MAPPING AND SPATIAL-TEMPORAL MODELING OF BROMUS TECTORUM INVASION IN CENTRAL UTAH

by

Zhenyu Jin

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Department of Geography

The University of Utah

December 2011

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Zhenyu Jin**

has been approved by the following supervisory committee members:

Richard R. Forster	, Chair	**28 July 2011** Date Approved
Bing Xu	, Co-Chair	**25 Oct 2011** Date Approved
George F. Hepner	, Member	**27 Oct 2011** Date Approved
Peng Gong	, Member	**28 Oct 2011** Date Approved
Zhiliang Zhu	, Member	**25 Oct 2011** Date Approved

and by **George F. Hepner**, Chair of the Department of **Geography**

and by Charles A. Wight, Dean of The Graduate School.

ABSTRACT

Cheatgrass, or Downy Brome, is an exotic winter annual weed native to the Mediterranean region. Since its introduction to the U.S., it has become a significant weed and aggressive invader of sagebrush, pinion-juniper, and other shrub communities, where it can completely out-compete native grasses and shrubs.

In this research, remotely sensed data combined with field collected data are used to investigate the distribution of the cheatgrass in Central Utah, to characterize the trend of the NDVI time-series of cheatgrass, and to construct a spatially explicit population-based model to simulate the spatial-temporal dynamics of the cheatgrass.

This research proposes a method for mapping the canopy closure of invasive species using remotely sensed data acquired at different dates. Different invasive species have their own distinguished phenologies and the satellite images in different dates could be used to capture the phenology. The results of cheatgrass abundance prediction have a good fit with the field data for both linear regression and regression tree models, although the regression tree model has better performance than the linear regression model.

To characterize the trend of NDVI time-series of cheatgrass, a novel smoothing algorithm named RMMEH is presented in this research to overcome some drawbacks of many other algorithms. By comparing the performance of RMMEH in smoothing a 16-day composite of the MODIS NDVI time-series with that of two other methods, which are the 4253EH, twice and the MVI, we have found that RMMEH not only keeps the

original valid NDVI points, but also effectively removes the spurious spikes. The reconstructed NDVI time-series of different land covers are of higher quality and have smoother temporal trend.

To simulate the spatial-temporal dynamics of cheatgrass, a spatially explicit population-based model is built applying remotely sensed data. The comparison between the model output and the ground truth of cheatgrass closure demonstrates that the model could successfully simulate the spatial-temporal dynamics of cheatgrass in a simple cheatgrass-dominant environment. The simulation of the functional response of different prescribed fire rates also shows that this model is helpful to answer management questions like, “What are the effects of prescribed fire to invasive species?” It demonstrates that a medium fire rate of 10% can successfully prevent cheatgrass invasion.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my supervisor, Dr. Bing Xu, for her encouragement, inspiration, direction, and help during my studies. I appreciated her constant availability, patience, and valuable advice. Her wide knowledge and her logical way of thinking have been of great value for me. Her understanding, encouragement, and personal guidance have provided a good basis for the present thesis. This thesis could not have been finished without her.

I am also thankful to Drs. George Hepner, Rick Forster, Peng Gong, and Zhiliang Zhu for serving as members of my committee. Their encouragement helped me realize the significance of my work, and added to my enjoyment of the process of scientific discovery.

Most importantly, I owe my loving thanks to my wife Jing Sun. She has lost a lot due to my research taking me out of town. Without her encouragement and understanding, it would have been impossible for me to finish this work.

This work is dedicated to my lovely son, Daniel Jin, and my parents.

CHAPTER 1

INTRODUCTION

Background

The invasion of plants poses significant threats to both natural and human managed ecosystems and causes huge ecological and economical consequences. It is estimated that the cost of removing invasive weeds on the rangelands in the United States is approximately 2 billion dollars annually (Ditomaso, 2000). Among more than 2000 nonnative invasive species that exist in the United States, *Bromus tectorum* causes the most severe problems in the Intermountain West (Peterson, 2005). It is estimated that its area of invasion reaches at least 40 billion hectares in the United States (Ditomaso, 2000).

Cheatgrass, or Downy Brome (*Bromus tectorum L.*), is an exotic winter annual weed that is native to the Mediterranean region (Leger, 2008). It may also act as a spring annual if too little fall moisture is available for germination (Ferrero et al., 2010). Introduced into the U. S. in packing materials and possibly as contaminant of crop seeds, cheatgrass was first found in Denver, Colorado, in the late 1800s (Leger, 2008). It spread far and wide by train, excessive livestock grazing, and wildlife. By the early 1900s, cheatgrass was present in much of its current range, though it was sparsely distributed (Mitich, 1999).

Early infestation of cheatgrass is usually found in wheat cropland and railroad. Once introduced, cheatgrass will spread into adjacent areas and will adapt to the local environments. As a result, the ecosystem was invaded and seriously altered. The expansion of cheatgrass was especially rapid in parts of the Intermountain West, from the Rockies to the Cascades and Sierra Nevada, north from central Utah, Nevada, and northeastern California to Canada. It is obvious that native plants will not evolve with this heavy grazing pressure (Mitich, 1999). In these areas, it became a significant weed and an aggressive invader of sagebrush, pinion-juniper, and other shrub communities. By the 1930s, cheatgrass had become the dominant grass in the Pacific Northwest and the Intermountain area, and the worst Western weed. It is estimated that the area of invasion reaches over 41 million hectares (101 million acres) in the Western states (Mitich, 1999). Approximately 5 million hectares of overgrazed rangeland in Idaho and Utah are covered with almost pure stands of cheatgrass (FICMNEW, 1997). Now widely distributed throughout North America (Whitson et al., 1991), it commonly grows along roadsides, waste areas, pastures, rangelands, and croplands.

Studies have demonstrated that the growth and expansion of cheatgrass is closely related to both environmental conditions (temperature, elevation, soil characteristic, etc.) and anthropogenic disturbance (cultivation, grazing, fire, etc.) (Pierson and Mack, 1990a, b; Rice et al., 1992). Therefore, perennial environmental variation results in considerable fluctuation in population and its attributes, such as recruitment, survivorship, and fecundity. Cheatgrass will grow in almost any type of soil, but it grows best in deep, loamy, or coarse-textured soils (Mack, 1981). The amount of cheatgrass growth commonly depends on the amount and timing of moisture received, varying widely from

year to year (Young, 2000). The seed production varies with plant density, time of germination, and environmental conditions (Piemeisel, 1951). Cultivation and subsequent land abandonment, excessive livestock grazing, overstory removal, and repeated fires can interact to proliferate cheatgrass (Bates et al., 1998). Cheatgrass can maintain dominance for many years on sites where native vegetation has been eliminated or severely reduced by grazing, cultivation, or fire (Concannon, 1978). Once cheatgrass is established, complete protection from grazing or other disturbances will not usually reduce cheatgrass abundance. Such invasions have had profound negative consequences on native species, which will be replaced. It has altered soil food webs, decomposition cycles, and soil nutrient availability (Upadhyaya et al., 1986).

Due to its invasion and dominance, recent extensive wildfires in the Great Basin have increased in frequently and extensively, due to the fact that cheatgrass ignites and burns easily. These cheatgrass-induced fires change native vegetation pattern and structure, and play an important role in postfire community succession. Surviving in unburned organic materials on a site, after fire releases some available nitrogen, cheatgrass utilizes this nutrition before perennial grasses and shrubs can use it. And the rapid growth and reproduction make it a pioneering species that will dominate the postfire land. This wildfire cycle prevents the reestablishment of other native species (Novak, 2001) and makes cheatgrass to be more widely distributed in its potential habitats, thus reducing the native plant biodiversities and increasing soil erosion.

Problem Statement

Combating this cheatgrass-induced debilitation of Western ecosystems will be aided by an improved understanding of the geographic distribution. At the same time, the spatial-temporal invasion of cheatgrass under different environmental conditions, management policies, and disturbance should also be investigated so that we can predict the long-term spread of cheatgrass. So far, there has been a lack of studies that could address the aforementioned issues. Essentially, my research will seek for answers to the following three key research questions about science, technique, and management:

1. Can we use remote sensing data to detect the abundance of cheatgrass? Can we use remote sensing data to extract invasive species' phenological curve.
2. What kind of spatial-temporal models are suitable for modeling cheatgrass spatial-temporal dynamics? How can we use existing remote sensing and collected field data to predict the spatial-temporal dynamics of cheatgrass invasion?
3. What are the effects of management to the spatial-temporal dynamics? What kind of prescribed fires is best for controlling or managing the cheatgrass invasion?

Objectives of Research

In relation to these three questions, three specific interrelated research objectives are established:

1. To develop and test techniques for cheatgrass abundance mapping using high resolution Landsat TM and/or ETM+ imagery with field data and ancillary data from LANDFIRE geospatial data sets.

2. To develop a technique to smoothen the 250 meter resolution MODIS NDVI time-series to investigate the difference of an NDVI curve between cheatgrass and other vegetation.
3. To develop a spatially explicit population-based model that can be used to project the cheatgrass infestation, using remote sensing data. This model will be used to investigate the spatial-temporal dynamics of cheatgrass under different conditions of prescribed fire.

In essence, the first two objectives attempt to determine the current extent of cheatgrass distribution and the characteristic of the cheatgrass VI time-series curve. The Landsat TM/ETM+ data of 30-meter resolution used in this research to map the percent cover of cheatgrass can supply a high resolution product, while the MODIS Vegetation Index can be used to produce a low resolution cheatgrass distribution product. The third objective addresses the spatial-temporal dynamics of cheatgrass invasion and answers the question of the effect of management on cheatgrass invasion; the traditional logistical population model in biology will be coupled with spatially explicit GIS data and the plant spatial dispersal mechanism to model the whole process of cheatgrass invasion.

Significance of Research

This research focuses on addressing questions of cheatgrass abundance prediction and spatial-temporal modeling; it contributes to the field of ecology, remote sensing, and GIS in the summary provided below. Chapters 2-4 describe the details of the findings:

1. This research proposes a method for predicting closure cover of invasive species using remote sensing data of different dates. Since different invasive

species have their own distinguished phenologies, the satellite images in different dates could be used to capture the phenology. Although only one invasive species, cheatgrass, is mapped in this research, this method could be easily expanded in application of mapping other invasive species with specific phenology based on the same strategy.

2. This research proposes a novel and fast smoothing technique to reconstruct the noisy MODIS NDVI time-series data. Compared with other widely used smoothing techniques, this new technique is very simple in theory and easy to implement, and it is very resistant to many different noises.

3. This research proposes a spatially-explicit population-based model to predict the spatial-temporal change of cheatgrass. It sets an example of extracting unknown parameters using constrained Monte-Carlo simulation. A scheme of investigating the effects of management is also developed; this model could be applicable to other studies involving monitoring the spread of invasive species.

Research Responsibilities

Although this research is a collaborative project between the University of Utah and UC Berkeley, the focuses for our responsibilities are different. The UC Berkeley team is responsible for 1) developing an atmospherically corrected model for Landsat TM images based on the parameter information from MODIS water vapor data, and 2) developing a meta-prediction data mining technique to predict the cheatgrass percentage using the atmospherically corrected TM images. Therefore, the UC Berkeley team focused more on the technical remote sensing aspects of the project. For example, a

component of the research they led is published in Photogrammetric Engineering & Remote Sensing (Clintion et al., 2009).

My responsibility mainly includes 1) use and comparison of linear and nonlinear regression analysis to predict cheatgrass with atmospherically and topographically corrected Landsat images (different techniques to remove atmospheric and topographic influence are used between our researches), 2) development of a combined Landsat MODIS VI time-series data smoothing technique, and 3) spatial-temporal modeling of cheatgrass concentration.

There are also some cooperative works between us. We took field trips together during June and July of 2005 and 2006 to collect field data of cheatgrass percentage. We discussed what imageries we should use in this research and then ordered the Landsat TM imageries listed in Chapter 2 from EROS and USGS. All the literature or reference about cheatgrass found from each group was sent to the other team, and was thus shared by both.

CHAPTER 2

CHEATGRASS ABUNDANCE PREDICTION USING TM DATA

Introduction

Early detection or prediction of the distribution of invasive weeds to reduce cost or improve treatment is essential to range management (Ditomaso, 2000). Field surveys of the distribution of invasive weeds can be extremely labor intensive and inefficient, and limited in space and time. Remote sensing techniques provides us the alternative to traditional ground surveys, in that they have potential to monitor and model the spread of invasive weeds in a large spatial and temporal scale. Since all the Earth's surface objects, like vegetation, have their own unique reflected and emitted spectral properties, they can be captured by the remote sensors and identified in remote sensing images according to these different spectral features; this technique is the physical basis of vegetation mapping using remote sensing.

For some weeds, remote sensing may provide a cheaper, more efficient method for mapping infestations than ground surveys. However, many weeds are not good targets for remote sensing because they are indistinguishable from other native plants, particularly during vegetative growth. Therefore, it is critical to assess the likelihood of adequate detection before initiating a remote sensing-based mapping project. In general, different weed species have their own special phenologies, such as bloom, peak, or

senescence. These phenological differences can usually be discriminated on the corresponding remote sensing data since the weed species in different phenology will show different color, which means the spectral differences for remote sensing data. By exploiting the spectral differences, we can use remote sensing to identify the specific weed species. Many authors have investigated the relationships between plant phenology and remote sensing data, which could be utilized for species identification. For example, Lopez-Granados et al. (2006) used remote sensing to successfully discriminate grass weed infestations from wheat, and suggested that mapping grass weed patches in wheat is feasible with high-resolution satellite imagery or aerial photography acquired 2 to 3 weeks before crop senescence. Likewise, Van Wagtendonk et al. (2001) used the Landsat TM Normalized Difference Vegetation Index (NDVI) time-series to detect the fuel information.

Cheatgrass has a distinguished biological trait in that it greens up early in the spring and senesces before other grasses, making it a good target for remote sensing at such times. Figure 2.1 and Figure 2.2 show how the reflectance of cheatgrass will show in Landsat TM images at two different seasons: green-up and senesce season.

It is clearly shown in Figure 2.1 that the reflectance of cheatgrass during green-up season is a standard vegetation reflectance curve, while the reflectance in senesce season (Figure 2.2) seems like a soil reflectance curve. If one single acquisition of Landsat TM image is used to predict the abundance of cheatgrass, the result will surely be over-predicted (including additional green vegetation cover information in green-up season and soil cover information in senesce season). So, to succeed in mapping cheatgrass, two acquisitions of imagery are necessary for producing superior results, compared to a single

acquisition. Peterson (2005) and Noujdina et al. (2008) both supported this point in their research in mapping cheatgrass percent cover using ETM+ and AVIRIS data.

Since cheatgrass usually mixes spectrally and structurally with such native plants as pinion-juniper and sagebrush in a Landsat TM/ETM+ pixel, it is difficult and inaccurate to apply traditional hard classification methods, such as K-means clustering, Minimum Distance Classification, Maximum Likelihood Classification (MLC), and Spectral Angle Mapper (SAM). Contrary to the "hard labeling" classification is the "soft labeling" classification (Chen, 1999). The three major "soft labeling" classification methods, which attempt to use remotely sensed data to quantify vegetation cover, are physically based models, spectral mixture analysis (SMA), and empirical models.

Physically based models are based on radiative transfer models (Li, 1992); such models are often complicated and generally immature for use in regional land cover applications (Kime, 1998). SMA is a technique that allows one to determine the fraction of each land cover type within a pixel (Adam, 1986); it assumes that several pure land cover classes (end-members) within a pixel linearly mix their spectra, with the intensity proportional to their surface area fractions. Although SMA has been successfully applied to predict fraction of green vegetation cover (Elmore, 2000; Xiao, 2005), its use is limited by two aspects. First, the physically based assumption of mixing linearity was found invalid in many studies (Borel, 1994; Ray, 1996). Second, spectral end-members identified in SMA, such as green vegetation, nonphotosynthetic vegetation, and shadow (Roberts, 1993), do not correspond to specific land cover components such as cheatgrass. Additional efforts are required to estimate the proportion of relevant land cover components (Adam, 1995).

While both physically based models and SMA depend on physical knowledge in the radiative transfer process, empirical models do not need such knowledge and have the advantage of giving results that are easy to interpret (Huang, 2003). Empirical models, such as linear and nonlinear regression models, and neural networks, are common strategies for predicting biophysical variables from remotely sensed data. In the case of predicting vegetation cover, an empirical model is firstly developed based on training data with known spectral reflectance and actual field vegetation cover. Once the model is validated, it is then used to predict vegetation cover for each pixel throughout the whole image.

