_base} = min\ (cloud\ object\ BT) \qquad (2.22)$$

Where,

$R = sqrt\ (total\ pixels\ of\ a\ cloud\ object\ /\ 2\pi)$

$If\ T_{cloud\_object} > T_{cloud\_base}$

$$T_{cloud\_object} = T_{cloud\_base} \tag{2.23}$$

Since within the cloud object the air is wet, Fmask assumes the lapse rate in the cloud is a constant of -6.5 K km^{-1}. Therefore, the cloud top height can be estimated based on the cloud base height and relative BT difference between cloud base and cloud top:

$$H_{cloud\_top} = H_{cloud\_base} + 6.5(BT_{cloud\_base} - BT_{cloud\_top})\ km \tag{2.24}$$

Finally, as the matched cloud shadow may have holes, Fmask buffers by 3 pixels in 8-connected directions for each of the matched cloud shadow pixel to fill those small holes. Moreover, as the potential cloud shadow layer produced previously includes all shadow areas, Fmask further refines the cloud shadow mask by only choosing the overlap between the potential cloud shadow layer and the matched cloud shadow objects. For cloud objects less than 3 pixels, Fmask excludes them from cloud mask and does not match cloud shadows for them as most of them are misidentification of small bright cold noncloud pixels.

The details of the cloud and cloud shadow matching algorithm are shown in Figure 2.3. Because snow pixels, cloud pixels, and cloud shadow pixels may overlap, Fmask sets cloud pixels to have the highest priority, cloud shadow pixels have the second highest priority, and snow pixels have the lowest priority. In this case, if the three classes overlap for a pixel, the class with the highest priority will be its label.

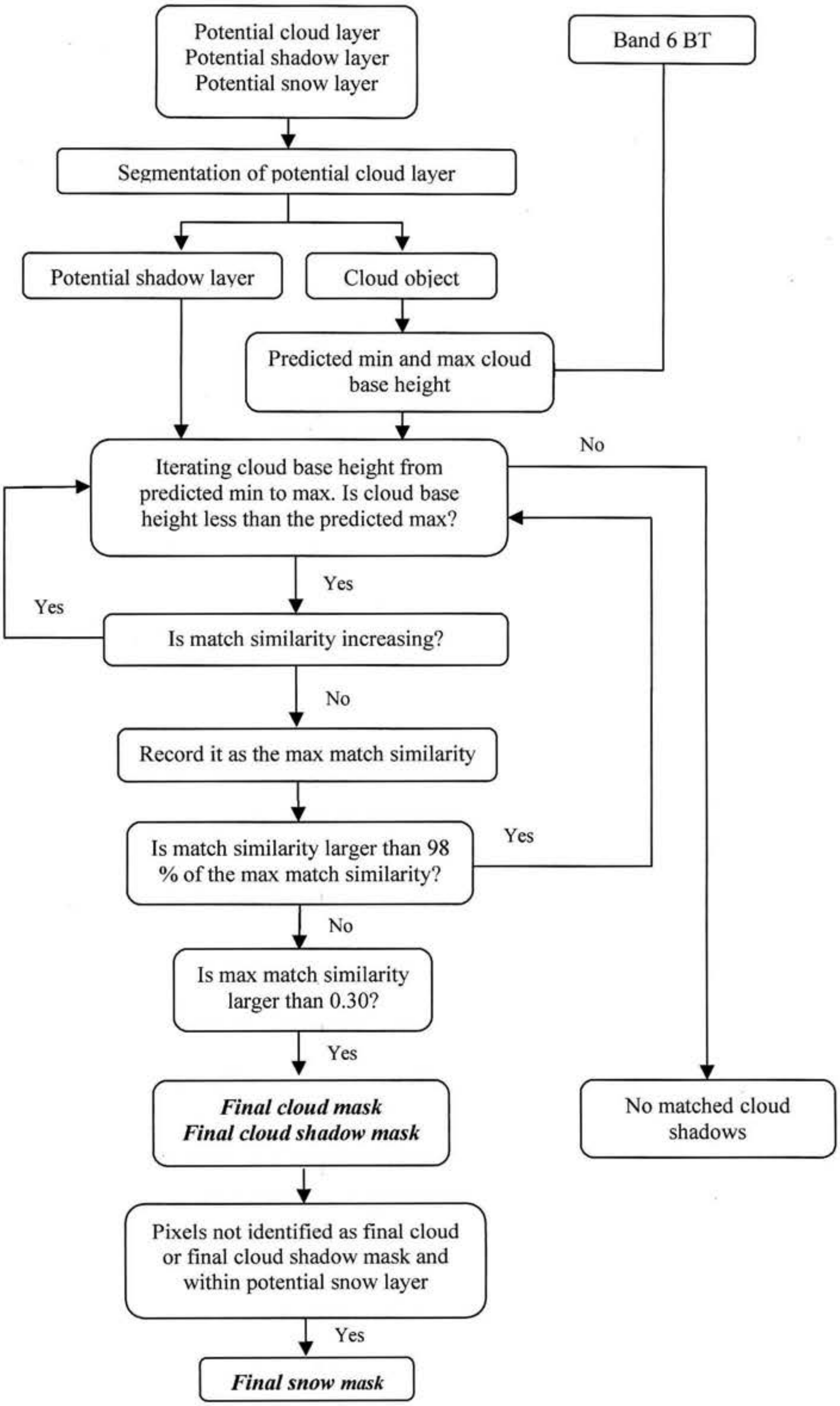


Figure 2.3. Flow chart of object-based cloud and cloud shadow match algorithm

2.2.3 Multitemporal masking of cloud, cloud shadow, and snow

Though the Fmask algorithm provides relatively accurate masks for clouds, cloud shadows, and snow, it is not perfect, as the same statistical threshold is used for each image. Moreover, there are other ephemeral changes such as thick aerosols, smoke, or flooding that may also contribute to sudden spectral response change that may be confused with the real land cover change signal. Therefore, I used an additional step to further remove noise from the time series data (Zhu, et al., 2012a; Zhu, et al., in preparation). "Clear pixels" previously identified by Fmask were used as inputs. A time series model (Equation 2.25) was estimated with the Robust Iteratively Reweighted Least Squares (RIRLS) method (Street et al., 1988; DuMouchel, 1989; O'Leary, 1990; Holland et al., 1977). The robust feature reduces the influence of outliers from Fmask algorithm.

$$RIRLS(i,x) = a_{0,i} + a_{1,i}\,cos(\frac{2\pi}{T}x) + b_{1,i}\,sin(\frac{2\pi}{T}x)$$

$$+a_{2,i}\,cos(\frac{2\pi}{NT}x) + b_{2,i}\,sin(\frac{2\pi}{NT}x) \qquad (2.25)$$

Where,

x: Day-of-year

i: The ith Landsat Band

T: Number of days per year (T = 365)

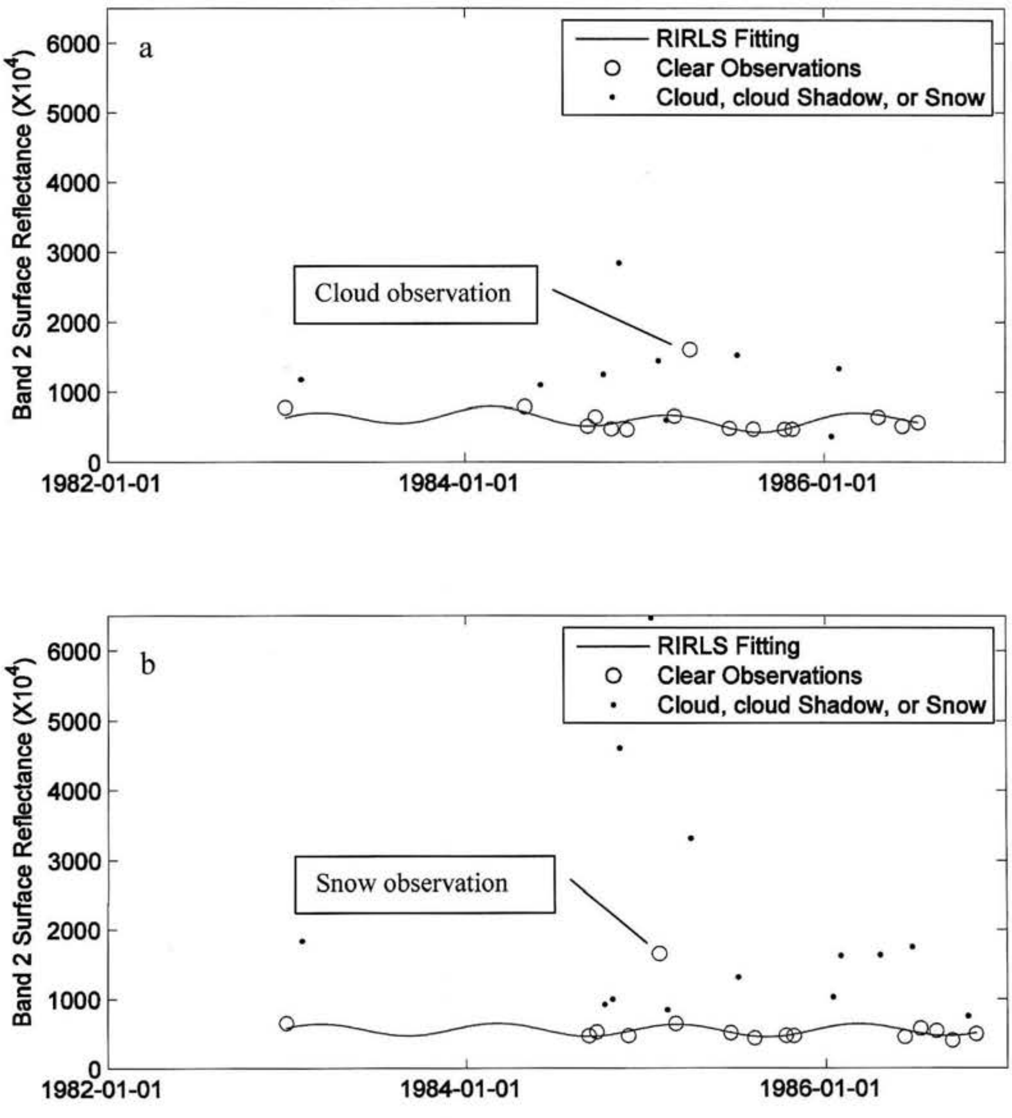
N: Number of years of time series Landsat data

$a_{0,i}$: Overall values for the ith Landsat Band

$a_{1,i}, b_{1,i}$: Coefficients for intra-annual change for the ith Landsat Band

$a_{2,i}, b_{2,i}$: Coefficients for inter-annual change for the ith Landsat Band.

Due to the factor that clouds and snow will make Band 2 brighter and cloud shadows and snow will make Band 5 darker, we fit a time series model for Band 2 and Band 5 separately. By comparing the observed and the model predicted values, it is easy to identify all clouds, cloud shadows, snow, and other ephemeral changes (Figure 2.4 & Equation 2.25).



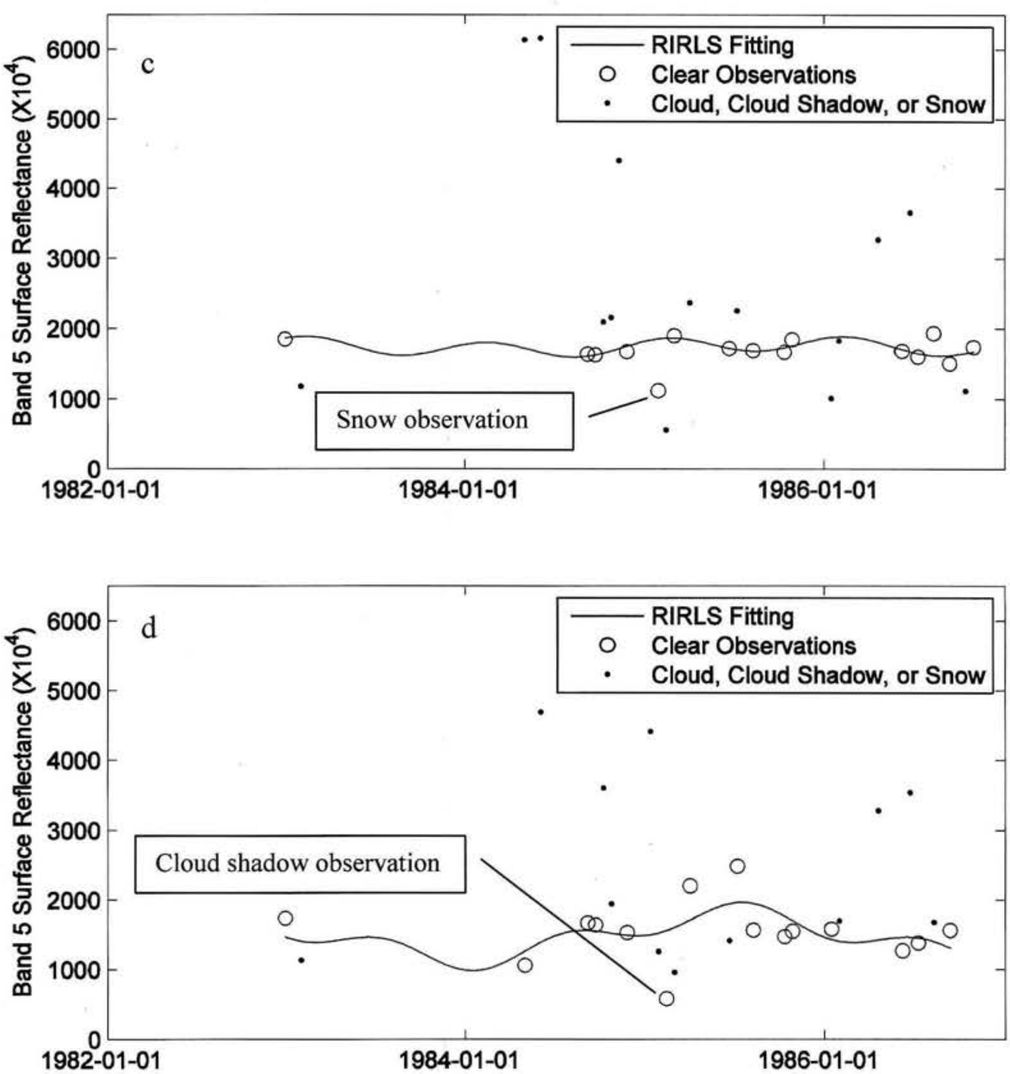


Figure 2.4. Illustration of multitemporal screening of cloud, cloud shadow, and snow. In these figure, the circle are pixels indentified as "clear" by Fmask and those highlighted are ones found by the multitemporal analysis to be clouds, cloud shadows, and snow. Based on the fact that clouds and snow will make Band 2 brighter (Figure 2.4a & Figure 2.4b) while cloud shadows and snow will make Band 5 darker (Figure 2.4c & Figure 2.4d), all clouds, cloud shadows, snow, and other

ephemeral changes will be easily identified by comparing clear observations with model predictions.

The model estimating starts when there are a total of 15 "clear" (determined by Fmask) observations, in which the first 12 of them are compared between observed and model predicted values to decide whether there are outliers or not. The last 3 clear observations are used to help the curve respond to land cover change so that changes happening at the end of the fitting will not be identified as outliers. The reason for picking the first 12 clear observations is that the continuous change detection and classification algorithm uses 4 coefficients to initializing the time series model (See section 4.2 for detail) and 3 times the number of coefficients helps make model fitting robust and accurate. The season for picking the last 3 clear observations is that the algorithm uses three more observations to determine if a pixel is has changed or not (See section 4.2 for detail) and 3 clear observations are enough to respond to land cover change. If any of the first 12 pixels are found to be possible cloud, cloud shadow, snow or other ephemeral changes, it will be removed from the clear pixels list.

$$Outliers\ if: Band\ 2(x) - RIRLS(2, x) > 0.04$$
$$OR\ Band\ 5(x) - RIRLS(5, x) < -0.04 \qquad (2.26)$$

Where,

x: Julian date

$Band\ i(x)$: The ith Landsat Band at the Julian date of x

$RIRLS(i, x)$: The RIRLS fitting for the ith Landsat Band at the Julian date of x.

2.3 Results

By comparing the results of Fmask with false color composites visually (Figure 2.5abcd), it appears to work well in identifying cloud (yellow), cloud shadow (green), and snow/ice (cyan). Figure 2.5a is one of the Sub-tropical South images with "well-behaved" clouds and cloud shadows over highly vegetated areas. Fmask was able to identify this kind of clouds (including some of the thin clouds) and their shadows. On the other hand, Figure 2.5b is a Sub-tropical North image with large variability in surface reflectances. Fmask works well in terms of identifying clouds in areas of very bright rock and has no problem in labeling cloud shadows over this bright surface. Furthermore, Figure 2.5c is one of the South Polar images with thick and thin clouds over very bright snow/ice. The snow/ice is accurately identified in the cyan color, and the clouds (both thick and thin) are separated well from the bright snow/ice. Finally, Figure 2.5d is a very difficult image, as it has extremely thin cirrus clouds (see red arrows) and bright turbid water (see yellow arrows). In the Fmask result, there are no commission errors of clouds from the bright turbid water and the extremely thin cirrus clouds are also identified with high accuracy. This sort of qualitative evaluation was an important part of the development of the algorithm. To more rigorously assess its accuracy, reference data are used.

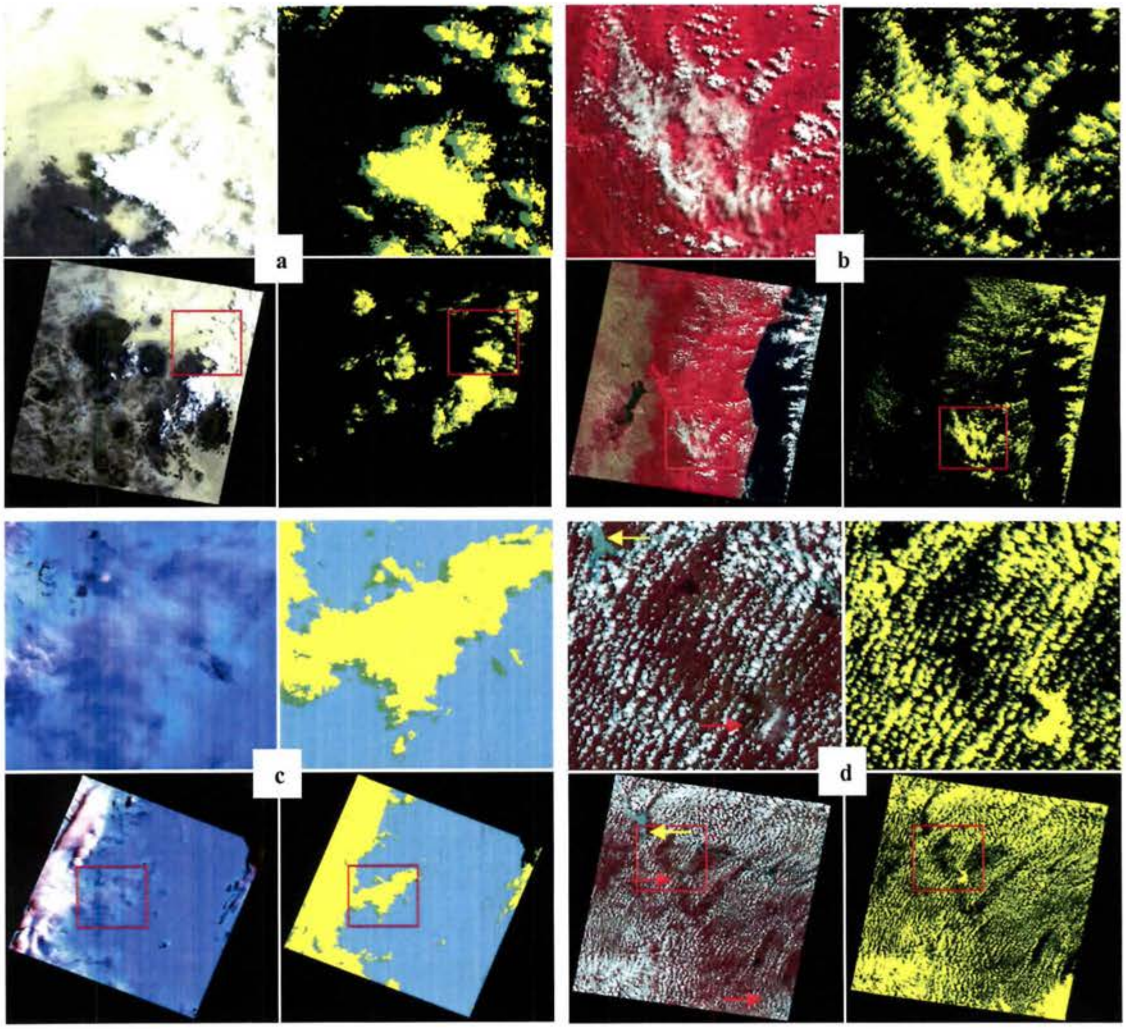


Figure 2.5. Fmask results of four Landsat scenes. (a). Results of a Sub-tropical South Landsat scene (p31_r43 & 20010615). (b). Results of a Sub-tropical North Landsat scene (p189_r47 & 20010805). (c). Results of a Polar South Landsat scene (p217_r107 & 20011215). (d). Results of a Tropical Landsat scene (p190_r54 & 20010929). In each Landsat scene, (lower left) shows an entire false color composited Landsat image (Figure 2.5a, b, and d are shown with Band 4, 3, and 2 composited; Fig 2.5c is shown with Band 5, 4, and 3 composited). (lower right) shows the corresponding Fmask cloud (yellow) and cloud shadow (green) mask for

the whole scene. Black pixels are clear. (upper left) and (upper right) images are enlargements of (lower left) and (lower right) images with a size of 60×60 km^2.

The ACCA reference scenes are the only sample available at present designed to systematically cover the full range of global environments and cloud conditions (Irish, et al., 2006). There are manual cloud masks for all reference scenes and a few of them have manual cloud shadow masks. A total of 188 Landsat Worldwide Reference System (WRS) locations in nine latitudinal zones were chosen (Figure 2.6).

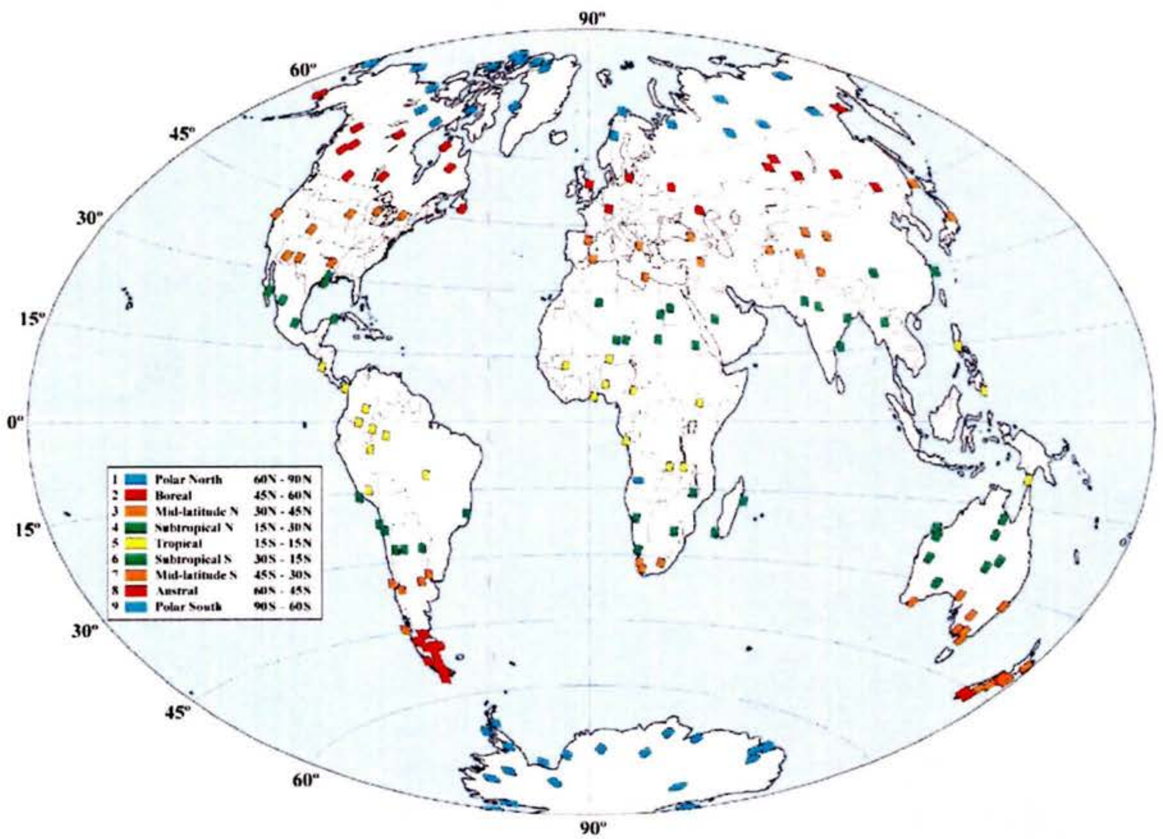


Figure 2.6. Locations of the Landsat scenes used in the reference dataset (Plate 7. in Irish, et al., 2006).

There are 212 reference scenes evenly distributed among nine latitude zones. The manual mask was derived by visual assessment of a full resolution scene in Adobe Photoshop using different combinations of bands (including overlay of the resampled thermal band if necessary) by three U. S. Geological Survey (USGS) image analysts. To obtain the approximate error of the manual masks, 11 scenes were examined by all three analysts, and the average difference was around 7% (Oreopoulos et al., 2001). Due to the difficulty in identifying cloud and cloud shadow, not all the reference masks are accurate enough for accuracy assessment of cloud identification at the pixel level. After carefully looking through the reference archive (by experts from Boston University and USGS), a total of 70 reference scenes were excluded, due to either low accuracy of the cloud manual mask or artifacts in the Landsat reference images. The remaining 142 reference scenes were used for accuracy assessment of Fmask results. The cloud shadow reference masks are not as accurate as the cloud reference masks as these reference scenes were originally interpreted to test estimates of percent cloud cover. In total, there are 26 scenes for accuracy assessment of cloud shadows. Five different accuracies were used to assess the accuracy of the algorithm results. Considering cloud and noncloud (including cloud shadow) as two classes, we have the following three accuracies for cloud accuracy assessment (Equation 2.25~2.27):

$$Cloud\ overall\ accuracy = \frac{agreement\ between\ manual\ mask\ and\ algorithm\ mask}{total\ pixels} \tag{2.25}$$

$$Cloud\ producer's\ accuracy = \frac{agreement\ of\ cloud}{agreement\ of\ cloud + omission\ of\ cloud} \tag{2.26}$$

$$Cloud\ user's\ accuracy = \frac{agreement\ of\ cloud}{agreement\ of\ cloud + commission\ of\ cloud} \tag{2.27}$$

On the other hand, considering shadow and nonshadow (including cloud) as two classes, we have the following two accuracies for cloud shadow accuracy assessment (Equation 2.28~2.29):

$$Cloud\ shadow\ producer's\ accuracy = \frac{agreement\ of\ shadow}{agreement\ of\ shadow + omission\ of\ shadow} \quad (2.28)$$

$$Cloud\ shadow\ user's\ accuracy = \frac{agreement\ of\ shadow}{agreement\ of\ shadow + commission\ of\ shadow} \quad (2.29)$$

The cloud shadow overall accuracy is not used for accuracy assessment, because cloud shadows are usually much smaller in size compared to clouds, and this would make the cloud shadow overall accuracy always high even if cloud shadows identification is totally wrong. We suggest that producer's accuracy is more important than user's accuracy, because errors of omission of clouds or cloud shadows are more serious than errors of commission. If clouds or cloud shadows are missed they will greatly undermine future analyses like change detection or image classification. However if clear areas are masked as clouds or cloud shadows, the only consequence is a little lost data.

In addition to the per-pixel accuracies described above, Fmask results are compared with ACCA in terms of percent cloud cover. ACCA mainly consists of two passes in which the second pass of ACCA is only used to improve the scene-wide cloud cover percent and the first pass is the only phase that creates a per-pixel cloud mask (Scaramuzza, private communication). Therefore, we compared the Fmask cloud cover percent with results of the second pass of ACCA, and Fmask cloud accuracies at the pixel level with results from the first pass of ACCA.

Estimates of percent cloud cover from Fmask are very accurate (Figure 2.7), with an R-square of more than 0.99. The slope of the regression line is 1.00, with a very small intercept (0.83%), and relatively small Root Mean Square Error (RMSE) (3.25%). ACCA estimates of percent cloud cover are also accurate, with an R-square of 0.95 and the slope of the regression line is 0.95 with an intercept of 0.39% and an RMSE of 6.56%. For the purpose of estimating percent cloud cover for a scene, Fmask appears to be an improvement over ACCA as except for the magnitude of intercept, it has a higher R-square value, lower RMSE, and less bias in the slope of regression line.

At the pixel scale, the average Fmask cloud overall accuracy is 96.41% with a small standard deviation of 3.2% (Figure 2.8). It is a significant increase compared with ACCA whose average overall cloud accuracy is 84.8% with a standard deviation of 11.9%. Cloud producers and users accuracies for images with cloud cover less than 5% were not analyzed here, as producers and users accuracies computed for clouds with very small size may be biased greatly because of the definition of cloud boundaries. The average Fmask cloud producers accuracy is 92.1% (Figure 2.9) with a standard deviation of 13.3% which is a significant improvement compared with ACCA whose average cloud producer's accuracy is 72.1% with a standard deviation of 26.5%. Moreover, the average Fmask cloud users accuracy is 89.4% (Figure 2.10) with a standard deviation of 9.8% which is similar as ACCA whose average cloud user's accuracy is 91.8% with a standard deviation of 12.3%. Considering that producers accuracy is more important (in our opinion) than users accuracy, the improvement of cloud identification in Fmask is significant compared with ACCA.

On the other hand, Fmask seems to overestimate cloud shadows, which is mainly caused by the 3 pixels buffering (in 8-connected neighborhood) for each cloud shadow pixel (Figure 2.11.). The average producer's accuracy for cloud shadow is larger than 70%, and its average users accuracy is around 50% (images with cloud shadows covering less than 1 percent of the image are not analyzed here). The lower accuracies are partly the result of errors in the manual cloud shadow masks and the relatively small size of cloud shadows compared to clouds in the scene. Even very small amounts of disagreement (differences in defining cloud shadow boundaries, mistakes in Fmask or reference shadow mask) reduce the cloud shadow user's and producer's accuracy greatly.

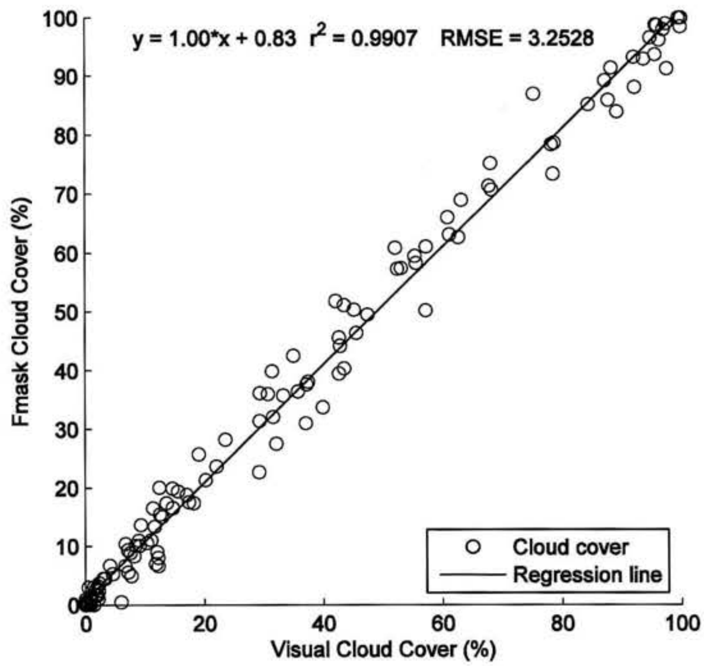


Figure 2.7. Visual cloud cover vs. Fmask cloud cover

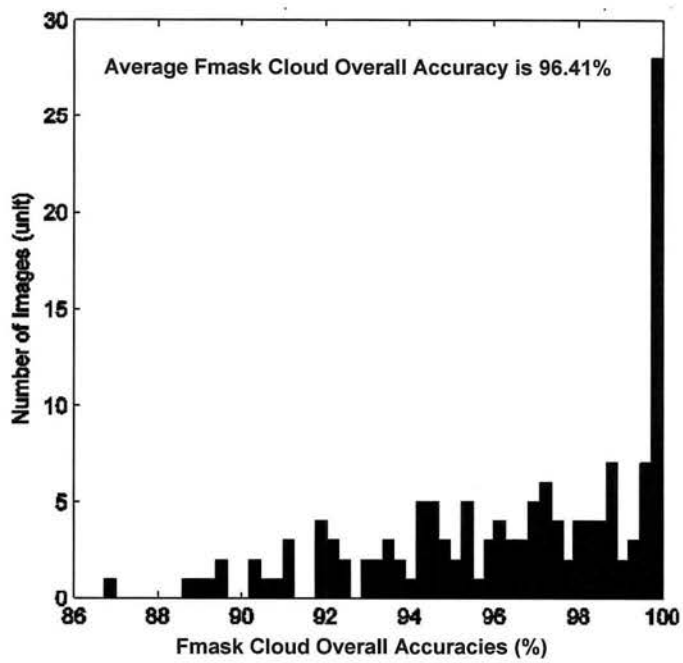


Figure 2.8. Histogram of Fmask cloud overall accuracies

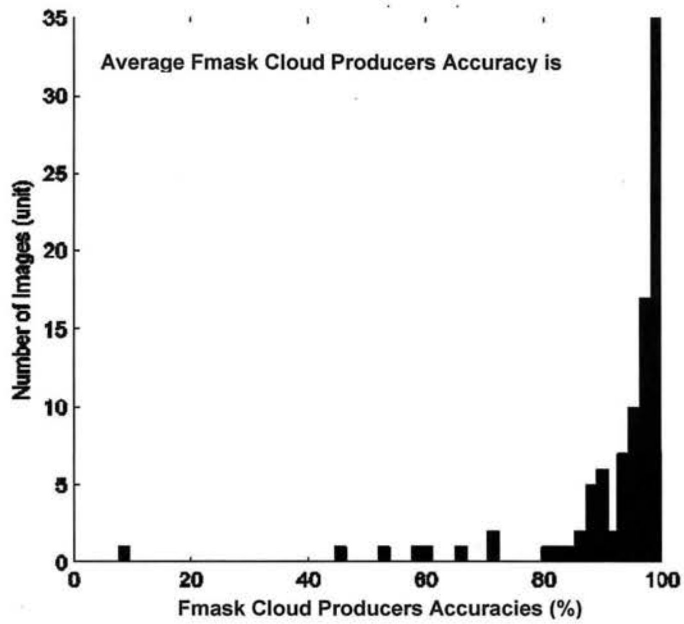


Figure 2.9. Histogram of Fmask cloud producer's accuracies

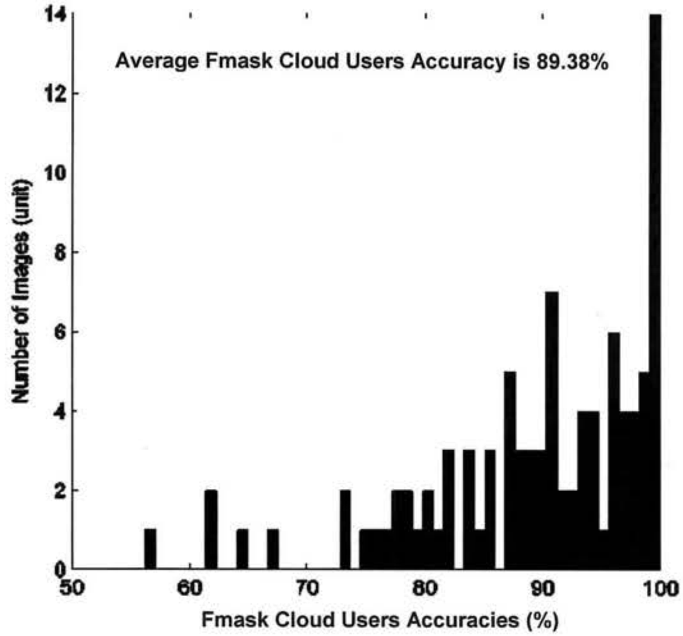


Figure 2.10. Histogram of Fmask cloud user's accuracies

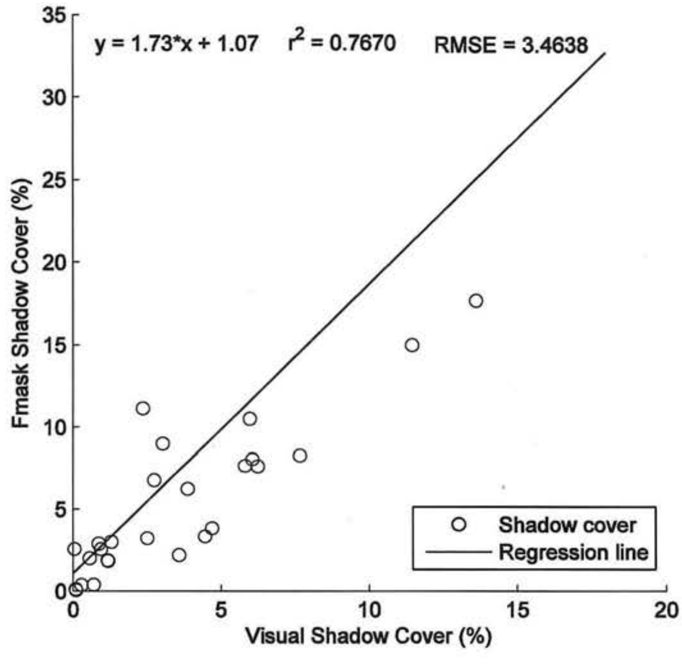


Figure 2.11. Visual cloud shadow cover vs. Fmask cloud shadow cover

The multitemporal cloud masking results are illustrated in Figure 2.12. The Image on the left is the original Landsat image. The images in the center and on the right show the semi-transparent cloud/cloud shadow masks overlaid on the original image. We dilate the clouds and shadows by 3 pixels in an attempt to exclude the edges of clouds and shadows that are often spectrally inseparable from the rest of the image. The clouds and cloud shadows missed by Fmask (Figure 2.12b) are identified by the multitemporal algorithm (yellow arrows).

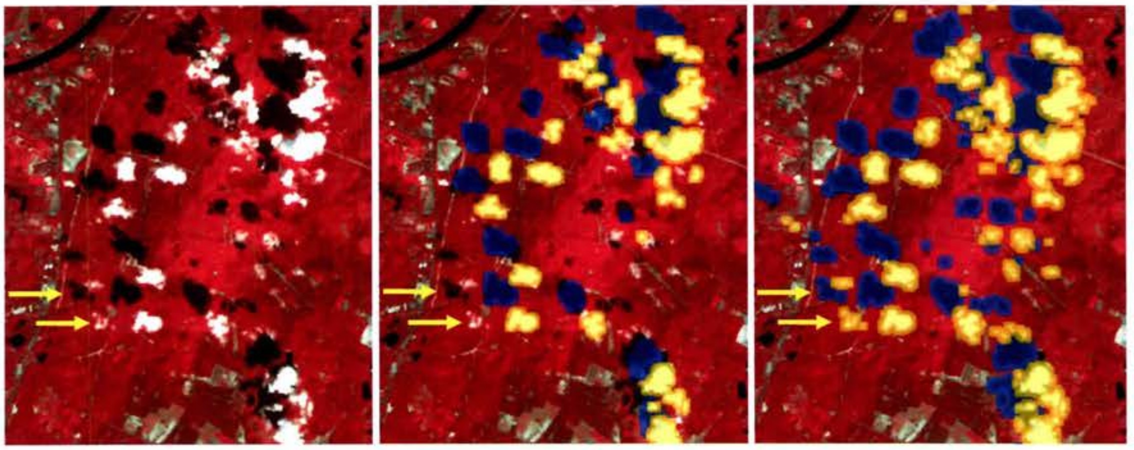


Figure 2.12. Illustration of the two-step cloud, cloud shadow, and snow masking results. Figure 2.12a shows a small piece of a Landsat image (shown with Bands 4, 3, and 2 in red, green, and blue). Figure 2.12b shows the results of the Fmask algorithm. Clouds are yellow and shadows are blue. Figure 2.12c shows the results after use of the multi-temporal approach. Notice that the cloud and cloud shadow missed in Figure 2.12b were found in Figure 2.12c.

2.4. Discussion and conclusions

The Fmask algorithm effectively finds clouds and cloud shadows, which helps with a wide assortment of remote sensing activities. The goal is to provide an automated method for screening clouds and their shadows such that time series of Landsat images can be easily compiled. The need for effective cloud and shadow screening has grown tremendously for two primary reasons. First, the Landsat L1T format now provides accurate enough registration of images that they can be compiled into a time series without significant attention to registration issues. And second, free access to the archive is changing what we consider a useful Landsat image. Now that images are free, it can be worth processing images even if substantial portions of the images are cloudy to extract the cloud free observations. As a result, more images with more clouds are being used and the need for automated cloud and cloud shadow screening is growing.

The estimates of percent cloud cover from Fmask are a slight improvement compared with ACCA estimates. The cloud masks generated from Fmask are significantly better than from the first pass of ACCA, with cloud overall accuracy of 96.41% (84.8% in ACCA), cloud producers accuracy of 92.1% (72.1% in ACCA), and cloud users accuracy of 89.4% (91.8% in ACCA). The cloud probability mask generated from Fmask will be beneficial for customizing cloud masking results, as instead of a binary mask, it can provide the probability of a pixel being cloudy. Users can make their own decisions in choosing the confidence level (e.g. 50%) for defining a cloudy pixel for their specific locations and applications. Fmask has achieved producer's accuracy for cloud shadow of more than 70% and user's accuracy higher than 50%. The reliability of these estimates is questionable due to the frequency of errors in the reference datasets. By examining each cloud shadow mask carefully, we find Fmask identified many cloud shadows that are not included in the cloud shadow reference masks. Therefore, more accurate reference data for cloud shadows are necessary for better assessment of cloud shadow detection algorithms.

There are some limitations in Fmask cloud detection. First, Fmask may fail to identify a cloud if it is both thin and warm (Figure 2.13 upper left and upper right images). These errors of omission may not be that important, as usually thin and warm "clouds" are actually haze or aerosols and they can be further removed by atmospheric correction (Vermote & Saleous, 2007). Second, Fmask may also identify other very bright and cold land features (salt pans, cold snow etc.) as clouds (Figure 2.13 lower left and lower right images). We think commission is better than omission in cloud detection, as this kind of error will only remove a few clear-sky pixels from subsequent remote sensing applications. Finally, as Fmask uses a scene-based threshold and applies this same threshold to all pixels in the image, it may not work well for some images with very complex surface reflectances.

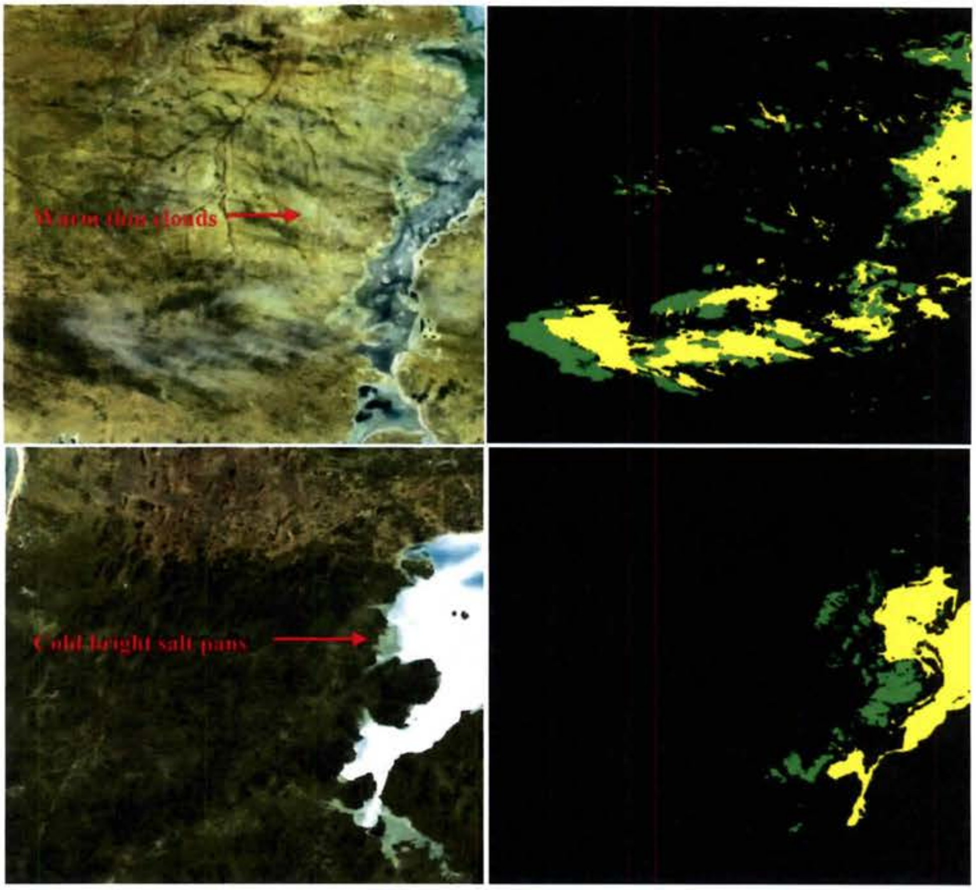


Figure 2.13 Two subset Landsat images from Mid-latitude South (p100_r82 & 20011212). (lower left) shows a subset of Landsat image where commission of clouds was observed (Band 4, 3, and 2 composited). (lower right) shows the corresponding Fmask cloud (yellow) and cloud shadow (green) mask. Black pixels are clear. (upper left) shows a subset of Landsat images where omission of clouds was observed (Band 4, 3, and 2 composited). (upper right) shows the corresponding Fmask cloud and cloud shadow mask. The red arrows point to the omission and commission errors of clouds in Fmask results as examples.

Fmask has three tuning parameters: the number of pixels dilated for cloud object, the number of pixels dilated for cloud shadow object, and the cloud probability threshold. For the following studies (Chapter 3 & Chapter 4), we dilate each cloud and cloud shadow by 3 pixels in 8 connected directions and choose a cloud probability of 12.5 for defining clouds. We used the cloud probability of 12.5 instead of the default value of 22.5 because preferring a more conservative strategy in cloud detection; omission of clouds is much more serious than commission of clouds as there are many Landsat images available. In Figure 2.14, we tested the Fmask algorithm with different cloud probability threshold against a globally set of distributed 142 reference sites (Zhu & Woodcock, 2012). Though the optimum overall accuracy is obtained at a cloud probability threshold of 22.5, the decreasing omission error can still balance most of the increasing commission error at a cloud probability threshold of 12.5, which is also demonstrated by similar cloud overall accuracies. When the cloud probability threshold is less than 12.5, the commission error increased significantly but the omission error only decreased slightly. Therefore, to reduce the false positive errors in land cover change detection, a cloud probability of 12.5 is used in the CCDC algorithm which only slightly reduced the cloud overall accuracy (still higher than 96%) while omission error decreased more than 30%.

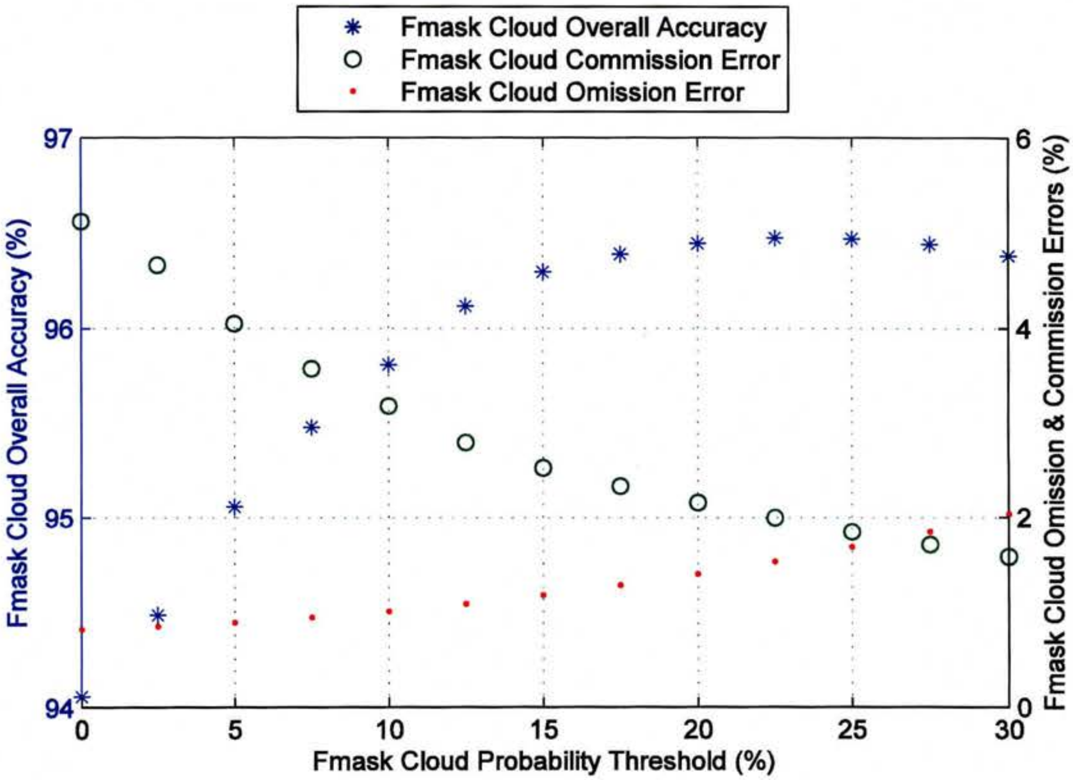


Figure 2.14. Fmask cloud probability threshold vs. cloud accuracies

By adding the time dimension, the second step is able to further improve cloud and cloud shadow detection. Moreover, only two optical bands are used, the thermal band is no longer a necessity in the second step. There are also limitations. First, in order to have the multitemporal fitting work, we need huge amount of data, which is difficult and time consuming. Second, there is a tradeoff between finding the abnormal pixels (cloud, cloud shadow, or snow) and false detection of the noises. For example, if we set the threshold too high, the false detection would be greatly reduced. At the same time, some of the thin clouds and not very dark cloud shadows may also be eliminated.

One issue needing further attention is the establishment of a standard definition for clouds and cloud shadows. For example, what category should we put pixels that are shadows from a high cloud falling on a lower cloud? Should we include smoke, thick aerosols, and haze in the cloud mask? Shall we still validate cloud shadows over water? Answers to these questions would facilitate future accuracy assessments and comparison of alternative methods.

In conclusion, the new two-step cloud and cloud shadow masking algorithm is able to provide an accurate cloud, cloud shadow, and snow mask which has limited remote sensing activities for a long time.

Chapter 3

3. Continuous monitoring of forest disturbance using all available Landsat data

3.1 Introduction

Change detection is one of the most difficult but also common problems in remote sensing activities. Landsat data has been widely used for change detection because of its long historic archive and high spatial resolution (Coiner, 1980; Coppin & Bauer, 1994; Jensen, et al., 1995; Cohen et al., 1998; Seto et al., 2002, Masek et al., 2008). Due to the large amount of carbon stored in forests, monitoring forest change is of great importance for understanding the global carbon budget (Dixon et al, 1994; Turner et al., 2004; Goward et al., 2008). Knowing where and when forest disturbance happens is crucial for forest management and carbon cycle modeling. Numerous forest change detection algorithms have been developed, validated, and applied to different parts of the world (see for example Collins & Woodcock, 1996; Hayes & Sader, 2001; Woodcock et al., 2001; Hostert et al., 2003; Healey et al., 2005; Healey et al., 2006; Kennedy et al., 2007; Goodwin, et al., 2008; Masek et al., 2008; Vogelmann et al., 2009; García-Haro et al., 2010; Hilker et al., 2009; Huang et al., 2010; Kennedy et al., 2010).

Most of the change detection algorithms are based on two dates of Landsat images (see for example Collins & Woodcock, 1996; Woodcock et al., 2001; Healey et al., 2005; Healey et al., 2006; Masek et al., 2008). These algorithms are simple to use, but not always applicable. One problem is that both images have to be at the same time of year to minimize phenology differences and Bidirectional Reflectance Distribution Function (BRDF) effects. The other problem is that the input images need to be cloud and snow free. Even under these conditions, these change detection algorithms can only provide the spatial pattern of the disturbance that occurred between the two images and it is impossible to know when the change occurred in the time between the two images. This is important because the time period between images is frequently as much as five or ten years.

Based on the idea that forest change can be better detected using many observations of a place and the increasing need for detecting changes as they are occurring, a number of methods for change detection using many dates of Landsat imagery have been developed (Kaufmann & Seto, 2001; Hostert et al., 2003; Kennedy et al., 2007; Goodwin, et al., 2008; Vogelmann et al., 2009; Kennedy et al., 2010; Hilker et al., 2009; Huang et al., 2010). These algorithms have been shown to be more automatic in identifying forest change and more robust to noise from registration, BRDF, and seasonal effects.

Nevertheless, these newly developed algorithms still have limitations in image selection, as all the images used should be within the growing season to minimize phenology and BRDF differences and at the same time they should be almost cloud and snow free to make multi-temporal image differencing possible. Though some new indices such as the Integrated Forest Z-score (IFZ) (Huang et al., 2010) and Disturbance Index (DI) (Healey et al., 2005; Healey et al., 2006; Masek et al., 2008) can reduce phenology and BRDF effects by normalizing the indices with a predefined forest sample, they may have problems when handling data from different seasons or within heterogeneous areas where both deciduous and evergreen forests exist. Therefore, cloud and snow free Landsat images from three of the seasons will not be able to be used. Sometimes images acquired during times other than the growing season, such as winter images, can be more useful than growing season images for detecting subtle disturbance (pest infestation) when the forest understories are dense during the growing season, which makes the mixed signal almost the same as healthy forest (Bolton and Woodcock, in preparation). Some studies even found that snow-covered Landsat imagery can be used for change detection, often allowing for a longer period of observed changes than the growing season (Takao, 2003). Due to the requirement for cloud and snow free images acquired during the growing season, most of the multi-temporal change detection algorithms can only provide annual or biennial change results.

Recently, Hilker et al. (2009) used both Moderate Resolution Imaging Spectroradiometer (MODIS) and Landsat to detect forest disturbance in Canada. By blending MODIS and Landsat data, a high temporal frequency (16 days) and fine spatial resolution (30m) disturbance map was produced. Though this algorithm can identify the time of forest disturbance at a high frequency, it may take a long time to acquire two clear Landsat images for the blending of MODIS and Landsat data. If we want to monitor immediate problems such as illegal logging or encroachment on protected area, we need an algorithm that can monitor changes as they are occurring (within a few weeks) and can update the results as soon as a new observation is available.

Because of the limitations of existing algorithms for fully utilizing the Landsat archive, we developed the Continuous Monitoring of Forest Disturbance Algorithm

(CMFDA), which uses all available Landsat images to find the location and timing of forest disturbance. CMFDA considers each pixel separately, taking advantage of any clear views for each pixel to track spectral trends over time. In this study we are only focusing on human-induced forest disturbance (forest clear-cut/thinning), though CMFDA may be able to identify other natural-induced forest disturbances that cause surface reflectance to deviate from expected values. CMFDA produces a map showing where and when the disturbance happens at high spatial resolution (30 m) and temporal frequency (a few weeks). Currently, the highest temporal frequency for CMFDA is 8 days when both Landsat 5 and Landsat 7 are used. Clouds and cloud shadows are flagged as part of the procedure. The following steps are necessary to implement and test CMFDA:

1. A two-step cloud, cloud shadow, and snow masking approach.

2. Estimate time series models of surface reflectance.

3. Define a stable forest mask

4. Predict the "next" clear observations to serve as a basis for comparison with new observations.

5. Detect forest disturbance with single-date and multi-date differencing algorithms

6. Test the disturbance map derived from the algorithms against an independently-derived reference map.

3.2. Study area and data

The study area (Figure 3.1) is located in the Savannah River Basin covering 2000×2000 Landsat pixels (60×60 km). The Savannah River is along the border between Georgia and South Carolina. There are a variety of land covers in this study area. Along the Savannah River, there is a large area of deciduous forest and wetland. Most of this study area is covered by evergreen forest and agriculture. Three urban areas (Sylvania, Allendale, and Estill) along the Savannah River are within the study area. Though there is no snow present in this study area, it is frequently cloudy. By applying a newly developed cloud and cloud shadow detection algorithm called Fmask (Function of mask) (Zhu and Woodcock, 2012) to all available Landsat ETM+ images (Path 17 and Row 37) from 2001 to 2002 (Figure 3.2), the expected frequency of cloudy observations for each pixel was approximately 50% and almost all the cloud free images were acquired at the beginning or end of the year (Figure 3.3). The cloud cover during the growing season was always heavy. There was not a single cloud free ETM+ image during the growing season for either 2001 or 2002.

We used a time series of Landsat TM and ETM+ images for Path 17 and Row 37. All available ETM+ L1T images (a total of 64) acquired from 2001 to 2004 were downloaded if the cloud cover was less than 90%. Because the year 2003 was the time that CMFDA was applied to find disturbances and also the year when Scan Line Corrector failed in Landsat 7, all available TM L1T images (a total of 12) acquired in 2003 were downloaded if cloud cover was less than 90% to help find the disturbance time with higher frequency.

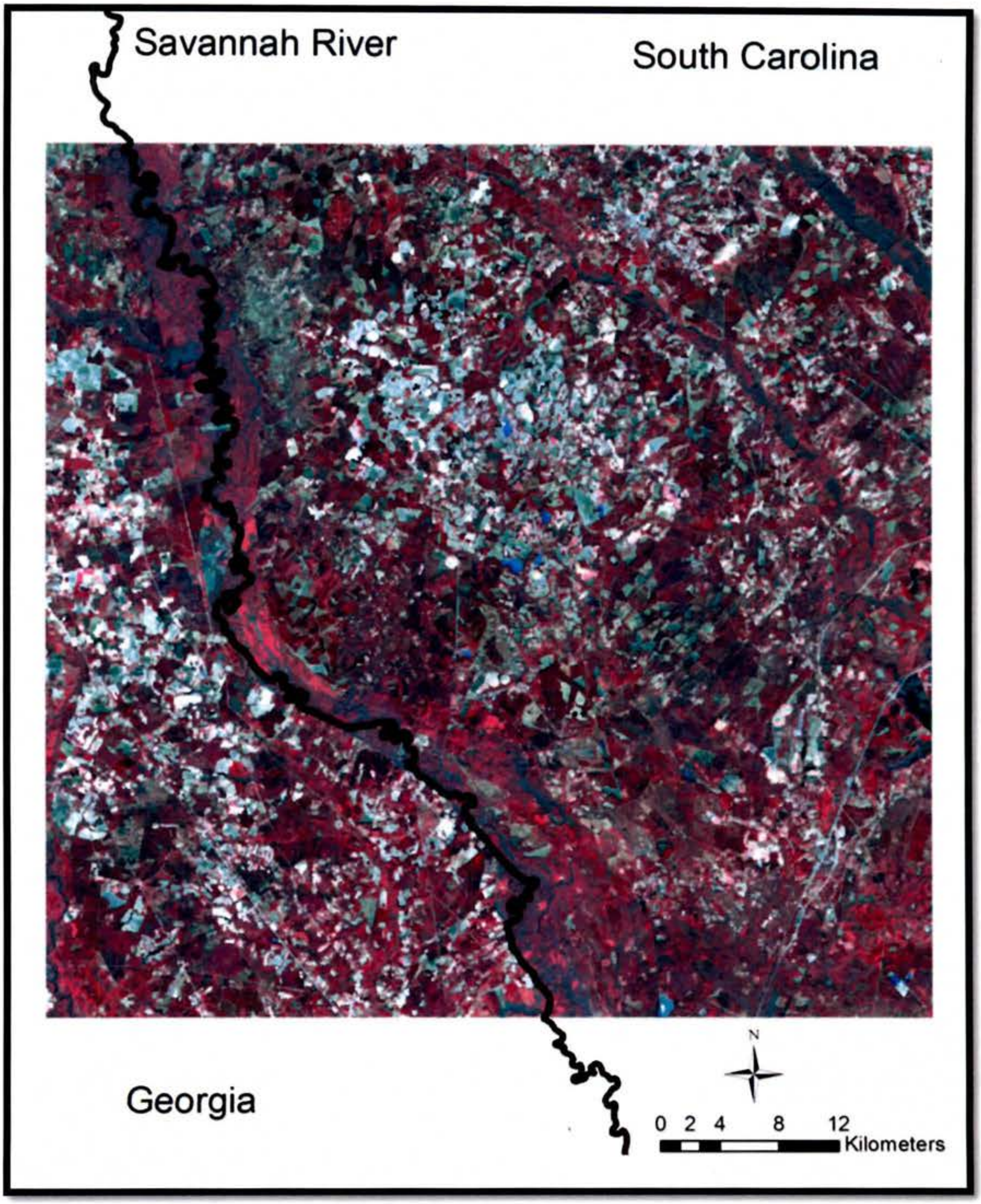


Figure 3.1. Study area (subset of November 23rd 2002 Landsat ETM+ image shown with Bands 4, 3, and 2 in red, green, and blue)

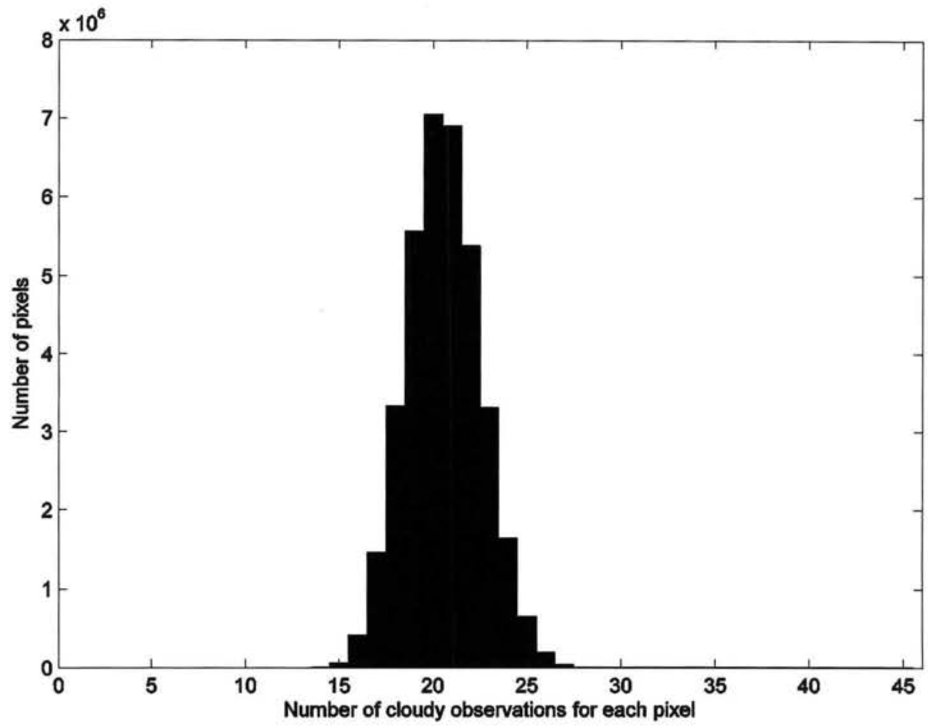


Figure 3.2. Histogram showing the frequency of cloudy observations from 2001 to 2002 for all available ETM+ images. There were 46 images from this time period, indicating that about half of all observations collected were cloudy. Notice that this is not the same as saying half the images were cloud free.