Among the empirical models, regression models are most commonly used to predict biophysical parameters, such as the Leaf Area Index (LAI), Crown Closure, Tree Age, and Vegetation Cover. Both direct (Butera, 1986; Larsson, 1993; Van, 1995; Xu, 2003; Peterson, 2005) and inverse regression models (Curran, 1986; Chen, 1996; Eastwood, 1997; Elmore, 2000; Mcmorrow, 2001) have been widely applied. Direct models have the advantage of incorporating more than one independent variable or spectral band. The inverse models, which take one spectral band or index as dependent variable, are unable to fully utilize multiple spectral dimensions of satellite images. However, the models do not violate the assumptions that spectral data collected are free of error and therefore acknowledge the inherent errors in remotely sensed data (Curran, 1986).

Vegetation Index (VI), a unitless measurement calculated from linear or nonlinear combination of red and near infrared bands of remotely sensed data, is sensitive to the green vegetation, and is thus incorporated in regression models to investigate the

vegetation properties. So far, a lot of VIs have been developed and investigated, such as the Simple Ratio Vegetation Index (SR) (Jordan, 1969), Normalized Difference Vegetation Index (NDVI) (Tucker, 1979), Perpendicular Vegetation Index (PVI) (Richardson, 1977; Baret, 1991), Soil Adjusted Vegetation Index (SAVI) (Huete, 1988), Transformed Soil Adjusted Vegetation Index (TSAVI) (Baret, 1991), Modified Soil Adjusted Vegetation Index (MSAVI) (Qi et al., 1994), Global Environmental Monitoring Index (Pinty, 1992), and Enhanced Vegetation Index (EVI) (Huete, 1997). Among these, SR and NDVI are the most widely used VIs. However, these two are sensitive to both soil background and atmospheric effects, while others lose sensitivity after reducing the influence of soil background and atmospheric effects.

Because of its simplicity and close relationship to green vegetation, VI has been intensively utilized by scientists to measure different vegetation cover from satellite data (Larsson, 1993; Eastwood, 1997; Purevdori, 1998; Huang, 2003). In these studies, VIs calculated from single-date reflectance data are used to correlate with field estimates of vegetation cover; then, regression analysis is used to predict these variables from the VIs. Previous research has shown that different VIs are appropriate for different vegetation cover estimation. Larsson (1993) showed that NDVI had the highest correlation with field-measured canopy cover of Acacia woodland; Eastwood (1997) compared different VIs for monitoring saltmarsh and concluded that MSAVI and GEMI were the best indices for quantitative mapping of saltmarsh vegetation cover; Purevdorj (1998) compared NDVI, SAVI, MSAVI, and TSAVI in estimating the green vegetation cover of grasslands using a second-order polynomial regression model, and demonstrated that TSAVI and NDVI gave the best estimates for a wide range of grass densities. Huang

(2003) developed a stepwise regression tree to estimate the subpixel forest cover and found that NDVI is the root and most important node in the regression tree.

One disadvantage of using single-date VIs to predict the vegetation cover is that they have not fully utilized the spectral dimensionality of remotely sensed data. Cohen (2003) argued that including VIs from multiple dates in regression analysis could be a significant strategic improvement over existing uses of regression analysis in remote sensing. DeFries (1997) found the strongest correlations between the percent forest cover and the mean annual NDVI of 30 multitemporal metrics derived from AVHRR data. Peterson's (2005) research supported this argument; he obtained Landsat ETM+ images for the same area on two different dates to capture the unusual early phenology of cheatgrass, and then found the difference of NDVI to be the most important factor in the tobit regression model. Bradley (2005, 2006), although he did not use the regression model, used multiple phenology-based mapping methods to predict the distribution of cheatgrass-dominated areas based on differences of NDVI in two different years; but, his method could only predict the presence or nonpresence of cheatgrass and could not quantify the percent of cheatgrass within every pixel. Noujdina et al. (2008) found that the high spatial and hyperspectral AVRIS data acquired in different seasons are more effective for detection of cheatgrass invasion, compared with single-date datasets.

Although VIs have shown a reasonable and significant correlation with vegetation cover in most of the above research, there have not been many attempts to evaluate which VI and which empirical model are most appropriate for cheatgrass mapping. Peterson (2005) mapped the percent cover of cheatgrass across Nevada using a direct tobit regression model from landsat ETM+ images. He tested NDVI and one direct regression

model in his research. There is no extensive comparison of the various empirical models for mapping cheatgrass abundance in the literature, therefore, there is a need to compare and evaluate different empirical models and VIs in mapping percent cover of cheatgrass. A new vegetation index which is suitable for mapping cheatgrass should also be investigated and constructed in future research. The problem of tradeoffs between spatial and temporal resolution of remote sensing data used to map cheatgrass also needs to be investigated.

Research Area and Data Sets

The research area is located in the central highland ecoregion of Utah covered by USGS map zone 16, which is spanned by seven Landsat scenes (Figure 2.3). The arid and semiarid climate of this region is continental with hot dry summers and cold winters; it consists of the 9,000-10,000 foot Wasatch Mountains on the north, the Uinta Mountains to the northeast, and a group of high plateaus that range from Spanish Fork and Price canyons toward the southwest corner of Utah. This region consists of several ecosystem types, including arid desert shrub communities, high elevation Douglas-fir forests, pinion-juniper woodland, and open grassland. Most of the annual precipitation in this region occurs as snowfall during winter months. The elevation of this region ranges from 500 to 4,000 meters.

Three main data sets are used for this research. First, because of the surface heterogeneity (cheatgrass density changes), fine-resolution Landsat-5 TM remote sensing data are used as a basis for cheatgrass mapping and analysis activities. Landsat-5 TM has provided repetitive, synoptic, global coverage with high-resolution multispectral imagery

observing the same 185 km ground swath every 16 days since 1984. A Landsat-5 TM scene has an instantaneous field of view (IFOV) of 30 meters by 30 meters (900 square meters) in bands 1 through 5 and band 7, and an IFOV of 120 meters by 120 meters (14,400 square meters) on the ground in band 6 (thermal band). The spectral resolution for the TM sensor is shown in Table 2.1. The characteristics of the TM bands were selected to maximize detection and the monitoring of different types of the Earth's resources. For example, TM band 2 is designed to detect green reflectance from healthy vegetation; TM band 3 is designed for detecting chlorophyll absorption in vegetation; TM band 4 data are ideal for detecting near-IR reflectance peaks in healthy green vegetation and for detecting water-land interfaces. These three bands are usually combined to make false-color composite images where band 4 displays the red, band 3 displays the green, and band 2 displays the blue portions of the electromagnetic spectrum. In such false-color composite images, the vegetation appears as shades of red; the soils show from white (sands) to greens or browns, depending on the moisture and organic matter content; water bodies will appear blue or black, depending on the sediment content; urban areas appear blue-gray and clouds and snow appear bright white. Clouds and snow are usually distinguishable from each other by the shadows associated with clouds.

The TM data used in this research are available to the project from the USGS EROS Data Center, and are already terrain corrected with the geometric error of less than 1 pixel. For elucidating temporal trends, the Landsat multispectral images were acquired in spring (April) and summer (June) of 2006 to fully capture the early phenology (green-up and senescence) of cheatgrass. These two dates corresponded to two life stages of cheatgrass; in early spring, the cheatgrass is already established while the native

vegetation is just beginning to emerge from winter dormancy. In summer, the lifecycle of cheatgrass is complete and it is senesced and brown, while the native vegetation is still green. Table 2.2 summaries the acquisition dates and path/row of every Landsat TM image.

The second main data set selected is a Digital Elevation Model (DEM) since the invasion of cheatgrass is usually related to the elevation. A study in northern Utah indicates that cheatgrass is most invasive on midelevation sagebrush and shad scale sites on benches where livestock grazing and fires are common, and is less invasive on the higher-elevation sites where rugged topography and low water availability have resulted in less disturbance both by fire and livestock (Zamora, 1999). Therefore, elevation should be an important variable that influences the prediction accuracy of the model. The National Elevation Data downloaded from the Utah Automated Geographic Reference Center (AGRC) are used here with the same spatial resolution (30 meters) as Landsat TM data.

The third main data set is the field data used for training and testing the models. The field data used were collected in the summer of 2006. To ensure that the field samples could represent the cheatgrass in the whole research area, a stratified random spatial sampling was performed. An interpretation of remote sensing imagery was performed first to classify the whole region into two groups: high-elevation forest, and others. Within every group, a spatial random point sampling process was performed. Since high-elevation areas have minimal cheatgrass, the number of points is less than that of the other region. However, because of this accessibility, some points have not been sampled and some were not sampled exactly at these points. More than two hundred

points were sampled in this region. This procedure guarantees that we do not waste a lot of time collecting many noncheatgrass samples in forests area, while still achieving a random sampling of the cheatgrass percent over the whole region. There were about 247 field plots collected (Figure 2.4) within the boundary of these Landsat TM scenes.

Every field plot represents a circular area with homogeneous cheatgrass percent cover; at each of these homogeneous polygons, the attribute of radius and cheatgrass percent is recorded (Figure 2.5). Each record was accompanied by several photographs for further reference. After differential correction, all of these plots are stored into a GIS database with an accuracy of within 1 meter.

The result of our stratified random sampling is shown by investigating the histogram of the field data. The whole data range is from 0 to 95%, because it is hard to get a sample where the percent of cheatgrass is 100%. Also, the distribution is likely a skewed distribution, in that the number of points with noncheatgrass is the largest. When the cheatgrass percent increases, the frequency decreases. There is only one place where the cheatgrass percent reaches 95%. This finding corresponds well with the truth that although many places in this region are invaded by cheatgrass, they are mostly occupied also by sage brush or other grasses and shrubs because of competition; the number of places where cheatgrass is the only vegetation is rare. More detailed discussions about spatial sampling are included in the section about spatial sampling strategy.

All of the above three datasets are projected into a GIS database with the NAD 83 and a UTM Zone 12 N projection, so that the pixel values of TM scenes and the DEM could be extracted with the corresponding field plots.

Representative Issue

The arid West of the United States is under attack by exotic, invasive species. These exotic invaders, either as a result of or in addition to modifying ecosystem processes, pose a significant threat to the biodiversity and persistence of native species in the ecosystems they infest (Blank et al., 2009). Among these invaders, cheatgrass is a highly invasive, winter annual that occupies or dominates a wide variety of ecosystem types in the Intermountain West (Monsen, 1994).

In this research, the study area consists of the central part of the state of Utah, U.S.A. This research area is selected as a study area because it embodies much of the problems of cheatgrass invasion within similar ecosystem settings that are typical for large areas in the Western United States (Nevada, Arizona, Idaho, etc.). This area not only has serious cheatgrass infestations, but has a wide range of ecosystem types typically found in the Intermountain West, including arid desert shrub communities, desert communities, high elevation conifer and broadleaf forests, pinyon-juniper woodland, grazed grassland, irrigated agriculture, etc. Cheatgrass infestation in this research area occurs mostly on those areas at lower elevations in different ecosystem types, primarily disturbed areas (Clinton, 2009).

Based on the above two points, it could be clearly seen that the study area of the central portion of Utah is a representative research area for the Intermountain West confronting problems of cheatgrass infestation. Through the investigation of cheatgrass invasion in such a representative research area, we can apply and test the developed mapping model in new research areas of similar ecosystem types in the Intermountain West. One limiting effect of choosing this field area is that it cannot represent all the

places with cheatgrass infestation. For instance, the land cover, ecosystems, and elevation are totally different in New York state compared with the Intermountain West; therefore, although the method itself can be used and tested, the model with calibrated parameters in this research will have to be modified to fit the spectral characteristics of other areas.

Spatial Sampling Strategy

Statistical analysis is based upon sample data and statistical inferences are made using information collected from samples. Usually one assumes that sample observations are taken randomly from some larger population of interest. When the population consists of all of the points or pixels in a geographical region of interest, there are several spatial sampling methods: random spatial sampling, stratified spatial sampling, and systematic spatial sampling (Rogerson, 2001). The random spatial sampling method is the simplest of the three and it is easy to implement; however, it cannot always ensure adequate coverage of the study area while the other two methods can. Haining (1990) discusses spatial sampling in detail and suggests that systematic random sampling is often slightly better than the other two methods because of several advantages: it not only ensures adequate coverage of the study area, but reduces more redundant information and provides better estimates of the variable's mean value.

Selecting representative sample sites in large regions is not a straightforward process since selection criteria are often subjective and rely on intuitions that are surely not made explicit. In this research, a revised stratified spatial sampling method, which is objective and repeatable, is implemented and described as follows:

Input data consist of the mosaic Landsat TM images for the whole research area. First, an initial regionalization is produced with the interpretation of remote sensing imagery, thus to classify the area into two exclusive subregions, including high-elevation forests, and others. This process is implemented due to the fact that cheatgrass infestation in this research area occurs mostly on those areas at lower elevations in different ecosystem types, primarily disturbed areas (Clinton, 2009). This process also ensures that the sampled pixels are adequate to cover the whole study area.

Next, within each area, a spatial random point sampling process is performed. Since high-elevation forest area has less cheatgrass, the number of points sampled is less than that of the other regions. This process assures that we do not collect relatively too many useless and dummy points in this research area. Subsequently, the final picked sample sites are chosen in the field trip based on the accessibility.

Although this whole process seems objective, several significant issues are discovered during and after the completion of the field trip. First, the limitation of transportation and labor, and the accessibility of some samples (especially at canyons or high peaks) cause many samples, either far from roads or not accessible, to be abandoned, while these points closer to roads have higher probability of being selected during the field trip. The result of this influence could be clearly seen as in Figure 2.4; most sampled points are along the network of highways, thus decreasing the randomness of the sample points. This disadvantage could possibly be avoided by increasing the number of surveyors to reach these points. In our case, we had only two surveyors for the field trip; therefore, many points had to be abandoned so that we could cover the whole research area during the field trip.

Second, studies have demonstrated that the growth and expansion of cheatgrass is closely related to both environmental conditions (temperature, elevation, soil characteristic, etc.) and anthropogenic disturbance (cultivation, grazing, fire, etc.) (Pierson and Mack, 1990a, b; Rice et al., 1992); Cheatgrass prefers to invade disturbed areas; therefore, more points should be sampled in those areas. Detailed maps of ecosystem types, temperature, elevation, soil characteristic, or disturbed/nondisturbed areas should be considered during the subregion stratification. The stratification strategy in this research is mainly limited to the factor of elevation, and high-elevation forests and others are the two only strata. However, if we attempt to achieve more complete, objective, and representative samples, these above sets of data could be combined to reach a better stratification.

Third, the simple random spatial sampling method is used within each stratum in this research. As noted above, this method cannot ensure adequate coverage of the study area, thus leading to a lack of samples which are not enough to cover the whole study area (Figure 2.4). A better strategy is proposed here for use in future research. Within each stratum, numerous subregions are stratified and evenly distributed, and then a systematic spatial random sampling is implemented by taking the same sample spatial configuration in each subregion. In this way, a complete set of samples could cover the whole study area, and compared with the simple random spatial sampling, this method could avoid creating redundant information by preventing the situation of closer sample locations.

The question of which spatial sampling strategy is best depends on the spatial characteristics of variability in the data. By solving these issues, the physical process

characteristics of cheatgrass could be fully utilized to complement the revised stratified random sampling strategy.

Imagery Standardization

As it is well known, raw digital numbers (DNs) derived from satellite optical systems cannot confidently be used for geophysical measurements and multitemporal studies, since they include effects derived from instrument calibration, as well as atmospheric and topographic influences. Image standardization is the process of normalizing image pixel values and converting satellite-recorded DN to surface reflectance (Chavez, 1996). Standardizing imagery improves the ability to compare imagery over time and mosaic adjacent imagery. Because some procedures of image standardization involve collecting in-situ atmospheric measurements during each satellite overflight and radiometric transfer code (RTC), they are difficult and more costly to implement. The procedures used here are image-based, requiring no additional information other than that provided by the imagery and Digital Elevation Model (DEM). The following are the main procedures:

1) Atmospheric Correction

Atmospheric correction implemented here is based on the image-based COST model proposed by Chavez (1996), since the corrections generated by this model are as accurate as those generated by the models that used in-situ atmospheric field measurements and RTC software. The equation is as follows:

$$\rho_\lambda = \frac{\pi((DN_\lambda * Gain_\lambda + Bias_\lambda) - (DarkDN_\lambda * Gain_\lambda + Bias_\lambda)) * D^2}{ESUN_\lambda * COS(90 - \theta) * \tau} \quad \ldots\ldots 2.1$$

where

ρ_λ Reflectance for Band with wavelength λ

DN_λ DN value for Band with wavelength λ

$Gain_\lambda$ The gain value for Band with wavelength λ

$Bias_\lambda$ The bias value Band with wavelength λ

$DarkDN_\lambda$ Digital Number representing Dark Object for Band with wavelength λ

D Earth-Sun distance in astronomical units

$ESUN_\lambda$ Solar Irradiance for Band with wavelength λ

θ Solar Elevation Angle in degrees

τ Atmospheric Transmittance expressed as $COS(90-\theta)$

The parameters for $Gain_\lambda$, $Bias_\lambda$, D, and $ESUN_\lambda$ are from Chander (2003), and θ is typically given in the header file of imagery.

2) Topographic Correction

Topographic correction refers to the compensation of the different solar illuminations due to the irregular shape of the terrain (Riano, 2003). This effect causes a high variation in the reflectance for similar vegetation types: sunny areas show more than expected reflectance. Therefore, the process of topographic normalization may be critical in areas of rough terrain as a necessary step to the application of multitemporal remote sensing such as in this research. These pixels have the same land cover but different

topography will have different Digital Numbers, thus biasing the prediction of cheatgrass percentages.

The C-correction method described by Teillet et al. (1982) is implemented using IDL to topographically correct the Landsat TM image used in this research. This method is implemented because it improves classification accuracy (Riano, 2003). Following are the main procedures:

$$\rho_T = m\cos(i) + b \quad \text{.....2.2}$$

The above equation corresponds to a regression line with the original reflectance on the y axis and cos(i) on the x axis; the b and m values are calculated using the least square method.

$$c = \frac{b}{m} \quad \text{.....2.3}$$

$$\rho_H = \rho_T \left[\frac{\cos(sz) + c}{\cos(i) + c} \right] \quad \text{.....2.4}$$

where: c = correction parameter; m = slope of the regression line; b = y - intercept of regression line; ρ_H = reflectance observed for horizontal surface; ρ_T = reflectance observed over sloped terrain; sz = sun's zenith angle; i = sun's incidence angle in relation to the normal on a pixel.

Methodology

The fundamental objective of this study is to build a model that is reasonably accurate in the prediction of cheatgrass abundance using remote sensing. The Multiple Linear Regression (MLR) model is a conventional approach used to make predictions of biophysical or biochemical parameters, such as LAI and biomass, in remote sensing. An insurmountable limitation of the linear approach is that the model can predict the cheatgreass percentages outside the range, i.e., values of < 0 or >1. Generally, all predicted negative values are forced to 0 and those values >1 are set to 1. Even without this modification, it is not meaningful to compare these conditions. For example, a prediction value of 150% cannot be interpreted as a higher cheatgrass percent than 101%, since the maximum percentage of cheatgrass is 100%.

Considering the above issue of the suitability of the MLR in predicting percentage data, the regression tree (Breiman, 1984) model, which is a nonlinear regression model devised to overcome just such a problem, as well as the problem of unknown nonlinear function is to be used. The regression tree model is built through an iterative process of splitting the data into partitions, and then splitting it up further on each of the branches.

This whole splitting process recursively partitions the observations into subsets such that the total residual sum square (RSS) of all subsets is minimized. The RSS of the kth subset is calculated as:

$$RSS_k = \sum_{i=1}^{N_k} (\hat{y}_i - y_i)^2 \quad \text{.....2.5}$$

where N_k is the number of samples in the kth subset, and $\hat{y}_i$ and y_i the predicted and actual values of the dependent variable. In this case, the dependent variable is the cheatgrass abundance within every pixel. The independent variables will include the spectral value of remote sensing data.

The result of this iterative binary partitioning is a model whose structure can be displayed as a tree-like graph (Figure 2.6), with each split in the tree labeled according to the threshold used to define the split. When using a regression tree to predict the value of the dependent variable, the estimated value is the mean value of the dependent variable of the rows falling in a terminal (leaf) node of the tree.

A simple example of a regression tree model for predicting the value of LAI using Landsat TM data is shown in Figure 2.6. In this figure, all observations that satisfy the criterion at a given split fall to the left-hand child node while those not meeting the criterion continue to the right. The number of observations and their average LAI value, which is equal to the model estimate, is shown for each terminal node. In this example, the dependent variable is LAI, and the independent variables are “band 3” and “band 4” of Landsat TM imagery. From the tree we see that if the value of the independent variable “band 4” is greater than 0.3, then the estimated (average) value of the LAI is 4; whereas if the value of “band 3” is less than or equal to 0.1, then the average value of the LAI is 2.

There are several reasons to choose this method in this research:

1. A regression tree is nonparametric. In other words, no assumptions are made regarding the underlying distribution of the dependent variables.

2. A regression tree is capable of approximating a nonlinear relationship and does not require prior knowledge on the nonlinear form.

3. A regression tree is simple to interpret; and unlike neural network, it is easy to interpret the importance of the independent variables.

My hypothesis is that the local Vegetation Index (VI) difference (ΔVI) between the green-up and senescence season are useful in distinguishing cheatgrass from native species. The magnitude of ΔVI should vary with different levels of cheatgrass invasion, ecosystem type, and amount of annual precipitation. Therefore, Landsat TM will be obtained for two time periods in order to use the unusual early phenology (green-up and senescence) of cheatgrass. The NDVI is calculated on the imagery of each time period and the difference in NDVI between two time periods is then used as a measure of phenology across the landscape.

In the satellite image layers described above, independent variables are the reflectance bands (after atmospheric correction and terrain correction), elevation, and ΔNDVI, while the dependent variable is the cheatgrass abundance. Within each field plot of a polygon that is produced by a buffering operation on every field point using the radius attribute, there is a corresponding set of pixels in each of the independent data layers, as shown in Figure 2.7. Figure 2.7 also demonstrates the principle of pixel extraction within each polygon, if the centre of a pixel is to the right of the left active edge's intersection with the scan line, and its center is to the left of the right active edge's intersection with the scan line. Then, this pixel is extracted for either training or testing samples.

Two thirds of the data are used to construct the regression tree, and the other one-third is to be used to test the model. The extracted pixels are first ordered by the value of percentage; then the first two of every three consecutive pixels are selected as training samples. The other one pixel is then selected as a testing sample unless all of the pixels are selected to construct or test the model (Figure 2.8).

The procedures for the whole mapping process are:

1. Image preprocessing (atmospheric and topographic correction) of all the Landsat TM images
2. Calculation of the different VIs (SR, NDVI, EVI) using the surface reflectance of Landsat images
3. Calculation of the VIs difference between images of different dates
4. Identification of the ground plots in the scenes and extraction of the remote sensing and elevation data for each of these plots using IDL
5. Development of a linear regression model and a regression tree model using 2/3 of data from all the scenes. The algorithm developed in this way will be more reliable than those developed for individual scenes, because the number of data points for individual scenes is generally small and the dynamic range of the cheatgrass abundance is generally small within one Landsat scene.
6. Production of a cheatgrass abundance map for each Landsat scene using the developed algorithm.
7. Accuracy assessment using the remaining 1/3 of ground plot data.

With the above set of methods developed, the current cheatgrass abundance over the entire central Utah could be mapped and validated.

Results and Conclusions

Atmospheric and topographic correction is a method used to reduce atmospheric and illumination effects on remote sensed data to retrieve physical parameters of the Earth's surface such as atmospheric conditions, thermal and atmospheric radiance and transmittance (Riano, 2003). Usually when images are used for classification or are enhanced for display effects or interpretation, this method is not required. In this research, multitemporal satellite images with different times of acquisition are used; therefore, this method is required to reduce the influence of atmosphere and topography so that the changes observed are due to different features on the Earth's surface rather than differences of the atmospheric or topographic condition.

Basically, there are three common types of atmospheric correction: Radiative Transfer Modeling (RTM), Dark Object Subtraction (DOS), and Empirical Line Method (Chavez, 1996). Among these three methods, the DOS is an easy and quick method, without the input of any extra real time atmospheric data, but it only explains the additive effect of atmosphere and does not explain the multiply effect; therefore, the accuracy is not high. The improved Dark Object Subtraction called COST, which incorporates the multiply effect of atmosphere, is appropriate in this research and implemented (Chavez, 1996). The C-correction method described by Teillet et al. (1982) is selected in this research because of its reliable performance in improving classification accuracy (Riano, 2003).

In order to test the results of the atmospheric correction, the field spectral reflectance data (Figure 2.9(a)) of dry cheatgrass measured using SPEC are spectrally resampled to match the response of the Landsat-5 TM instrument (Figure 2.9(b)). Figure

2.10 demonstrates the comparison between the at-sensor reflectance, atmospherically corrected reflectance, and the resampled Landsat-5 TM reflectance. The results show that before the atmospheric correction, the reflectance data of dry cheatgrass in Landsat TM imagery are higher than the measured reflectance due to the influence of atmospheric path radiance over the all six bands (Figure 2.10(a)); the mean difference between the measured reflectance and the reflectance before atmospheric correction is about 0.05. While after atmospheric correction, the reflectance curve of dry cheatgrass is approaching the actual measured reflectance curve, the mean difference between the measured reflectance and the atmospherically corrected reflectance is less than 0.02 (Figure 2.10(b)). This demonstration shows that the DOS method used to reduce the atmospheric correction is not only easy to implement, but very reliable.

To validate and evaluate the two models used for cheatgrass percentage prediction, both scatter plot and quantitative analysis are used by comparing the observed cheatgrass percentage with the predicted cheatgrass percentage. Scatter plot is a type of mathematical diagram using Cartesian coordinates to display values for two variables for a set of data. Scatter plot is usually selected to visualize the relationship between two variables because of its powerful ability in disclosing both linear and nonlinear relationship. The quantitative analysis used to evaluate the model performance in this research includes the coefficient of determination (R^2) and root-mean-square error (RMSE). The R^2 is used to determine whether the regression between predictions and observations could successfully explain a significant portion of the variation. RMSE, defined as the square root of the variance between two variables, is a frequently used

measure of the differences between values predicted by a model and the values actually observed. The F-test is performed to test whether the R^2 is significant.

Figure 2.11 shows the scatter plots of predicted cheatgrass percentage versus the observed percentage, with Multiple Linear Regression and Regression Tree separately. The red dotted line describes the Y = X line. It could be clearly seen from the scatter plot that there is a linear relationship between the predicated and observed cheatgrass percentage for both models. An interesting finding in the scatter plot is that in the cheatgrass percentage range of [0-10%], both models significantly overestimate. One possible explanation is that there are numerous sample points in the study area without cheatgrass.

Table 2.3 shows that the values of R^2 between predicted and observed cheatgrass percentage for both models are significant ($p<0.05$). Therefore, the regressions in Figure 2.11 can both successfully explain a significant portion of the variation. The RMSE for both models are around 10%; this illustrates relatively low differences between predicted and observed cheatgrass percentage and means a close relationship between predicted and observed values. Therefore, the quantitative analysis, as well as the scatter plots, strongly supports that both Multiple Linear Regression and Regression Tree models can approach the reality of cheatgrass percentage modeling.

Although the results of both models show significant linear relationship between predicted and observed cheatgrass percentage and they provide similar fit, Table 2.3 also shows that the Regression tree model outperforms the simple Multiple Linear Regression model in that it has higher significant R^2 and lower RMSE. This means that the regression tree analysis develops a better model; this is mainly because it could describe

the complicated nonlinear relationship between the cheatgrass cover and NDVI, while Multiple Linear Regression assumes that a linear relationship exist between dependent and independent values. Further research comparing the other nonlinear modelling techniques like neural network may help to explain the advantage of nonlinear modelling over linear modelling.

These results show that, combined with the field observations and application of statistical analysis, a remote sensing technique could supply an easy and quick method for invasive species abundance prediction. With the development of Hyper-spectral remote sensing, more variables will be added into remote sensing modelling. Nonlinear modelling technique, as an alternative method to traditional linear modelling technique, may benefit this process in that it could model the more complicated nonlinear relationship between numerous variables.

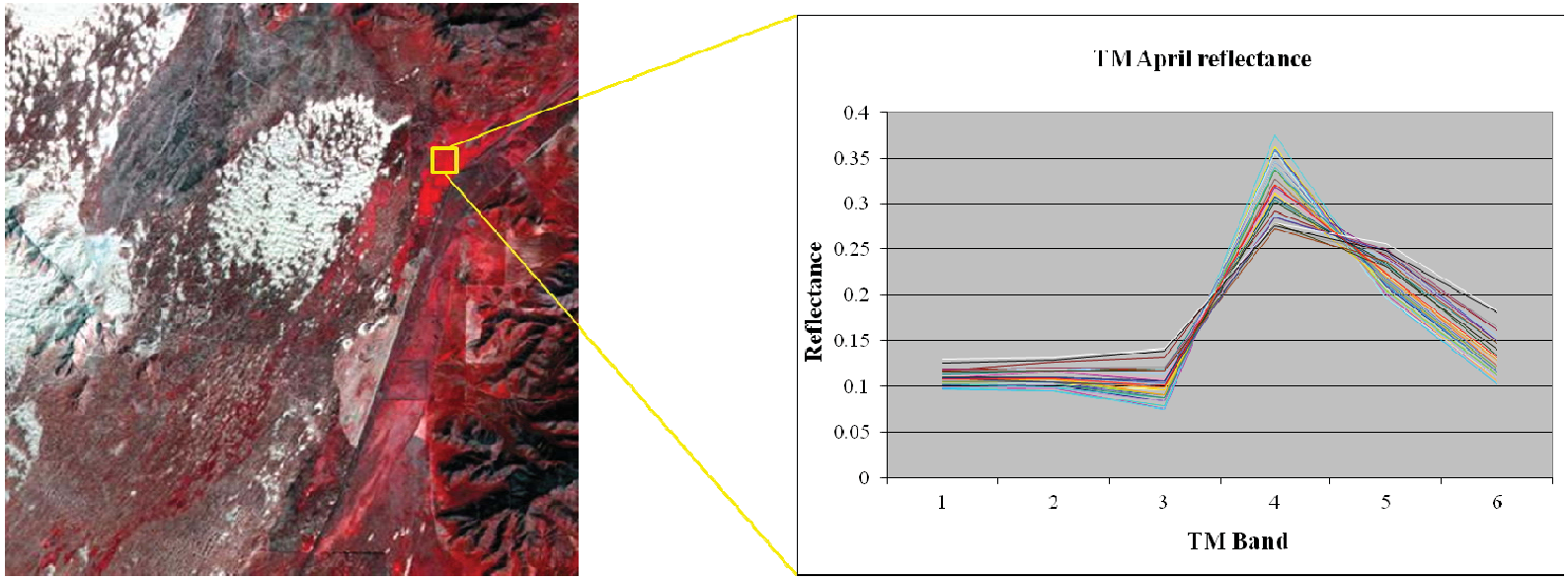


Figure 2.1. Cheatgrass reflectance in Landsat TM April image
(Band 4: R; Band 3: G; Band 2: B)

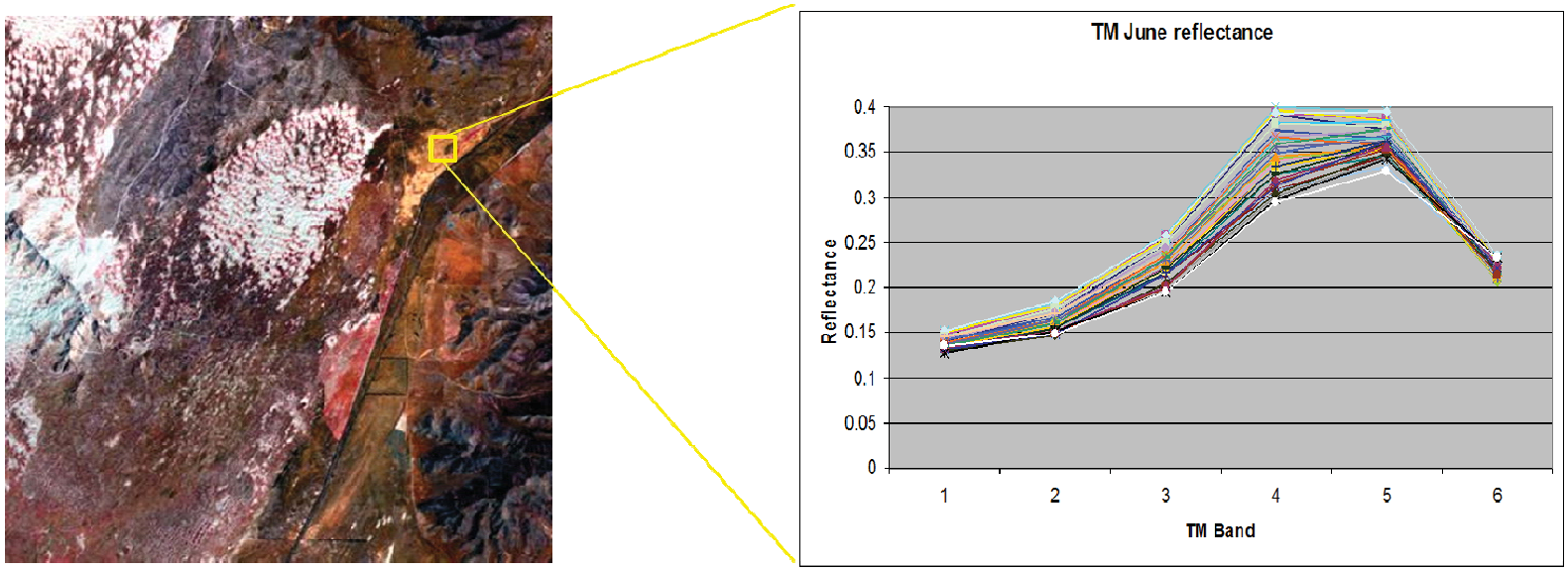


Figure 2.2 Cheatgrass reflectance in Landsat TM June image
(Band 4: R; Band 3: G; Band 2: B)