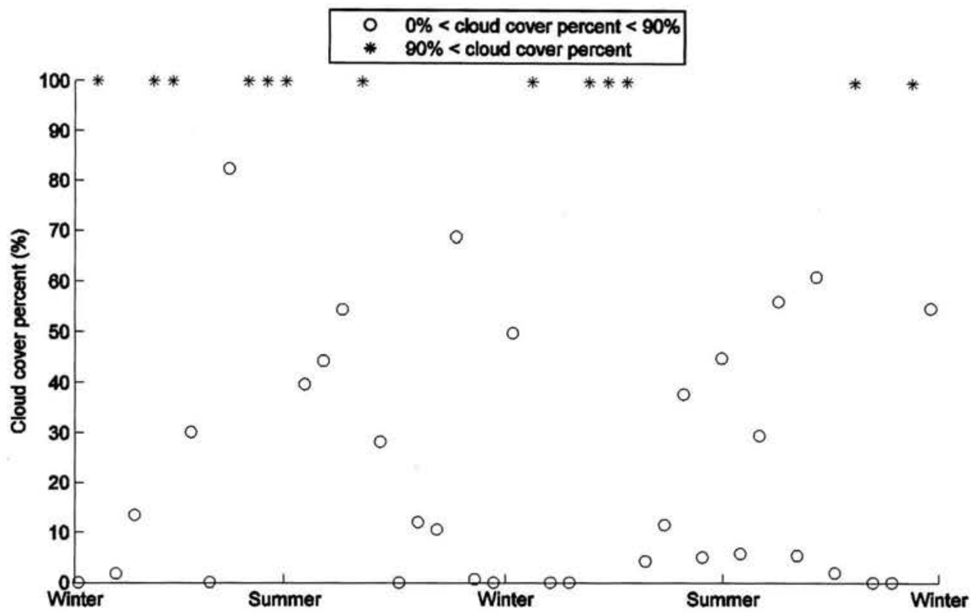


Figure 3.3. Cloud cover percent of all available ETM+ images from 2001 to 2002. Notice that during the peak growing season (summer), not a single ETM+ image was "cloud free" (meaning zero percent cloud cover).

3.3. Methodology

CMFDA has many component parts, including: image preprocessing; single-date masking of clouds, cloud shadows, and snow; multi-temporal masking of clouds, cloud shadows, and snow; estimation of surface reflectance models; mapping of stable forest; predicting the "next" Landsat observations, and identification of forest disturbance (Figure 3.4).

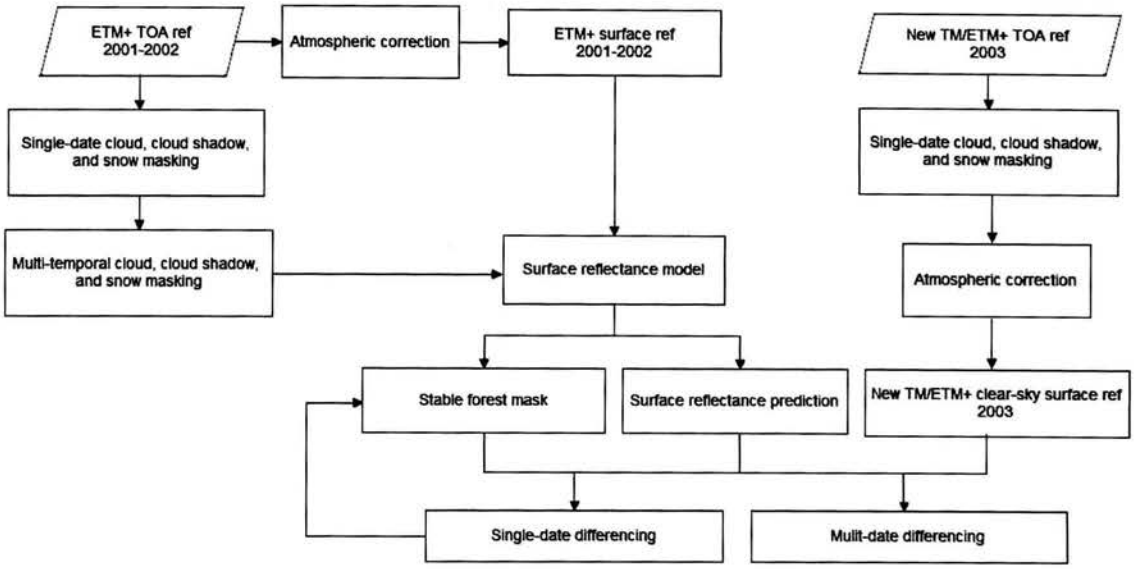


Figure 3.4. Flow chart of continuous monitoring of forest disturbance algorithm

3.3.1. Preprocessing

Geometric registration and radiometric normalization are important steps in change detection, facilitating comparison of change indices across time and space. We assume Landsat L1T images are already precisely registered. All the Landsat images were atmospheric corrected with the LEDAPS, using the 6S radiative transfer approach (Vermote et al., 1997; Masek et al., 2006). Clouds, cloud shadows, and snow were screened by the two-step method introduced in Chapter 2.

3.3.2. Estimating the surface reflectance model

After removing observations covered by clouds, cloud shadows, and snow, CMFDA uses the remaining clear Landsat observations to estimate surface reflectance models for each pixel using the Ordinary Least Square (OLS) algorithm. At this stage, OLS fitting is used rather than RIRLS simply because it is faster and any significant outliers have been removed. The surface reflectance model is a function of sines and cosines shown in Equation 3.1. It includes a two term harmonic (Fourier) model (Rayner, 1971; Davis, 1986) and an inter-annual change model newly developed here. The two term harmonic model (i=1 & N+1) is used to capture the seasonality and BRDF effects of the annual surface reflectance. The harmonic model is chosen due to the advantage of requiring estimation of fewer coefficients (fewer necessary clear observations) and being less sensitive to short term data variations and inherent noise (e. g., missed clouds, cloud shadows, snow, and image misregistration). The inter-annual change model (i=2, 3, 4…N) is composed of sines and cosines that represent variation that occur on an i-year cycle, which mostly result from land cover change:

$$f(x) = a_0 + \sum_{i=1}^{N} \left(a_i \cos(\frac{2\pi}{iT}x) + b_i \sin(\frac{2\pi}{iT}\mathrm{x})\right) + a_{N+1}\cos(\frac{2\pi}{0.5T}x)$$

$$+b_{N+1}\sin(\frac{2\pi}{0.5T}\mathrm{x}) \tag{3.1}$$

Where,

x: Day-of-year.

N: Number of years.

T: Number of days per year (T = 365).

a_0: Overall surface reflectance.

a_i, b_i: Coefficients that capture the changes of surface reflectance for the i^{th} year.

a_{N+1}, b_{N+1} : Coefficients that capture the bimodal variations of surface reflectance for each year.

Two years (2001 and 2002) of Landsat ETM+ images (33 images) were used to estimate the surface reflectance model for the Savannah River site. In this case, there are 7 parameters for each surface reflectance model. The last two parameters a_3 and b_3 are used to capture the bimodal variations for each year, which mostly occurs in agricultural areas due to an initial period of growth in the spring that is followed by plowing and a second period of growth. The parameters a_1 and b_1 are used to capture the annual change caused by phenology and BRDF effects. The inter-annual change is captured by the two parameters a_2 and b_2. The mean overall surface reflectance for the two years is represented by a_0. To estimate these 7 parameters, at least 7 clear observations are necessary in two years. To strengthen the robustness of the fitting, CMFDA only estimates a pixel if the number of clear observation is more than one and a half times of the number of total parameters to be estimated, that is a total of 11 clear observations. Considering the 23 observations per year from Landsat 7 and 23 observations per year from Landsat 5 per year at this U.S. site, 11 clear observations in two years are easily obtained even though there were very few cloud free images. In fact, when only using the ETM+ images, the highest number of cloudy observations at the pixel level is 28 (Figure

3.2), meaning there would be at least 18 clear observations for each pixel between 2001 and 2002.

3.3.3. Defining a forest mask

Some land uses exhibit abrupt changes in surface reflectance that do not represent land use change. For example, agricultural fields are plowed, resulting in reflectance changes that do not represent land use change. To limit our change detection efforts to forests, we created a mask of the forest areas and only assessed these pixels for change. The forest mask was produced automatically using the estimated coefficients derived from the 33 Landsat ETM+ images acquired between 2001 and 2002 based on the fact that forests are observed to have high NDVI values (Masek et al., 2008) and low reflectance in the SWIR bands (Kennedy et al., 2007; Huang et al., 2010). As compared to Band 5, Band 7 is more robust to different atmosphere conditions, so we chose Band 7 as our SWIR band for extracting forests. From the previous surface reflectance models we have the overall surface reflectance represented by $a_{0,i}$ (i=1, 2, 3, 4, 5, and 7), where i stands for the Landsat TM or ETM+ band number. Therefore, the overall NDVI values can be calculated with Band 4 and Band 3 overall surface reflectance model coefficients ($a_{0,3}$ and $a_{0,4}$) and the overall Band 7 surface reflectance ($a_{0,7}$). We define a possible forest pixel if it meets the criteria that overall NDVI is larger than 0.6 and overall Band 7 reflectance is less than 0.1. To better illustrate how the two thresholds work, we plotted ten samples of the time series for each land cover class for Band 7 and NDVI (Figure 3.5). In Figure 3.5, the overall forest NDVI values are always above 0.6, but sometimes other vegetation types like crop, grass, and shrub can also have overall forest NDVI values above 0.6. As forests are usually dark in SWIR bands compared to other vegetation types, a threshold of 0.1 in overall Band 7 surface reflectance excludes other vegetation types that may have high NDVI values.

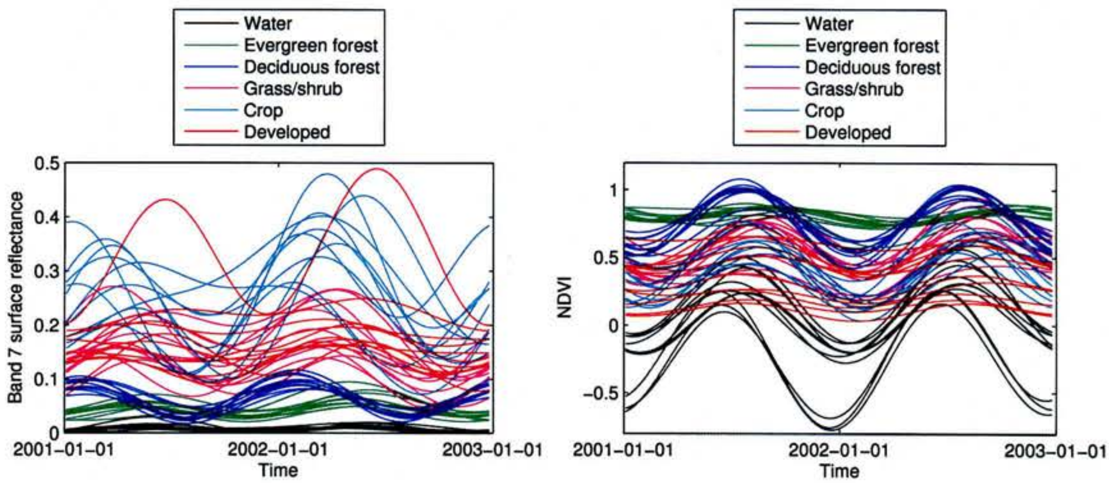


Figure 3.5. Estimated Band 7 and NDVI temporal trajectories of different land cover classes between 2001 and 2002. For each land cover class, ten time series of samples were estimated for their surface reflectance models.

If land cover change occurs within the estimating period (2001-2002), future use of the resulting model will be problematic. Therefore, we excluded these pixels that were changed during the estimating period from the forest mask. As parameters a_2 and b_2 capture the inter-annual differences, their amplitude represents the degree of changes (excluding phenology and BRDF differences) in surface reflectance within the estimating period. In this study, surface reflectance parameters computed from Band 7 were used for detecting changes that occurred in the estimating period because of their robustness to atmospheric influences and sensibility to forest disturbance. A threshold for pixels where $\sqrt{a_{2,7}^2 + b_{2,7}^2}$ is larger than 0.02 worked well in identifying the pixels that changed during the estimating period. However, if change occurs at the end of the estimating period, one or two changed observations may not contribute enough to the inter-annual change parameters, making $\sqrt{a_{2,7}^2 + b_{2,7}^2}$ still less than 0.02.

As the model cannot capture forest disturbance that happened at the end of the estimating period, the predicted values from the model will still have similar values as if no change occurred. In this case, these disturbed pixels can be easily identified by comparing the last clear observations with the predicted values (see section 3.3.4.1.1 for detail). The final forest map is created by combining all these criteria above in Equation 3.2 and the final forest mask is shown in Figure 3.6.

$$forest\ mask = \frac{a_{0,4}-a_{0,3}}{a_{0,4}+a_{0,3}} > 0.6\ AND\ a_{0,7} < 0.1\ AND\ \sqrt{a_{2,7}^2 + b_{2,7}^2} < 0.02$$

$$AND\ stable\ in\ the\ last\ clear\ observation \qquad (3.2)$$

The use of the magnitude of the interannual change parameters (a_2 and b_2) to find change represents a new method for finding change retrospectively. Since we are pursuing methods for monitoring forest change as it is occurring, we used this approach here only to find changes during the estimating period so that they won't be confused with forest change in the testing period.

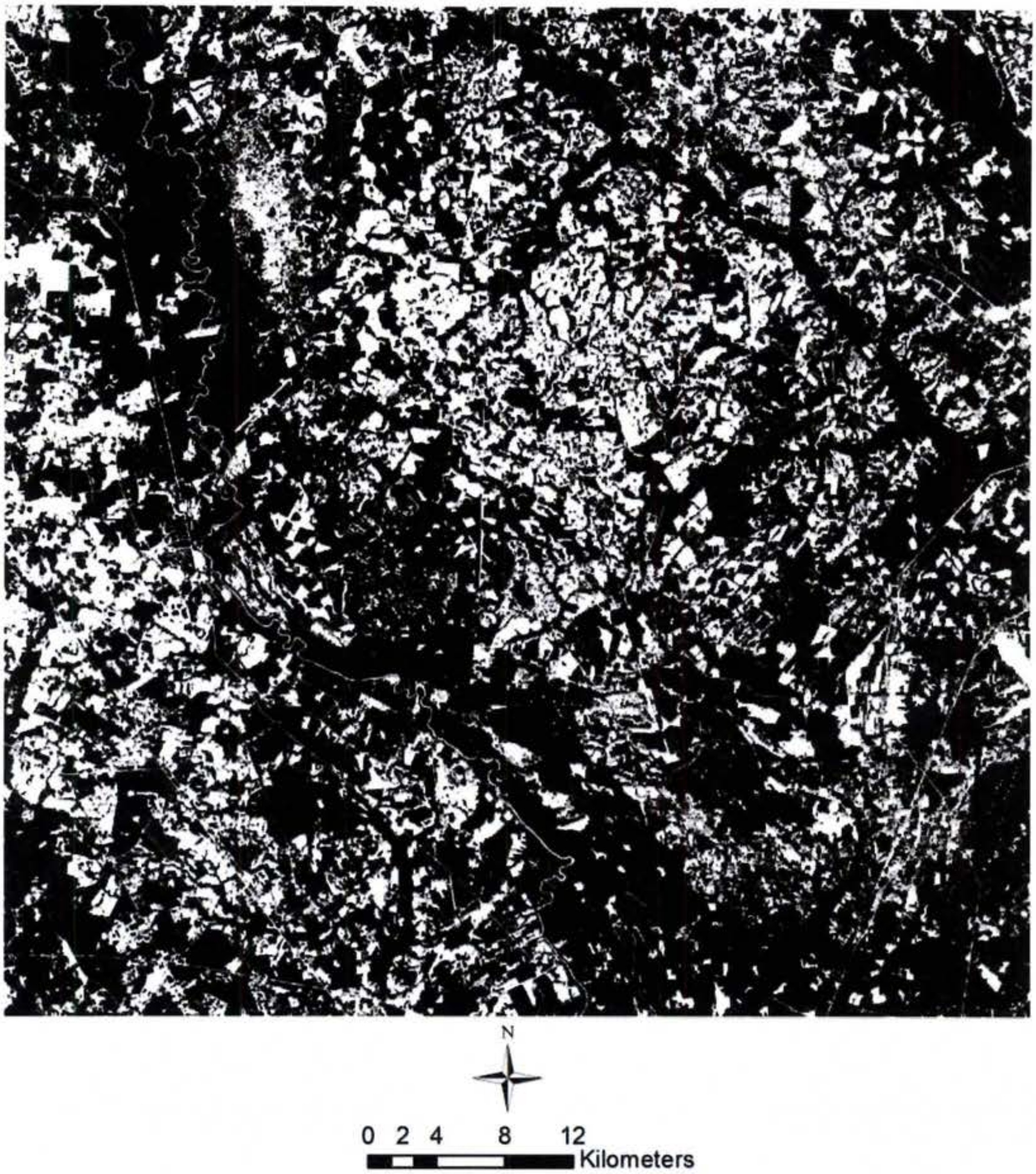


Figure 3.6. Forest mask for the study area (white for disturbed or nonforested areas; black for stable forested areas)

3.3.4. Predicting the "next" Landsat image

Assuming no land cover change has occurred, Equation 3.3 is used to predict the surface reflectance for each pixel and each spectral band at any time of the year, or the equivalent of the "next" Landsat image. The parameters were previously estimated using Equation 3.1 based on images between 2001 and 2002. Removing the inter-annual change parameters, the rest of the variables capture all kinds of influences including phenology, BRDF, topographic shadowing, etc. A Landsat image at any date can be predicted as:

$$predict(x) = a_0 + a_1 \cos(\frac{2\pi}{T}x) + b_1 \sin(\frac{2\pi}{T}x) + a_2 \cos(\frac{2\pi}{0.5T}x)$$
$$+b_2 \sin(\frac{2\pi}{0.5T}x) \tag{3.3}$$

Where,

x: Day-of-year.

T: Number of days per year (T = 365).

a_0: Overall surface reflectance

a_1, b_1: Annual changes of surface reflectance

a_2, b_2 : Bimodal variations of surface reflectance for each year.

After estimating the models for each pixel and for each spectral band, it is possible to predict what the "next" Landsat image will look like at any location and any date if there is no snow, cloud, or cloud shadow (Figure 3.7).

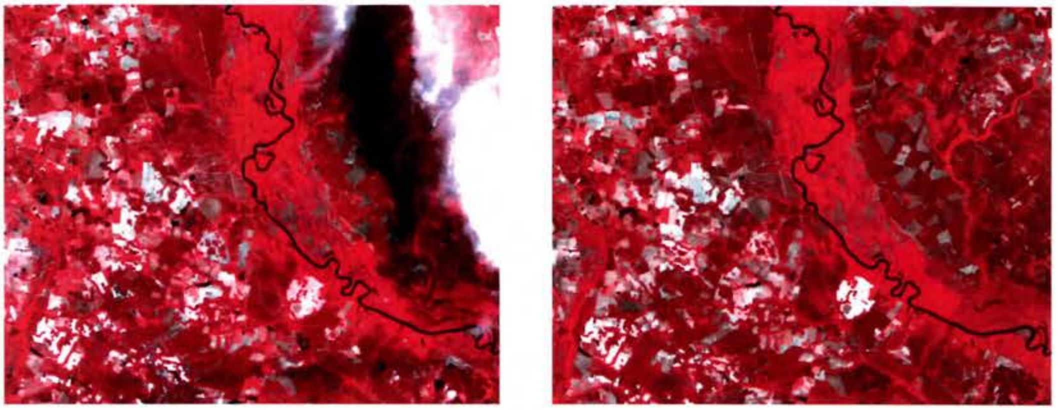


Figure 3.7. (left) Observed and (right) predicted Landsat surface reflectances at a subset of the study area (shown with Bands 4, 3, and 2 in red, green, blue)

One question that arises is how well we predict future Landsat images. We test this by predicting images for dates of future acquisitions and comparing them with real images. For these comparisons it is important to exclude pixels that have undergone land cover change. To select only "stable pixels" (i.e. no land cover change), we again use the parameters from the time series model ($\sqrt{a_{2,7}^2 + b_{2,7}^2} < 0.02$ & stable in the last clear observations). Several metrics have been used to assess the error of predicted Landsat images as compared to observed Landsat images (Table 3.1). Root Mean Square Error (RMSE) and Residual normalized by Mean reflectance (RM) has been used for error assessment separately by Gao et al. (2006) and Roy et al. (2008). We also use a conventional R-square metric. Both Gao and Roy's algorithms use multi-temporal MODIS data and Landsat data to predict Landsat observations. Our CMFDA prediction only uses multi-temporal Landsat data and shows very small errors. Four clear Landsat ETM+ images acquired in different seasons were compared with the predicted images for the study area (Table 3.1). The prediction errors for autumn and winter images are smaller than those in spring and summer images. The spring and summer images are more difficult to predict because phenological change is large during these times.

Table 3.1. R2, RM, and RMSE for predicted Landsat images at different seasons. For each spectral band, the highest R2, smallest RMSE and RM are in bold.

Time	Spring (04/26/2001)			Summer (06/16/2002)			Autumn (10/06/2002)			Winter (01/04/2001)		
Metrics	R^2	RM	RMSE	R^2	RM	RMSE	R^2	RM	RMSE	R^2	RM	RMSE
Band 1	**0.89**	0.19	0.006	0.84	0.16	0.006	0.83	**0.13**	**0.004**	**0.89**	**0.13**	**0.004**
Band 2	0.91	0.2	0.009	0.87	0.16	0.01	0.9	**0.08**	**0.004**	**0.92**	0.09	0.005
Band 3	**0.94**	0.22	0.009	0.89	0.26	0.014	0.89	0.17	**0.006**	**0.94**	**0.11**	0.007
Band 4	**0.91**	0.04	0.016	0.89	0.07	0.027	0.89	**0.03**	**0.012**	0.9	0.06	0.014
Band 5	0.95	0.1	0.018	0.91	0.08	0.018	0.94	**0.05**	**0.011**	**0.97**	0.06	0.013
Band 7	0.95	0.14	0.014	0.91	0.17	0.019	0.91	**0.1**	0.01	**0.96**	**0.09**	0.011

3.3.4. Change detection algorithms

The basis of our methods is comparison of the predicted images with observed images to find change. Since we can make these comparisons for any date that has Landsat acquisitions, we are faced with a question regarding how many dates, or comparisons, to use. Ideally, a single comparison would be definitive. However, there is sufficient noise in the system due to factors like atmospheric haze, missed clouds or cloud shadows, that when using a single date for comparison, there are numerous false positives (false identification of forest disturbance). One approach to try to minimize this effect is to process a set of dates together as a group, as noise factors tend to be ephemeral in nature, but forest disturbance is persistent through time.

3.3.4.1. Single-date differencing algorithm

Ideally, if clouds, cloud shadows, and snow are well screened, the difference between the predicted and the observed image at the same day-of-year should be land cover change. However, omission of clouds and thick aerosols can lead to comparably large rates of false detection of forest disturbance. Fortunately, most of surface reflectance change caused by missed thin clouds or heavy aerosols behaves different spectrally from changes caused by forest disturbance. Though, in both of these situations, Band 1 and Band 7 surface reflectance will increase, the magnitude differs greatly. When forest is clear-cut or thinned, Band 7 will be strongly influenced while Band 1 will be only slightly changed (Healey et al., 2006), which is the opposite of the effect of thin cloud and thick aerosols. A ratio between increases in Band 7 and Band 1 can separate the noise effects (thin clouds and thick aerosols) from forest disturbance. Based on sensitivity analysis, a threshold of 3.0 for the ratio was used. Pixels with ratio values less than this threshold are ignored even if they show large difference in change indices.

Many different change indices have been developed for detecting forest disturbance. The simplest change index is the original surface reflectance. Healey et al. (2006) suggested the red band and SWIR bands are more sensitive to forest disturbance than the other Landsat optical bands. The most commonly used indices are from the Tasseled Cap Transformation (Crist & Cicone, 1984; Crist, 1985) which reduces the six

Landsat optical bands into three orthogonal indices - Brightness (B), Greenness (G), and Wetness (W), capturing the three major axes of spectral variation across the solar reflective spectrum. Wetness is particular useful in forest disturbance detection (Collins & Woodcock, 1996; Franklin et al., 2000). The Normalized Difference Vegetation Index (NDVI) (Tucker, 1979) derived from TM Band 3 and Band 4 has been used extensively in many kinds of vegetation change detection algorithms (Zhan et al., 2002; Jin & Sader, 2005; Lunetta et al., 2006). The Normalized Burn Ratio (NBR) has been used to assess the burn severity in Landsat images using Band 4 and Band 7. This index provides the best difference between health and burned vegetation (Howard et al., 2002). Recently, the newly developed Disturbance Index (DI) (Healey et al., 2005; Healey et al., 2006) has been used for large area Landsat forest disturbance detection (Masek et al., 2008). It is based upon the observation that cleared forest stands usually have a higher Brightness value and lower Greenness and Wetness values than forest stands. The DI transformation is a linear combination of the three Tasseled Cap indices re-scaled by the mean and standard deviation of the scene's forest value. As CMFDA models the phenology and BRDF effects in the data, we do not need to re-scale the DI by the mean and standard deviation. We simply used a linear combination of B-(G+W) as one of the tested change indices. We tested most of the indices discussed above, including: Band 3 surface reflectance, Band 7 surface reflectance, NDVI, NBR, Wetness, and B-(G+W). Equation 3.4 was applied to every pixel for the observed image and predicted image of the same day and this process was repeated for all available dates of Landsat data to identify forest disturbance. The B-(G+W) change index was used as it performed the best among all the tested change indices when a threshold of 0.18 was used for our study area (see Section 3.4.1.2. for detail).

$$forest\ disturbance = stable\ forest\ AND\ clear\ observation\ (Fmask)$$

$$AND\ \frac{\text{obs B7}-\text{pred B7}}{\text{obs B1}-\text{pred B1}} > 3$$

$$AND\ (B-(G+W))_{obs} - (B-(G+W))_{pred} > 0.18 \qquad (3.4)$$

As we are identifying the disturbance pixels using a single observation, the only criterion that determines whether a pixel has changed or not is the change magnitude of the index and a fixed Band 7/Band 1 ratio. Therefore the single-date differencing algorithm is sometimes affected by ephemeral noise and causing relatively large commission errors in change detection. Moreover, at different locations and for different forest types, the magnitude of the change threshold may differ, and the Band 7/Band 1 ratio test may fail, leading to lower accuracy for the single-date differencing algorithm. A method based on multi-date observations may solve these problems.

3.3.4.2. Multi-date differencing algorithm

One way to help reduce the effect of noise factors on commission errors (false forest change) is to use multiple observations through time. If a pixel is observed to change in multiple successive images, it is more likely to be forest disturbance. The multi-date differencing algorithm determines a disturbance pixel by the number of times that observed and predicted images differ more than a threshold in successive images. Pixels showing change for one or two times will be flagged as "probable change". If a third consecutive change is found, the pixel is assigned to the "change" class. Optimal results were obtained when B-(G+W) was used as the change index with a threshold of 0.12 (see section 3.4.2.2. for detail). The details of the multi-date differencing algorithm are presented in Equation 3.5.

$$forest\ disturbance = stable\ forest\ AND\ clear\ observation\ (Fmask)$$
$$AND\ (B-(G+W))_{obs} - (B-(G+W))_{pred} > 0.12\ three\ consecutive\ times \quad (3.5)$$

Notice that its optimal threshold is 0.06 less than the single-date differencing algorithm. With a lower threshold, very subtle changes (forest thinning) will be identified. Thanks to the addition of the temporal dimension, most of the commission errors were excluded by the need for consecutive observations to exceed the threshold (see section 3.4.2.2 for details). Also, the empirically derived Band 7/Band 1 relationship is no longer necessary. Considering all these benefits from using consecutive observations, the multi-date differencing algorithm was chosen as the final CMFDA result, and the single-date differencing algorithm was used only for finding changes in the last clear observation in defining stable forest mask.

3.3.5. Accuracy assessment

3.3.5.1. Reference map

Maps derived from remotely-sensed imagery should always be assessed for accuracy against an independent dataset that is closer to the truth. This independent dataset is usually from *in situ* field work or manual interpretation of finer resolution images like IKONOS or QuickBird. In this study, not only do we need to determine where disturbance occurs but also when. As there are no independent datasets available that have finer spatial resolution and higher temporal frequency than Landsat images, the reference data were derived from manual interpretation of the original Landsat images (Cohen et al., 2010). High spatial resolution images from Google Earth (http://earth.google.com/) were used to help the manual interpretation. Though the high spatial resolution images in Google Earth cannot provide the same temporal frequency as Landsat TM/ETM+, their high spatial resolution is helpful in separating forest, nonforest, and disturbance at longer time intervals. Two types of forest disturbance (clear-cut and thinning) were included in this reference dataset. The partial cuts and clear-cuts are quite easy to identify in the high spatial resolution images in Google Earth, as the details of the individual trees can be clearly seen. False color composites of Landsat Band 4, Band 3, and Band 2 surface reflectances were used to visualize the different types of disturbances (see Figure 3.1 for example). In these images, mature forests appear dark red, while clear-cut areas are bright white and the partial cut locations are less dark red. We chose 21 rectangular areas that contain forest disturbance patches of different sizes and include other land cover classes to train and evaluate the algorithm. All chosen rectangular areas, each with width and length larger than 3 km, were carefully interpreted to determine precisely the location and timing of forest disturbance.

Two steps were used to produce the final disturbance reference map. First, an annual disturbance map was generated by visually comparing the last clear Landsat image in 2002 and the last clear Landsat image in 2003. Forest disturbance that occurred in 2003 should be captured in this annual disturbance map. If there was confusion in comparing the two Landsat images, high spatial resolution images before and after 2003 (can be a few years apart) from Google Earth were used to help determine what was happening at the specific locations. In the worst case, if both high spatial resolution images from Google Earth and the Landsat images pairs do not support a confident decision, the time series of surface reflectances was used to better identify the disturbed pixels (Figure 3.8).

The timing of the disturbance was derived by careful interpretation of all available Landsat TM/ETM+ images acquired in 2003 (a total of 24 images with cloud cover less than 90%). Within each rectangular area, the interpreter sorted through all the TM/ETM+ images carefully. The disturbance date is the first time when forest changes are found and it is determined by visually comparing each pair of consecutive images. The result is a set of reference rectangles that show the location of forest disturbance that occurred in 2003, labeled with the date when the disturbances were first observed (Figure 3.9). These sites then serve as our reference data to train and evaluate CMFDA.

We divided the reference rectangles into two groups: one group used for training CMFDA, that is, to find the optimal change index, threshold, and number of consecutive observations; one group used for evaluating CMFDA accuracy. The reference rectangles were sorted and ranked by size and the odd number ranked rectangles (in blue) were used for help training CMFDA, and the even number ranked rectangles (in red) were used for evaluating CMFDA (Figure 3.9). The goals of this approach were to roughly divide the reference data in half for training and testing, and to avoid bias by making sure that entire polygons were either training or testing, but not both (Friedl et al., 2000).

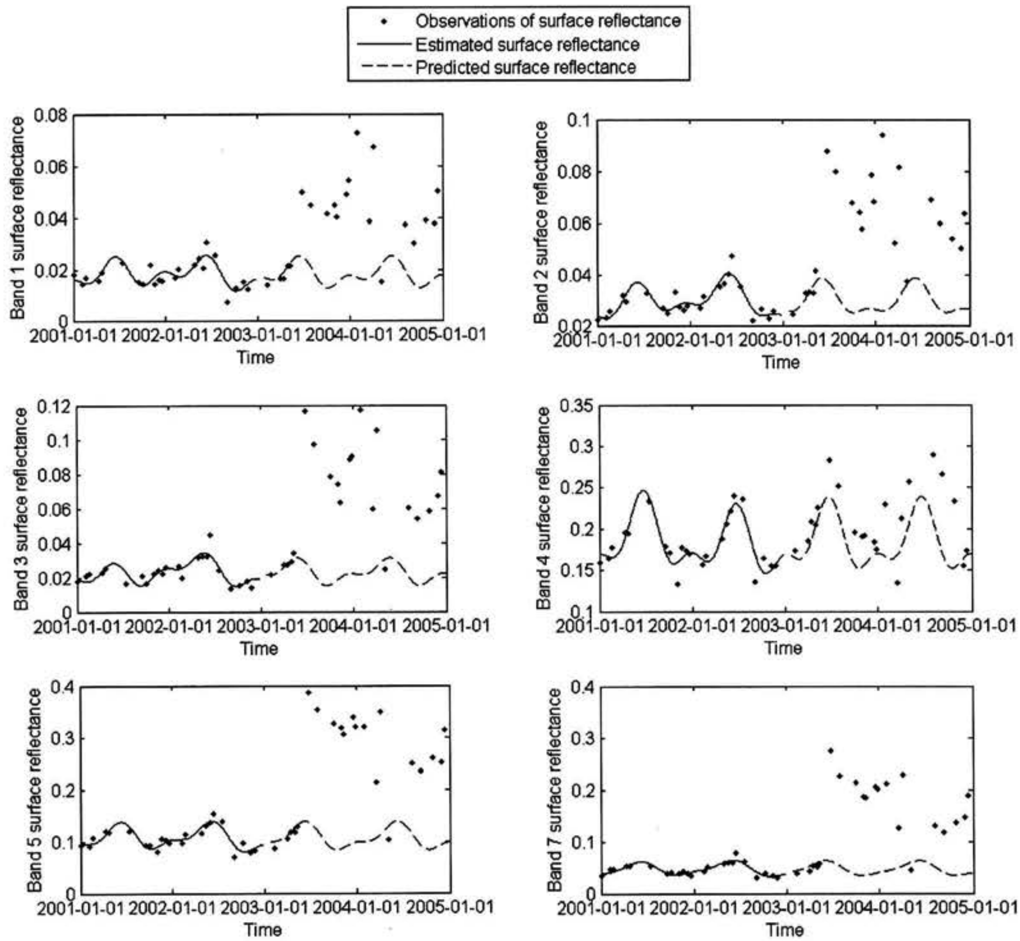


Figure 3.8. Time series data of a typical disturbed deciduous forest pixel for all 6 optical Landsat surface reflectance bands. Persistent changes are easily identified by comparing the predicted surface reflectances and the observed clear surface reflectances.