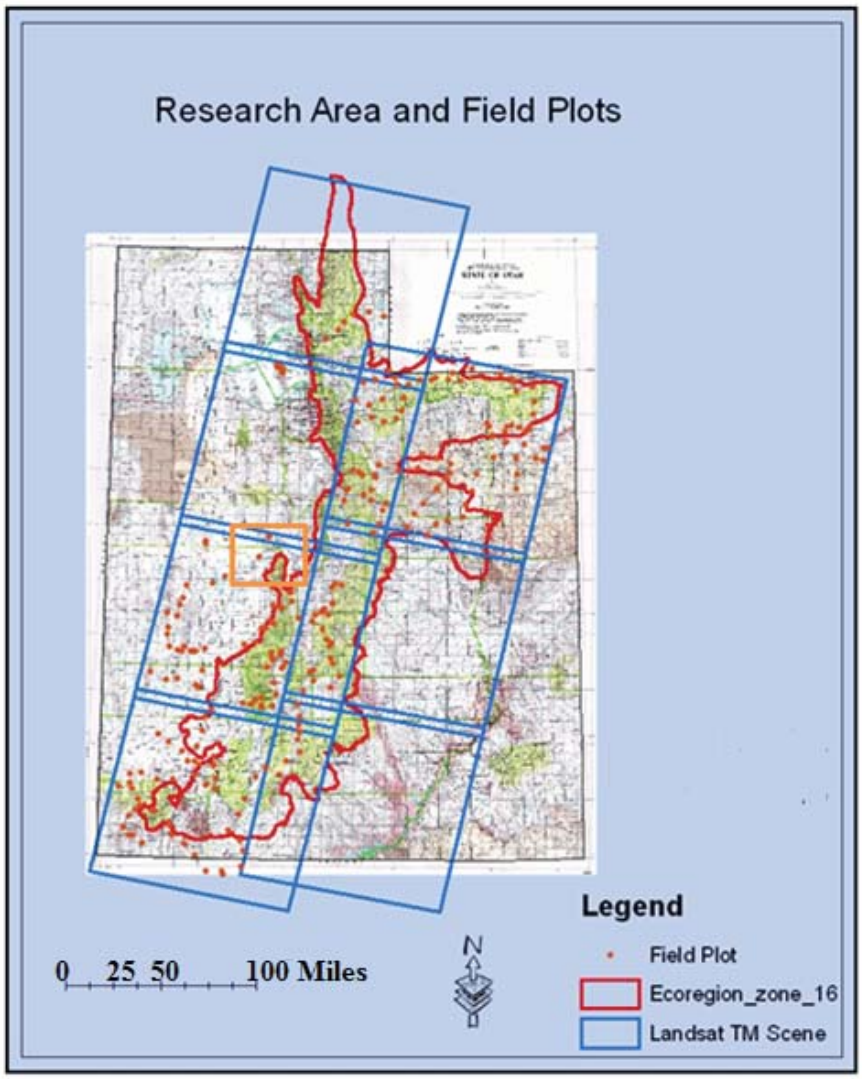


Figure 2.3. Study area and Landsat TM coverage

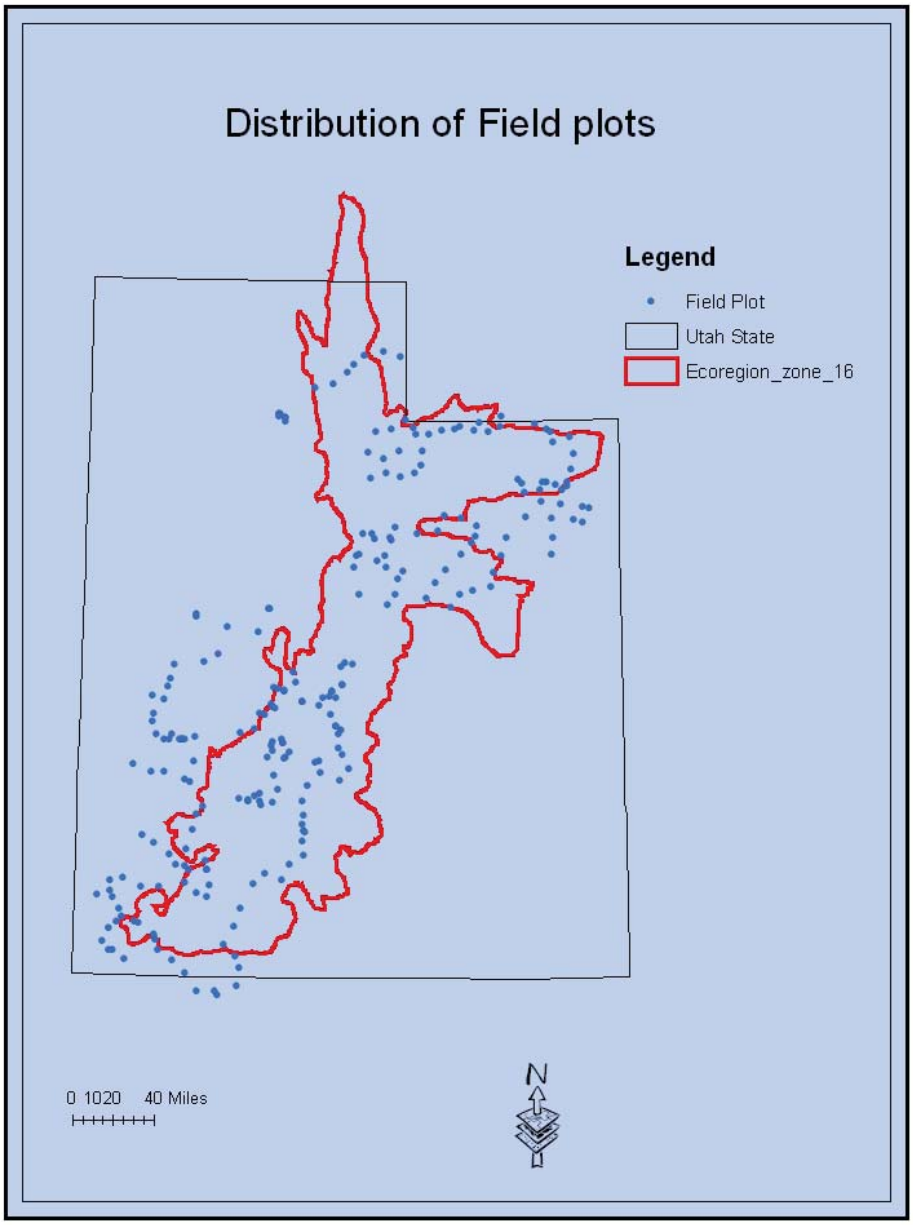


Figure 2.4. Distribution of field plots

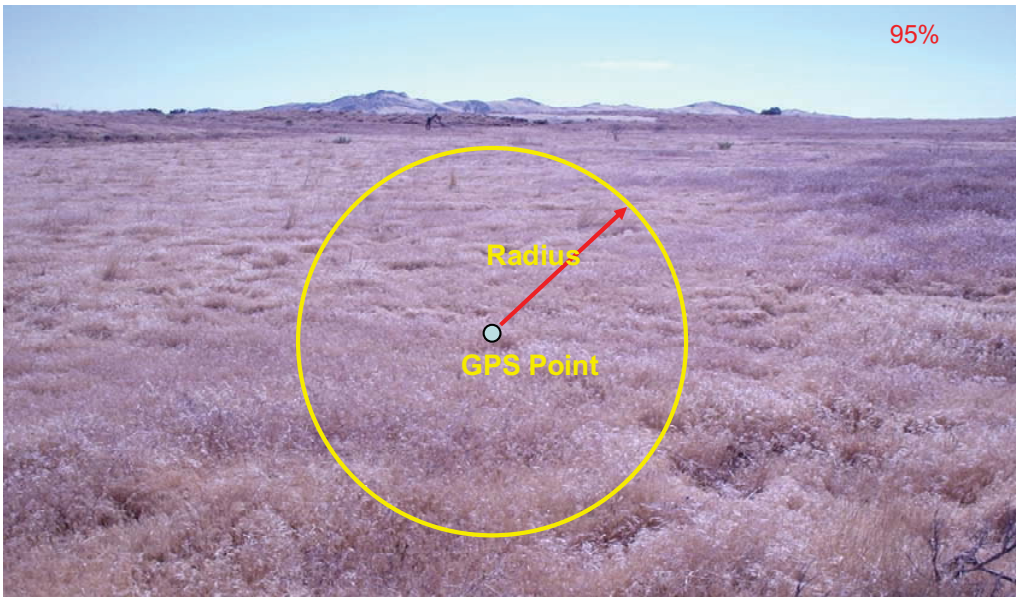


Figure 2.5. The field plot survey

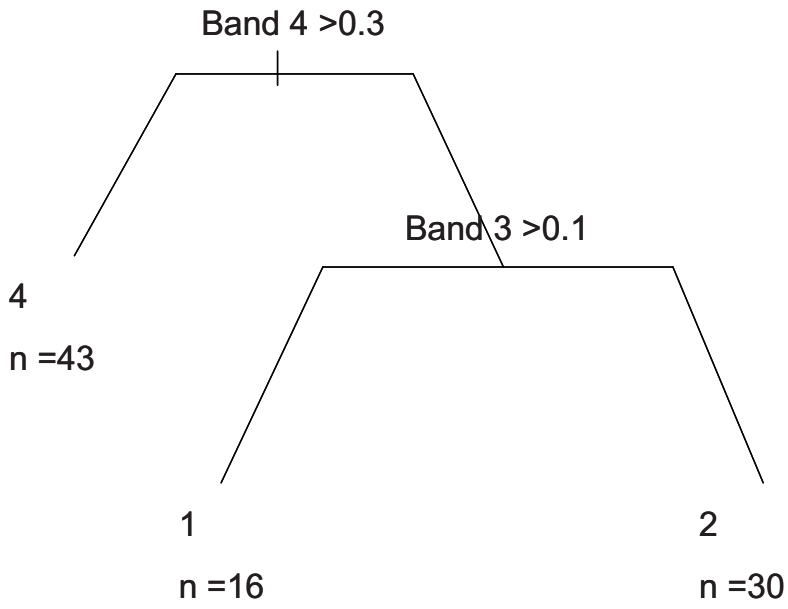


Figure 2.6. Regression tree model for LAI prediction

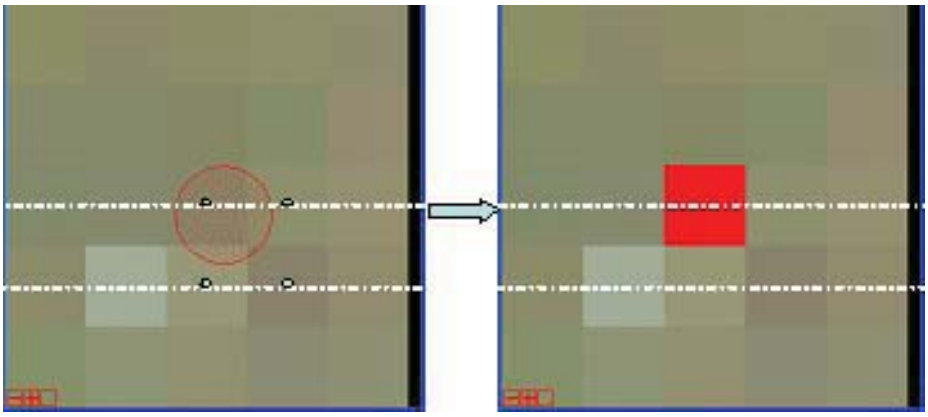


Figure 2.7. The principle of pixel extraction

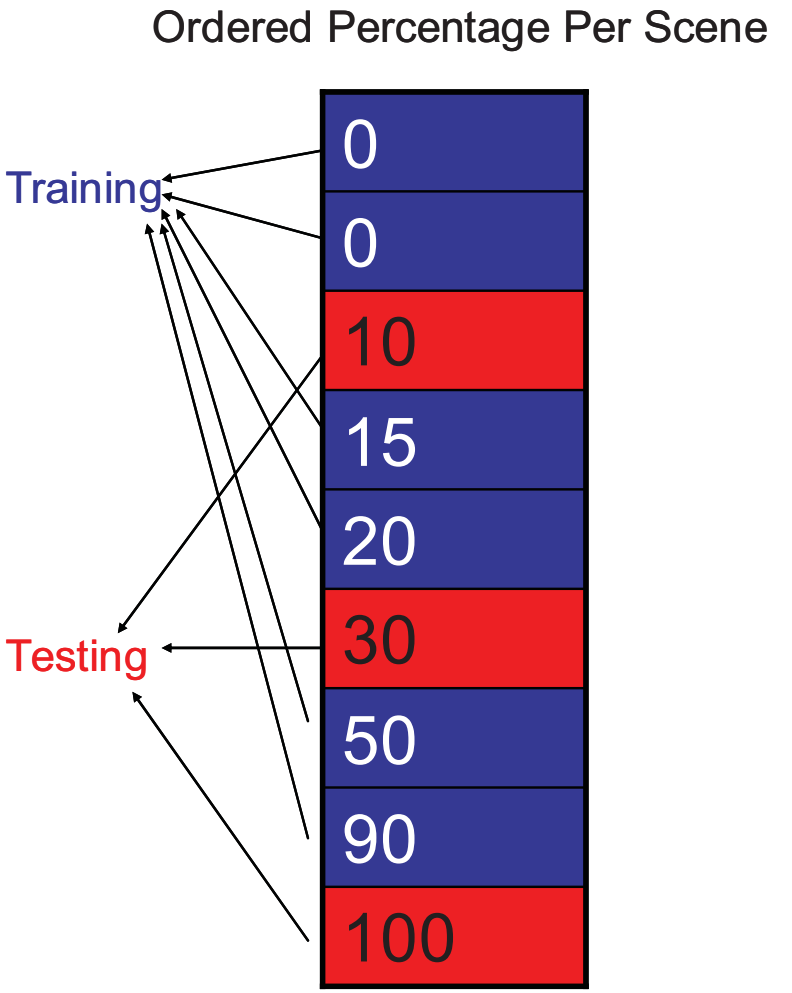
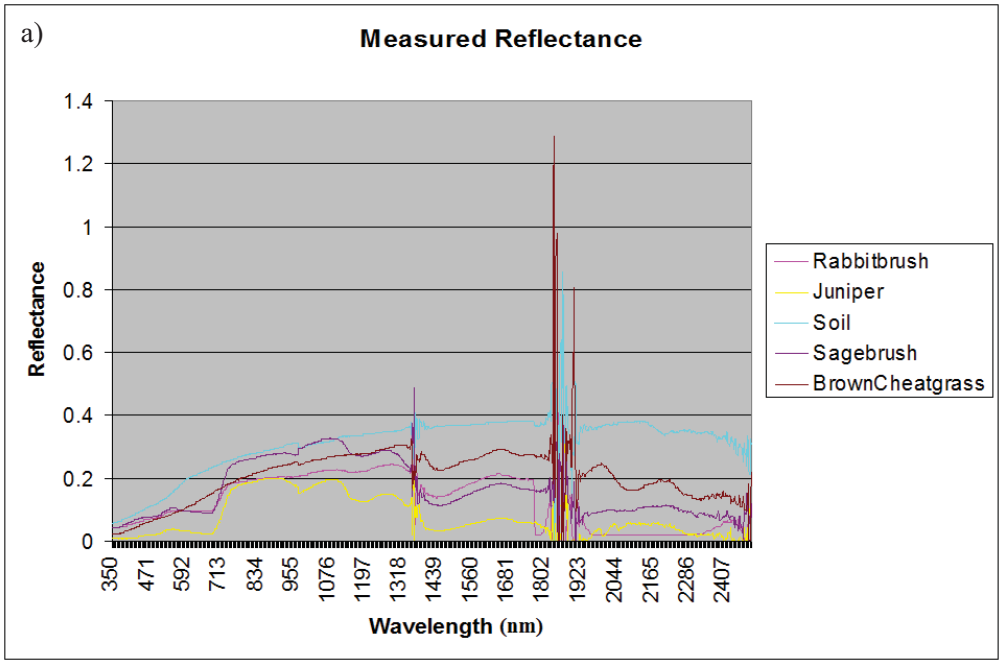


Figure 2.8. Selection of training and testing pixels



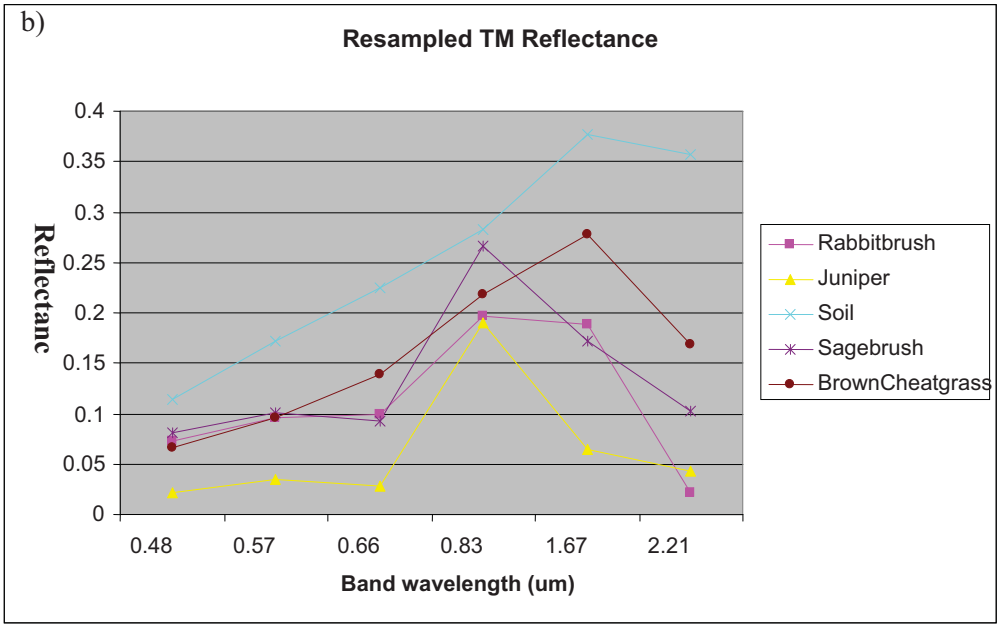


Figure 2.9. The measured and resampled reflectance of cheatgrass (a). The measured reflectance of brown cheatgrass. (b). The TM reflectance resampled/simulated from measured reflectance.