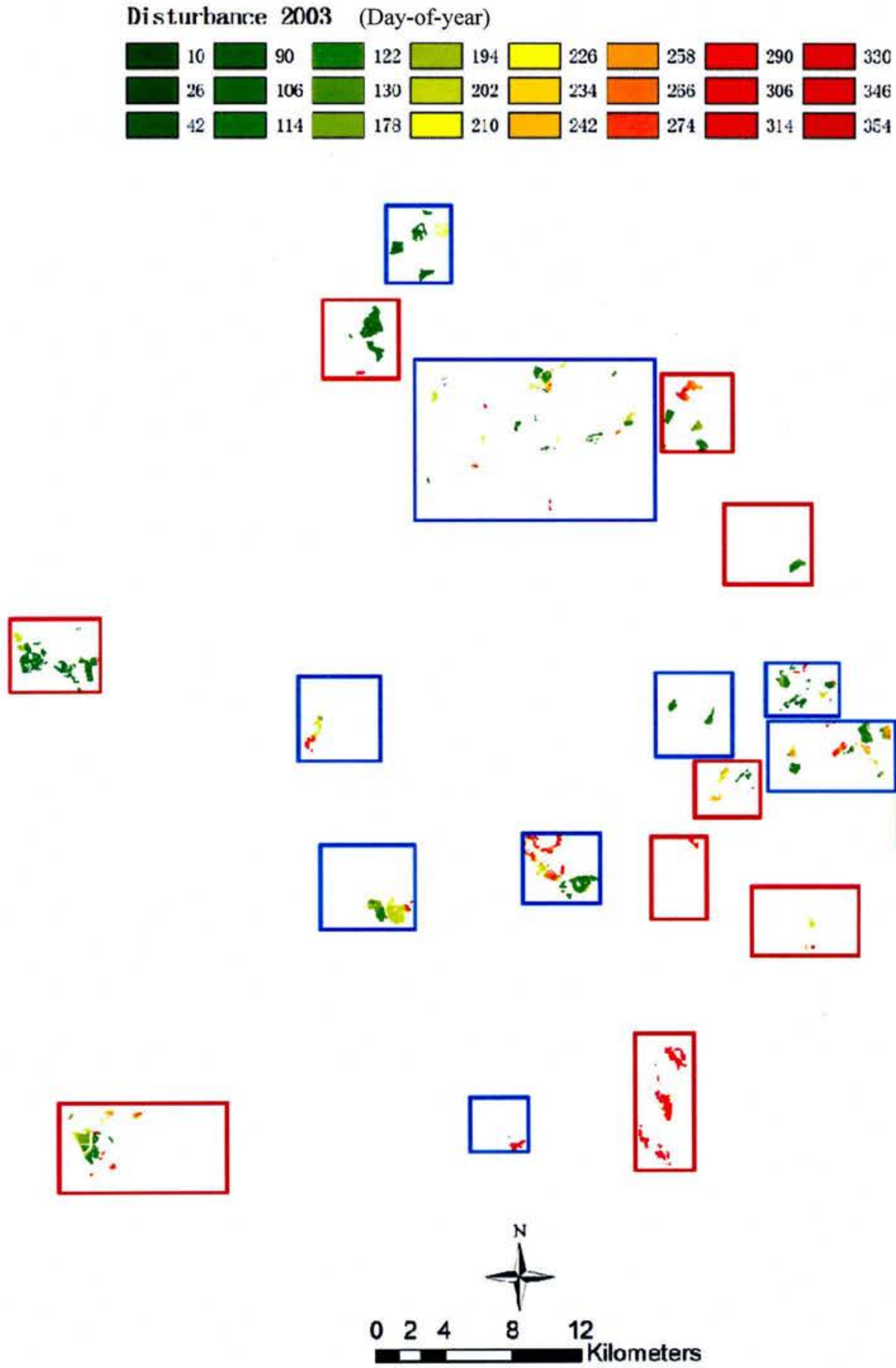


Figure 3.9. Reference map for forest disturbance in 2003 at study area. The blue reference rectangles were used for training and the red ones were used for evaluating the accuracy of CMFDA

3.3.5.2. Definitions of accuracy

Considering the misregistration errors in the Landsat images and especially the ambiguity in validating disturbance at the edge of the change patches, pixels located on the edges of the disturbance polygons (within 1 pixel of the border) were not included in accuracy assessment. As the disturbance map has both spatial and temporal information, we assessed the spatial and temporal map accuracies separately. We think the spatial accuracy is more important than temporal accuracy, as omission and commission of disturbance is more serious than finding disturbance later than the reference map. Overall map accuracy is not a very useful measure in this case as the proportion of forest change is small and therefore the accuracy of the forest change class would not contribute significantly to the overall accuracy. Instead, the producer's and user's accuracies for the forest disturbance class in Equation 3.6 and 3.7 were more important for evaluating the algorithm. Basically, with higher producer's accuracy, there will be fewer omission errors and with higher user's accuracy, fewer commission errors (Congalton, 1991).

$$producer's\ accuracy = \frac{num\ of\ correctly\ identified\ disturbance\ pixels}{num\ of\ disturbance\ pixels\ in\ reference\ map} \tag{3.6}$$

$$user's\ accuracy = \frac{num\ of\ correctly\ identified\ disturbance\ pixels}{num\ of\ disturbance\ pixels\ in\ algorithm\ map} \tag{3.7}$$

Temporal accuracy is evaluated for the forest disturbance pixels that are correctly identified spatially. With the temporally dense Landsat images, forest thinning may be observed before a forest clear-cut as clear-cutting may need a few weeks to finish. It is difficult to determine the disturbance time for this kind of subtle change before a clear-cut. The reference map labels a disturbance time when the disturbance is initially observed with high confidence by the interpreter. However, CMFDA can find very subtle changes at the very beginning of the disturbances. Therefore, the algorithm occasionally finds disturbances earlier than the reference map, which is not considered a mistake, but rather the limitation of manual interpretation in defining subtle changes. We think the algorithm is correct temporally if the disturbance time found by CMFDA is earlier or equal to the disturbance time in the reference map and the temporal accuracy is calculated with Equation 3.8.

$$temporal\ accuracy = \frac{num\ of\ pixels\ (algorithm\ time \leq reference\ time)}{num\ of\ correctly\ identfied\ disturbance\ pixels} \quad (3.8)$$

The producer's/user's accuracies in the spatial domain and the temporal accuracy were used to find the best change index, its optimal threshold, and the number of successive clear observations to use for change identification.

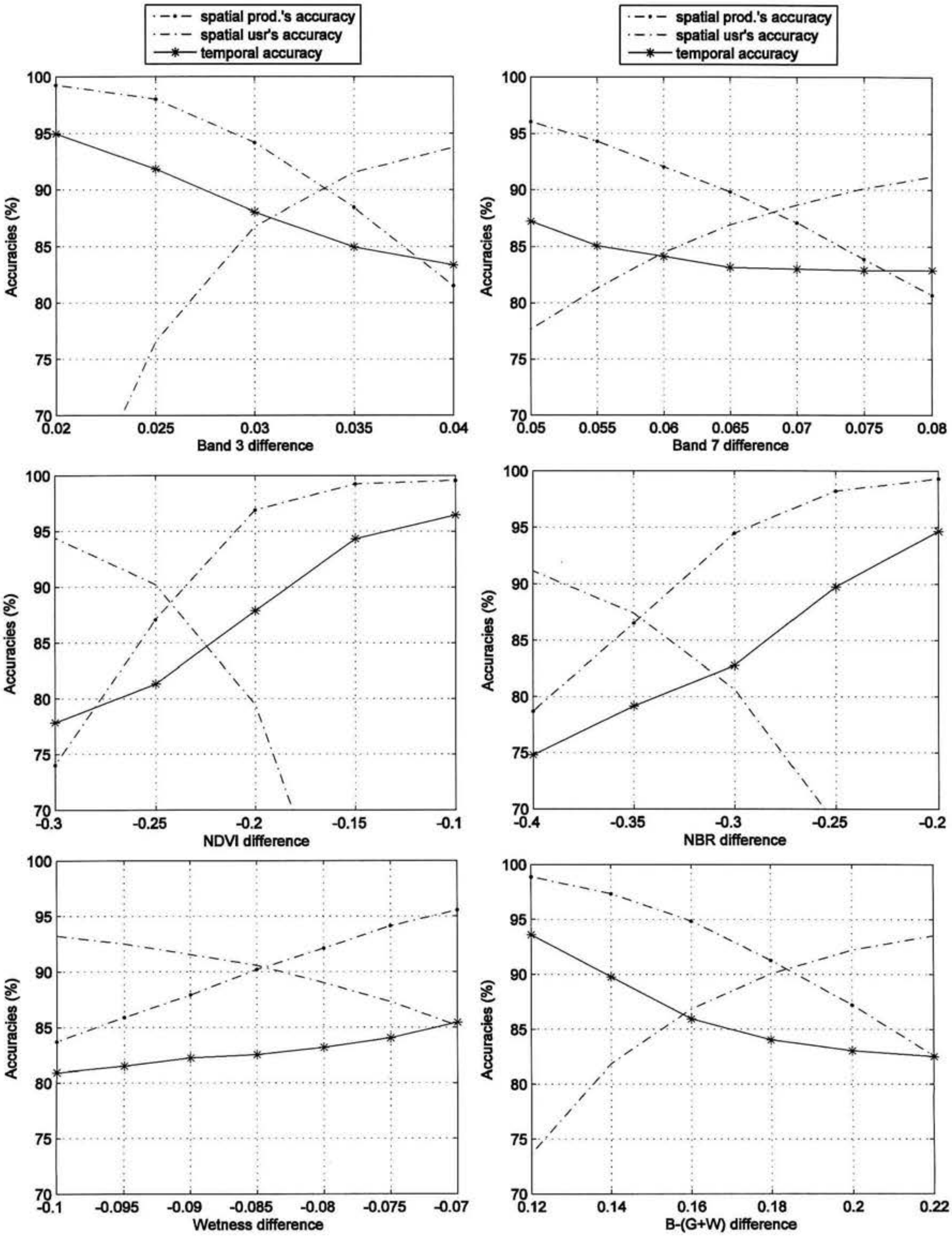
3.4. Results

3.4.1. Single-date differencing algorithm results

3.4.1.1. Selecting a change index and thresholds

The odd number ranked rectangles were used for helping select a change index and thresholds for the single-date differencing algorithm. In Figure 3.10 the spatial user's and producer's accuracies and the temporal accuracy are plotted as a function of the threshold used for different change indices. We use the intersection of the producer's and the user's accuracies as the "best" threshold, as it balances errors of omission and commission. The Disturbance Index B-(G+W) performed the best among all the tested change indices.

When the threshold of 0.18 was used for the change in the Disturbance index, both the producer's and user' accuracies were around 90%, and the temporal accuracy is around 85%. The spatial accuracies of the other five tested indices are slightly lower, but are all above 88%. The temporal accuracies are also related to the thresholds used for defining change, usually the higher the threshold, the later the captured change, which leads to lower temporal accuracies. For most of the indices the temporal accuracies are around 85%, except for the two indices (NDVI and NBR) that use the NIR band as its main input. The NIR surface reflectance varies significantly with vegetation phenology, which may induce problems in determining the time of detected changes and reduce the temporal accuracies.



spatial prod.'s accuracy
spatial usr's accuracy
temporal accuracy
Accuracies (%)
Band 3 difference
Band 7 difference
NDVI difference
NBR difference
Wetness difference
B-(G+W) difference

Figure 3.10. Accuracies for the single-date differencing algorithm for the six different change indices.

3.4.1.2 Testing on unseen reference data

The even number ranked rectangles were used to evaluate the accuracy of the single-date differencing algorithm with B-(G+W) as its change index and a threshold of 0.18. The confusion matrix of spatial accuracies is shown in Table 3.2. Both the user's and producer's accuracies for forest disturbance are more than 93%, and producer's accuracy is slightly higher than the user's accuracy. The temporal accuracy of those spatially corrected identified forest disturbance pixels is about 90%.

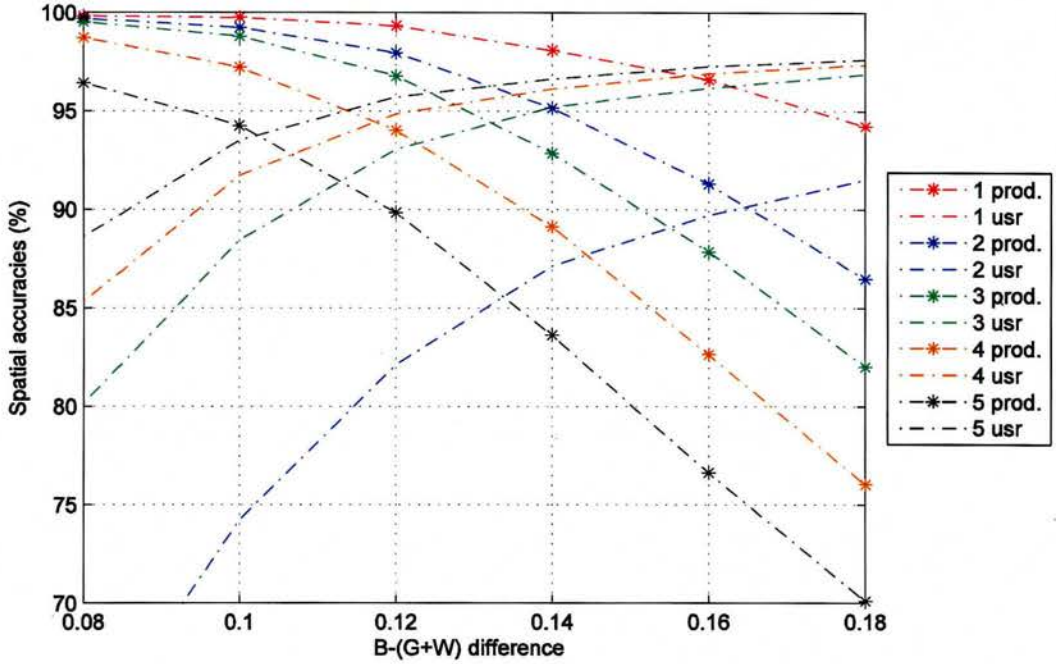
Table 3.2. Confusion matrix for the accuracy assessment of the single-date differencing algorithm. The overall accuracy results are not terribly revealing, as after excluding the edges of the change polygons, the change pixels left are only about 3% of the total interpreted pixels.

Reference data				
Single-date differencing	Forest disturbance	Others	Total	User's (%)
Forest disturbance	7422	523	7945	93.42
Others	492	241969	242461	99.80
Total	7914	242492	250406	
Producer's (%)	93.78	99.78	Overall (%)	99.59

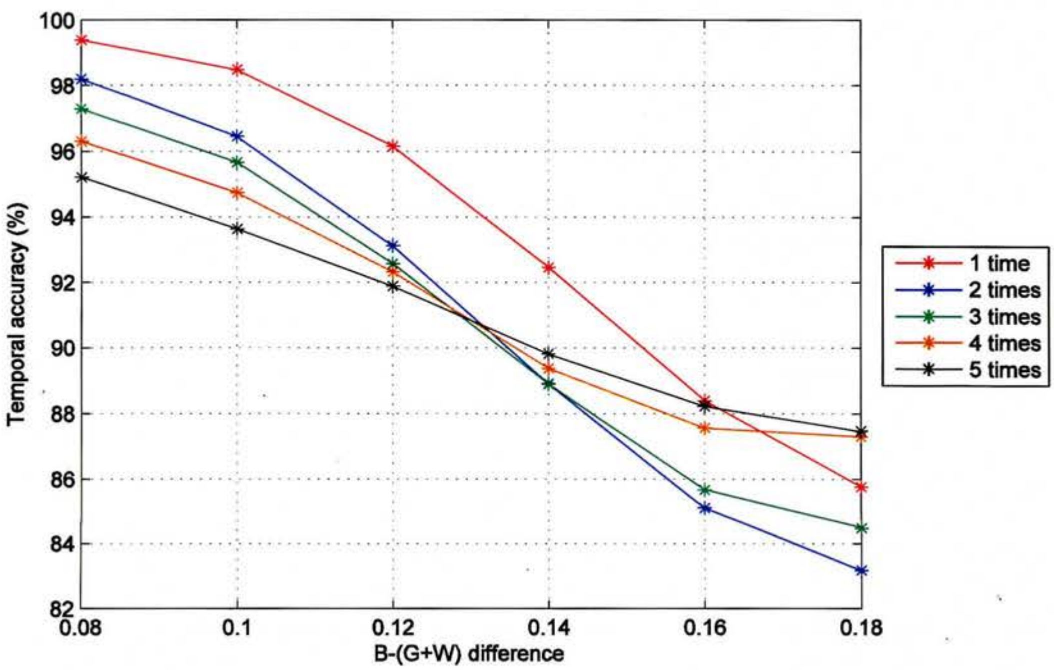
3.4.2. Multi-date differencing algorithm results

3.4.2.1 Selecting the number of successive observations and thresholds

The effects of using multiple consecutive dates as part of the change detection with B-(G+W) as the tested change indices are shown in Figure 3.11. Figure 3.11 (top) shows the spatial accuracies (both user's and producer's) for different change thresholds and different lengths of successive identifications of change for the same pixel. The highest accuracy is achieved from three successive clear observations with a threshold of 0.12. In this best scenario, the producer's and user's accuracies in the spatial domain are around 95% and the temporal accuracy is approximately 93%. Due to the relatively large number of commission errors when using only one clear observation, the user's accuracy is too low to be shown in Figure 3.11 (top). Though the optimal threshold is 0.06 less than the single-date differencing algorithm, its commission error is lower (higher user's accuracy). The spatial accuracies are relatively robust to this optimal threshold when three consecutive clear observations are used. The temporal accuracies are related to the change thresholds and the number of successive observations, usually the higher the threshold or the larger number of successive observations, the later the captured change, and lead to lower temporal accuracies. Generally, the temporal accuracies for multi-date differencing algorithms are all high (more than 85%) when the threshold varies greatly (from 0.08 to 0.16) and they are not very sensitive to the number of successive observations used.



Spatial accuracies (%)
B-(G+W) difference
100
95
90
85
80
75
70
0.08
0.1
0.12
0.14
0.16
0.18
1 prod.
1 usr
2 prod.
2 usr
3 prod.
3 usr
4 prod.
4 usr
5 prod.
5 usr



Temporal accuracy (%)
B-(G+W) difference
100
98
96
94
92
90
88
86
84
82
0.08
0.1
0.12
0.14
0.16
0.18
1 time
2 times
3 times
4 times
5 times

Figure 3.11. (top) Spatial and temporal (bottom) accuracies of the multi-date differencing algorithm for different combinations of numbers of succession observations of change and change index thresholds. Notice that the intersection of user's and producer's accuracies get maximum for 3 consecutive observations.

3.4.2.2 Testing on unseen reference data

Similarly, the even number ranked rectangles were used to evaluate the accuracy of the multi-date differencing algorithm with B-(G+W) as its change index, a threshold of 0.12, and three consecutive observations. Both the user's and producer's accuracies for forest disturbance are more than 95%, and producer's accuracy is also higher than the user's accuracy. The temporal accuracy of the correctly identified forest disturbance pixels is almost 94%.

Table 3.3. Confusion matrix for the accuracy assessment of the multi-date algorithm. The overall accuracy results are not terribly revealing, as after excluding the edges of the change polygons, the change pixels left are only about 3% of the total interpreted pixels.

	Reference data			
Multi-date differencing	Forest disturbance	Others	Total	User's (%)
Forest disturbance	7653	333	7985	95.83
Others	261	242159	242420	99.89
Total	7914	242492	250406	
Producer's (%)	96.70	99.86	Overall (%)	99.76

The map derived from multi-date differencing algorithm shows locations and dates of forest disturbance during in 2003 (Figure 3.12). The colors of the polygons represent the first date the forest disturbance is captured by the algorithm using all available Landsat images. Within all the 21 reference rectangles, the multi-date differencing results and the reference map are very similar. However, at this 60×60 km scale, it is difficult to find any significant differences between the reference map and the map derived by the algorithm. Looking closer at three reference rectangles used for testing (Figure 3.13), the disturbance map derived from the algorithm agrees closely with the disturbance found in the reference map both spatially and temporally (green color). The three types of disagreements (blue, violet, and red) are all distributed at the edges of disturbance patches. These false identifications are mainly caused by the misregistration in the image stack and problems in interpreting forest disturbance at the boundaries of patches.

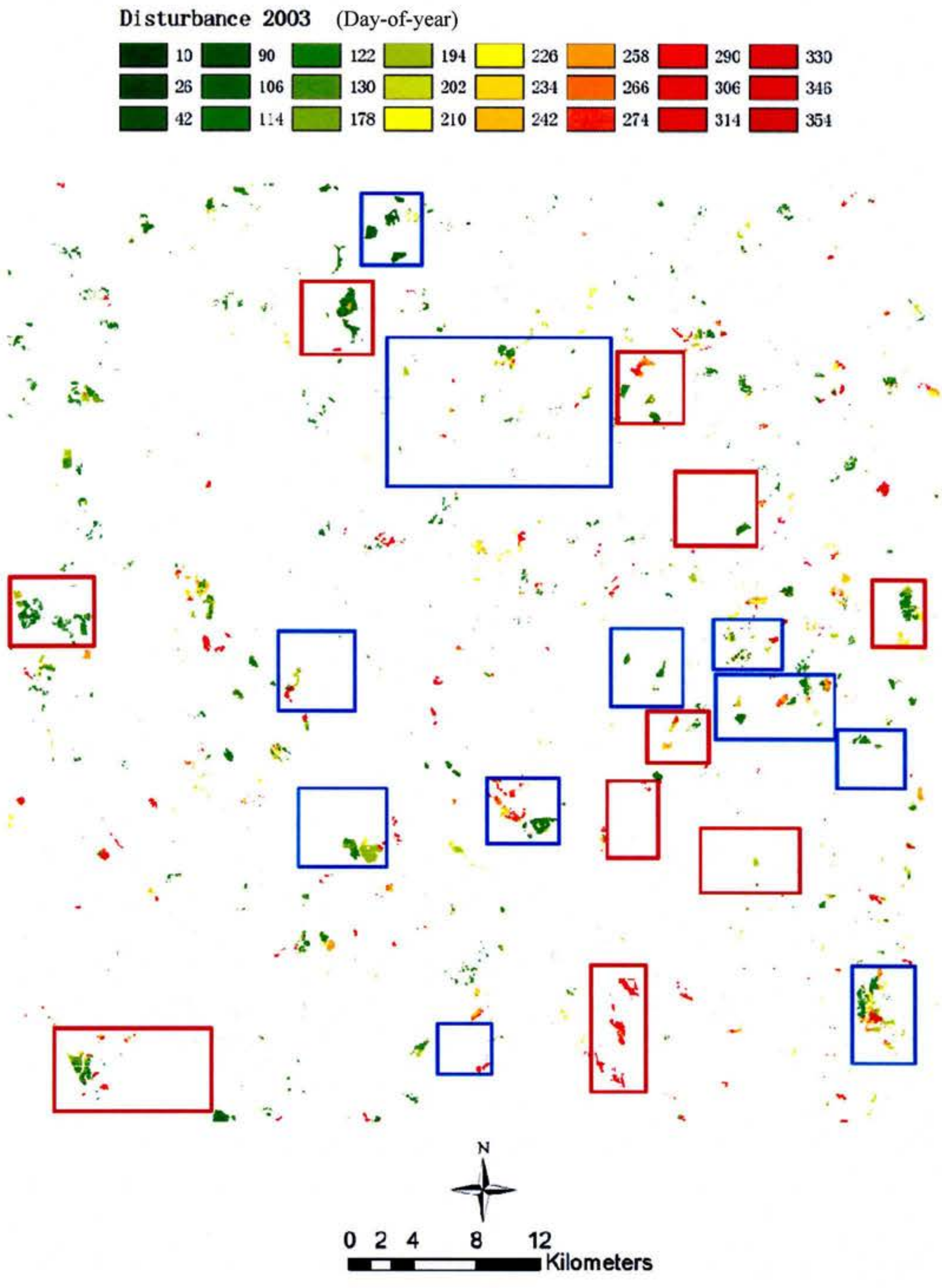


Figure 3.12. Multi-date differencing aglorithm map for forest disturbance in 2003 at study area

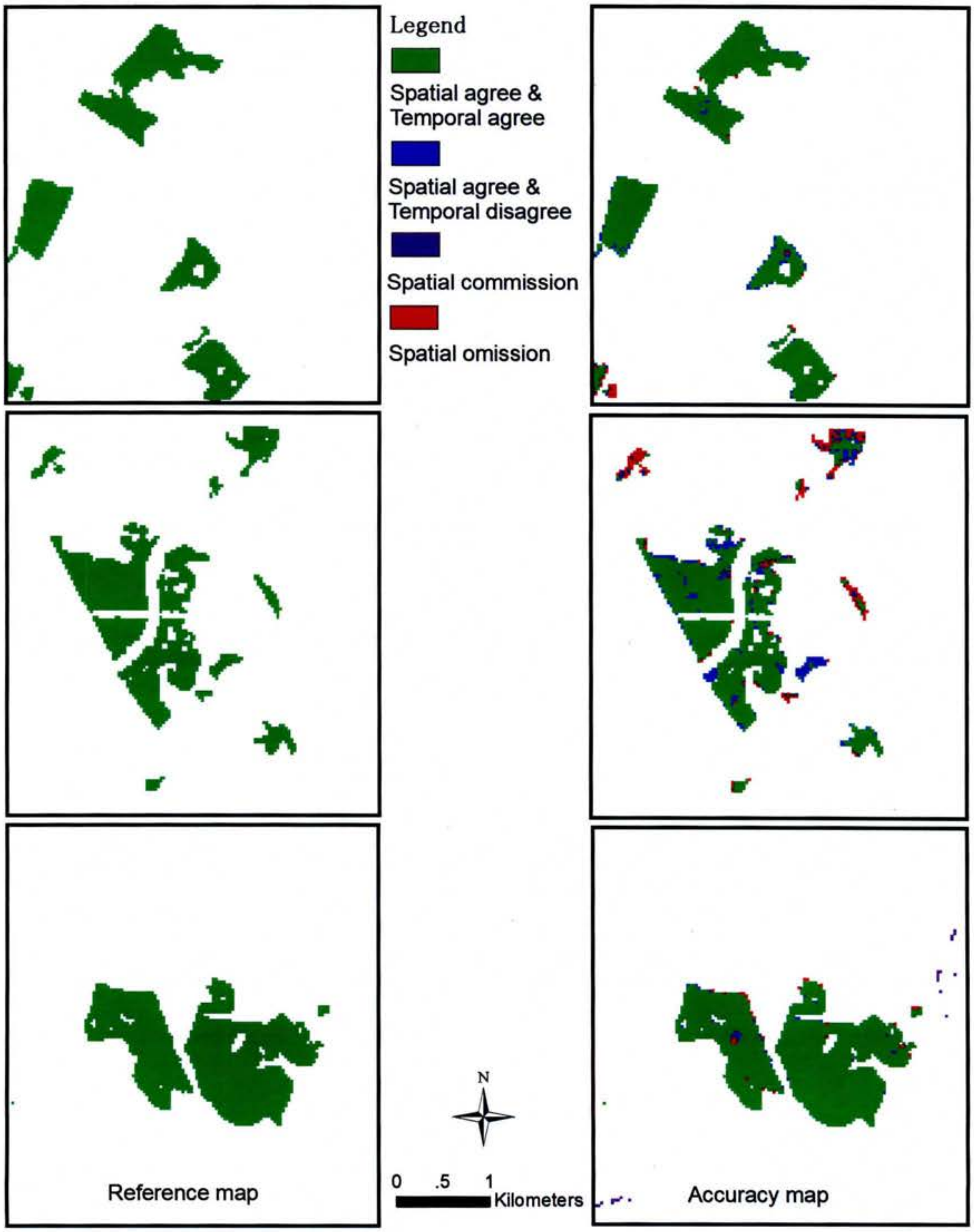


Figure 3.13. Zoom in of reference and error maps. The colors in accuracy map show different types of errors (see legend of this figure).

3.5. Discussions and conclusions

In this study, we developed a new algorithm for continuous monitoring of forest disturbanc. This approach also allows construction of a history of forest disturbance. Using all the available Landsat ETM+ images in two years, models using sines and cosines are fit for each pixel and each spectral band. These models can predict Landsat images at any date assuming there is not any land cover change. CMFDA flags forest disturbance by differencing the predicted and observed Landsat images. We tested two algorithms called single-date and multi-date differencing for detecting forest disturbance. The multi-date differencing algorithm was chosen as the final change detection method for CMFDA due to its higher spatial and temporal accuracies. It uses B-(G+W) as its change index and a threshold of 0.12 for defining "change". It determines a disturbance pixel by the number of times "change" is observed consecutively. Pixels showing "change" for one or two times are flagged as "probable change". If a third consecutive "change" is found, the flag will be mapped as forest change. The reference map revealed that the CMFDA result is accurate in detecting forest disturbance both spatially and temporally at the tested Savannah River site, with producer's and user's accuracies higher than 95% and temporal accuracy of approximately 94%.

CMFDA has many advantages. It can be fully automated and has the potential of monitoring forest disturbance continuously as new images are collected. Though a predefined change threshold is used at this specific study site, sensitivity analysis suggests this algorithm is relative robust to this threshold when three consecutive clear observations are used. The continuous character of the monitoring makes the algorithm capable of identifying disturbance soon after Landsat observations become available. Therefore, how fast the CMFDA is able to find change accurately is solely dependent on the frequency of available clear observations. The potential to use the methods presented here for monitoring surface change will improve as the frequency of high resolution images from sensors like Landsat become more available (Arvidson et al., 2006). The first major step forward in this domain was the opening of the US archive so that the vast holding of Earth Resources Observation and Science (EROS) Data Center can be used. With regard to future observations, the launch of the Landsat Data Continuity Mission (LDCM) should greatly increase the frequency of available observations as the duty cycle for LDCM is larger than any of the previous Landsat satellite. More importantly, when the two Sentinel 2A/2B satellites are launched, they will have a repeat time of every five days. When combined with LDCM data, there would be as many as 8 high resolution observations per month, which will greatly improve the availability of observations such that we will be able to begin to monitor change in near real time.

By considering each pixel separately, CMFDA can overcome most of the limitations of conventional approaches. By using any clear observations for each pixel to track spectral trends over time, CMFDA expands the use of Landsat images to any time of year and to all kinds of conditions (e. g., cloud, snow, heavy aerosols). As CMFDA fits models for each pixel, it can work in heterogeneous forest areas that are reported to be problematic for the scene-based normalized change indices such as IFZ and DI (Masek et al., 2008; Huang et al., 2010). The problem caused by the failure of the Scan Line Corrector (SLC) in Landsat 7 is not nearly as significant for CMFDA as compared with more conventional approaches. The scan line gaps are treated just like clouds that remove observations from images and the available good observations are used. One area of future research will be to integrate observations from adjacent Landsat images in the zone of "side lap". This approach will further minimize the effects of Landsat 7 SLC-off gaps as they are most pronounced in these areas of side lap. The same is true for images with partial cloud cover, as they have many useful observations. As a result, it would be highly desirable if the Landsat satellites of the future collected all possible observations, as even partially cloudy images have value in analysis systems like CMFDA.

CMFDA also has limitations. First of all, CMFDA works better with larger number of observations. However, areas outside the U.S. may not have enough observations, particularly for some years in the 1980s and 1990s (Goward et al., 2006). Luckily, the new acquisition strategies for Landsat 7 (Arvidson et al., 2001; Arvidson et al., 2006) provide more frequent observations. The years between 1999 and early 2003 might be considered as the "golden years" for estimating models for CMFDA, as during this time both Landsat 5 and Landsat 7 are functioning normally. Therefore, we could build models for each pixel during this period and flag future changes. Similarly, we can detect changes that occurred in the 1980s and 1990s as long as there are data available using the models calibrated in the "golden years", by running the time series analysis backward in time. The second limitation of CMFDA is the computation time associated with creating models to predict future surface reflectance. The two-step cloud/cloud shadow, and snow masking is critical, as including noise factors in the data undermines the entire process. However, when the prediction models are ready, CMFDA is able to update a disturbance map as soon as a new Landsat observation is acquired. This process is very fast and does not require reanalysis of the historical data. Such an approach points to the possibility of processing images in CMFDA as part of the process for ingesting new images, paving the way for monitoring land cover change as close as possible to when it happens using Landsat data. A third limitation is that the methods proposed above are all based on the assumption that land cover change only occurs once in the detection period which is not true if the detection period is longer than some of the "permanent change" like forest disturbance. Masek et al. (2008) suggested that the highest forest disturbance cycle time is approximately 5 years. To include changes of this frequency, CMFDA needs to re-estimate the surface reflectance models using the newest data available at 5 years intervals. Finally, though CMFDA identifies forest disturbance much quicker than the conventional approaches, the expected time to find "probable change" and "change" in CMFDA is still too long to monitor changes as they are occurring. Assuming cloud probability of 50%, CMFDA will typically need at least half a month to find "probable change" and one and a half months to find "change" in places with the most frequent observations like United States. It will take longer in other parts of the world due to less frequent Landsat observations. To achieve the goal of global near-real time monitoring of land cover change, using more Landsat-like sensors or fusion with higher temporal frequency sensors like MODIS (Hilker et al., 2009) are choices in future studies.

In this first use of this approach we estimated the models on the years 2001 and 2002 and applied the model to 2003. When looking retrospectively (to reconstruct the history of forest disturbance) it will be possible to look at much longer time series of images and select a set of years that do not exhibit change for calibrating the surface reflectance models that are then applied to other years. In this case, there is no reason to constrain all pixels in an area to being estimated on the same set of years. This approach will in some ways simplifying both the estimating of the surface reflectance models and their usage as they won't be complicated by change near the beginning or end of the estimating period. Moreover, CMFDA has the potential of monitoring other land cover changes if a specific predefined land cover mask can be derived accurately. For instance, it is possible to identify wetland loss by finding changes within a predefined wetland map, or if we are looking for changes in agriculture land use, we may be able to monitor agriculture abandonment. Further studies are necessary for detecting other land cover changes using algorithms similar as CMFDA.

Chapter 4

4. Continuous change detection and classification of land cover using all available Landsat data

4.1. Introduction

Land cover monitoring and mapping has been widely recognized as a key scientific goal (Anderson, 1976; Tucker et al., 1985; Gopal et al., 1999; Hansen et al., 2000; Loveland et al., 2000; Friedl et al., 2010). Each land cover has unique physical characteristics such as albedo, emissivity, roughness, photosynthetic capacity, and transpiration that significantly influence the energy balance, carbon budget, and hydrologic cycle. Land cover change can be natural or anthropogenic, but with human activities increasing, the Earth surface has been modified significantly in recent years by various kinds of land cover change such as deforestation, agriculture expansion and intensification, urban growth, and wetland loss (Jensen et al., 1995; Coppin & Bauer, 1996; Woodcock, et al., 2001; Seto et al., 2002; Galford et al., 2008). Knowledge of land cover and land cover change is necessary for modeling the climate and biogeochemistry of the Earth system and for many management purposes. Satellite remote sensing has long been used to assess Earth surface because of repeated synoptic collection of consistent measurements (Lambin & Strahler, 1994).

The Landsat datasets are one of the most important sources for studying the different kinds of land cover change due to their long historical record and fine spatial resolution (Kennedy et al., 2007; Wulder et al., 2008; Pflugmacher et al., 2012). One of the drawbacks of Landsat data is their low temporal frequency. For each Landsat sensor, it can only measure the Earth surface repeatedly in 16 days and more than half of the time the signal is blocked by clouds (Zhang et al., 2004). Therefore, most of the change detection algorithms are developed by comparing two dates of clear Landsat images acquired at the same growing season (Collins & Woodcock, 1996; Healey et al., 2005; Masek et al., 2008). Though these kinds of algorithms are relatively simple to implement, they are not always applicable. They may take a few years to find an ideal pair of Landsat images that are cloud-free and snow-free and acquired at the same growing season. To find change faster, some more advanced change detection algorithms using many dates of Landsat data are appearing in the literature (Hostert et al., 2003; Kennedy et al., 2007; Goodwin et al., 2008; Vogelmann et al., 2009; Huang et al., 2010). However, these newly developed algorithms still have limitations in selecting "good" images. For example, to minimize the phenology and BRDF effects, the "good" images should be within the same growing season. Moreover, tough they can use images that are partly covered by clouds, they still need most of the "good" images cloud free. To satisfy all these requirements with Landsat data, the best these multitemporal change detection algorithms can provide is the annual or biennial change results. Recently, MODIS time series data have been extensively explored for monitoring varies kinds of land cover change (Jin & Sader, 2005; Roy et al., 2005; Lunetta et al., 2006; Galford et al., 2008; Eklundh et al., 2009; Verbesselt et al., 2010) because of its much higher temporal frequency. Though the MODIS time series data can find change much faster than Landsat, they are limited by low detection accuracy, especially when the changed patches are small (Jin & Sader, 2005). The coarse spatial resolution of the MODIS data limits its ability for detecting small changes or changes that are occurring in parts of a pixel (Jin & Sader, 2005), which are very common for anthropogenic changes disturbances. Also, the gridding artifacts of MODIS have introduced errors spatially, making multitemporal comparison difficult (Tan et al., 2006). To monitor land cover change as they are occurring and be able to include small footprint changes such as encroachment of protected area or illegal logging, the remote sensing community needs an algorithm that uses fine spatial resolution data such as Landsat and uses as many observations as possible to detect land cover change accurately and quickly.

Land cover classification is one of the most studied remote sensing topics and land cover maps provide the basis for many scientific studies like modeling of carbon budget, management of forests, estimation of crop yield, etc. (Wolter et al., 1995; Lark & Stafford, 1997; Jung et al., 2006; Rogan et al., 2010). While it is relatively easy to generate a land cover map from remotely sensed data, it is not easy to make it accurate. Using multitemporal data as inputs are reported to help improve classification accuracy (Wolter et al., 1995; Guerschman et al., 2003), especially for classification of vegetation, because of the unique phenology characteristics linked with different vegetation type. To achieve higher classification accuracy, most of the current land cover products are generated using multitemporal imags as their main inputs (Tucker et al., 1985; Gopal et al., 1999; Hansen et al., 2000; Friedl et al., 2010). Nevertheless, the use of multitemporal images also causes problems for the conventional land cover classification algorithms.

First, they need all the multitemporal images without cloud and snow to classify every pixel in the images, which are very unlikely to find, especially for sensors with relatively low temporal frequency like Landsat. In this case, for some cloudy locations, we may need to wait a few years to get a few clear multitemporal Landsat images. Therefore, most of the Landsat-based land cover maps are produced at the interval of five or ten years which would greatly reduce the time availability of the land cover maps. Second, when making land cover maps with multitemporal images, we are assuming there is no land cover change occurring in the time interval of the multitemporal images. This assumption is not always valid, especially when images from long time intervals are used as input or for areas that are changing frequently (Rogan, et al., 2002). Moreover, the land cover maps produced from conventional methods cannot be used directly for identifying land cover change because frequently the magnitude of the error in classification is much larger than the amount of land cover change (Fuller et al., 2003; Friedl et al., 2010). If we compare land cover maps produced at different times for defining change, the errors from classification will show up as change and this would cause serious problems for places where change area is small. Therefore, the remote sensing community needs an classification algorithm that increases the time availability of land cover maps, works for areas where land cover change is common, and makes land cover maps comparable for identification of change.

The opening of the Landsat archive in 2008 (Woodcock et al., 2008) has led a big change in the use of Landsat images. Previously, a single Landsat image would cost hundreds of U.S. dollars. To minimize costs, most researchers chose to use only a few

Landsat images without any clouds or snow for analysis. After the free access of Landsat archive, a large amount of Landsat data has been used for different kinds of studies. In this study, we proposed to use all available Landsat data not only because of the free data policy, but also due to the large number of clear observations contained in the Landsat images with high percentage of clouds. Figure 4.1 is derived from analysis of all available Landsat data at Path 12 Row 31 between 1982 and 2011. Based on this information, if we only use Landsat images with cloud cover less than 10% like most of the conventional methods do, we would omit more than 50% percent of the total clear observations. Even Landsat images with more than 40% cloud cover contain almost 20% of the total clear observations. The use of all available Landsat dataset has opened a door for many studies that cannot imagine before, such as study phenology at Landsat scales (Eli et al, submitted) and detecting forest disturbance continuously at high spatial resolution and high temporal frequency (Zhu et al., 2012a).

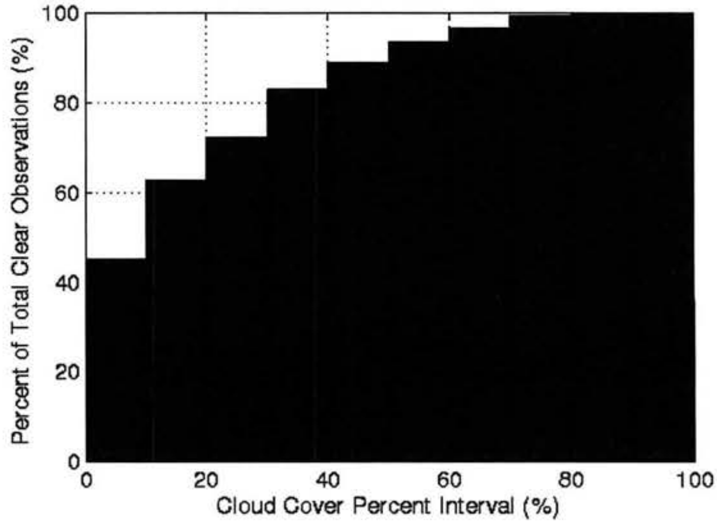


Figure 4.1. Percent of total clear observations vs. cloud cover percent interval based on all available Landsat TM/ETM+ images from 1982 to 2011 at Path 12 Row 31.

By using all available TM/ETM+ observations from Landsat 4, 5, and 7, I developed a new Continuous Change Detection and Classification (CCDC) algorithm that solves the problems raised in the conventional change detection and classification methods. This new algorithm is capable of detecting many kinds of land cover change at high temporal frequency and at the same time providing land cover maps for any given time. The change detection and classification maps are updated continuously when more clear observations are available. To implement the CCDC algorithm, the following steps are necessary:

1) an accurate cloud, cloud shadow, and snow screening approach;

2) estimation of time series models that can be used to predict the "next" observation;

3) detection of land cover change by comparing the observed and predicted observations;

4) classifying land cover categories based on the coefficients of the time series model and Root Mean Square Error (RMSE) from model fitting;

5) update the time series models continuously as new observations are available.

4.2. Study area and data

4.2.1 Study area

The study area is located on the Northeastern United States coast (Figure 4.2). It includes all of Rhode Island, as well as much of Eastern Massachusetts, and parts of

Eastern Connecticut. It has been selected because: (a) it includes Boston making field visits easy; (b) it includes a wide variety of environments and land uses that provide examples of many of the primary kinds of land cover change occurring in the United States, including: extensive urbanization (three major metropolitan areas – Boston, Providence, and Worcester), abandonment of agricultural fields, and forest clearing; and (c) it is rare to find a cloud-free image of this scene, making it an outstanding place to test the robustness of our new algorithm.

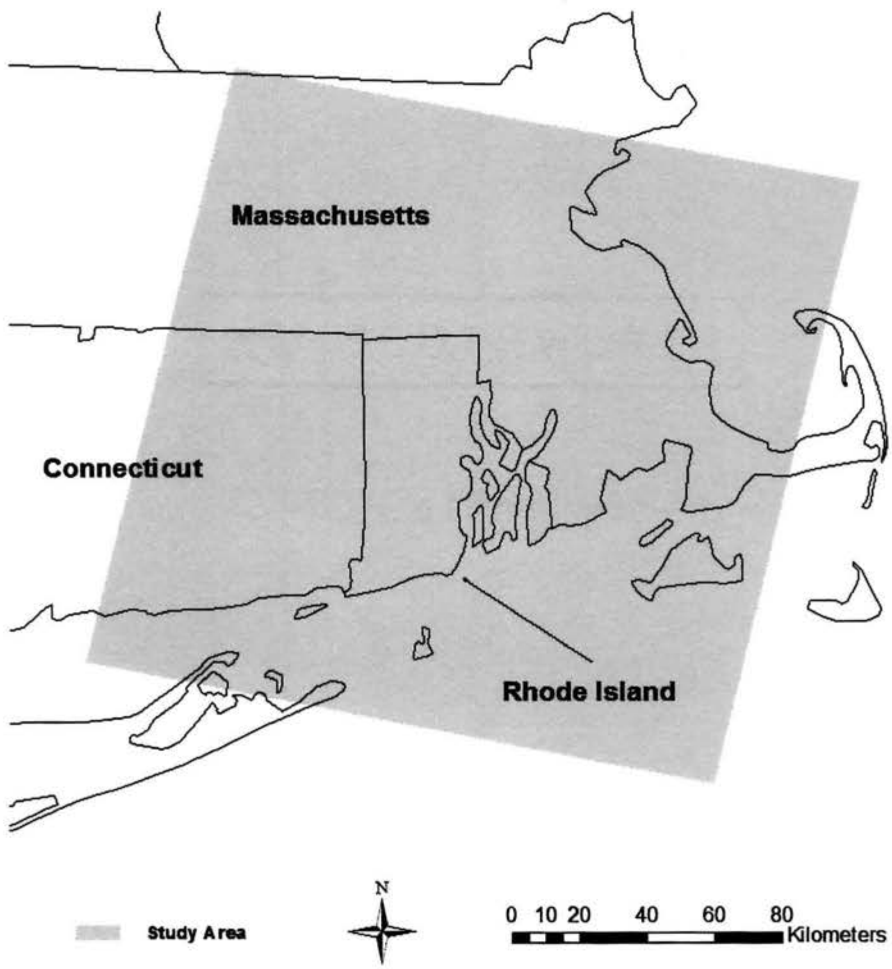


Figure 4.2. Study area

4.2.2 Landsat data

All available Landsat TM/ETM+ images for Worldwide Reference System (WRS) Path 12 and Row 31 were used (Figure 4.2). The "all available" refers all L1T Landsat images with cloud cover less than 80%. A total of 519 images from Landsat 4 (TM), 5 (TM), and 7 (ETM+) between 1982 and 2011 were used in this location.

4.2.3. Land cover reference data

The land cover reference data were previously used to calibrate the HERO Massachusetts Forest Monitoring Program (MaFoMP) 2000 land cover product (Rogan et al., 2010). They were created with the aid of aerial photographs and many field visits between 2005 and 2007. All reference sites were 60 x 60 m in dimension, and were distributed throughout Massachusetts to capture variation in reflectance values within the study area. In the original data, water is divided into two land cover types (*shallow water* and *deep water*). To simplify, we combined the *shallow water* and *deep water* into one land cover – *water*. There are a total of 8,220 reference sites with 16 categories of land cover in the study area (Table 4.1).

Table 4.1. 16-categories land cover description.

Class	Number of sites	Description
Orchards	234	Managed plantation of fruit trees, primarily apples
Cranberry Bogs	265	Managed bog containing cranberry bushes, seasonally flooded
Pasture/ Row Crops	541	Open and cultivated agricultural grasslands
Deciduous Forest	570	Forested land ≥ 80% broadleaved deciduous canopy cover
Conifer Forest	582	Forested land ≥ 80% needleleaved evergreen canopy cover
Mixed Forest	702	Forest land > 20% conifer and <80% deciduous canopy cover
Golf Course	486	Highly managed open grasslands
Grassland	502	Grassland dominated open spaces
Low Density Residential	511	Residential land with equal parts impervious surface & vegetation
High Density Residential	466	Residential land minimally vegetated, > 60% impervious surface
Commercial/ Industrial	613	Impervious surface
Water	1088	Standing water present > 11 months
Wetland	513	Vegetated lands with a high water table
Salt Marsh	436	Tidal saltwater rivers/ mudflats & surrounding herbaceous cover
Sand Quarry	374	Sand & gravel mining pits
Bare Soil	337	Bare land sparsely vegetated, > 60% soil background

4.3. Methodology

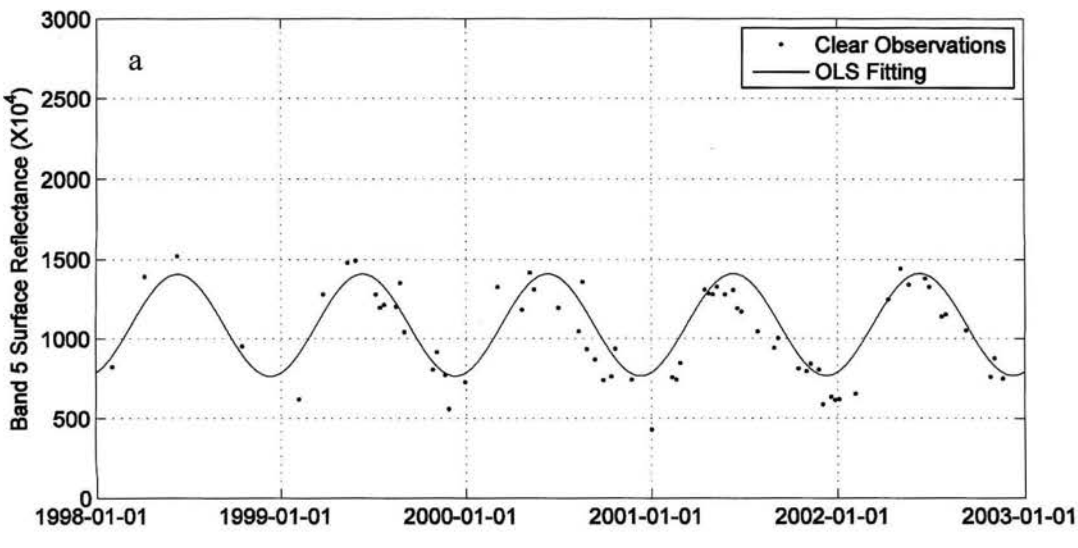
The CCDC algorithm has many components, including: image preprocessing, screening of cloud, cloud shadow, and snow, time series model initialization, continuous change detection, and continuous land cover classification.

4.3.1. Image preprocessing

Geometric registration and radiometric normalization are critical in change detection, as they facilitate comparison of images across time and space. In this research, we assume all Landsat L1T images are already precisely registered and sub-pixel misregistration will not influence our analysis. Radiometric normalization is done using the LEDAPS atmosphere correction tool (Vermote et al., 1997; Masek et al., 2006). It converts the raw DN values into surface reflectance and Brightness Temperature (BT). At the same time, the thermal band was resampled to 30 m to match the surface reflectance bands by LEDAPS. Clouds, cloud shadows, and snow were screened by the two-step method introduced in Chapter 2.

4.3.2. Initialization of the time series model

Generally, land surface change can be divided into three categories: (1) intra-annual change, the vegetation phenology driven by annual temperature and precipitation or vegetation types (Figure 4.3a); (2) inter-annual change, gradual change that caused by climate variability, vegetation growth or gradual change in land management or land degradation (Figure 4.3b); and (3) abrupt change, caused by deforestation, floods, fire, insects, urbanization and so on (Figure 4.3c). Therefore, we proposed a time series model that has components of seasonality, trend, and breaks that captures all three categories of changes on the surface of Earth (Equation 4.1). The model coefficients were estimated by the Ordinary Least Square method based on the remaining clear Landsat observations.



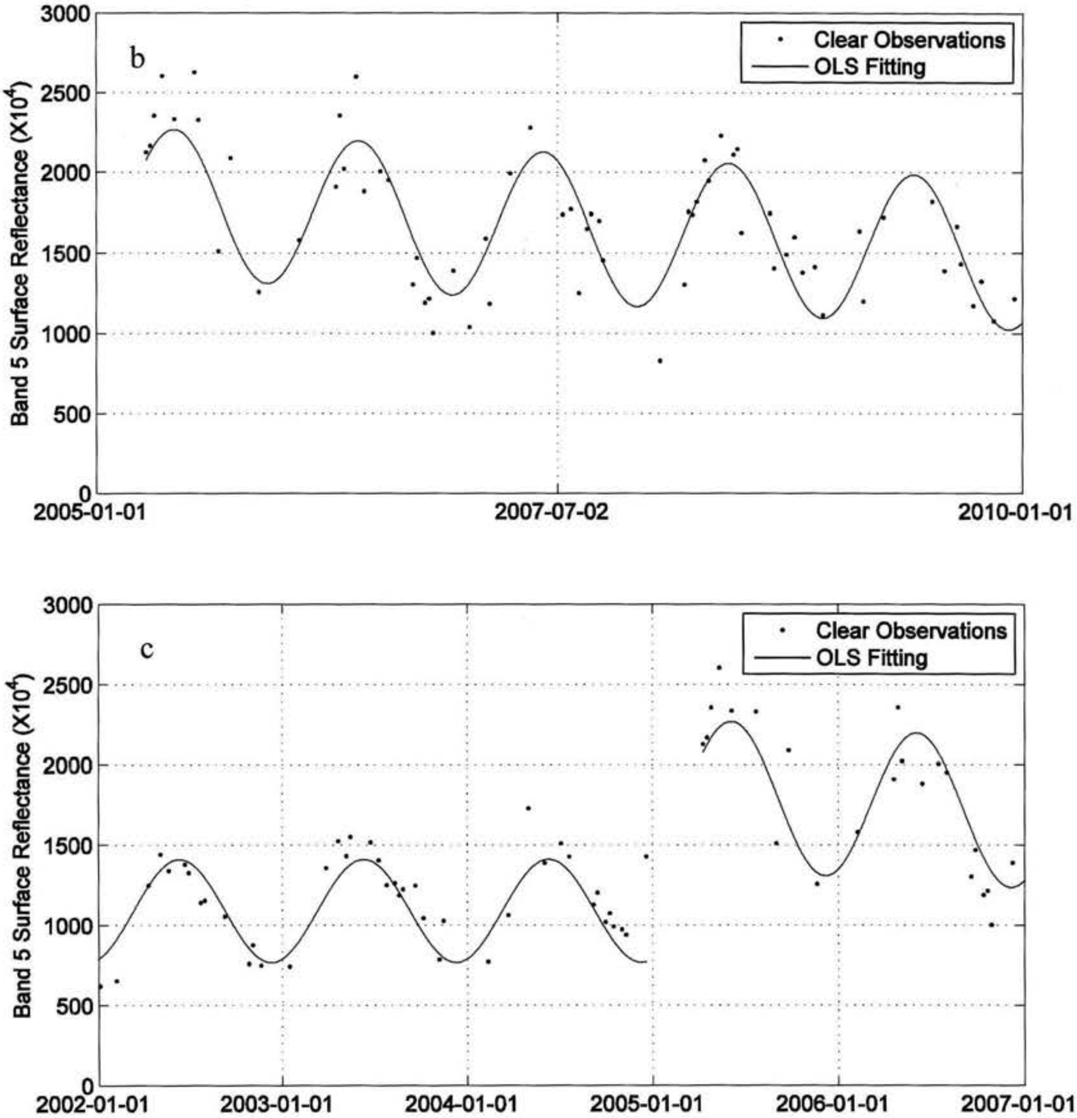


Figure 4.3. Three categories of land surface change: (1) intra-annual change (Figure 4.3a); (2) inter-annual change (Figure 4.3b); and (3) abrupt change (Figure 4.3c)

$$OLS(i,x) = a_{0,i} + a_{1,i}\,cos(\frac{2\pi}{T}x) + b_{1,i}\,sin(\frac{2\pi}{T}x) + c_{1,i}x \tag{4.1}$$

$$\{\tau^*_{k-1} < x \leq \tau^*_k\}$$

Where,

x : Julian date

i: The ith Landsat Band

T: Number of days per year (T = 365)

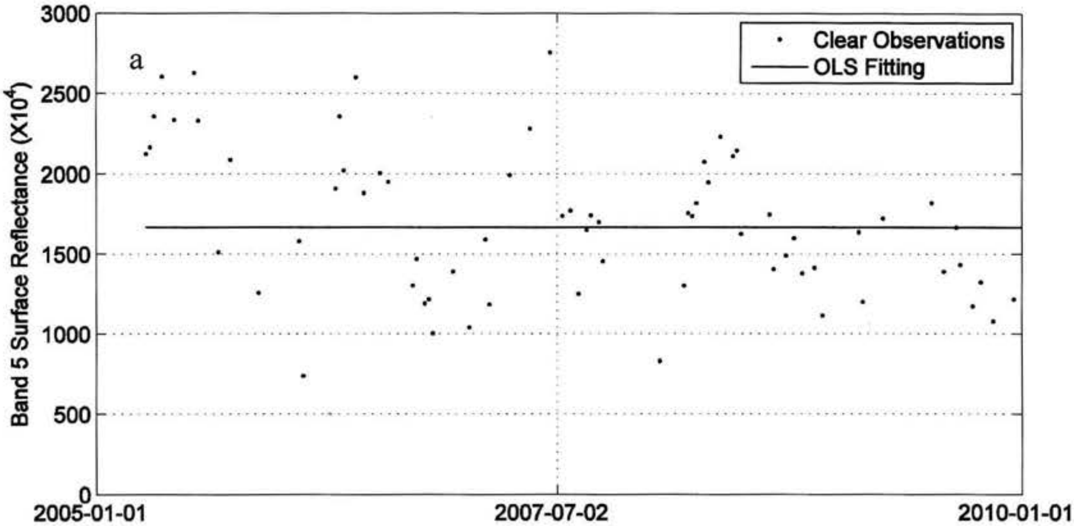
$a_{0,i}$: Overall mean value for the ith Landsat Band

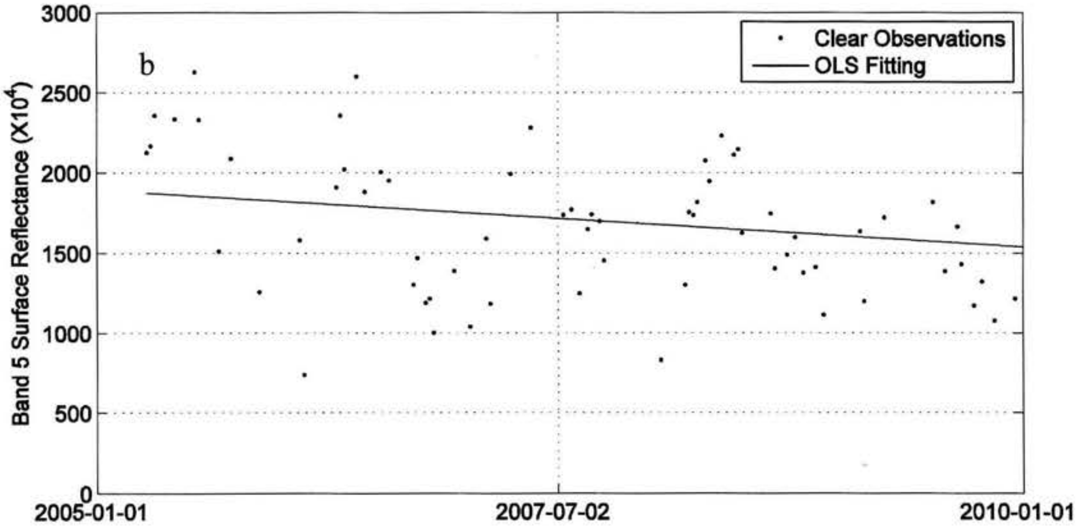
$a_{1,i}, b_{1,i}$: Coefficients for the intra-annual change for the ith Landsat Band

$c_{1,i}$: Coefficients for the inter-annual trend for the ith Landsat Band

τ_k^*: The kth break points.

The overall value for the ith Landsat Band is captured by $a_{0,i}$. Coefficients $a_{1,i}$ and $b_{1,i}$ are used to estimate the intra-annual changes caused by phenology and BRDF effects for the ith Landsat Band. The inter-annual trend for the ith Landsat Band is captured by $c_{1,i}$. Figure 4.4 shows the estimation results by including different components for the time series model. If we only use a single constant coefficient ($a_{0,i}$.), it can capture the overall reflectance while all the intra- and inter- annual variability is lost (Figure 4.4a). By including the inter-annual trend ($c_{1,i}$), the time series model is able to capture the inter-annual decreasing trend, however, losing all intra-annual variability (Figure 4.4b). The best result comes when all three components are used (Figure 4.4c).


a
Clear Observations
OLS Fitting
Band 5 Surface Reflectance (X10^4)
3000
2500
2000
1500
1000
500
0
2005-01-01
2007-07-02
2010-01-01


b
Clear Observations
OLS Fitting
Band 5 Surface Reflectance (X10^4)
3000
2500
2000
1500
1000
500
0
2005-01-01
2007-07-02
2010-01-01

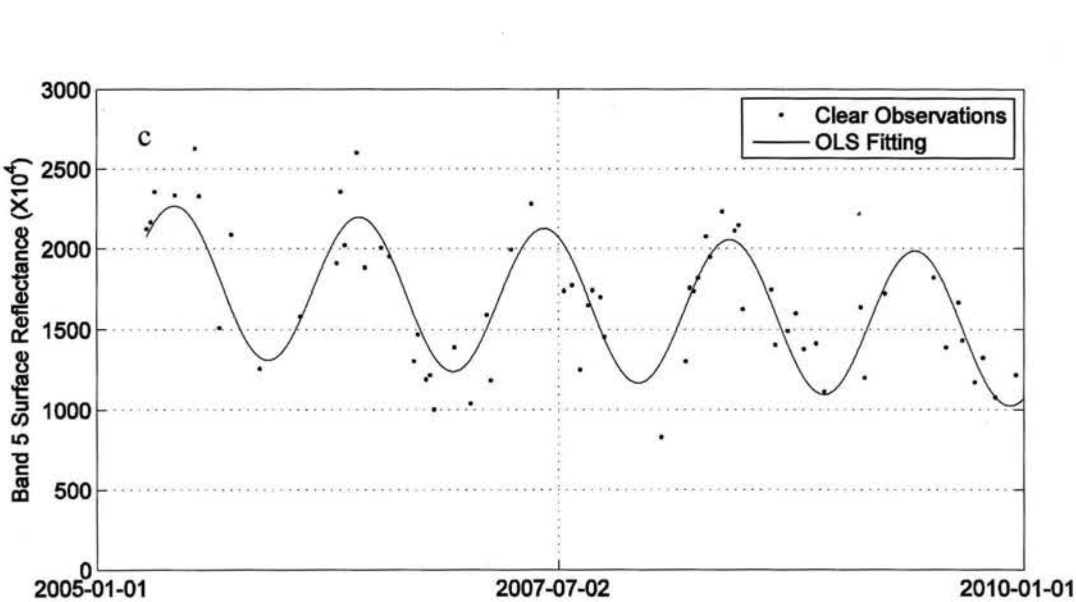


Figure 4.4. Illustration of OLS fitting results by including different components of the time series model. Figure 4.4a shows the results of only including a single constant coefficient. Figure 4.4b shows the results of adding the inter-annual trend to the time series model. Figure 4.4c shows the results of adding all three components.

Ideally, the more the parameters estimated, the more accurate the model will be. However, when there are too many parameters to estimate, it may start to fit to noise. To make the model robust to noise, the CCDC algorithm only uses two parameters to capture the intra-annual change. Model initialization is started when the number of clear pixels is equal to 12 and there is no land cover change during model initialization. Land cover change happened within the time of model initialization can make estimated time series model prediction biased. Therefore, if there is possible land cover change occurred during the process of model initialization, the CCDC algorithm will remove the first clear observation and add one more clear observation and this process will continue until no possible change is detected within the initialization time period. The possible change is detected in three ways in model initialization: abnormal slope, abnormal first observation, and abnormal last observation.