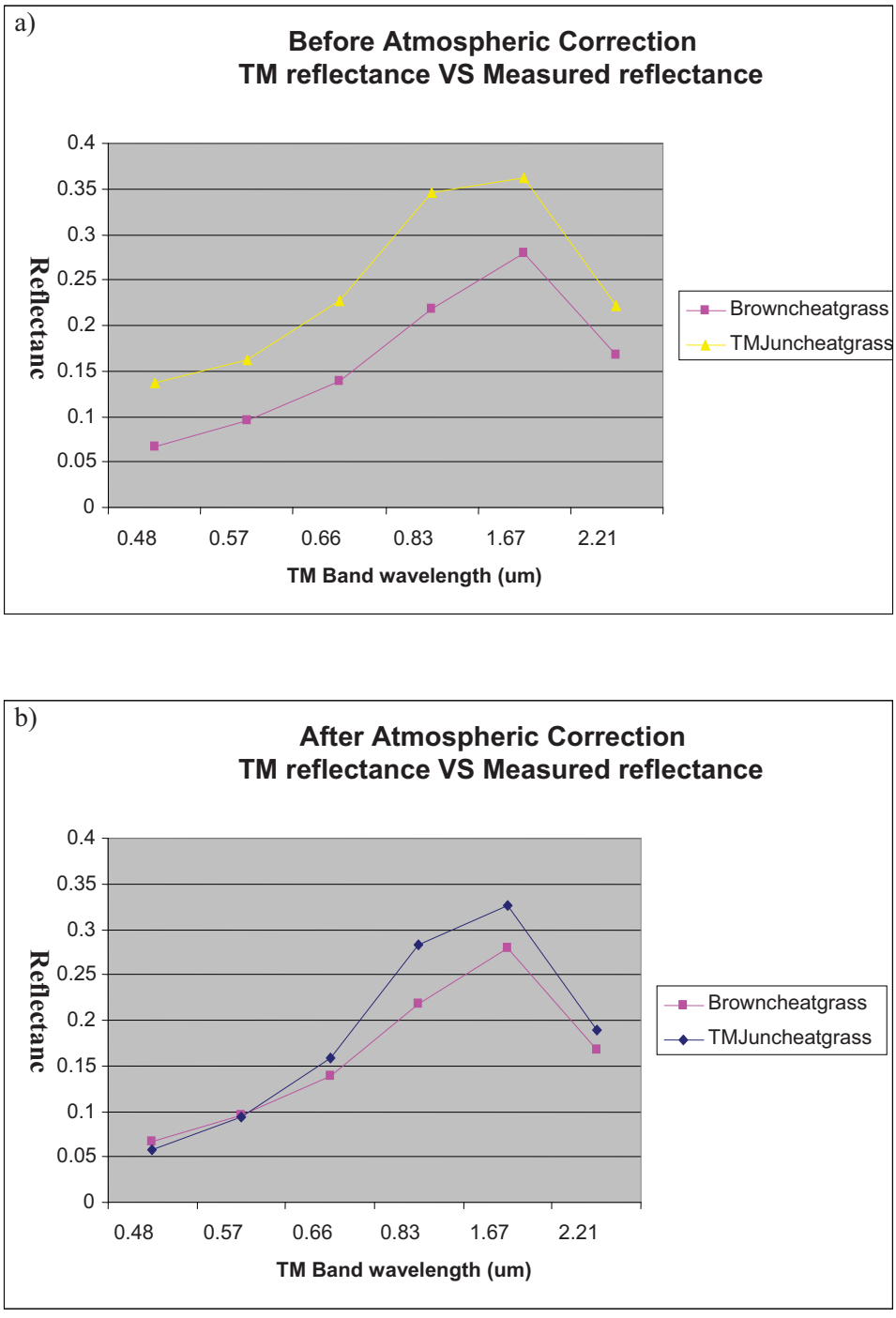
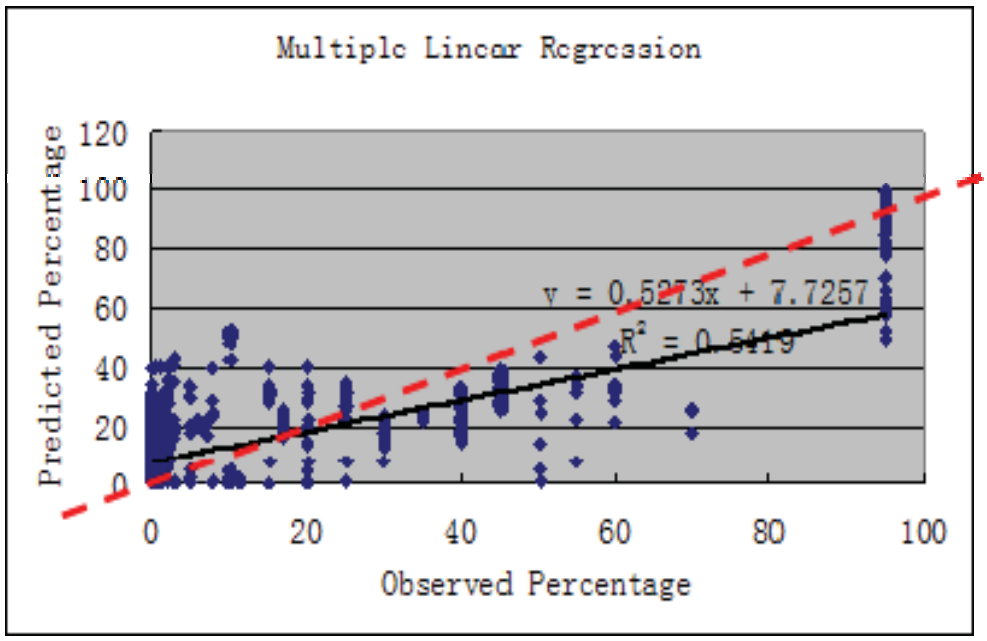


Figure 2.10. Comparison of reflectance before and after atmospheric correction. (a). TM At-Sensor reflectance vs. resampled TM reflectance. (b). TM surface reflectance vs. resampled TM reflectance.



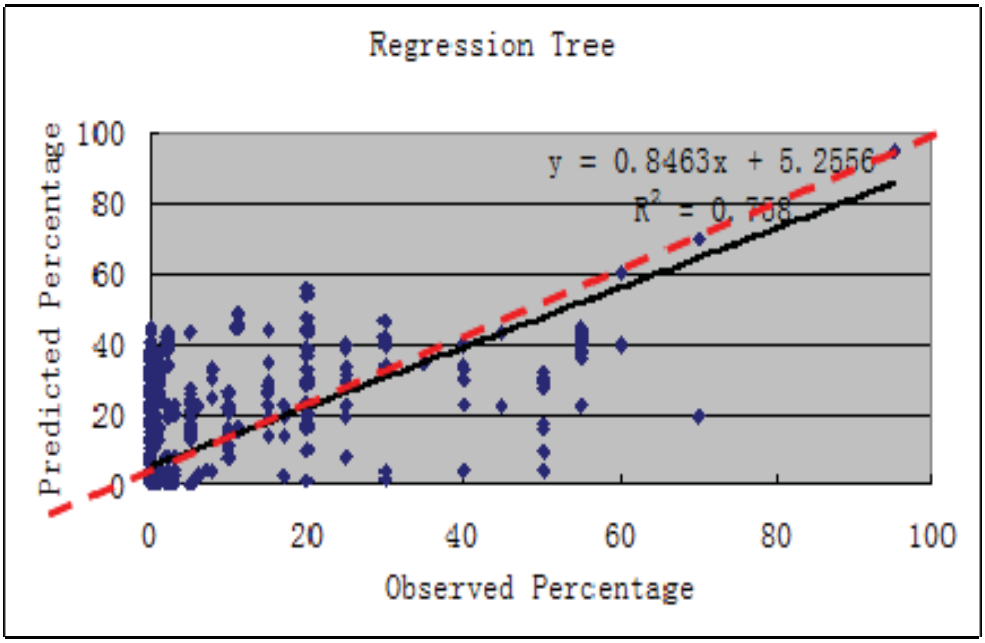


Figure 2.11. Scatter plots of predicted vs. observed percentage
(Red dotted line: Y = X).

Table 2.1 Spectral and spatial resolution of Landsat-5 TM

Band Landsat-5 TM	Spectral Resolution (um)	Spatial Resolution (m)
1	0.45 - 0.52	30
2	0.52 - 0.60	30
3	0.63 - 0.69	30
4	0.76 - 0.90	30
5	1.55 - 1.75	30
6	10.40 - 12.50	120
7	2.08 - 2.35	30

Table 2.2 Landsat TM data description.

Sensor	Date	Path/Row	Processing
TM	04/19/2006	38_33	terrain corrected
TM	06/22/2006	38_33	terrain corrected
TM	04/19/2006	38_34	terrain corrected
TM	06/22/2006	38_34	terrain corrected
TM	04/19/2006	38_32	terrain corrected
TM	06/22/2006	38_32	terrain corrected
TM	04/19/2006	38_31	terrain corrected
TM	06/22/2006	38_31	terrain corrected
TM	04/28/2006	37_32	terrain corrected
TM	07/01/2006	37_32	terrain corrected

Table 2.3 Quantitative evaluation of MLR and regression tree

Method	R Squared	RMSE
MLR	0.54*	14%
Regression Tree	0.76*	11%

* Significance ($P < 0.05$)

CHAPTER 3

RMMEH — A NOVEL COMPOUND SMOOTHER TO RECONSTRUCT MODIS NDVI TIME-SERIES

Introduction

The term "phenology" is generally described as "the study of the timing of recurring biological phases, the causes of their timing with regard to biotic and abiotic forces, and the interrelation among phases of the same or different species" (Lieth, 1974). For the phenology of vegetation, it is concerned with the determination of the dates of critical phases during the growth and development of plants, such as the start, length, or end of the growing seasons. Monitoring these phenological dynamics of vegetation is essential for climate change study, since climatic processes such as temperature and precipitation mainly drive the phenology of a region.

Traditional phenological studies have been conducted by limited ground-based observations. The remote sensing measurements, however, have been adopted as an alternative to monitor vegetation phenology, since less field work is needed and larger areas at landscape scale could be investigated readily (Clinton, 2010).

During the last two decades, several studies have characterized vegetation phenological dynamics at continental scales using time-series of vegetation indices (VIs),

mostly Normalized Difference Vegetation Index (NDVI), obtained from different satellite data. Biweekly composite NOAA/AVHRR NDVI data from 1989 to 1992 over the conterminous United States were used to extract a set of twelve phonological metrics, including the onset of greenness, time of peak NDVI, maximum NDVI, rate of green-up, rate of senescence, and integrated NDVI (Reed et al., 1994). Zhang et al. (2003) used a series of fitted piecewise logistic functions to represent the intra-annual vegetation dynamics and identified four transitional phenological dates by detecting maximal curvature.

Therefore, the accuracy of these phenological measurements is highly dependent on the quality of the time-series VIs products. Although these products, such as NOAA Global Vegetation Index (GVI) and MODIS/Terra 250m 16-day composition Vegetation Indices (MOD13Q), have been preprocessed with techniques like atmospheric correction or maximum value composite (MVC, Holben, 1986) to eliminate spuriously low values caused by atmospheric effects and cloud contamination, there remained significant residual effects of subpixel clouds, persistent cloudiness, presence of snow, variable illumination and viewing geometry, remnant geometric errors, and other negative effects. These effects tend to reduce the NDVI value and disturb the temporal profile, thus affecting the algorithms that are measuring the vegetation phenology (Reed et al., 2003). Therefore, multiple smoothing algorithms with a varying range of complexity have been proposed to reduce noise and reconstruct high-quality time-series NDVI data sets, such as polynomial and compounded smoothers (4253EH, twice, Dijk et al., 1987), Best Index Slope Extraction (BISE, Vivoy et al., 1992), nonlinear running median smoother (Reed, et al., 1994), Fourier analysis (Sellers, 1994;), Splines (White et al., 1997), Weighted

Least Squares (Swets, 1999), Asymmetric Gaussian function fitting (Jonsson et al., 2002), Savitzky-Golay (SG) Filter (Savitzky, 1964; Chen et al., 2004) and Mean Value Iteration (MVI, Ma & Veroustraete, 2006). Chen et al. (2004) pointed out that each of the abovementioned methods has its own advantages and drawbacks that limit its use, and presented a robust method based on SG filter. However, this method requires an ancillary data of cloud flag band in which the uncertainty exists, and it suffers from an edge effect that has an adverse effect on the eight endpoints.

In this study, our objective is to develop an algorithm to reduce noise from NDVI time-series fast and efficiently with the following two constraints: 1) Ancillary data is not required; 2) The whole procedure must be applicable automatically without expert support. That means each parameter in the method is fixed rather than land cover or research area dependent. We present a simple automated compound smoother called RMMEH to fulfill the above criteria. This new method involves several operations such as running medians (RM) smoother, arithmetic average, maximum (MAX) operation, and weighted moving average (WMA). This method is tested with the 16-day composite MODIS NDVI time-series. A window covering part of the state of Utah was also used for this study.

Data and Methodology

Many remote sensors, such as AVHRR, MODIS, and SPOT-VEGETATION, are currently used to derive spatially and temporally consistent global Vegetation Index (VI) products. The NDVI time-series data from MODIS/Terra 250m 16-day composition Vegetation Indices (MOD13Q1, Version 4) products were utilized in this study. MODIS

NDVI is computed from atmospherically corrected level 2 daily MODIS surface reflectance (MOD09 series) and composited over 16-day time intervals (Huete et al., 2002). These VI products are available to the public through the EOS Data Gateway and distributed in HDF (Hierarchical Data Format)-EOS format. Quality assurance (QA) and stage 1 validation of the products are finished and the products are ready for use in scientific publications (http://edcdaac.usgs.gov/modis/mod13q1v4.asp).

The concept of our method assumes that: (1) a generalized vegetation NDVI temporal profile without any contamination changes smoothly and gradually following seasonal cycles of growth and decline. Stated more explicitly, near NDVI points in the time domain are more related than distant points. This is valid because vegetation canopy changes are small with respect to a composite period of 16 days; and that (2) all contaminations, among which are cloud contamination, atmospheric variability and bi-directional effects, usually depress the NDVI value. Most of these smoothing algorithms also based on these two assumptions (Reed et al., 1994; Jonsson et al., 2002; Chen et al., 2004).

Therefore, the best NDVI estimation at any time point should not be less than the raw NDVI value, in light of these assumptions. The RMMEH method was developed to derive a best “upper-NDVI-point” for each point; the whole procedure is described as flowing steps and is shown in Figure 3.1.