Usually if land cover change occurs during model initialization, the slope of the time series model will be much larger than its normal magnitude, because the slope will mostly respond to land cover change instead of gradual inter-annual changes. The slope for the ith band is firstly normalized by $3 \times RMSE_i/t_{model}$ (see Equation 4.2 for details), and if the average value of the normalized slope for all bands is larger than 1, it will be detected as an abnormal slope magnitude and be identified as a possible change within the model initialization time. The reason for normalizing the slope by $3 \times RMSE_i/t_{model}$ is based on the factor that if land cover change happens, the signal will usually deviate more than 3 times the RMSE (see Chapter 4.4.3 for detail) and this will make the slope of the time series model larger than $3 \times RMSE_i/t_{model}$.

On the other hand, if land cover change occurred at the start or the end of model initialization, there may not have been enough land cover change observations to make the magnitude of slope abnormal, but the land cover change observations may still influence model estimation. In this case, the CCDC algorithm will compare Landsat observation with the model prediction for the first and the last observations during model initialization, as usually a few change observations at the very beginning or end of the model initialization time will not influence model prediction significantly and they can still be detected by comparing the observations with model predictions. Therefore, the CCDC algorithm also calculates the difference between observations and model predictions for the first and the last observations. The difference for each Landsat band is also normalized by $3 \times RMSE_i$ (see Equation 4.2 for details). If the average normalized difference value of all Landsat bands is larger than 1, it will be identified as a possible change within the model initialization time. As soon as the model initialization is finished, it will be used as basis for continuous change detection and classification.

$$Possible\ Land\ Cover\ Change\ if\text{:}\ \frac{1}{7}\sum_{i=1}^{7}\frac{abs\big(Band\ i(x_1)-OLS(x_1,i)\big)}{3\times RMSE_i}>1$$

$$OR\ \frac{1}{7}\sum_{i=1}^{7}\frac{abs\big(Band\ i(x_n)-OLS(x_n,i)\big)}{3\times RMSE_i}>1\ OR\ \frac{1}{7}\sum_{i=1}^{7}\frac{abs\big(c_{1,i}(x)\big)}{3\times RMSE_i/t_{model}}>1 \tag{4.2}$$

Where,

x: Julian date

x_1: The Julian date for the first observation during model initialization

x_n: The Julian date for the last observation during model initialization

t_{model}: The total time used for model initialization

$Band\ i(x)$: The ith Landsat Band at the Julian date of x

$OLS(i,x)$: The OLS fitting for the ith Landsat Band at the Julian date of x

$c_{1,i}$: Coefficient for the trend for the ith Landsat Band from Equation 4.1

$RMSE_i$: Root Mean Square Error for the ith Landsat Band from Equation 4.1

4.3.3. Continuous change detection

The continuous change detection is started when the time series model has been initialized. The basis of our methods is comparison of the time series model prediction with satellite observation to find change. Ideally, a single date comparison would be definitive for detecting change. However, there is sufficient noise in the system due to factors like atmospheric haze, smoke, clouds, cloud shadows, snow, and changes in soil wetness, that when using a single date for comparison, there will be numerous false detections of land cover change. As noise factors tend to be ephemeral in nature, but land cover change is more persistent through time, the CCDC algorithm minimizes this effect by processing a set of dates together as a group for defining change. That is, if a pixel is observed to change in multiple successive dates of images, it is more likely to be land cover change. Based on previous studies (Zhu et al., 2012a), change identified in three successive dates showed the best results. Pixels showing change for one or two times will be flagged as "possible change". If a third consecutive change is found, the pixel is assigned to the "change" class.

If we are only focusing on one kind of land cover change such as forest disturbance, a single change index is sufficient. To find many kinds of land cover change, it is necessary to use all the spectral bands of the data as different kinds of land cover change may be responsible for different change magnitudes in different spectral bands. Therefore, the CCDC algorithm uses all seven spectral bands (including the thermal bands) and a data-driven threshold to detect many kinds of land cover change. After model initialization, there will be enough clear observations for estimating the time series model, and the CCDC algorithm will add new clear observations one at a time continuously for model estimating and detecting change. The OLS method (Equation 4.1) is applied to all seven Landsat bands and RMSE is computed for each band. This process continues when more clear observations are available. The difference between observation and model prediction for each Landsat band is normalized by $3 \times RMSE_i$. We used three times of RMSE due to the fact that when land surface changes, the spectral signals will usually deviate from model prediction for more than three times of RMSE.

Figure 4.5 illustrates how the three times of RMSE is used for detecting change for a deforestation pixel in Band 5 with all available Landsat data at different time. When there is no land cover change, the three next clear observations are always within the range of model prediction -/+ $3 \times RMSE_i$ (Figure 4.5a & Figure 4.5c). Figure 4.5b shows how this first change is detected by comparing the next three observations with model predictions and their range of three times of RMSE. In Figure 4.5 we only showed one Band to illustrate the algorithm, actually all the seven bands were used to detect change, as if land cover change occurs, all the spectral bands will deviate from their original trajectories.

In Figure 4.6, the same deforestation pixel was used and when deforestation happened, all spectral bands changed significantly. Therefore, CCDC averages the normalized difference value of all Landsat bands, and if it is larger than 1 for the next three consecutive clear observations, it is determined as change, otherwise, it will be identified as ephemeral changes and the next first clear observation will be flagged as outlier. The CCDC algorithm updates the time series model as soon as a new clear observation becomes available, adding a dynamic character to the process that proves more accurate. Equation 4.3 explains the details of how land cover change is identified.

$$Change(x) = \frac{1}{7}\sum_{i=1}^{7} \frac{\mathrm{abs}(\mathrm{Band\ i(x)} - \mathrm{OLS_i(x)})}{3 \times \mathrm{RMSE_i}} > 1 \ successively\ three\ times \qquad (4.3)$$

Where,

x: Julian date

i: The ith Landsat band

Band i(x): The ith Landsat band at the Julian date of x

$\text{OLS}_\text{i}(\text{x})$: The OLS fitting for the ith Landsat band at the Julian date of x

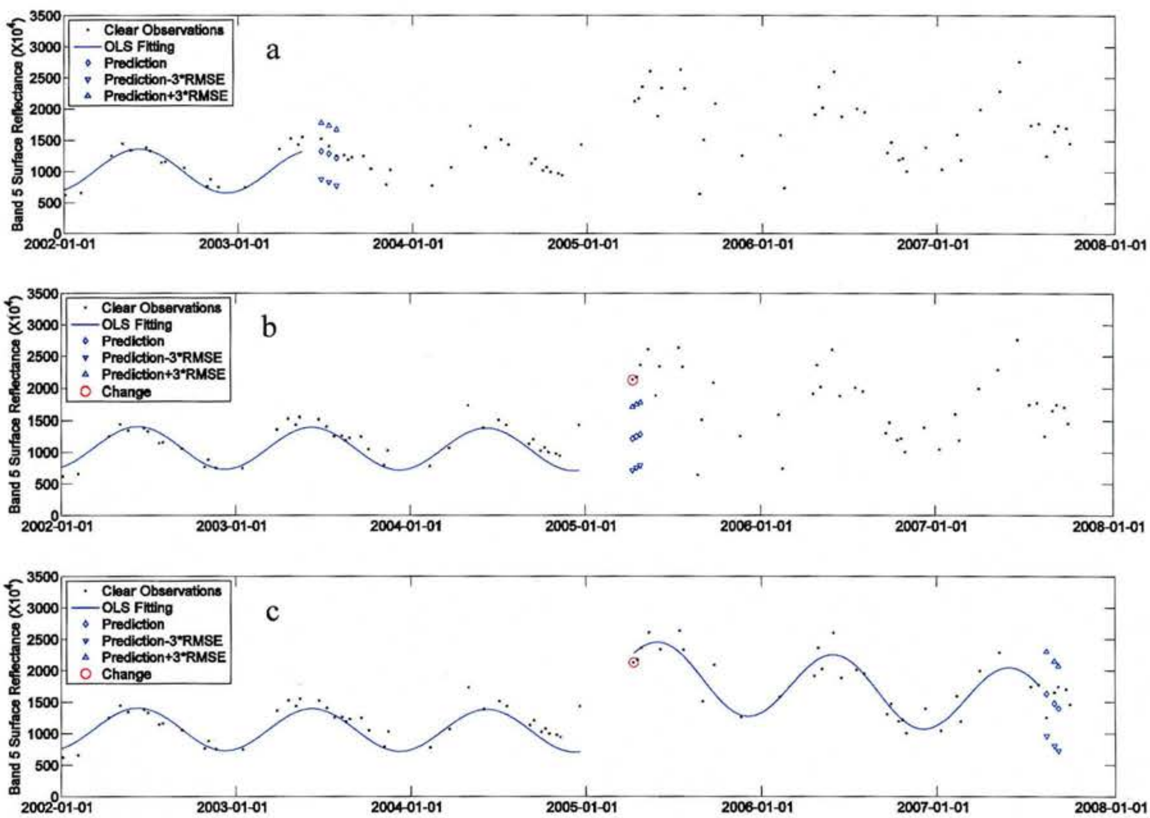
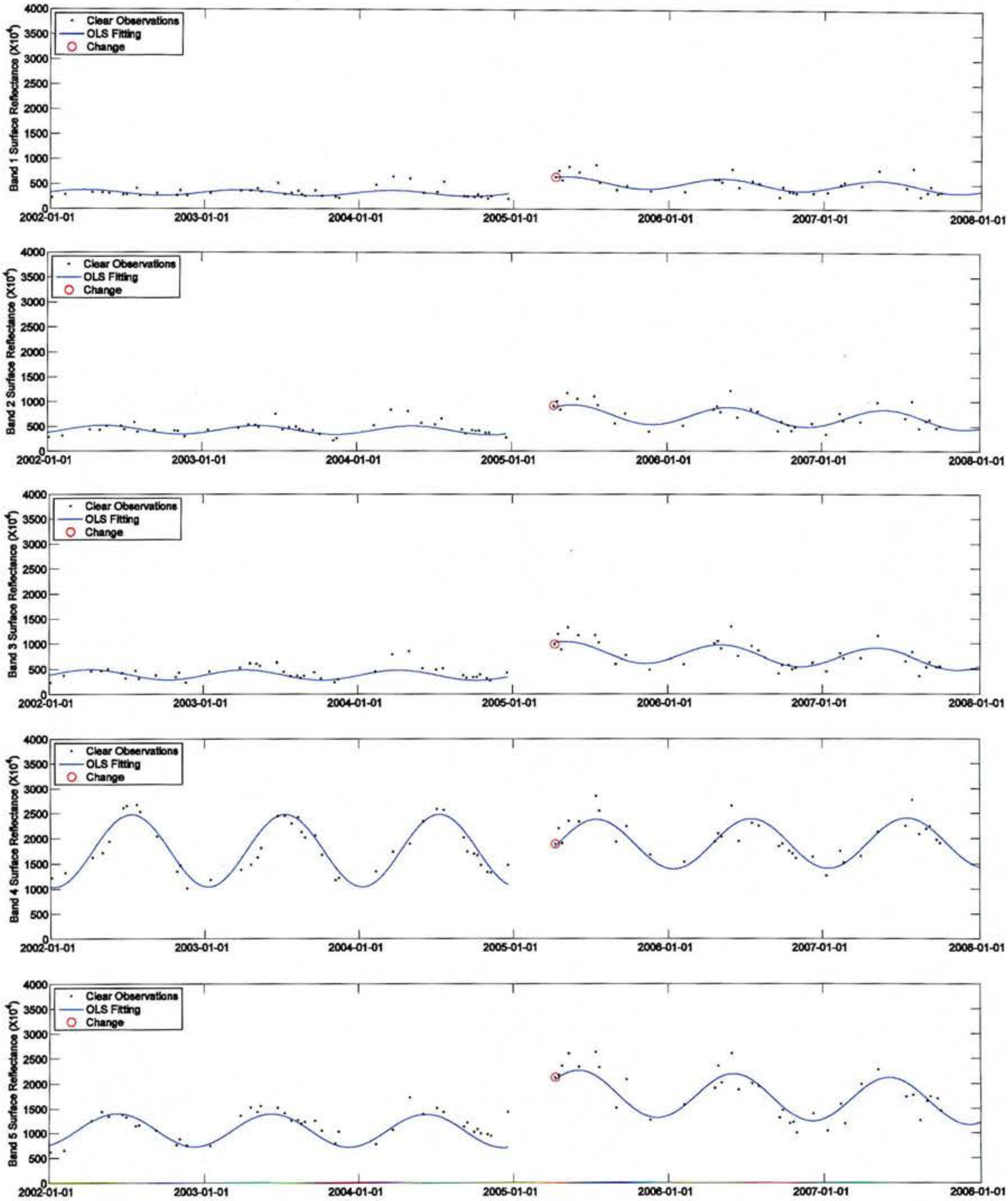


Figure 4.5. Illustrating the reason for using three times of RMSE for change detection. Figure 4.5a shows the model prediction and three times of RMSE before change happens. Figure 4.5b shows the model prediction and three times of RMSE when change happens. Figure 4.5c shows the model prediction and three times of RMSE after change happens.


Clear Observations
OLS Fitting
Change
Band 1 Surface Reflectance (X10^4)
Band 2 Surface Reflectance (X10^4)
Band 3 Surface Reflectance (X10^4)
Band 4 Surface Reflectance (X10^4)
Band 5 Surface Reflectance (X10^4)
0
500
1000
1500
2000
2500
3000
3500
4000
2002-01-01
2003-01-01
2004-01-01
2005-01-01
2006-01-01
2007-01-01
2008-01-01

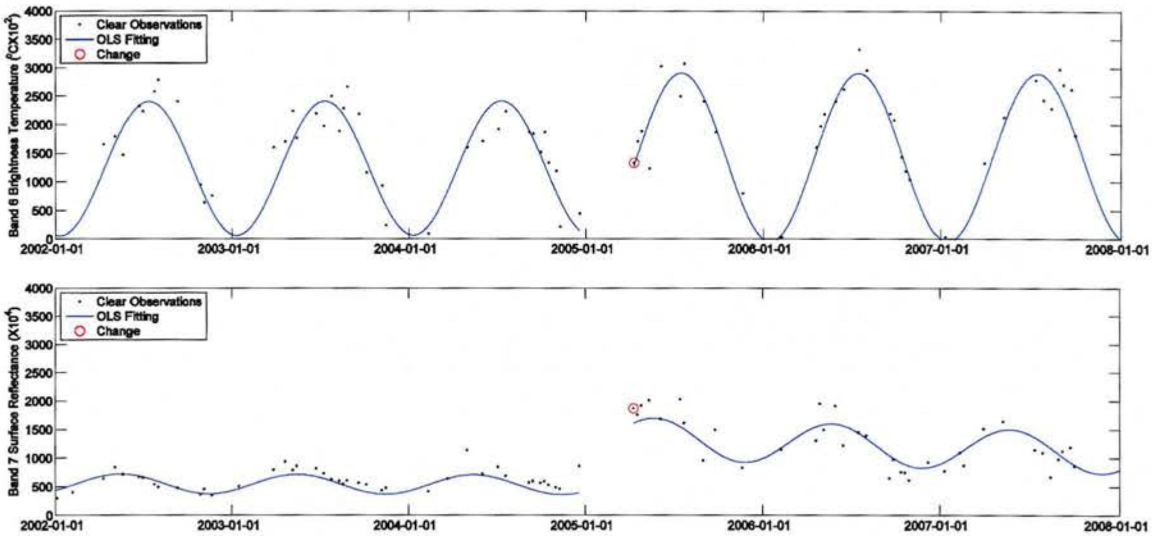


Figure 4.6. Illustration how the seven Landsat spectral bands (Band 1-7) deviate from their original trajectories when deforestation occurred

4.3.5. Continuous land cover classification

Finding land cover changes is important, but it will be more beneficial if we can know the land cover categories before and after change. Instead of classifying the original Landsat observations as the conventional methods do, the CCDC algorithm uses coefficients of time series models as inputs for land cover classification. After a change is detected, each pixel will have a time series before and after the changes represented by a few time series models. By classifying the time series model based on its corresponding coefficients, this algorithm can provide the land cover type at the time intervals of each time series model. Figure 4.7 shows examples of the estimated time series models for all seven Landsat bands for different kinds of land cover change in the study area. Taking Figure 4.7 *forest* to *developed* for instance, by classifying the first time series model, the

CCDC algorithm is capable of providing land cover class (*forest*) of this pixel between 2001 and 2002. Similarly, the classification results of the second time series model can provide the land cover class (developed) for this pixel between the middle of 2005 and 2006. The gaps in the middle of the two models are classified as *disturbed* in the land cover maps, because of the extremely large variability of the data during the transition time that prevent model initialization.

When *forest* is changed to *developed*, the time series models show completely different shapes, especially in Band 4 and Band 6. The reduction of Band 4 reflectance is easy to understand: vegetation reflects strongly in Band 4. The increase of Band 6 is mostly due to reduced evapotranspiration and urban heat island effects. When *forest* is changed to *barren*, the most significant changes are observed in Band 5 and Band 7, as forest is usually low in SWIR but barren is always high in these spectral bands. For the pixel that undergone changed from *forest* to *grass*, there is not much difference in the time series models in the visible bands, but the NIR, SWIR, and thermal bands are different. For the pixel that changed from *forest* to *agriculture*, the time series model of Band 4 show the biggest difference. Therefore, the time series model contains information that is very helpful for land cover classification.

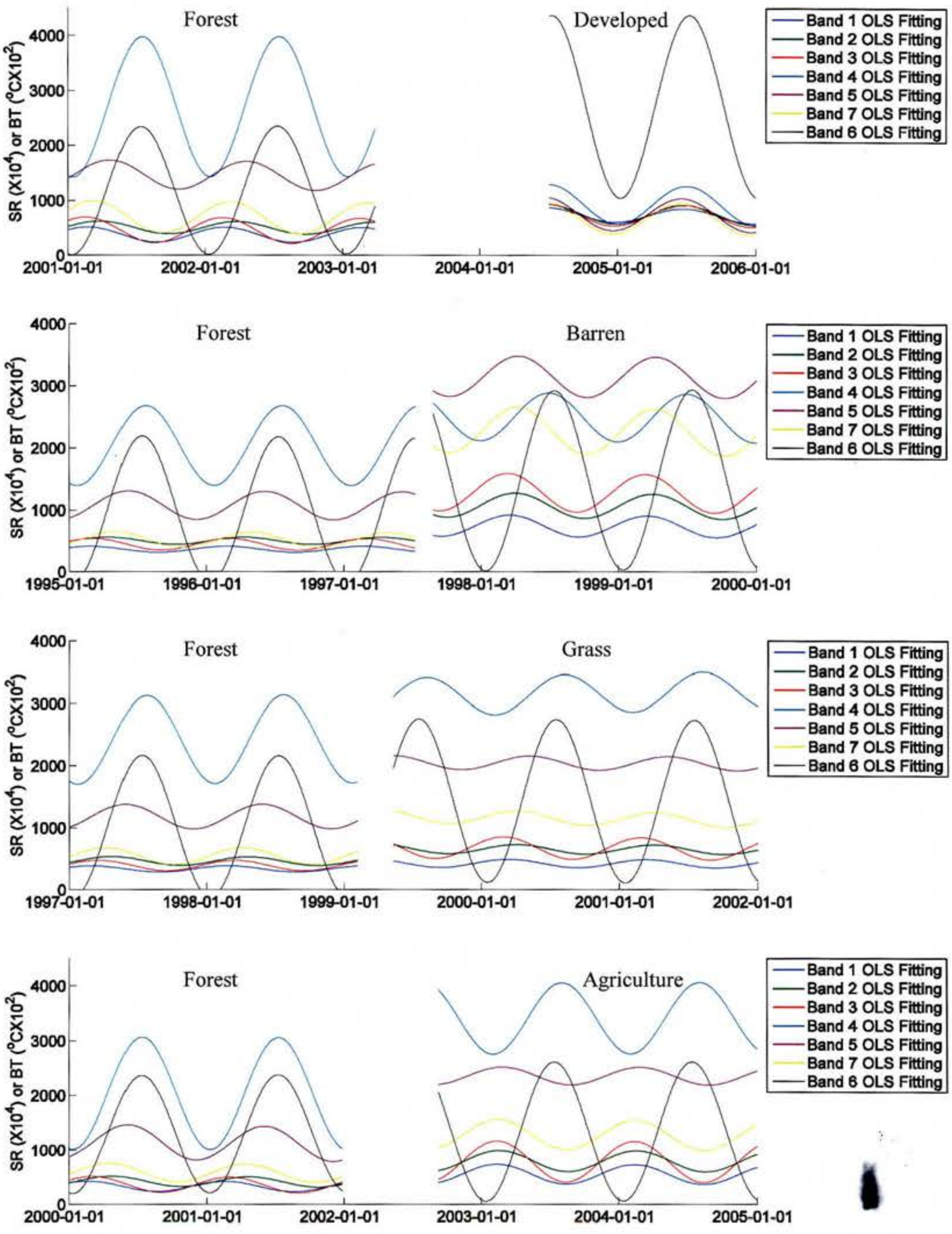


Figure 4.7. Examples of how the continuous land cover classification is done for the four most common kinds of land cover change in the study area.

The coefficients defining the time series model are the main inputs for this continuous classification. The coefficients of the time series model provide rich information in the spectral and temporal dimensions. The coefficient ($a_{0,i}$) represents overall value for the ith Landsat band at Julian date of zero is primarily responding to the spectral character of the data. However, it is meaningless to use it alone for land cover classification as zero Julian date is more than 2000 years ago. The CCDC algorithm converts this constant coefficient to an overall mean value at the center of the curve (central overall value) by combing the constant and trend coefficients together (Equation 4.4). The coefficients that capture annual variations (2 variables) and trend of the time series (1 variable) provide temporal information. The RMSE computed from the model fitting for each Landsat band is also used as one of the inputs for land cover classification, because of its rich information in describing data fluctuation not captured by the time series model. Figure 4.8 shows how the different variables (RMSE, central overall value, trend, cosine, and sine) separate the different land cover types. This information was generated by using estimated time series model coefficients from all reference data and then averaged for each land cover. It is clear that different land covers show quite different shapes in the plots of the five variables, and this information can be very helpful for discriminating a variety of land cover types.

$$Central\ Overall\ Value_i = a_{0,i} + c_{1,i} \times \frac{t_{start}+t_{end}}{2} \tag{4.4}$$

Where,

x : Julian date

i: The ith Landsat band

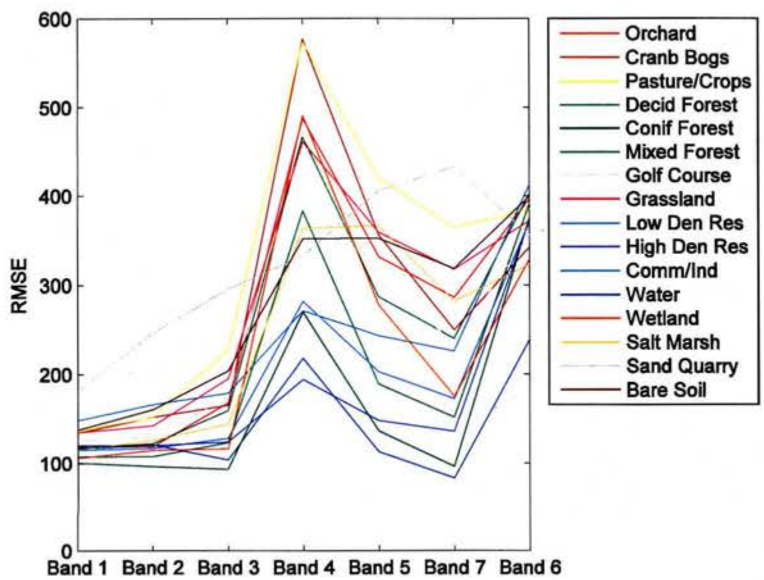
$a_{0,i}$: Coefficient represents overall value for the ith Landsat band at zero Julian date

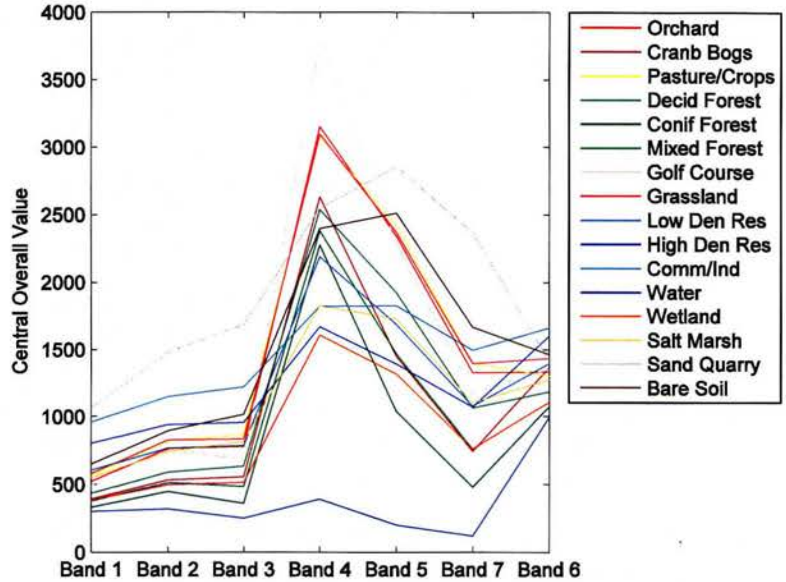
$c_{1,i}$: Coefficient capture the trend of Earth surface for the ith Landsat band

t_{start}: The time when model initialization started

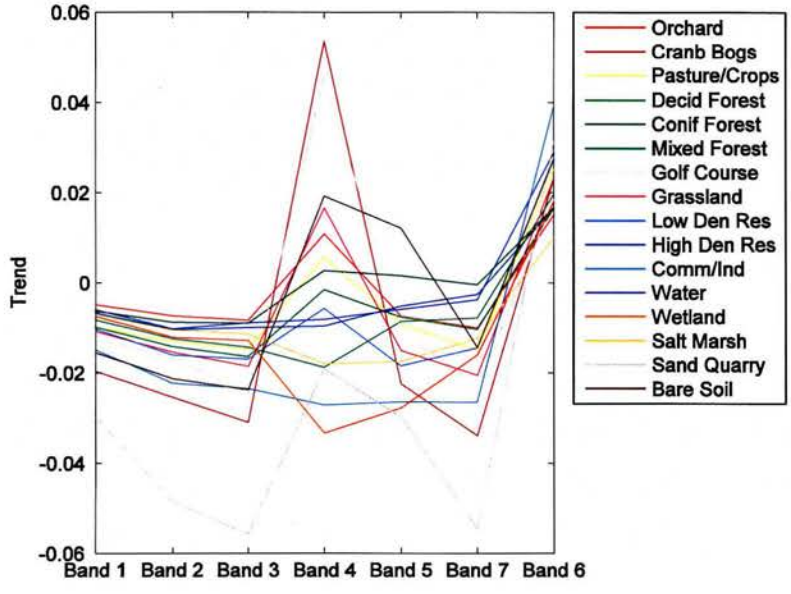
t_{end}: The time when model initialization ended

Therefore, there will be 5 variables for each of 7 bands, or 35 variables used as inputs for classification. As the reference data are derived between 2005 and 2007, the estimated coefficients of the reference pixels that were not detected to have land cover change between 2005 and 2007 are used as inputs for training the classifier. The Random Forest Classifier (RFC) is used to perform land cover mapping because of its relatively high accuracy and computational speed.





4000
3500
3000
2500
2000
1500
1000
500
0
Central Overall Value
Band 1 Band 2 Band 3 Band 4 Band 5 Band 7 Band 6
Orchard
Cranb Bogs
Pasture/Crops
Decid Forest
Conif Forest
Mixed Forest
Golf Course
Grassland
Low Den Res
High Den Res
Comm/Ind
Water
Wetland
Salt Marsh
Sand Quarry
Bare Soil



0.06
0.04
0.02
0
-0.02
-0.04
-0.06
Trend
Band 1 Band 2 Band 3 Band 4 Band 5 Band 7 Band 6
Orchard
Cranb Bogs
Pasture/Crops
Decid Forest
Conif Forest
Mixed Forest
Golf Course
Grassland
Low Den Res
High Den Res
Comm/Ind
Water
Wetland
Salt Marsh
Sand Quarry
Bare Soil

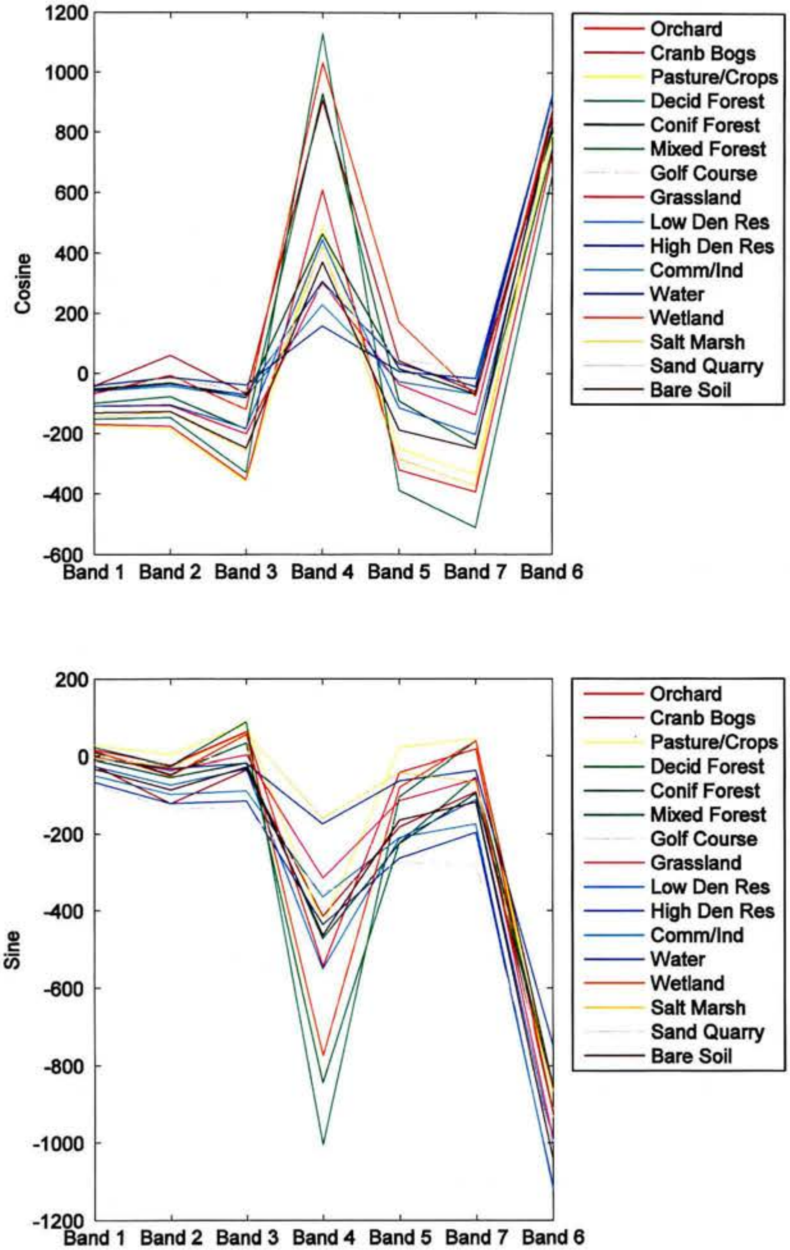


Figure 4.8. Illustrating the value of the five variables for land cover classification. The figure was generated by averaging the estimated time series model coefficients from all reference data for each land cover.

4.4 Results

4.4.1. Results of CCDC algorithm

The CCDC algorithm is capable of providing land cover change and a land cover classification map continuously with newly collected images. We used a very small (5.6km×2.8km) and a relatively large (90km×45km) area of Landsat data to better illustrate this algorithm and its results (Figure 4.9).

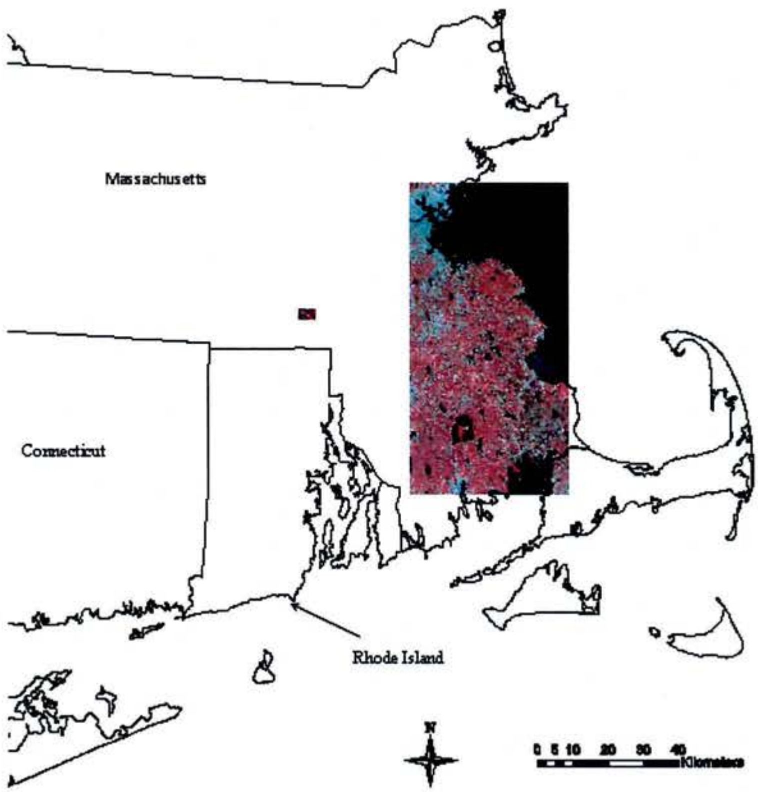


Figure 4.9. Two pieces of Landsat data used for illustration of the CCDC algorithm and its results. The small one (left) is used in Figure 4.10. and the larger one (right) is used in Figure 4.11.

Figure 4.10 illustrates the almost 30 year time series data for 3 different pixels in Band 4 that ultimately underwent some kind of land cover change. The top three panels are a small piece of a Landsat image acquired in July 16^{th} 2011 (left), a map showing the timing and location of land cover change (middle), and the land cover at the end of the time series (right). In each case, the changes in the pixels are obvious when viewed from the perspective of the entire time series. This approach allows the timing of each change to be identified, as well as the kind of change. When the time series has been built for a pixel and analyzed for change, it is possible to use the estimated models between the changes to identify the land cover class for the pixel. In Figure 4.10 for pixel located at site 1, the estimated model preceding the change in 1990 can be used to classify the land cover for that time period. Similarly the estimated model subsequent to the change can be used to identify what land cover came after the change in 1990. The shape of the time series model can be very helpful in land cover classification which is evident in Figure 4.10, as initially both pixels located at site 1 and 3 were conifer forest and pixel located at site 2 was a hardwood forest, and they are readily distinguishable by the big difference in the amplitude of the time series.

Figure 4.11 shows updated land cover change and classification maps for a larger area for the Boston scene. It can also provide new kinds of information about what kinds of land cover change occurred on a yearly basis for the entire scene. To make it simple, I collapsed the 16-categories land cover classes into 7-categories classes that have *forest*, *wetland*, *agriculture*, *barren*, *water*, *grass*, and *developed*. In Figure 4.12, the histogram provides the timing and nature of land cover change on a yearly basis. Moreover, it can generate information on one kind of land cover change, such as annual forest net loss (Figure 4.13).

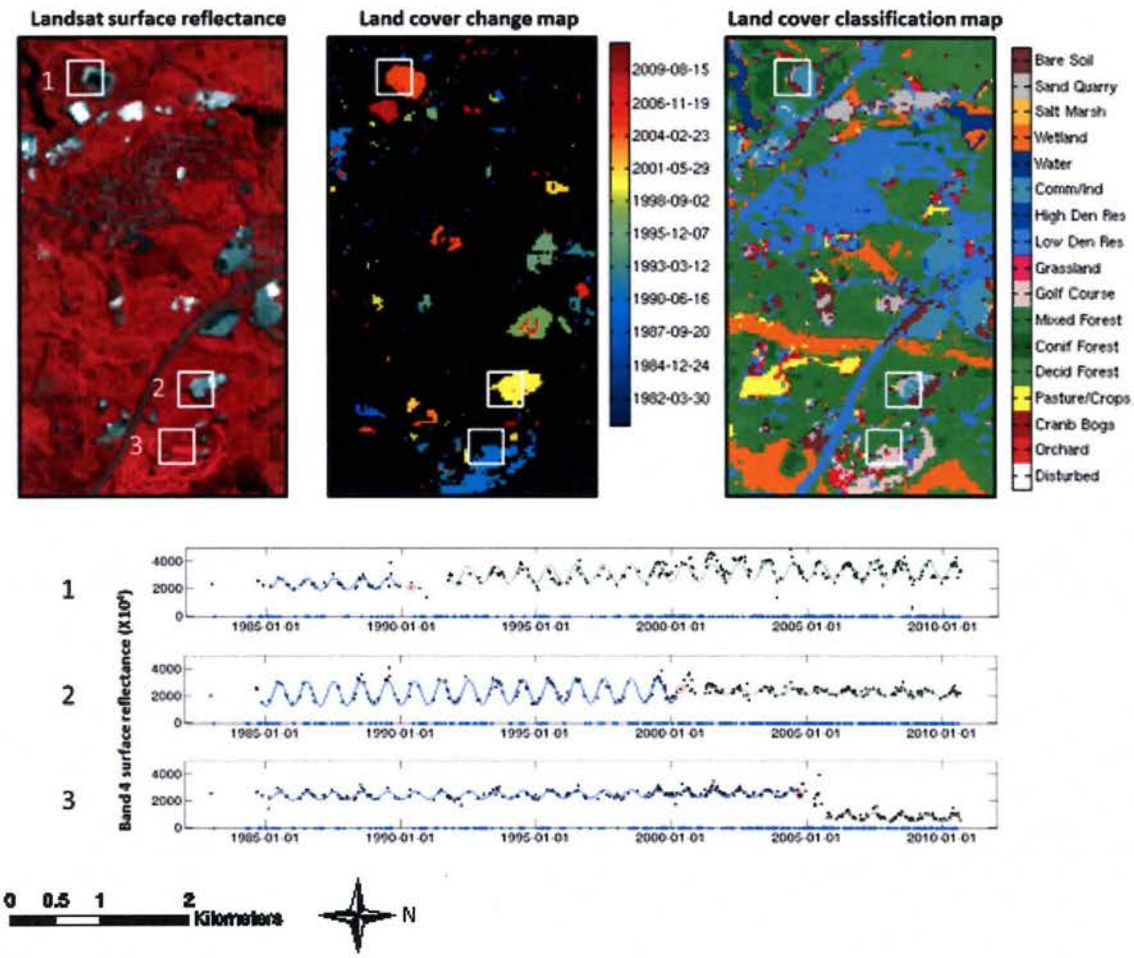


Figure 4.10. The smaller piece of Landsat data used for illustration of the CCDC results. The top three panels are a piece of Landsat image acquired in July 16th 2011 (left), a map showing the timing and location of land cover changes over the history of the Landsat in the TM and ETM+ eras (middle), and the land cover at the end of the time series (right). The three graphs at the bottom show the time series for the 3 pixels in sites 1, 2, and 3.

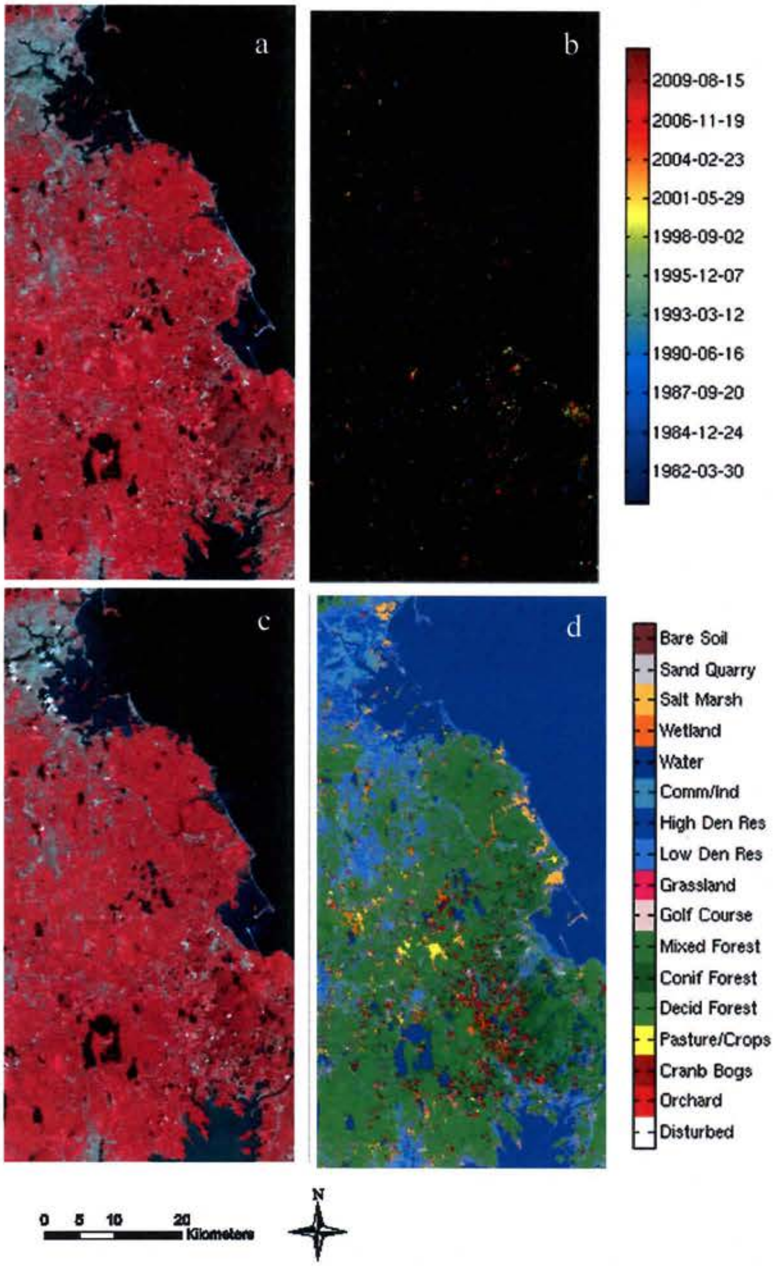


Figure 4.11. The larger piece of Landsat data used for illustration of the CCDC results. Figure 4.11a is the Sept. 7th 1984 Landsat image. Figure 4.11c is the July 16th 2011 Landsat image. Figure 4.11b is the land cover change map from 1982 to 2011. Figure 4.11d is the most recent classification map at the end of the time series.

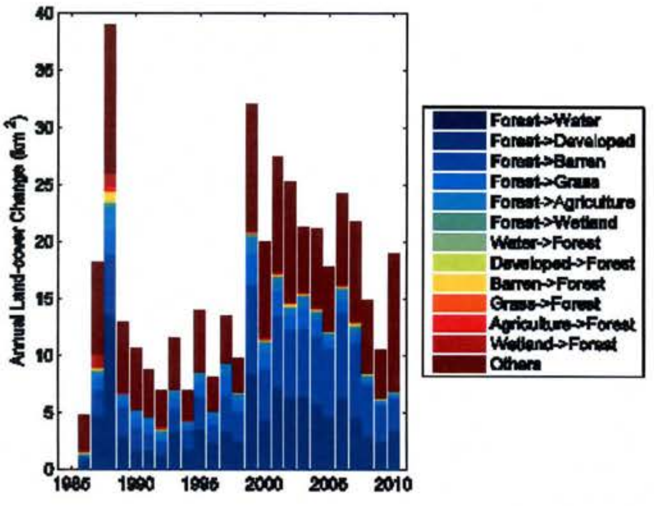


Figure 4.12. This graph shows the annual amounts (km^2) of different kinds of land cover change for the Landsat TM and ETM+ era

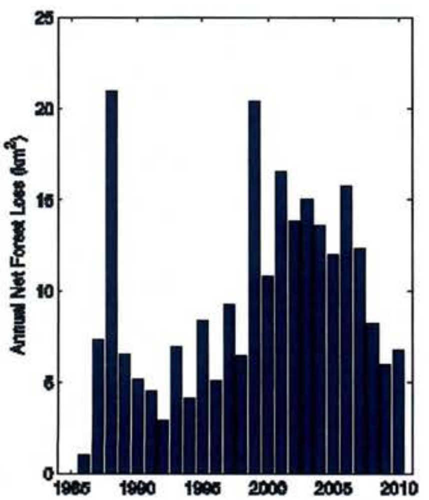


Figure 4.13. This graph shows the annual amounts (km^2) of net forest loss for the Landsat TM and ETM+ era

4.4.2. Accuracy assessment

4.4.2.1. Accuracy assessment for change detection

As this CCDC algorithm can not only provide land cover change maps at very high temporal frequency but also produce land cover maps at any given time. It would be very hard to get reference data that can thoroughly assess its accuracy both spatially and temporally. As there are no independent datasets available that have finer spatial resolution and higher temporal frequency than Landsat images, to know where and when the land cover change occurs, the only source for reference data is the original Landsat images (Cohen et al., 2010). High spatial resolution images from Google Earth (http://earth.google.com/) can be used to help manual interpretation. Though the high spatial resolution images in Google Earth cannot provide the same temporal frequency as Landsat data, their high spatial resolution is helpful in determining land cover change at longer time intervals. A random stratified sample design was used for assessing the change detection accuracy. A total of 500 reference pixels were selected, in which 250 pixels were selected within areas where land cover was persistent throughout the time and 250 pixels were selected within areas where there was land cover change identified in the time series analysis. By carefully examining the time series data for all seven bands, it can be quite easy to identify land cover changes and when they occur. If there is confusion in determining a change or when it occurred, I looked at the Landsat images before and after the possible change time or used high spatial resolution images from Google Earth to help determine what was happening at that time for that specific location. If there are multiple land cover changes within one pixel, only the first change is used for reference.

The accuracy assessment shows that the CCDC algorithm results were accurate for detecting change spatially, with producer's accuracy of 97.72% and user's accuracies of 85.60% for changed pixels, and overall accuracy of 91.80% (Table 4.2). The relative lower user's accuracy indicates more commission errors than omission errors in detected changes. A higher threshold or longer consecutive observations may better balance the commission and omission errors.

Table 4.2. Confusion matrix for the accuracy assessment of the CCDC algorithm in the spatial domain.

Reference data (spatial)				
CCDC algorithm	Changed pixels	Stable pixels	Total	User's (%)
Changed pixels	214	36	250	85.60
Stable pixels	5	245	250	98.00
Total	219	281		
Producer's (%)	97.72	87.19	Overall (%)	91.80

The omission errors are mostly due to the following two reasons: 1) partial change; 2) change occurs too early, before the model is initialized. The partially changed pixels are always hard to detect, as the change magnitude is mostly dependent on the proportion of change within that pixel. For example, in Figure 4.14, there was a partial forest cut around 2000 and we can see at that time the Band 5 surface reflectance variability changed significantly, but the magnitude of this change is relatively small. On the other hand, if land cover changed at the very beginning of model initialization, the CCDC algorithm will not be able to detect any kind of change as there is not enough data to initialize the time series model (Figure 4.15).

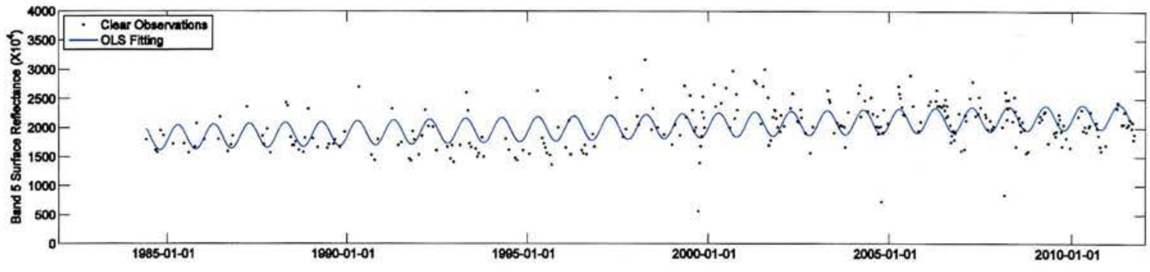


Figure 4.14. Omission problem – partial forest cut shown in Band 5 surface reflectance.

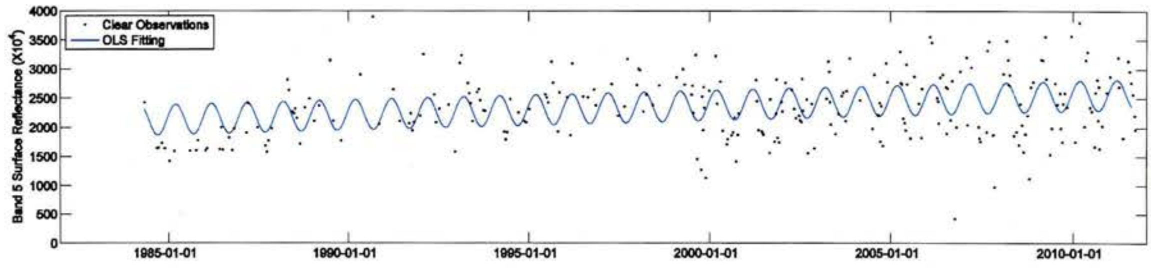


Figure 4.15. Omission problem– change happened too early shown in Band 5 surface reflectance

The commission errors are mostly the results from the following four reasons: 1) overfitting; 2) missed clouds consecutively; 3) scarce data; 4) small RMSE. Over fitting may cause serious problems. Though we are using very simple models, overfitting can still happen if data is always missing for a certain time of year. In Figure 4.16, overfitting has happened for Band 4 at the beginning of model estimating, due to constant snow cover during the winter. On the other hand, if clouds are missed three consecutive times, it will also be identified as land cover change (Figure 4.17). As the CCDC algorithm will initialize as long as there are 12 clear observations, scarce data may also cause problems in change detection, especially for pixels located at the edges of Landsat images (Figure 4.18). The last reason for false detected change is because the very small RMSE used for thresholding. When the RMSE is very small, very slightly change caused by the atmosphere can be easily identified as change and this is very common for water pixels, considering their dark character (Figure 4.19).

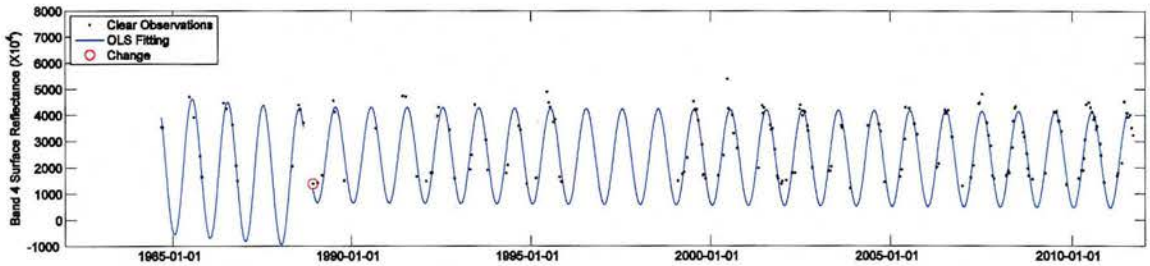


Figure 4.16. Commission problem– overfitting shown in Band 4 surface reflectance.

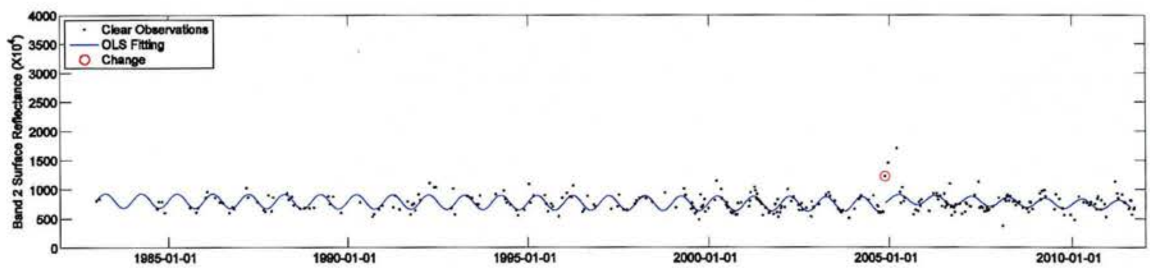


Figure 4.17. Commission problem– missed clouds three times consecutively shown in Band 2 surface reflectance.

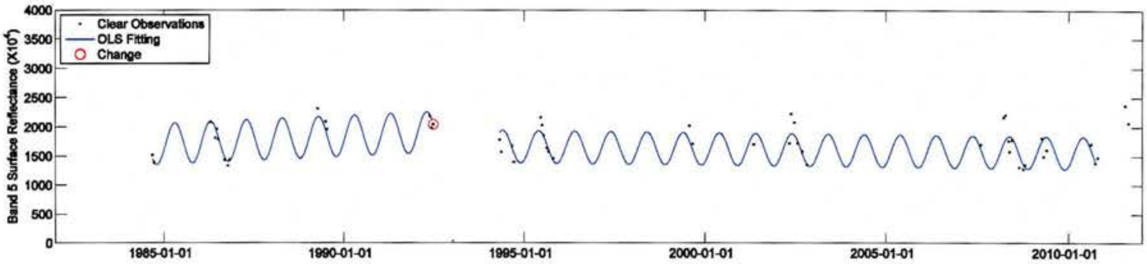


Figure 4.18. Commission problem– scarce of data shown in Band 5 surface reflectance.

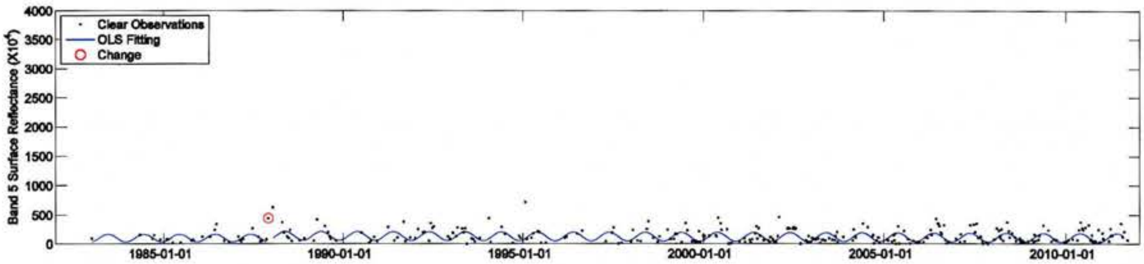


Figure 4.19. Commission problem– small RMSE shown in Band 5 surface reflectance.

The temporal accuracy of this change detection algorithm was assessed for all the pixels randomly generated within land cover change areas that are correctly identified in the spatial domain. The proportion of the pixels that have the same /before/after change time between the algorithm results and reference data is used to assess the temporal accuracy. The algorithm tends to find change later than the reference data and there are not any pixels found to have changed earlier than the reference data. The proportion of the pixels that have the same change time between the algorithm results and reference data is 79.91%, which means most of the detected change times are correct and for

20.09% of the pixels, change is found after the reference data and approximately 93% of the pixels change is found within 32 days of the first date when a change is observable.

Table 4.3. Table for the accuracy assessment of the CCDC algorithm in the temporal domain.

	Reference data (temporal)			
CCDC algorithm	Same	$0 < \text{Late} \leq 32$ days	Late > 32 days	Total
Changed pixels	171	28	15	214
Proportion (%)	79.91	13.08	7.01	100.00

The temporal errors are mostly due to the fact that at the very beginning of change, the pixel may only have partially changed, and the spectral signals are not large enough (less than three times of RMSE) to be identified as change by the CCDC algorithm. In this case, the change time detected by the algorithm may be later than the reference time for one or two clear observations (Figure 4.20). In Figure 4.20, before the red circle (break), there is one observation that deviates from model prediction slightly that was caused by a partial cut, but this deviation is still relatively small compared to the observation in the red cycle.

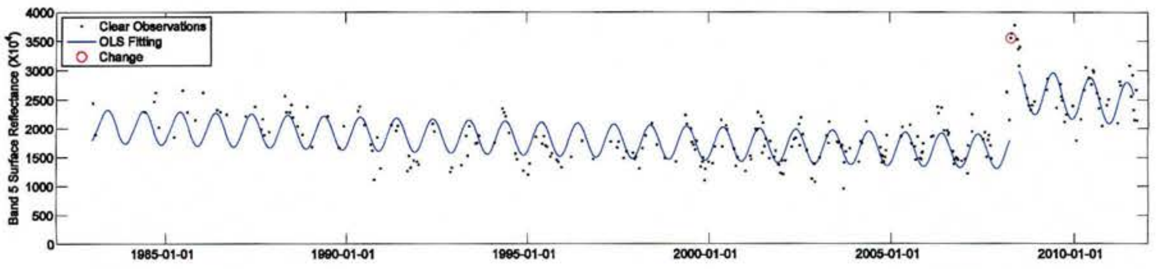


Figure 4.20. Temporal error problem – find change later than observed

4.4.2.2. Accuracy assessment for land cover classification

The land cover reference data (Rogan et al., 2010) was used as the basis for assessing the accuracy of the land cover classification. As the land cover reference data were collected between 2005 and 2007, we will only use the pixel if there is no change at this period of time and assess the accuracy of its land cover type at the same time. We performed a fifty-fold cross-validation analysis with the training database. A total of 80% of the ground reference data were randomly selected to train the classifier, and the remaining 20% were used to assess map accuracy (Fielding & Bell, 1997). This process was repeated 50 times and we use the average user's accuracy, average producer's accuracy, and average overall classification accuracy to assess the land cover classification accuracy of the CCDC algorithm, such that each reference pixel is used many times for both training and assessing, but never both for a single trial.

The 16-categories land cover classification results from CCDC also showed high accuracy with an average overall accuracy of 90.2% which is almost the same accuracy when using all available dimensions of Landsat data (Zhu et al., 2012b). The average producer's and user's accuracies for the different land cover types are also quite high, with mostly of them near 90% (Table 4.4). The largest confusions are from *Mixed Forest* and *Conifer Forest, Deciduous Forest and Mixed Forest, Low Density Residential and High Density Residential, and Low Density Residential and Commercial/Industrial.*

Table 4.4. Confusion matrix for 16-categories land cover classification derived from the CCDC algorithm.

	O	CB	PC	DF	CF	MF	GC	G	LD	HD	CI	W	WT	SM	SQ	BS	Use.
O	1,986	0	128	1	0	0	43	71	22	0	0	0	0	0	0	33	87.0
CB	0	2,152	2	0	0	3	8	0	0	0	0	5	0	0	0	2	99.1
PC	135	2	3,550	3	0	0	39	93	14	0	0	0	12	0	17	106	89.4
DF	3	0	10	3,619	11	209	34	5	26	0	9	0	21	0	0	14	91.4
CF	0	0	0	0	3,643	431	10	0	8	0	0	1	0	0	0	0	89.0
MF	1	8	23	136	379	4,686	23	12	39	0	73	20	42	14	15	39	85.1
GC	12	0	37	0	0	42	2,450	83	3	0	0	0	0	0	0	24	92.4
G	51	11	60	0	0	0	109	2,967	91	0	20	0	0	0	17	46	88.0
LD	32	13	110	0	34	21	24	117	2,880	101	266	0	16	8	2	50	78.4
HD	0	0	0	0	0	0	0	13	185	1,494	19	0	0	0	0	0	87.3
CI	0	0	0	0	0	0	0	61	44	136	3,108	0	0	7	114	59	88.1
DW	0	0	0	0	0	0	8	0	51	0	0	8,311	65	3	0	0	98.5
W	0	1	36	30	2	12	1	0	11	0	8	112	3,431	42	5	9	92.7
SM	2	0	0	0	0	0	3	0	6	0	0	0	21	2,987	0	0	98.9
SQ	0	0	0	46	0	32	0	34	5	16	165	13	0	0	2,635	159	84.9
BS	6	22	72	24	3	0	0	66	39	0	51	12	1	0	87	2,327	85.9
Pro.	89.1	97.4	88.1	93.8	89.5	86.3	89.0	84.2	84.1	85.5	83.6	98.1	95.1	97.6	91.1	81.1	90.2

Note: O=Orchards, CB=Cranberry Bogs, PC=Pasture/Row Crops, DF=Deciduous Forest, CF=Conifer Forest, MF=Mixed Forest, GC=Golf Course, G=Grassland, LD=Low Density Residential, HD=High Density Residential, CI=Commercial/Industrial, W= Water, WT=Wetland, SM=Salt Marsh, SQ=Sand Quarry, BS=Bare Soil

4.5 Discussions and conclusions

In this study, we developed a new algorithm for continuous land cover change detection and classification at high temporal frequency using all available Landsat data. This approach also allows reconstruction of the history of the Earth's surface. Using all available Landsat TM and ETM+ images, models using sines and cosines are estimated for each pixel and each spectral band. These models can predict Landsat images at any date assuming there is no land cover change. The CCDC algorithm flags land cover change by differencing the predicted and observed Landsat data. It determines a disturbance pixel by the number of times "change" is observed consecutively. Pixels showing "change" for one or two times are flagged as "possible change". If a third consecutive "change" is found, the flag will be mapped as change. The estimated coefficients (also including RMSE) were used for land cover classification. The reference data revealed that the CCDC results were accurate for detecting land cover change, with producer's accuracy of 98% and user's accuracies of 86% in the spatial domain and temporal accuracy of 80%. The CCDC classification results also showed high overall accuracy of 90%.

The CCDC algorithm has many advantages. It is fully automated and is capable of monitoring many kinds of land cover change continuously as soon as new images are collected. Moreover, there are no empirical and global thresholds used in change detection. The thresholds are generated through the original observations and model estimation. In this study, three times the RMSE is recommended for thresholding, but more subtle changes can be captured if two times of RMSE is used, which may also include more false detection of land cover change. The continuous character of the monitoring makes the algorithm capable of using as many images as possible. Therefore, how fast the CCDC algorithm is able to find change and its corresponding land cover type is solely dependent on the frequency of available clear observations. This algorithm will definitely improve as the frequency of high resolution images from sensors like Landsat increases. The opening of the archive from Earth Resources Observation and Science (EROS) Data Center is the first major step. In the near future, the launch of the

Landsat Data Continuity Mission (LDCM) should further increase the frequency of available observations considering the much larger duty cycle for LDCM compared with the previous Landsat satellites. Moreover, the launch of the two European satellites Sentinel 2A/2B will greatly increase Landsat like observations as they will have a repeat time of every five days. More importantly, when the two Sentinel 2A/2B satellites are launched, they will have a repeat time of every five days. By combining observations from all these sensors, we will be able to monitor land cover change in near real-time at Landsat scales.

By considering each pixel separately, the CCDC algorithm can overcome most of the limitations that the conventional approaches have. By using any clear observations for each pixel to estimate the time series model for each spectral band, this algorithm expands the use of Landsat images to any time of year and to all kinds of conditions (e. g., cloud, snow, heavy aerosols). It also can work in very heterogeneous areas which are reported to be problematic for the conventional methods (Masek et al., 2008; Huang et al., 2010). Moreover, the CCDC algorithm does not need to perform relative normalization for each image like the conventional methods do (Kennedy et al., 2007; Huang et al., 2010), as the time series model already includes the phenology and BRDF effects. By using many observations for model estimation, this algorithm is more robust to noise and the estimated data will be more stable compared to the original observations. Figure 4.21 illustrates the model estimated Landsat observations at the same time of year and the original Landsat observations during growing season through the TM and ETM+ era. It is clear that the model estimated observations are much stable compared to the original observations and this feature would reduce false positive errors in change detection significantly.

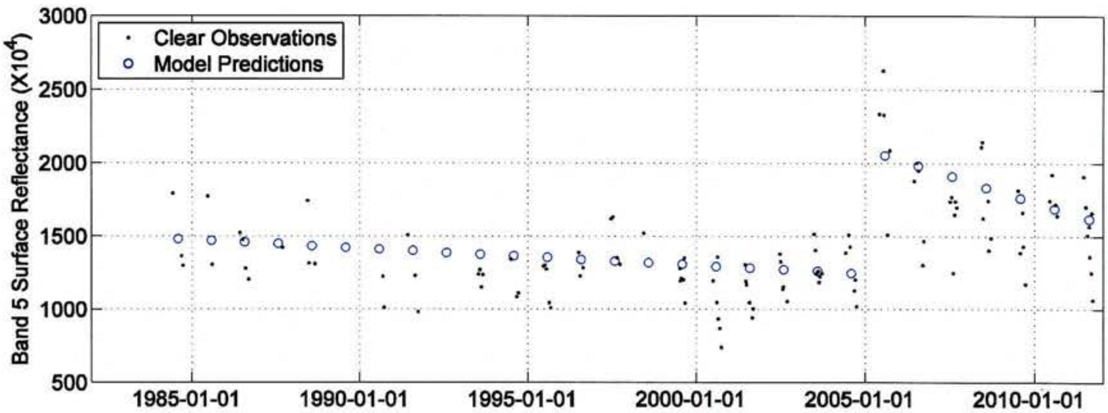


Figure 4.21. Clear observations from Jun. to Sept. and model predictions for Aug. 1st every year between 1984 and 2012. Forest clearing was occurred in 2005 for this pixel.