A first operation involves smoothing the original NDVI temporal profile with a nonlinear running-median smoother. Running-median smoothers are simple nonlinear smoothers and are more resistant to outliers than classical linear digital filters such as moving average; so, linear filters are usually applied safely after spikes have been

removed by nonlinear smoothing (Velleman, 1981). The most widely used odd-span running median of span v (the number of data values summarized by each median) is named with the digit "v" and defined by

$$z_t = med\{y_{t-u}, \cdots, y_t, \cdots, y_{t+u}\} \quad \text{.....3.1}$$

where $v = 2u + 1$, $u = 1,2,3,4,\ldots\ldots n$;

med represents the median operation;

y_t represents original time-series data (a sequences of y values ordered by time t);

z_t represents the smoothing result of running median of span v;

t represents ordered time-series, $t = 1,2,3,\ldots.n$;

The selection of the appropriate span usually relates to the duration of spikes. Generally, median smoothers with larger spans can resist more outliers, but they also suffer from both the long computation time and the flattening problem that departs from the original signal too much. Specifically, a span-3 median will be unaffected by any single outlier, but it cannot correct for two consecutive outliers, while a span-5 median will be completely resistant to a 2-point spike; a span-7 median, although it can correct a 2-point spike or even a 3-point spike, is more time-consuming and the result will flatten the original valid peak values.

In this study of MODIS 16-day composite NDVI time-series smoothing, a span of 5 is chosen due to the fact that the chance for more than three NDVI spikes in a row, which means the probability that NDVI for each of 48 consecutive days is contaminated,

is extremely small. For medians of 5, since the two points at each end of the profile cannot be smoothed, the method proposed by Velleman (1981) is used: we do not modify the end values, but use a running median of 3 to smooth the second and next-to-last values. This first step is described by the function

$$R_t = med\{NDVI_{t-2}, NDVI_{t-1}, NDVI_t, NDVI_{t+1}, NDVI_{t+2}\} \quad \text{.....3.2}$$

where $NDVI_t$ and R_t are the raw and running-median smoothed NDVI value at time t, respectively.

Although the running-median smoothing method was effective at eliminating much of the extremely low-value contaminated data, it also reduced some NDVI peaks in the profiles that are presumed to be valid, uncontaminated NDVI values (Reed et al., 1994). Therefore, further improvements are necessary to address this issue to retain valid high values and eliminate only invalid low values. In this study, arithmetic average calculation and maximum operation are combined with the first-step 5-span running median to produce compounded smoothers tailored to solving this problem.

Since the ideal vegetation NDVI temporal profile changes gradually as stated, each NDVI point should approach closely the mean value of its previous and next point. This assumption was valid and successfully utilized in the MVI filter method to reconstruct the 10-day composition pathfinder AVHRR land (PAL) NDVI datasets (Ma & Veroustraete, 2006).

Here, the second step is to calculate a new NDVI time-series, in which each NDVI point is calculated as the arithmetic average of its two neighbors around it. This step is also operated on the original NDVI time-series.

$$M_t = (NDVI_{t-1} + NDVI_{t+1})/2 \quad \text{.....3.3}$$

where M_t is the result of the arithmetic averaging operation and letter M here represents the mean value.

The three points including $NDVI_t$, R_t, and M_t are generally not located at the same place. An example of this situation is illustrated in Figure 3.2. Among these three points, the M_t has the highest value and the raw $NDVI_t$ itself has the lowest value. It is clearly seen that there is a sudden drop for $NDVI_t$ and it is obviously a contaminated point.

The next step is the most important part and is called maximum operation. As stated above, the best NDVI estimation at any time point should be higher than or as high as the original NDVI value; thus the "upper NDVI envelope" strategy is adopted to derive the best estimate by taking the maximal value of $NDVI_t$, R_t, and M_t.

$$RMM_t = MAX\{NDVI_t, R_t, M_t\} \quad \text{.....3.4}$$

where RMM_t represents the best estimate and MAX represents the maximum operation. In Figure 3.2, after the maximum operation, the best estimate RMM_t will be equal to M_t,

the maximum value among the three points, and taken as input for the next step. Here it is clearly seen that M_t is a good approximation to the true value according to the profile trend.

In Figure 3.3 is an example of a raw single-year MODIS NDVI time-series for annual grass land cover and the results obtained by using the above procedure. Figure 3.3a shows that the original NDVI temporal profile, which exhibits a clear temporal vegetation evolution, is contaminated by low-value noises caused by persistent cloud contamination, presence of snow, atmospheric variability, or other effects. The noises in this example exist in three different forms: a single-point narrow spike in March, a 3-point oscillated spike from end of June to beginning of September, and a 2-point narrow and flat spike during November.

It is clearly shown in Figure 3.3b that the arithmetic average is very sensitive to any outlier and it will reduce original valid values, but it can still be used to extract the correct value of any spike that has two valid neighbors. Although the 5-span running median smoothing can remove most of the outliers, it is not resistant to oscillated spikes and reduces the valid NDVI value in August. Thus, the 5-span running median and arithmetic average alone is not enough to resist all of these noises, while combining the maximum operation to them significantly improves the result and works well for a 1-point spike, 2-point spike, and oscillated spikes, as is shown in Figure 3.3b.

The two final steps include endpoints processing and weighted moving averaging. Since the initial and final value of the NDVI time-series are not surrounded by enough other values, they cannot be smoothed in the same way as the above method. The strategy of extrapolation (Velleman, 1981) is used here to achieve the new end values.

Weighted moving averaging (WMA), a traditional linear smoother, is safely implemented as the last step since the outliers have been removed in the first three steps. The basic principle of WMA is to replace each data value with the weighted average of the points around it. Here the classical smoother hanning (Tukey, 1977) with weights – 1/4, 1/2, 1/4 - is used.

The whole procedure, including the above five steps, is denoted by RMMEH, in which R represents running median, the first M represents the mean value, the second M represents maximum operation, E represents endpoints processing, and H represents hanning. This method ensures that the reconstructed time-series is smooth and of high quality.

Results

The data of Figure 3.3a are used to test the sensitivity to different kinds of noise of RMMEH and two other smoothing algorithms: 4253EH, twice and MVI. 4253EH, twice is a compound smoother that combines several elementary smoothers by both resmoothing and reroughing (Velleman, 1981). The 4253EH, twice, as a revised and improved version of T4253H (Pallant, 2011), has been successfully used to smooth and reconstruct AVHRR NDVI time-series (Dijk et al., 1987). The ability of the 4253H, twice filter to handle complex time-series while eliminating spurious drops and spikes was also praised by both Velleman (1981) and van Dijk et al. (1987). Hird and McDermid (2009) compared six popular smoothing techniques and found that the 4253EH, twice performs excellently, especially when no distinct negative bias is present. MVI is based on the assumption that any marked increase or decrease of NDVI value is

identified as noise and is smoothed by the mean value of the previous and following point (Ma & Veroustraete, 2006); this smoothing technique is not only simple, but quite effective in reconstructing temporally and spatially more homogeneous NDVI time-series. These two cutting edge smoothing techniques are selected to compare with RMMEH in this research based on their simplicity and great performance. All of these three algorithms are coded in IDL and embedded in ENVI as a smoothing package.

Figure 3.4 shows the result of the comparison among the three methods. For the MVI methods, the threshold for identifying spurious spikes is set as 0.1. Of the three methods, the 4253EH, twice can produce the smoothest profile and resist the 1-point spike, but it is very sensitive and not resistant to oscillated spikes. Most original valid points in June, August, and September are heavily influenced and show large deviation from the correct values. Furthermore, the oscillated spikes are not correctly removed. The MVI has the least smoothed profile because of the relative high threshold value of 0.1. Generally, the lower the threshold value, the smoother the result, but the result will deviate from the correct curve, and is also extremely time-consuming when a low threshold is set. The MVI is effectively preserving the original valid points and also successfully identifying and removing most of the spikes. Compared with RMMEH, both 4253EH, twice and MVI are not as resistant to a 2-point flat spike as RMMEH, although they improve the result a little. It can be seen that RMMEH is the least sensitive to the three kinds of noises. The result of RMMEH successfully keeps the original valid points and corrects the spurious spikes, and the result is smoother than that of MVI.

A pixel level comparison of these three methods is carried out for a more detailed assessment. We apply the new method to the 16-day composite MODIS NDVI time-

series data of year 2005, which included 23 layers and covered the central part of the state of Utah. This arid and semiarid region consists of a lot of ecosystem types, including arid desert shrub communities, high-elevation Douglas-fir forests, pinion-juniper woodland, and open grassland. Most of the annual precipitation in this region occurs as snowfall during winter months. The result of a field survey in 2005 was used to select pixels with different land covers.

Figure 3.5 illustrates the original and smoothed NDVI time-series for 10 pixels with different land covers, which include some dominant classes: annual grass, cropland, forests, pinion juniper, and sagebrush. The results show that all three methods are effective at removing most noisy points and produce similar smooth curves in most pixels, and the pattern of temporal evolution is apparent.

As stated above, 4253EH, twice produces the smoothest curves for each land cover, but it suffers from the problem of high sensitivity to oscillated spikes, which causes high deviation from original NDVI values. This problem is clearly illustrated by comparing the two pixels of annual grass; the two pixels are the same species and should have similar temporal evolution. For the pixel without much noise (in the right column of Figure 3.5), all three methods produce almost the same results. For the pixel with oscillated noises (in left column in Figure 3.5), the results of both MVI and 4253EH, twice cannot successfully approach the true temporal curve, while the result of RMMEH approaches the true temporal curve. This phenomenon is also obvious in the two pixels of grassland.

The MVI method still produces the least smoothed curve for each pixel. This method is sensitive to both the threshold value and the extremely low spikes. In the case

of cropland pixel, since the threshold value of 0.1 is not high enough, the values of some original valid points which neighbor to spikes are reduced greatly. In short, the performance of this method highly depends on the threshold value while the optimal threshold value is difficult to estimate.

For the RMMEH method, the reconstructed NDVI time-series approaches the upper NDVI envelope, and most of the noisy points are identified and corrected. The result shows that RMMEH not only keeps the original valid NDVI points and effectively remove these spurious spikes, but produces a smoother curve than MVI.

Conclusions

The NDVI time-series data are widely used in large-scale research in monitoring vegetation phenology, land cover/use change analysis, and biophysical and biochemical parameters extraction. Therefore, the quality of the NDVI time-series products is of great importance, and a smoothing algorithm is required to reduce the residual noise and reconstruct high-quality smooth NDVI time-series. Although many smoothing algorithms have been proposed, each has its own advantages and drawbacks that limit its use. A novel smoothing algorithm called RMMEH is presented in this paper to overcome some drawbacks of the other algorithms. This new method, as well as the two other algorithms 4253EH, twice and MVI, is coded and embedded as a smoothing package in ENVI.

By comparing the performance of RMMEH in smoothing 16-day composite MODIS NDVI time-series with two other methods, we have found the following advantages of this new method: (1) Since no ancillary data is required, we can avoid the possible uncertainties that lie in ancillary data; (2) RMMEH is applicable automatically

and no parameter is required in the whole procedure, so we avoid the difficulty finding optimal parameters like MVI algorithms; (3) RMMEH is very resistant to different noises, while the two other algorithms are sensitive to more kinds of noise; and, (4) RMMEH is essentially a compounded smoother that combines several elementary smoothers and operations, so it is very simple in theory and easy to implement.

Results show that RMMEH not only keeps the original valid NDVI points, but effectively removes these spurious spikes. The reconstructed NDVI time-series of different land covers are of higher quality and have smoother temporal evolution. Testing the potential improvement of using this improved NDVI time-series for invasive species (cheat grass) mapping will be described in a future study.

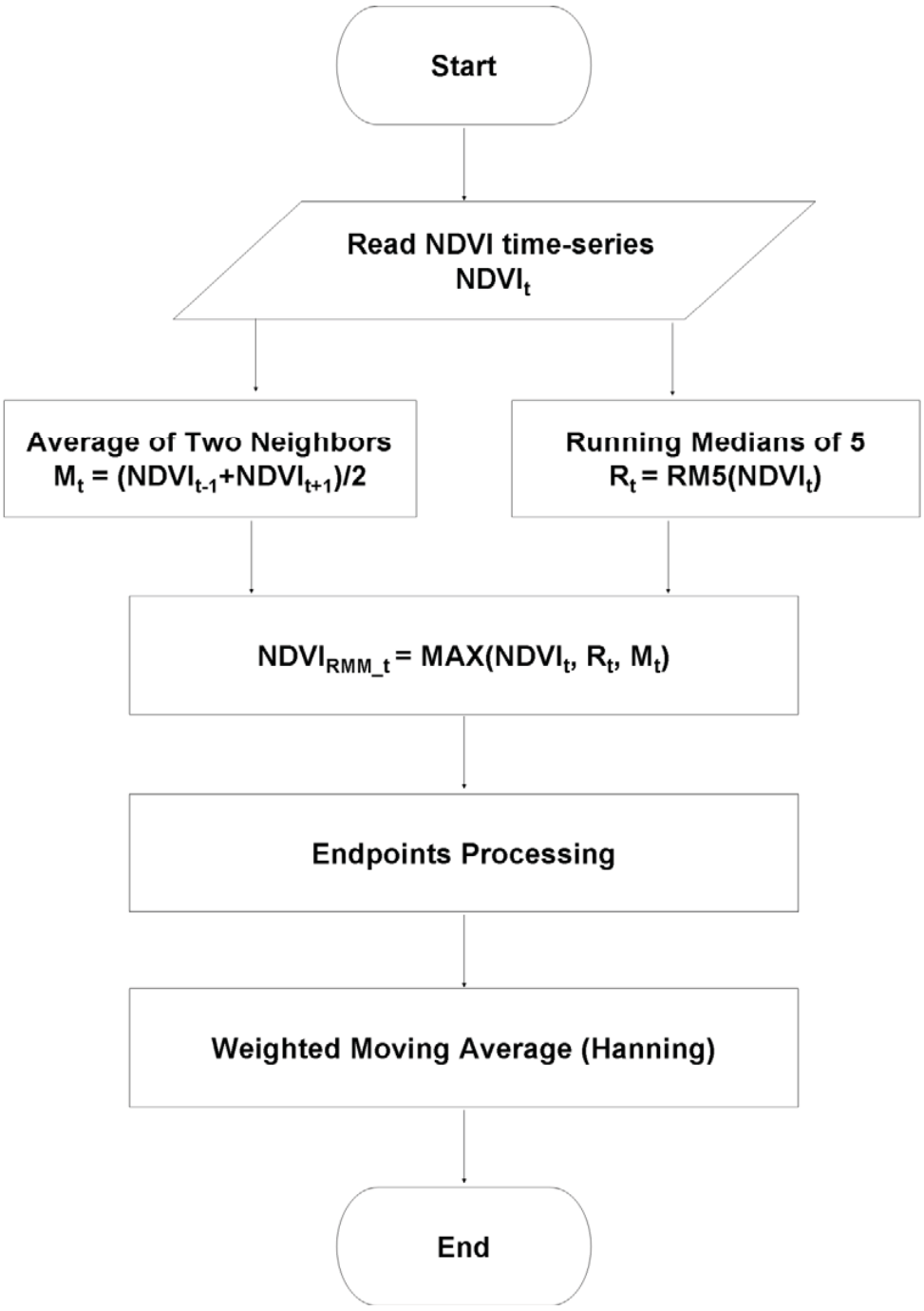


Figure 3.1. Flowchart of RMMEH method.

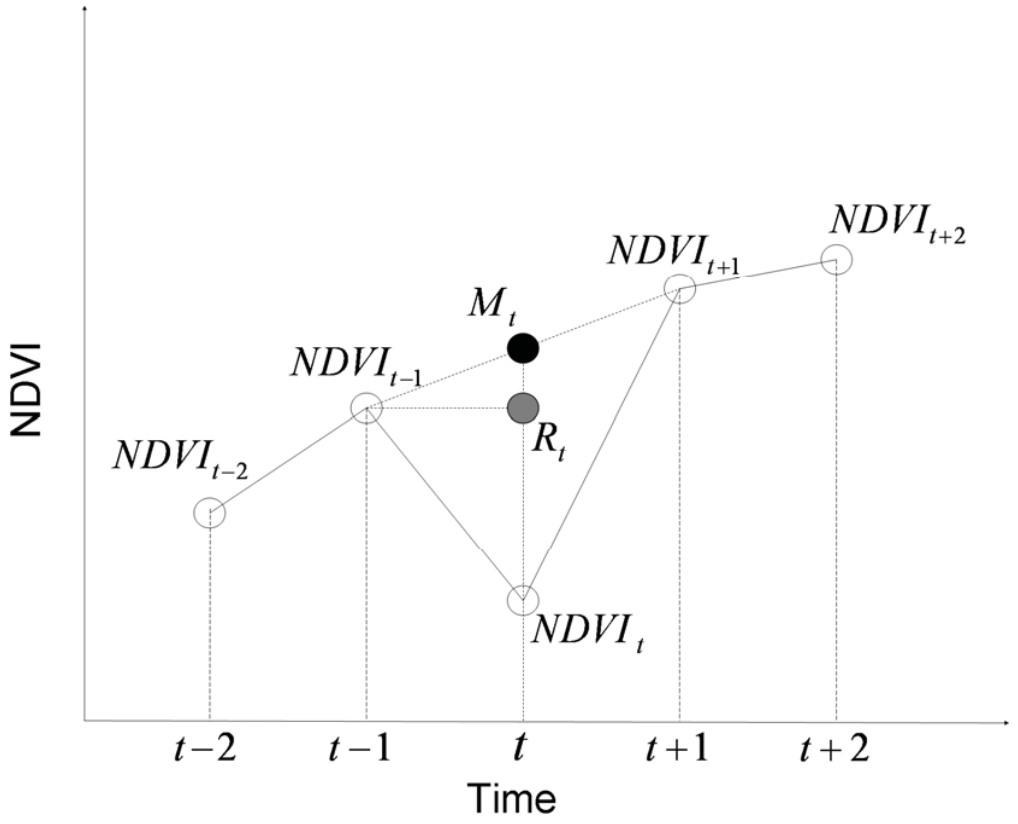


Figure 3.2. Geometric interpretation of $NDVI_t$, R_t, and M_t.

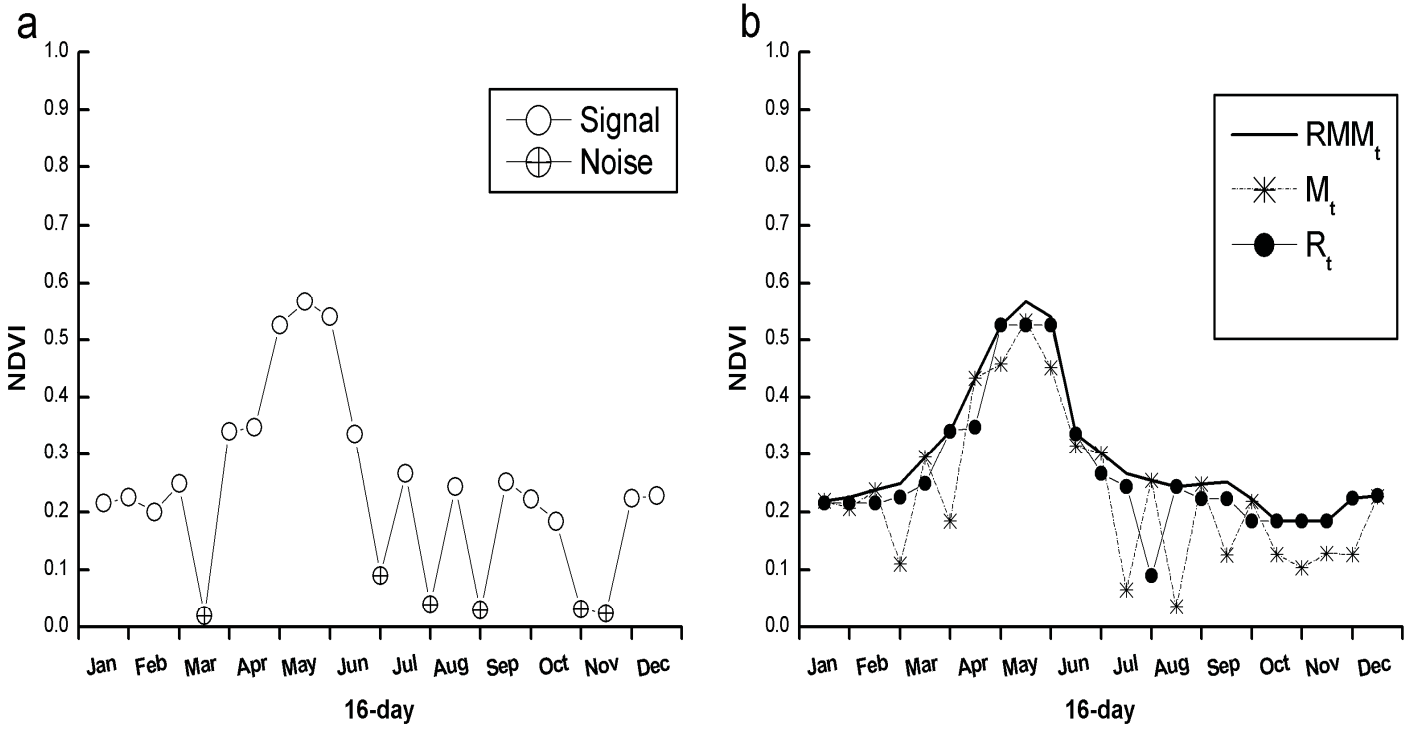


Figure 3.3. One-year NDVI time-series over an annual grass pixel with cloud contamination.

(a) Raw NDVI temporal profile, (b) Results of R_t, M_t, and RMM_t.

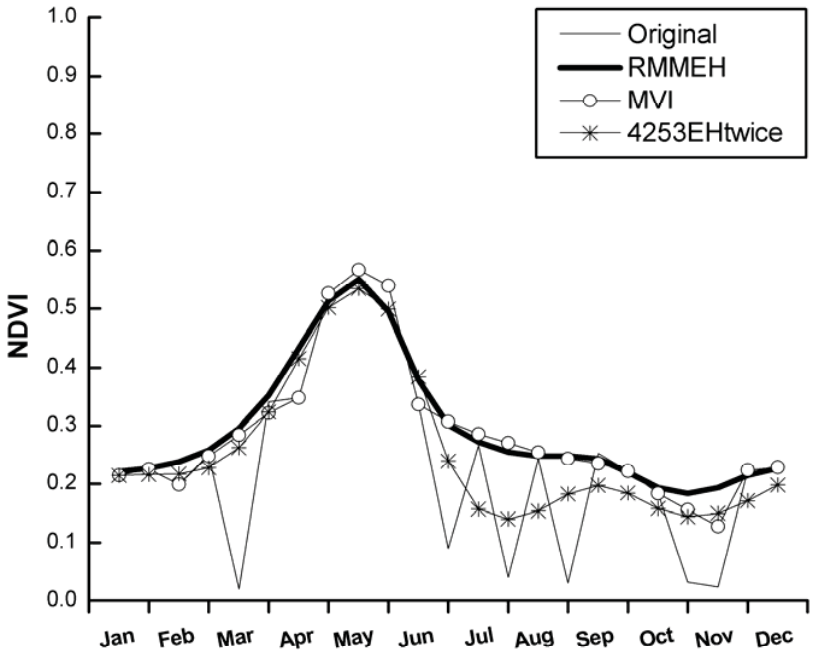


Figure 3.4 Sensitivity analyses of three methods.

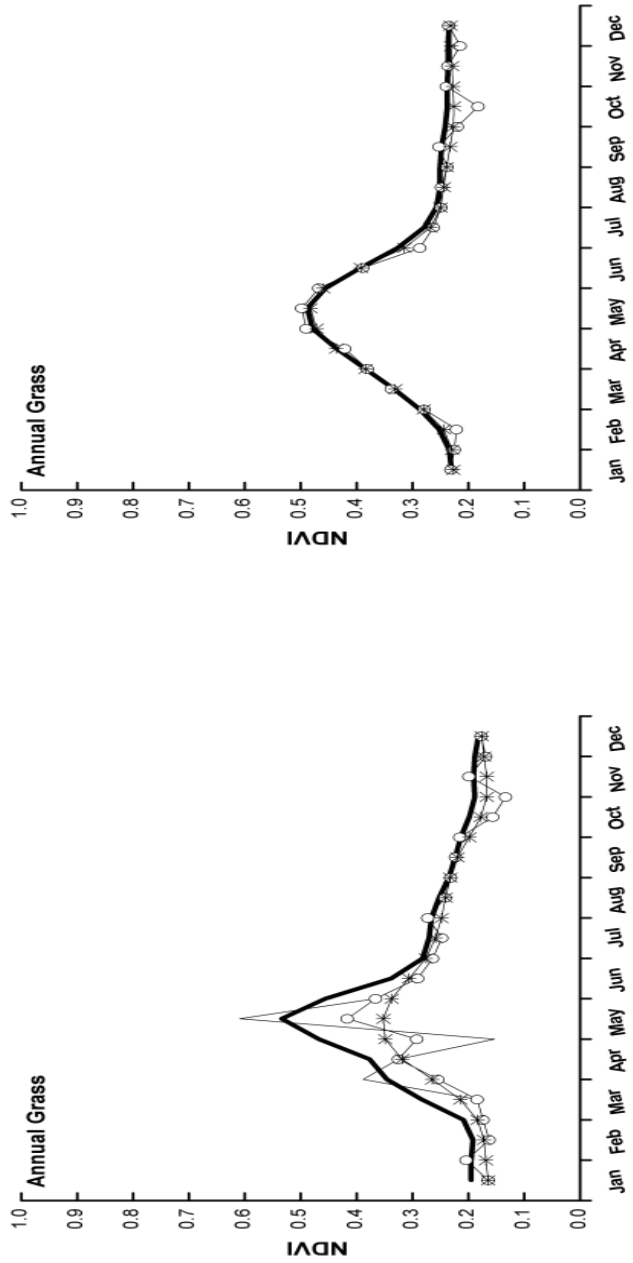


Figure 3.5. Comparison of three methods with different land covers.

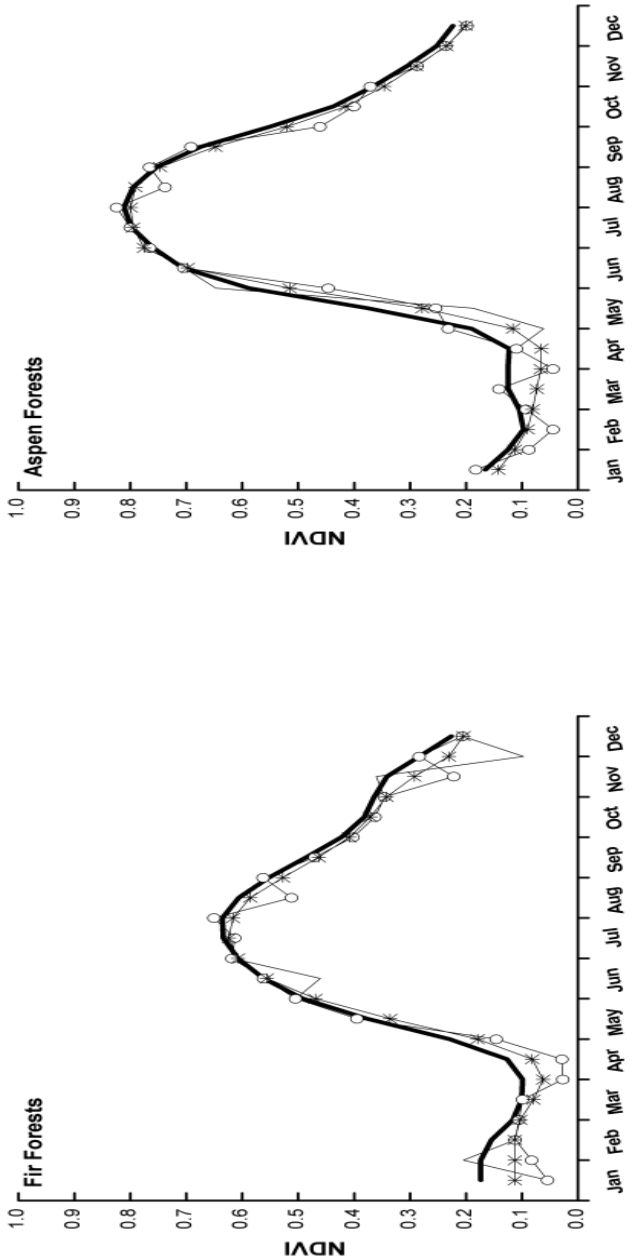


Figure 3.5. Continued.

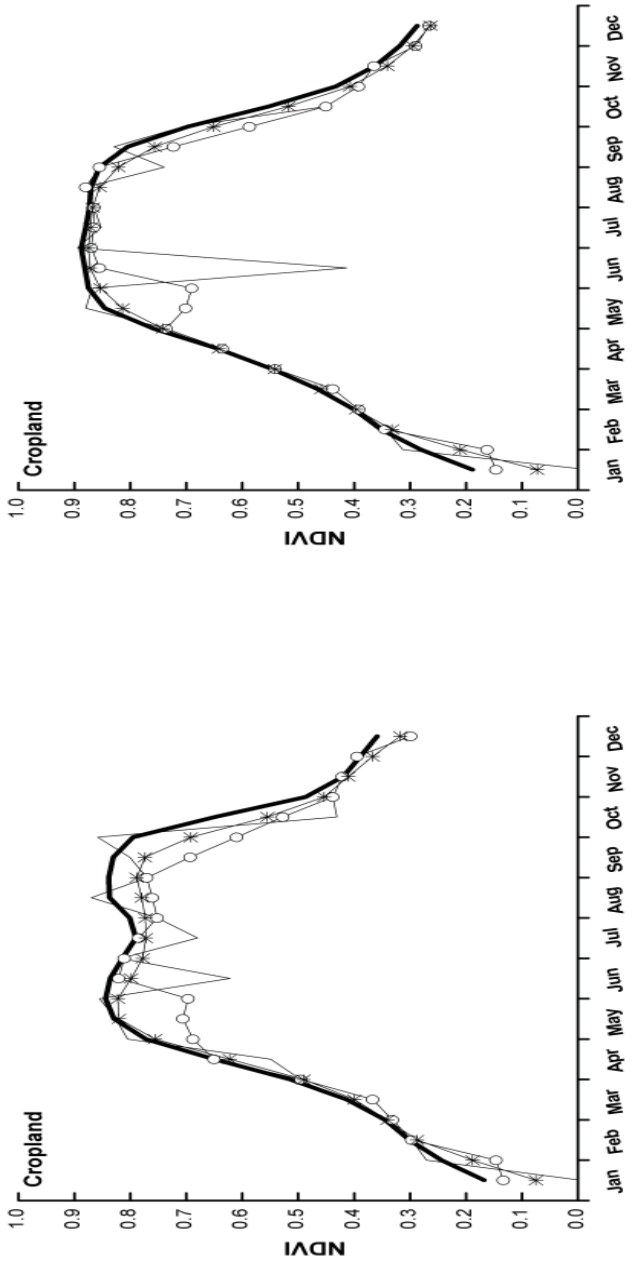


Figure 3.5. Continued.

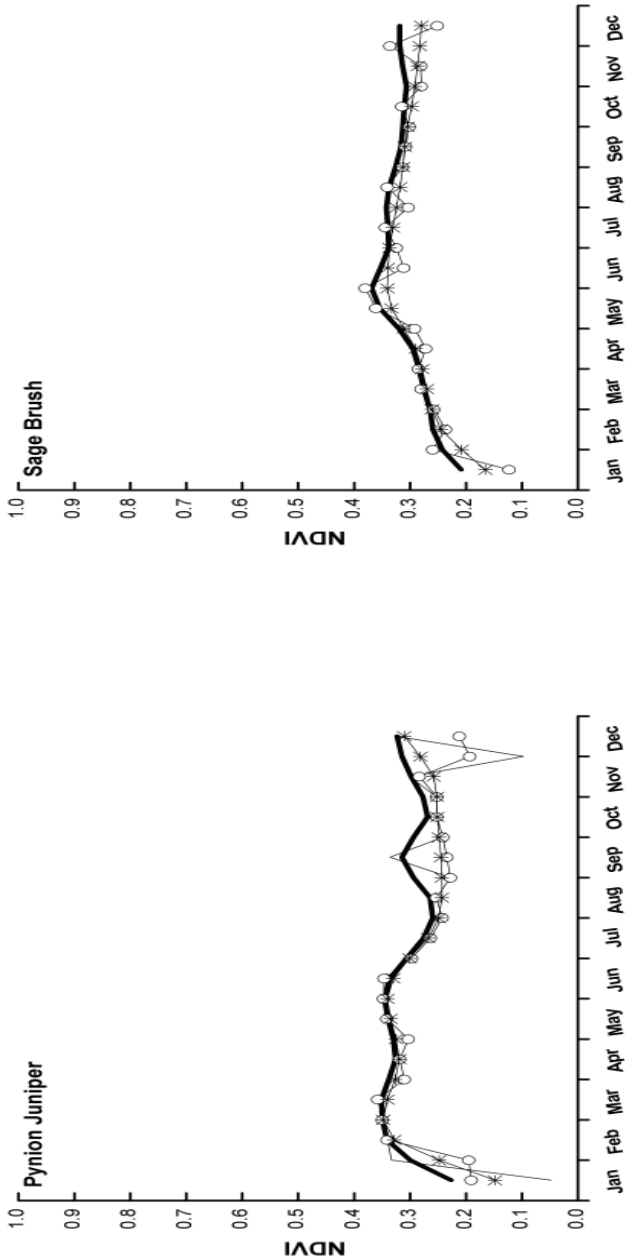


Figure 3.5. Continued.