The failure of the Scan Line Corrector (SLC) in Landsat 7 will not cause problem for this CCDC algorithm. The scan line gaps are treated just like clouds or other things that remove observations from images and the available good observations are used. One area of future research will be to integrate observations from adjacent Landsat images in the zone of "side lap". This approach will further minimize the effects of Landsat 7 SLC-off gaps as they are most pronounced in these areas of side lap. The same is true for images with partial cloud cover, as they have many useful observations. As a result, it would be highly desirable if the Landsat satellites of the future collected all possible observations, as even partially cloudy images have value in analysis systems like the CCDC algorithm.

Additionally, land cover maps from any time period in the history of the Landsat TM and ETM+ era can be generated. The continuous character makes it possible to provide the most recent land cover maps and update the map as soon as new observations become available. And, this algorithm can provide maps of land cover change over any specified time period and give information about the land cover classes before and after change occurs by simply differencing the two land cover maps at different times. Usually, it is very dangerous to compare two land cover maps at different times to find change, as the areas of land cover change are small compared to the large magnitude of error in land cover classification maps. This algorithm should avoid this problem when comparing two land cover maps to find change, as the land cover classification algorithm here is based on the results of change detection. Moreover, this algorithm can also provide information about inter-annual changes via the trend coefficients. It is possible to provide new land cover categories that are unique temporally, for example, the forest class can be separated into growing forest, mature forest, and dying forest. This kind of information is important for studying the health conditions of vegetation, but difficult to derive with the conventional land cover classification methods.

Chapter 5

5. Concluding remarks

Land cover change detection and classification is a difficult problem in remote sensing, especially if we want it accurate and timely. Fully use of the temporal domain of the data and time series analysis will improve the future of remote sensing. To use lots of Landsat images for automated analysis, I first developed a two-step algorithm for screening of cloud, cloud shadow, and snow. Next, I used time series model for detecting forest disturbance in one calendar year. Finally, I extended the algorithm to detect many kinds of land cover change and find changes throughout the TM and ETM+ era.

The first part is a two-step cloud, cloud shadow, and snow screening algorithm. The first step is a new algorithm called Fmask developed for cloud and cloud shadow detection for single-date Landsat images. Landsat TOA reflectance and brightness temperature are used as inputs. Fmask first uses rules based on cloud physical properties to separate potential cloud pixels and clear-sky pixels. Next, a normalized temperature probability, spectral variability probability, and brightness probability are combined to produce a probability mask for clouds over land and water separately. Then, the PCPs and the cloud probability mask are used together to derive the potential cloud layer. The darkening effect of the cloud shadows in the NIR Band is used to generate a potential shadow layer by applying the flood-fill transformation. Subsequently, 3D cloud objects are determined via segmentation of the potential cloud layer and assumption of a constant temperature lapse rate within each cloud object. The view angle of the satellite sensor and the illuminating angle are used to predict possible cloud shadow locations and select the one that has the maximum similarity with the potential cloud shadow mask. If the scene has snow, a snow mask is also produced. For a globally distributed set of reference data, the average overall cloud accuracy is as high as 96.4%. Next, a multitemporal method for automatically identifying clouds, cloud shadows, and snow is provided. With the long history of Landsat data in the archive, we can build up a cloud, cloud shadow free dataset. To build this dataset, we first use a single-image based method (Fmask) to exclude most of the "noisy" pixels and then use the robust linear least square fitting method to get the predicted cloud, cloud shadow, and snow free dataset. With this dataset, we can improve the previous single-image based mask results and use it in future dataset analysis.

The second part is the new change detection algorithm for continuous monitoring of forest disturbance at high temporal frequency. Using all available Landsat ETM+ images in two years, time series models consisting of sines and cosines are estimated for each pixel for each spectral band. Dropping the coefficients that capture inter-annual change, time series models can predict surface reflectance for pixels at any location and any date assuming persistence of land cover. The Continuous monitoring of forest disturbance algorithm flags forest disturbance by differencing the predicted and observed Landsat images. Two algorithms (single-date and multi-date differencing) were tested for detecting forest disturbance at a Savannah River site. The map derived from the multi-date differencing algorithm was chosen as the final CMFDA result, due to its higher spatial and temporal accuracies. It determines a disturbance pixel by the number of times "change" is observed consecutively. Pixels showing "change" for one or two times are flagged as "probable change". If the pixel is flagged for the third time, the pixel is determined to have changed. The accuracy assessment shows that CMFDA results were accurate for detecting forest disturbance, with both producer's and user's accuracies higher than 95% in the spatial domain and temporal accuracy of approximately 94%.

The third part of this research is the new algorithm for continuous change detection and classification of land cover using all available Landsat data. This algorithm is capable of detecting many kinds of land cover change continuously as new images are collected and providing land cover maps for any given time. A time series model that has components of seasonality, trend, and break estimates the surface reflectance and brightness temperature. The time series model is updated dynamically with the newly acquired observations. Due to the large difference in spectral response for various kinds of land cover change, the CCDC algorithm uses a data-driven threshold derived from all seven Landsat bands. When the difference between observed and predicted exceeds the thresholds three consecutive times, a pixel is identified as land surface change. Land cover classification is done after change detection. Coefficients from the time series models and the RMSE from model fitting are used as classification inputs for the RFC. We applied the CCDC algorithm for one Landsat scene at Worldwide Reference System (WRS) Path 12 and Row 31. All available Landsat images (a total of 519) acquired between 1982 and 2011 were processed. The accuracy assessment shows that CCDC results were accurate for detecting land surface change, with producer's accuracy of 98% and user's accuracies of 86% in the spatial domain and temporal accuracy of 80%. At the same time, the 16-categories land cover classification map from the CCDC algorithm also showed high accuracy with an overall accuracy of 90%.

The CCDC algorithm has potential for many applications. The first application is monitoring and assessing disasters. For instance, it can be applied for monitoring oil spills in the ocean. Just by labeling reference pixels of oil spills and clean ocean water at the time when oil spills occurred, it is possible to detect changes happened in ocean and classify them as oil spills or other kinds of changes. It can also be used for monitoring earthquakes and assessing their impacts. After the earthquake, the satellite signal of the Earth surface would certainly change and this change can be easily captured by the CCDC change detection component. By labeling pixels that have been affected by earthquake at different levels (serious, modest, and slight), the CCDC classification component is able to classify the levels of earthquake affected areas and this information will be very helpful for the decision makers as they can arrange rescues and supports based on the most recent earthquake influence map.

The second potential application is from the trend coefficient of the time series model. Most of this research is focusing on abrupt changes detected by the breaks in the time series model, but actually there are many kinds of gradual changes that are occurring over longer time periods. Take gradual changes in forests for example, the trend of the time series model may contains information such as forest growth, decline due to diseased forest, forest condition change because of drought, and so on. It is possible to generate a map of forest growth rate by using the magnitude of the trend coefficient from the time series model. By including the trend information into the classification component, this algorithm may be able to provide new land cover categories that are unique temporally, for example, the *forest* class can be separated into *growing forest*, *mature forest*, and *declining forest*. This kind of information is important for the study of vegetation health conditions and used for carbon modeling, but difficult to derive from the conventional land cover classification methods.

The third potential of this CCDC algorithm is that not only can it monitor changes in land cover, it can also detect changes caused by many other factors. For example, in forests, changes in species composition, tree density, forest succession, and background (rocks/soil) can also make satellite signals change. These kinds of change can be identified if it is larger than the CCDC threshold, though they may still be the same land cover type. If we have training data related to these kinds of change, the CCDC will have the ability to classify this non-categorical change which are also important in forest management and carbon modeling.

One area of future work will be improving cloud, cloud shadow, and snow detection algorithm. The first-step Fmask algorithm is quite simple, as it used a single global optimum cloud probability threshold and only Landsat data are used. In the future, instead of using the same cloud probability threshold, it can change this cloud probability threshold based on the location of the image, as different parts of world may have their local optimum threshold that can perform better than a simple global threshold. Moreover, ancillary data can be very helpful in improving the Fmask results. The most valuable ancillary data for Fmask would be Digital Elevation Model (DEM) and land/water mask. For example, the cloud probability threshold can change with the DEM, as the higher the altitude, the colder the temperature which may need higher cloud probability threshold to define true clouds. Land and water mask can be used to better identify clouds for land and water separately due to the fact that land and water have totally different characteristics in temperature and surface reflectance. The current Fmask algorithm detects clouds over land and water separately based on a land /water mask generated from the Landsat image itself. However, for areas that are covered with thick clouds, it would be impossible to know what is underneath the clouds. Therefore, accurate land and water mask can improve places that have large areas of water. Due to the complexity of the algorithm and the large amount of input data, the second-step has only been tested in two locations in this research. More tests of the multitemporal cloud, cloud shadow, and snow masking algorithm at other places are needed.

The potential to use the methods presented here for monitoring land cover will improve as the frequency of high resolution images from sensors like Landsat become more available (Arvidson et al., 2006). The opening of the Landsat archive, launch of the Landsat Data Continuity Mission (LDCM), and the launch of two sentinel 2A/2B satellites will greatly improve the availability of Landsat-like observations. Moreover, in this research the CCDC algorithm only applied to Landsat data, but it is also applicable to other sensors that have enough temporal frequency. For example, it may be able to apply to MODIS images that have more frequent observations and cover much larger areas. Considering the daily observations from MODIS sensor, it is possible to provide near real-time change detection and that can be very helpful for monitoring and responding to immediate problems. In tropical areas where cloud cover is high, time series radar data would be ideal for this CCDC algorithm as they have clouds penetration abilities. In the future, we can try to combine all possible dataset from different sensors into this CCDC system and this would definitely increase the detection accuracy and detect change events much faster.

This study is a "prototype" for continuous change detection and classification using all available Landsat data. The robustness of this approach has not been tested in other areas. Therefore, there is still much work needed. Expansion of the study to other regions will undoubtedly result in improvements to the approach because of the differences in land cover types. One future topic will be to broaden the variety of models used for the time series data. While combinations of sines and cosines worked well in this situation, there will be a need for other models in other locations. The new modeling paradigm in the statistical literature called Functional Data Analysis (FDA; Ramsay, 2005) does not assume any specific structural form or distribution for the data time series. Instead, it uses families of different "basis functions" to characterize the functional behavior of time series in a very flexible fashion. FDA methods have the additional advantage of being quite robust to missing data. That said, FDA methods are relatively new and are untested in remote sensing, and more traditional methods of time series analysis (e.g., autoregressive models, linear systems models, stat space models, panel data analysis techniques, etc.) may also be valuable for the analysis. There is no reason a whole family of models couldn't be tested for each pixels and the best model selected.

Though the data-driven threshold used in CCDC algorithm is able to handle many kinds of land cover change for the Boston scene, it may have problems for places that have large inter-annual variations. For some semi-arid areas, the green up of grass is highly dependent on the timing of the first rain, which can change significantly for different years. In this case, the data will fluctuate more at some times of the year and at this time a higher threshold is needed. In the future, the CCDC algorithm should be able to use a threshold that changes temporally to further improve the change detection accuracy.

Bibliography

Ackerman, S. A., Strabala, K. I., menzel, W. P., Frey, R. A., Moeller, C. C., & Gumley, L. E. (1998). Discriminating clear sky from clouds with MODIS. *Journal of Geophysical Research*, 103(24), 32141-32157.

Anderson, J. R. (1976). A land use and land cover classification system for use with remote sensor data (Vol. 964). *US Government Printing Office*.

Arvidson, T., Goward, S., Gasch, J., & Williams, D. (2006). Landsat-7 long-term acquisition plan: development and validation, *Photogrammetric Engineering & Remote Sensing*, 72(10), 1137-1146.

Arvidson, T., Gasch, J., & Goward, S. N. (2001). Landsat-7's long-term acquisition plan-an innovative approach to building a global imagery archive. *Remote Sensing of Environment*, 78(1-2), 13-26.

Asner, G. P. (2001). Cloud cover in Landsat observation of the Brazilian Amazon. *International Journal of Remote Sensing*, 22(18), 3855-3862.

Asner, G. P., Knapp, D. E., Balaji, A., & Páez-Acosta, G. (2009). Automated mapping of tropical deforestation and forest degradation: CLASlite. *Journal of Applied Remote Sensing*, 3, 033543.

Berendes, T., Sengupta, S. K., Welch, R. M., Wielicki, G. A., & Navar, M. (1992). Cumulus cloud base height estimation from high spatial resolution Landsat data: a Hough transform approach. IEEE Transactions, *Geoscience and Remote Sensing*, 30, 430-443.

Bolton, D. K., and Woodcock, C. E., Mapping past and present hemlock woolly adelgid change in the northeast US with a dense timeseries of Landsat data. *In preparation.*

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

Cohen, W. B., Fiorella, M., Gray, J., Helmer, E., & Anderson, K. (1998). An efficient and accurate method for mapping forest clear cuts in the Pacific Northwest using Landsat imagery. *Photogrammetric Engineering and Remote Sensing*, 64(4), 293-300.

Cohen, W. B., Yang, Z., & Kennedy, R. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync – Tools for calibration and validation. Remote Sensing of Envrionment, 114(12), 2911-2924.

Coiner, J.C. (1980). Using Landsat to monitor changes in vegetation cover induced by desertification processes. Symposium on *Remote Sensing of the Environment*, 3(14), 1341-1347.

Coppin, P. R., Bauer, M. E. (1994). Processing of multitemporal Landsat TM imagery to optimize extraction of forest cover change features. *IEEE Transactions on Geoscience and Remote Sensing* 32, 918-927.

Coppin, P. R., & Bauer, M. E. (1996). Digital change detection in forest ecosystems with remote sensing imagery. *Remote sensing reviews*, 13(3-4), 207-234.

Collins, J. B., & Woodcock, C. E. (1996). An assessment of several linear change detection techniques for mapping forest mortality using multitemporal Landsat TM data, *Remote Sensing of Environment*, 56:66-77.

Congalton, R. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*(1), 35-46.

Crist, E., & Cicone, R. C. (1984). A physically-based transformation of thematic mapper data – the TM Tasseled Cap. *IEEE Transactions on Geoscience and Remote Sensing*, 22, 256-263.

Crist, E. P. (1985). A TM tasseled cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, 17, 301−306.

Davis, J. C. (1986). Statistics and Data Analysis in Geology, Second Edition, *J. Wiley and Sons*, New York, N. Y., 646.

Derrien, M., Farki, B., Harang, L., LeGleau, H., Noyalet, A., Pochic, D., & Sairouni, A. (1993). Automatic cloud detection applied to NOAA-11/AVHRR imagery. *Remote Sensing of Environment*, 46(3), 246-26

Dixon, R., Brown, S., Houghton, R. A., Solomon, A. M., Trexler, M. C., & Wisniewski, J. (1994). Carbon pools and flux of global forest ecosystems.*Science*, 263(5144), 185-189.

Dozier, J. (1989). Spectral signature of alpine snow cover from the Landsat Thematic Mapper. *Remote Sensing of Environment*, *28*, 9–22.

DuMouchel, W. H., & O'Brien, F. L. (1989). Integrating a robust option into a multiple regression computing environment, *Computer Science and Statistics*: Proceedings of the 21st Symposium on the Interface, Alexandria, VA: American Statistical Association.

Eklundh, L., Johansson, T. & Solberg, S. (2009). Mapping insect defoliation in Scots pine with MODIS time-series data. *Remote Sensing of Environment*, 113(7):1566-1573.

Eli, K. M., Friedl, M. A., & Zhu, Z. submitted for publication. Monitoring interannual variation in deciduous broadleaf forest phenology using Landsat data. *Remote Sensing of Environment*.

Fielding, A. H., & Bell, A. J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24(1), 38-49.

Franklin, S. E., Moskal, L. M., Lavigne, M. B., & Pugh, K. (2000). Interpretation and classification of partially harvested forest stands in the Fundy model forest using multitemporal Landsat TM digital data. *Canadian Journal of Remote Sensing*, 26 (4), 318– 333.

Friedl, M. A., Menashe, D. S., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote sensing of Environment*, 114, 168-182.

Friedl, M., Woodcock, C.E., Gopal, S., Muchoney, D., Strahler, A.H., and C. Barker-Schaaf, 2000. A note on procedures used for accuracy assessment in land cover maps derived from AVHRR data, *International Journal of Remote Sensing*, 21(5):1073-1077.

Fuller, R. M., Smith, G. M., & Devereux, B. J. (2003). The characterisation and measurement of land cover change through remote sensing: problems in operational applications?. *International Journal of Applied Earth Observation and Geoinformation*, 4(3), 243-253.

Galford, G. L., Mustard, J. F., Melillo, J., Gendrin, A., Cerri, C. C., & Cerri, C. E. P. (2008). Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sensing of Environment*, 112(2):576-587.

Gao, B. C., & Kaufman, Y. J. (1995). Selection of the 1.375-μm MODIS channel for remote sensing of cirrus clouds and stratospheric aerosols from space. *American Meteorological Society*, 52, 4231-4237.

Gao, B. C., Kaufman, Y. J., Han, W., & Wiscombe, W. J. (1998). Correction of thin cirrus path radiance in the 0.4–1.0 μm spectral region using the sensitive 1.375 μmc cirrus detecting channel. *Journal of Geophysical Research*, 103(24), 32169–32176.

Gao, B. C., Yang, P., Han, W., L, R. R., & Wiscombe, W. J. (2002). An algorithm using visible and 1.38-μm channels to retrieve cirrus cloud reflectances from aircraft and satellite data, *IEEE Transactions Geoscience and Remote Sensing*, 40(8), 1659-1668.

Gao, F., Masek, J., Hall, J., & Schwaller, M. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 44(8), 2207-2219.

GarcÍa-Haro,F. J., Gilabert, M. A., & MeliÁ. (2001). Monitoring fire-affected areas using Thematic Mapper data. *International Journal of Remote Sensing*, 22(4), 533-549.

Gomez-Chova, L., Camps-Valls, G., Galpe-Maravilla, J., Guanter, L., & Moreno, J. (2007). Cloud-screening algorithm for ENVISAT/MERIS multispectral images. *Geoscience and Remote Sensing*, 45(12), 2, 4105-4118.

Goodwin, N. R., Coops, N. C., Wulder, M. A., Gillanders, Schroeder, T. A., & Nelson T. (2008). Estimation of insect infestation dynamics using a temporal sequence of Landsat data. *Remote Sensing of Environment*, 112, 3680-3689.

Gopal, S., Woodcock, C. E., & Strahler, A. (1999). Fuzzy neural classification of global land cover from a 1 degree AVHRR Data Set. *Remote Sensing of Environment*, 67, 230-243.

Goward, S., Arvidson, T., Williams, D., Faundeen, J., Irons, J., & Franks, S. (2006). Historical record of Landsat global coverage: mission operations, NSLRSDA, and international cooperator stations, *Photogrammetric Engineering & Remote Sensing*, 72(10), 1155-1169.

Goward, S. N., Masek, J. G., Cohen, W., Moisen, G., Collatz, G. J., Healey, S., et al. (2008). *Forest disturbance and North American carbon flux.* Eos, 89(11).

Guerschman, J. P., Paruelo, J. M., Bella, C. D. I., Giallorenzi, M. C., & Pacin, F. (2003). Land cover classification in the Argentine Pampas using multi-temporal Landsat TM data. *International Journal of Remote Sensing*, 24(17), 3381-3402.

Hall, D. K., Riggs, G. A., & Salomonson. (2001). Algorithm theoretical basis document (ATBD) for the MODIS snow and sea ice-mapping algorithms.

Hansen, M. C., Defries, R. S., Townshend, J. R. G. and Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21(6 & 7), 1331-1364.

Hartmann, D., L. (1994). Global physical climatology. (1st ed.). San Diego: Academic Press, (Chapter 3).

Hayes, D. J., & Sader, S. A. (2001). Comparison of change-detection techniques for monitoring tropical forest clearing and vegetation regrowth a time series. *Photogrammetric Engineering & Remote Sensing*, 67(9), 1067-1075.

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

Healey, S. P., Yang, Z., Cohen, W. B., & Pierce, D. J. (2006). Application of two regression-based methods to estimate the effects of partial harvest on forest structure using Landsat data. *Remote Sensing of Environment*, 101(2006), 115-126.

Hégarat-Mascle, S. L., & André, C. (2009). Use of Markov Random Fields for automatic cloud/shadow detection on high resolution optical images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(2009), 351-366

Hilker, T., Wulder, M. A., Coops, N. C. Linke, J., McDermind, 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.

Holland, P. W., and Welsch R. E. (1977). Robust regression using iteratively reweighted Least-Squares, *Theory and Methods*, 813-827.

Hostert, P., Röder, Achim, & Hill, J. (2003). Coupling spectral unmixing and trend analysis for monitoring of long-term vegetation dynamics in Mediterranean rangelands. *Remote Sensing of Environment*, 87, 183-197.

Howard, S.M., Ohlen, D.O., McKinley, R.A., Zhu, Z. Kitchen, J. (2002). Historical fire severity mapping from Landsat data. Pecora 15/Land Satellite Information IV/ ISPRS Commission I Symposium, Proceedings. *American Society for Photogrammetry and Remote Sensing.*

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.

Hutchison, K. D., Mahoney, R. L., Vermote, E. F., Kopp, T. J., Jackson, J. M., Sei, A., & Iisager, B. D., 2009, A geometry-based approach to identifying cloud shadows in the

VIIRS cloud mask algorithm for NPOESS, *Journal of Atmospheric and Oceanic Technology*, 26(7), 1388-1397

Irish, R. (2000). Landsat-7 automatic cloud cover assessment algorithms for multispectral, hyperspectral, and ultraspectral imagery. *The International Society for Optical Engineering*, 4049, 348-355.

Irish, R., Barker, J. L., Goward, S. N., & Arvidson, T. (2006). Characterization of the Landsat-7 ETM+ Automated Cloud-Cover Assessment (ACCA) algorithm. *Photogrammetric Engineering & Remote Sensing*, 72(10), 1179-1188.

Jensen J. R., Rutchey, K., Koch, M. S., & Narumalani. (1995). Inland wetland change detection in the everglades water conservation area 2A using a time series of normalized remotely sensed data. *Photogrammetric Engineering & Remote Sensing*, 61(2), 199-209.

Jin, S., & Sader S. A. (2005). MODIS time-series imagery for forest disturbance detection and quantification of patch size effects. *Remote Sensing of Environment*, 99(4):462-470.

Jung, M., Henkel, K., Herold, M., & Churkina, G. (2006). Exploiting synergies of global land cover products for carbon cycle modeling, *Remote Sensing of Environment*, 101(4), 534-553.

Kaufmann, R. K., & Seto, K. C. (2001). Change detection, accuracy, and bias in a sequential analysis of Landsat imagery in the Pearl River Delta, China: econometric techniques. *Agriculture, ecosystems & environment*, 85(1), 95-105.

Kennedy, R. E., Cohen, W. B., & Schroeder, T. A. (2007). Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sensing of Environment*, 110, 370-386

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr – Temporal segmentation algorithms. *Remote Sensing of Environment* (in press).

Lambin, E. F & Strahler, A. H. (1994). Change-vector analysis in multitemporal space: a tool to detect and categorize land-cover change processes using high temporal-resolution satellite data. *Remote Sensing of Environment*, 48:231-244.

Lark, R. M. & Stafford, J. V. (1997). Classification as a first step in the interpretation of temporal and spatial variation of crop yield. *Annals of Applied Biology*, 130(1), 111-121.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L. W. M. J., & Merchant, J. W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data.*International Journal of Remote Sensing*, 21(6-7), 1303-1330.

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

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

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

Masek, J. G., Vermote, E. F., Saleous, N.,Wolfe, R., Hall, E. F., Huemmrich, F., et al. (2006). A Landsat surface reflectance data set for North America, 1990–2000. *Geoscience and Remote Sensing Letters*, 3, 68−72.

O'Leary, D. P. (1990). Robust regression computation using iteratively reweighted least squares, *Society for Industrial and Applied Mathematics*, 11(3), 466-480Oreopoulos, L., Wilson, M., & Vármai, T. (2011). Implementation on Landsat data of a simple cloud mask algorithm developed for MODIS land bands. *IEEE Transactions on Geoscience and Remote Sensing*, 8(4), 597-601.

Pearce, W. A. (1985). Cloud shadow effects on remote sensing, *Geoscience and Remote Sensing*, 23(5), 634-639.

Pflugmacher, D., Cohen, W. B., & E Kennedy, R. (2012). Using Landsat-derived disturbance history (1972–2010) to predict current forest structure. *Remote Sensing of Environment*, 122, 146-165

Ramsay, J. (2005). *Functional data analysis*. John Wiley & Sons, Ltd.

Rayner, J. N. (1971). An Introduction to Spectral Analysis, *Pion Ltd.*, London, 174.

Riggs, G. and Hall, D. K., 2002, Reduction of cloud obscuration in the MODIS snow data product, 59th EASTERN SNOW CONFERENCE.

Rogan, J., Bumbarger, N., Kulakowski, D., Christman, Z. J., Runfola, D. M., & Blanchard, S. D. (2010). Improving forest type discrimination with mixed lifeform classes using fuzzy classification thresholds informed by field observations. *Canadian Journal of Remote Sensing*, 36(6), 699-708.

Rogan, J., Franklin, J., & Roberts, D. A. (2002). A comparison of methods for monitoring multitemporal vegetation change using Thematic Mapper imagery. *Remote Sensing of Environment*, 80, 143-156.

Roy, D. P., Jin, Y., Lewis, P. E., & Justice, C. O. (2005). Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data. *Remote Sensing of Environment*, 97(2):137-162.

Roy, D. P., Ju, J., Lewis, P., Schaaf, C., Gao, F., Hansen, M., & Lindquist, E. (2008). Multi-temporal MODIS-Landsat data fusion for relative radiometric normalization gap filling, and prediction of Landsat data. *Remote Sensing of Environment*, 112(3), 3112-3130.

Saunders, R. W., & Kriebel, K. T. (1998). An improved method for detecting clear sky and cloudy radiances from AVHRR data. *International Journal of Remote Sensing*, 9(1), 123-150.

Scaramuzza, P. (Nov. 2010). United States Geological Survey Earth Resources Observation Systems (EROS) Data Center (EDC), Sioux Falls, SD, *private communication*.Seto, K.C., Woodcock, C.E., Song, C., Huang, X., Lu, J., &

Kaufmann, R.K. (2002). Monitoring land-use change in the Pearl River Delta using Landsat TM. *International Journal of Remote Sensing*, 23(10):1985-2004.

Seto, K.C., Woodcock, C.E., Song, C., Huang, X., Lu, J., & Kaufmann, R.K. (2002). Monitoring land-use change in the Pearl River Delta using Landsat TM. *International Journal of Remote Sensing*, 23(10):1985-2004.

Simpson, J. J., & Stitt, J. R. (1998). A procedure for the detection and removal of cloud shadow from AVHRR data over land, *Geoscience and Remote Sensing*, 36(3), 880-890.

Simpson, J. J., Stitt, J. R., & Jin, Z. (2000). Cloud shadow detection under arbitrary viewing and illumination conditions. *IEEE Transactions Geoscience and Remote Sensing,* 36, 880-897

Soille, P. (1999). Morphological image analysis: principles and applications, *Springer-Verlag*, 173-174.

Soille, Pierre, Vogt, Jugen, and C. R. (2003). Carving and adaptive drainage enforcement of grid digital elevation models. *Water Resources Research*, *39*(12), 1-13.

Street, J. O., Carroll, R. J. & Ruppert, D. (1988). A note on computing robust regression estimates via iteratively reweighted least squares, *American Statistical Association*, 42(2), 152-154.

Takao, G. (2003). Indices from snow-covered Landsat imagery for boreal forest change detection. *Journal of the Japan Society of Photogrammetry and Remote Sensing*, 42(6), 52-66.

Tan, B., Woodcock, C. E., Hu, J., Zhang, P., Ozdogan, M., & Huang, D. et al. (2006). The impact of gridding artifacts on the local spatial properties of MODIS data: implications for validation, compositing, and band-to-band registration across resolution. *Remote Sensing of Environment*, 105(2), 98-114.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environments*, 8, 127-150.

Tucker, C. J., Townshend, J. R.G., & Goff, T. E. (1985). African land-cover classification using satellite data. *Science*, 227(4685), 369-375.

Turner, D. P., Guzy, M., Lefsky, M. A., Ritts, W. D., Van Tuyl, S., & Law, B. E. (2004). Monitoring forest carbon sequestration with remote sensing and carbon cycle modeling. *Environmental Management*, 33(4), 457-466.

Verbesselt, J., Hyndman, R., Newnham, G., & Culvenor, D. (2010). Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment*, 114(1):106-115.

Vermote, E. (Jan. 2010). *Landsat Science Team Meeting*. Mountain View, California, USA.

Vermote, E. F., El Saleous, N., Justice, C. O., Kaufman, Y. J., Privette, J. L., et al. (1997). Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: Background, operational algorithm, and validation. *Journal of Geophysical Research*, 102, 17131−17141.

Vermote, E. & Saleous, N. (2007). LEDAPS surface reflectance product description, Version 2.0.

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

Wang, B., Ono, A., Muramatsu, K., & Fujiwara, N. (1999). Automated detection and removal of clouds and their shadows from Landsat TM images, *Ieice Trans. Inf. & Syst.*, E82-D, 2.

Wildt, M. D. R. D., Seiz, G., & Gruen, A. (2007). Operational snow mapping using multitemporal Meteosat SEVIRI imagery, *Remote Sensing of Environment*, 109, 29-41.

Wolter, P. I, Mladenoft, D. S., & Crow, T., R. (1995). Improved forest classification in the northern lake states using multi-temporal Landsat imagery. *Photogrammetry Engineering & Remote Sensing*, 61(9), 1129-1143

Woodcock, C.E., Macomber, S.A., Pax-Lenney, M., & Cohen, W.B. (2001). Large area monitoring of temperate forest change using Landsat data: Generalization across sensors, time and space. *Remote Sensing of Environment*, 78(1-2):194-203.

Woodcock, C. E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., et al. (2008). Free access to Landsat imagery. *Science*, 320(5879), 1011-1011.

Wulder, M. A., White, J. C., Goward, S. N., Masek, J. G., Irons, J. R., Herold, M., et al. (2008). Landsat continuity: Issues and opportunities for land cover monitoring. Remote Sensing of Environment, 112(3), 955-969.

Zhan, X., Sohlberg, R. A., Townshend, J. R. G., DiMiceli, C., Carroll, M. L., Eastman, J. C., ... & DeFries, R. S. (2002). Detection of land cover changes using MODIS 250 m data. *Remote Sensing of Environment*, 83(1), 336-350.

Zhang, Y., Rossow, W. B., Lacis, A. A., Oinas, V., & Mishchenko, M. I. (2004). Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data. *Journal of Geophysical Research*, 109, 19, 19105

Zhang, Y., Guindon, B., Cihlar, J. (2002). An image transform to change characterize and compensate for spatial variability in thin cloud contamination of Landsat images. *Remote Sensing of Environment*, 82(2-3), 173-187.

Zhu, Z, & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118 (15), 83-94.

Zhu, Z. & Woodcock, C. E., Continuous change detection and classification of land cover using all available Landsat data. *In preparation.*

Zhu, Z., Woodcock, C. E., & Olofsson, P. (2012a). Continuous monitoring of forest disturbance using all available Landsat imagery. *Remote Sensing of Environment*, 122, 75-91.

Zhu, Z., Woodcock, C. E., & Olofsson, P., Automated cloud and cloud shadow screening based on multi-temporal Landsat imagery. *In preparation.*

Zhu, Z., Woodcock, C. E., Rogan, J., & Kellndorfer, J. (2012b). Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peri-urban land cover classification using Landsat and SAR data. *Remote Sensing of Environment*, 117, 72-82.

Curriculum Vitae

[redacted]

[redacted]

[redacted]