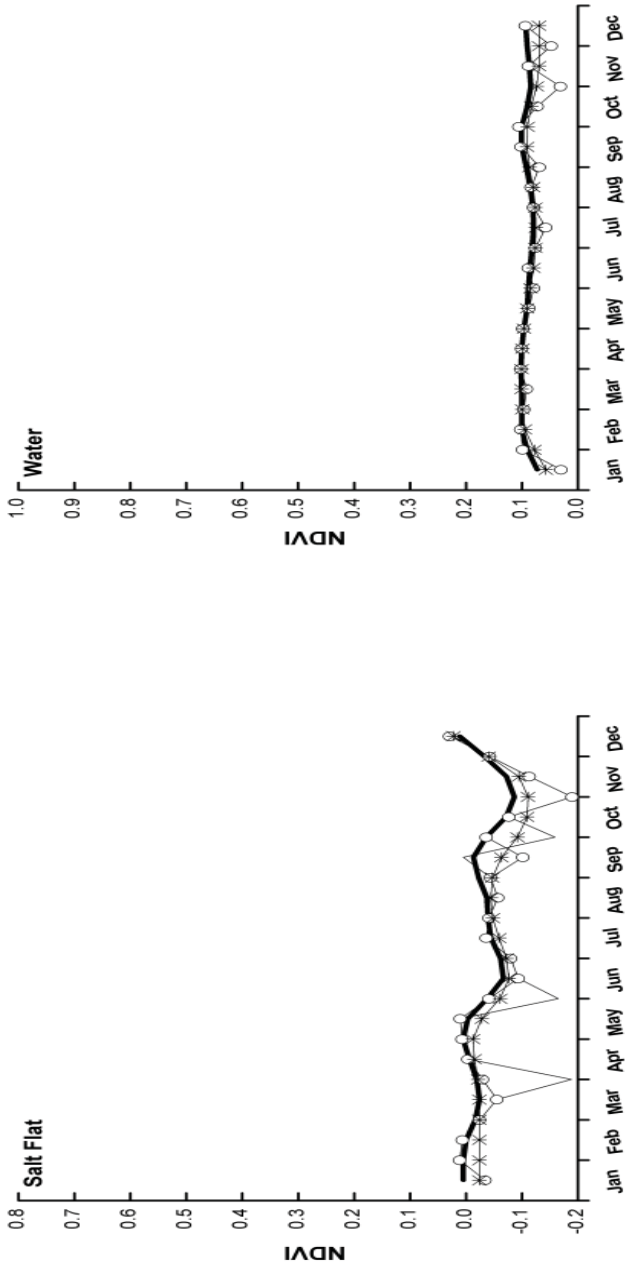


Figure 3.5. Continued.

CHAPTER 4

SPATIALLY EXPLICIT POPULATION-BASED MODELING OF CHEATGRASS

Introduction

Biological invasions are characterized by remarkable spatial-temporal dynamics, with many species having extended their distribution range from within a single region to much of the globe within the last century (Weber, 2003). It is estimated that invasive species are a major cause of biodiversity loss, accounting for as much as 20% of recent extinctions (Vitousek et al., 1997). Growing awareness of this problem has stimulated research on the establishment and expansion of alien species, and on the environmental consequences of invasive species introductions (Carroll, 2007). The aim of this study is to increase the understanding of the cheatgrass invasion in Utah by modeling this phenomenon at both the spatial and temporal level.

Models are an important tool for the analysis of complex interactions in ecological systems. Typically, models are a simplified and abstract representation of a system or a process. Most authors agree that we are simply not able to get the complete information of all processes within an ecological system. So models are necessary to simplify the reality and bring it into manageable form for investigation. It is commonly

agreed upon that models are in the first place useful to create and test scientific hypotheses by comparing real data with model results.

There are two classical model types that are widely used: statistical models and processed-based model. The statistical models usually attempt to construct a relationship between the dependent variables and the independent variables based on collected data using multivariate analyses, these models are data-driven models (Crawley 2007). A linear regression, for example, can be used to estimate the leaf area index with satellite vegetation index data. Process-based models are either detailed or simplified mathematical models such as differential or difference equations, which are based on representation of basic principles of the physical process to be investigated.

In modeling biological invasion, a variety of models have been developed during the last decades. In ecology, Skellam's (1951) seminal paper stimulated the application of diffusion approximations of population expansion, while the limitations in obtaining analytical solutions of this description led to development of new paradigms. Spatially explicit population models are becoming increasingly useful tools for population ecologists, conservation biologists, and land managers (Dunning et al., 1995). Models are spatially explicit when they combine a population simulator with a landscape map that describes the spatial distribution of landscape features. In general, spatially explicit models are of interest because these models give ecologists a technique for studying ecological processes that operate over different scales, and are useful in predicting population and community responses to phenomenon such as climate change or land-cover and land-use change.

Spatially explicit models can be either individual-based or population-based (Wadsworth et al., 2000). In the individual-based models, the location of each individual across the landscape is monitored, and individuals acquire fitness characteristics associated with the cell type they occupy. The status of each individual is followed through an entire simulation. In models with an annual time step, individuals can undergo an annual cycle of breeding, dispersal, and mortality (Wadsworth et al., 2000). This technique demands extensive knowledge of the behavior mechanism of each individual and the interaction between individuals, which may not be practical. In population models, each cell contains a population. This may be the most appropriate modeling strategy for some invasive species like cheatgrass in models where it may be difficult to follow each individual in a large population. In these models, patch-specific reproduction and mortality are measured by population growth rates, while movement between patches is measured by immigration and emigration rates.

In this research, because little quantitative data are available on the cheatgrass species ecology, a spatially explicit population-based model was developed to simulate the spatial-temporal dynamics of cheatgrass in a cheatgrass-dominant grassland ecosystem; the interspecific competition was neglected by selecting such simple ecosystems without competition. This model mainly includes two parts: a spatial and a temporal part. The temporal part is a difference equation explaining how the population of cheatgrass evolves in each grid over time. The rules and parameters of this model are based on the available information from previous experimental work, as well as on new experimental data particularly collected for this model development. Although the complete quantitative parameters required for model are unavailable, the constrained

Monte Carlo method (Purkis, 2008) is used to achieve the distribution for some unknown parameter. The second part is the spatial part which models the process of how cheatgrass seeds disperse with the wind, and should answer the question of how cheatgrass spread in space; the model was then evaluated with cheatgrass percentage data from two different years. Using this model, it should be possible to determine the relative importance of various factors for the spatial-temporal dynamics of cheatgrass, such as disturbance and climate change.

Spatially Explicit Population-based Model

Logistic Population Modeling

Population modeling is the application of mathematical models to the study of population dynamics (Machault, 2011). It has been successfully used to model animals, viruses, disease, vegetation, etc. The major aim of this section is to construct a population model to determine the abundance of cheatgrass in cheat grass-dominant grassland within each plot or cell. In doing so, the life cycle of cheatgrass is investigated in Figure 4.1.

Figure 4.1 shows that the cheatgrass only reproduces using seed. There are three main stages during the life cycle of cheatgrass: seed production, seed germination, and seedling growth. These three stages of cheatgrass can be simulated using classical logistic population models as described by equation 4.1:

$$N_{t+1} = N_t * P_s * f * g * e * (1 - N_t/K) + N_t * P \quad \ldots\ldots 4.1$$

N_{t+1}: Number of adult cheatgrass in time t+1

Nt: Number of adult cheatgrass in time t

f: Number of seeds per cheatgrass plant

Ps: The probability of seeds surviving

g: The probability of seed germination

e: The probability of seedling to adult cheatgrass

K: The carrying capacity for cheatgrass

P: The probability of an adult surviving next year

Since cheatgrass is an annual grass, the probability of an adult cheatgrass surviving next year is almost zero, although in some extreme cases some cheatgrass will live 2 years (Pellant et al., 1999; Pierson et al., 1990). So the equation 4.1 could be simplified as equation 4.2:

$$Nt+1 = Nt * Ps * f*g*e*(1-Nt/K) \quad \text{.....4.2}$$

Estimates of the population density of cheatgrass over a several-year period made various estimates of fecundity and survivorship during the seed and vegetative stages of the cheatgrass life-cycles (Young et al., 1969). These are listed in Table 4.1. Combing these values allows the finite rate of population increase (R) to be calculated (i.e., R = Ps*f*g*e). We provide an example of how theoretical and experimental data may be used to examine the various influences that affect the population dynamics of a plant.

In equation 4.2, the number of populations, N, at time t+1, is related to those at time t by the finite rate of population increase (R) and density dependent K.

The carrying capacity of a biological species in an environment is the maximum population size of the species that the environment can sustain indefinitely, given the food, habitat, water, and other necessities available in the environment (Hui, 2006). Usually, the carrying capacity may vary for different species and may change over time due to a variety of factors, including food availability, water supply, environmental conditions, and living space. Many biophysical and biochemical parameters limit its value. In the case of cheatgrass, climate, elevation, soils, and some other biophysical factors influence the value of carrying capacity. There is a clear inverse relationship between the cheatgrass population size and the probability of cheatgrass seedling to adults. This relationship is due to the intraspecific competition for various resources; for instance, when the population size is above the carrying capacity, there will be the highest competition among the cheatgrass seedlings attaining the available resources like water and nitrogen; because the space and resources determine the carrying capacity and limit the final population size, the probability of cheatgrass seedling to adults has to be decreased so that the cheatgrass population size can finally reach the population equilibrium.

Population Model Calibration

Although we have a range of values for most of the parameters, the probability of a seedling growing to adult cheatgrass and carrying capacity for cheatgrass is not available; therefore, a constrained Monte Carlo simulation is developed to achieve the distribution of these two parameters.

The observed cheatgrass percentage in eight cheatgrass-dominant plots in northern Utah collected by Utah Range Trend Studies (http://wildlife.utah.gov/range/) in the years 2001 and 2006 (Figure 4.2) are used to calibrate the model using constrained Monte Carlo simulation. These plots are located south of Ogden County, Utah. Each plot number identified in Utah Range Trend Studies and its position in UTM projection are described in Table 4.2. Dominant vegetation species in all eight plots are cheatgrass only, so we do not need to involve the interspecific competition in our model. In Figure 4.2, we could clearly see that compared with other plots, the cheatgrass percentages of the test plots 3, 4, and 5 do not change much between 2001 and 2006. Since these test points are all located in south Ogden County and the distance between each one is a given range of distances, we could easily neglect the difference of influence of the physical factors like temperature and precipitation. Therefore, more details of the descriptions of these three plots (http://wildlife.utah.gov/range/) without large cheatgrass percentage change are investigated and alternative explanations are proposed: First, two of these points have really low cheatgrass percentages which are less than 1%, and although the absolute difference is small for these two plots between 2001 and 2006, the relative difference is similar compared with the other plots; Second, the remaining one plot is placed in a burn, therefore it is reasonable that the cheatgrass percentage difference does not change significantly between 2001 and 2005, even if it decreases.

In Figure 4.2, the cheatgrass percentages of 8 field plots with a size of 3x3 meters are recorded in the years 2001 and 2006. These plots are mostly occupied by cheatgrass and soil, so the interspecific competition could be neglected here. It is clearly shown that within 5 years, the invasion of cheatgrass occurred in every plot to different extents.

Combined with equation 4.2 and the experimental data, the constrained Monte-Carlo model was developed as the following steps to approximate the distribution of these two parameters:

1. The value for unknown parameter e and K was randomly taken from two uniform distributions. The distribution for the probability of seedling to adult cheatgrass is [0, 1.0] and a [500, 1500] uniform distribution is used to extract the value for carrying capacity at each time-step.
2. The initial value Nt used in this equation is the percentage values of the 8 field observations in 2001. Starting with this value, this model ran for 5 generations because of the 5-year interval from 2001 to 2006.
3. A total of 100,000 iterations are calculated and those values which satisfied the following two criteria are accepted as reasonable approximations of the two parameters:
    a. Correlation coefficient between the predicted and observed percentage values of year 2006 for the 8 plots is larger than 0.8
    b. The mean absolute error (MAE) between the predicted and observed percentage values of year 2006 is less than 3%

$$\sum_{i=1}^{n} \left| predicted_i - actual_i \right| / n < 3\%$$

The correlation coefficient measures the strength of the linear association between two variables, and the square of the correlation coefficient determines whether the regression between predictions and observations could successfully explain a significant portion of the variation. A threshold value of 0.8 for the correlation coefficient is selected

in this research, mainly because a significant correlation coefficient value of 0.8-1 between the observed and predicted variable is usually acceptable and satisfactory for a good fit of the model, either in vegetation modeling or remote sensing modeling (Peterson, 2005; Clinton et al., 2009; Laura et al., 2009; Frankenberg et al., 2011).

Figure 4.3 shows the results of the space that satisfies these two criteria. It seems that about 50 out of 100,000 iterations meet the two criteria and the two parameters cluster in the center of (0.137, 1000).

Figure 4.4 shows the comparisons predicted with observed percentage values in 2006 for the 8 field plots. It can be clearly seen that the predicted values have a good fit with observed values. For the 8 field plots, all of these predicted have a high correlation (R = 0.89) with observed values, although the cutting-off point value in the parameter selections is only 0.80 (Figure 4.5).

The constrained Monte-Carlo approach was successfully developed in this research to address a common situation in the study of biological invasion: the ecology of the species has not been quantified properly. The present approach yielded a statistically valid explanatory model for a grid-based population calculation, providing insight into this complex pattern.

Spatial Dispersal of Cheatgrass Seeds

Plant dispersal has been addressed using a variety of theoretical approaches, including reaction-diffusion, integro difference, random-walk, and simulation models, leading to a rich theory for estimating rates of spread (With, 2002; Higgins et al., 2003). Models and detailed biological studies have shown that it is the characteristics of the

dispersal kernel that determine the rate at which plants can spread spatially when introduced into new environments (Hengeveld, 1994).

Most cheatgrass seeds fall to the soil surface near the parent plant, or are spread short distances by wind or water. Long-distance dispersal is facilitated by humans and wild and domestic animals (Hulbert, 1955). Usually, the spatial dispersal of plants is modeled by equally dispersing the seeds into all directions according to species' specific exponential function (Hengeveld, 1994). In this research, a fixed disperse kernel is used to determine the rate of cheatgrass spread with a fixed spatial resolution of 3 meters assuming that the cheatgrass seeds fall into the neighborhood within a distance of about 6 meters (Figure 4.6).

In Figure 4.6, d is denoted as the percentage of seeds dispersed into the 8 neighbor pixels and each pixel accepts d/8% of cheatgrass seeds. So 1-d describes the percentage of seeds staying within the center pixel. The d value is set as 0.4 according the description of Pierson (1990a).

Figure 4.7 describes how the cheatgrass seeds within a cell in the center of the 100x100 cells are dispersed into its neighbor cells. It can be seen that it will take about 70 years for cheatgrass to fill the space. This is somehow contrary to the field observation in which the cheatgrass invades faster with the other long-distance dispersal by humans and animals mostly.

Model Evaluation

In this part, the spatial-temporal model is evaluated by comparing the model prediction with ground truth of cheatgrass percentage in a field site in central Utah.

Two maps of cheatgrass percentage in 2006 and 2009 produced with National Agricultural Imagery Program (NAIP) natural color aerial photography in 1-meter resolution were used in this research. The imagery was collected in summer (Http://gis.utah.gov/aerial).

The study area is located in central Utah. It covers an area of 300mx300m from approximately 39°9'40.20"N, 112°8'15.40"W to 39°9'30.71"N, 112°8'2.76"W. This area was intensively investigated during a field trip in June 2006, and was mostly occupied by cheatgrass with different densities. An unsupervised classification is used to classify these two images into two classes: cheatgrass and open land. The open land may also consist of some dead wood within the area of interest. It is defined as a land cover type consisting of recently burned areas and barren soil/rock.

Figure 4.8 shows the flowchart of model evaluation. Methods for the model evaluation were investigated using cheatgrass percentage images of 3-meter resolution, derived by resampling the 1-meter resolution classification results of aerial photography (Figure 4.9.a).

An IDL program was used to produce new image files for cheatgrass area fractions within each new 3x3m pixel, based on the known pixel information from the classification map of the aerial photography of the years 2006 and 2009 (Figure 4.9.b). This step is most vital since the whole algorithm could not be derived without the knowledge of cheatgrass percentage.

The calibrated spatial-temporal model was then used in the cheatgrass percentage product of year 2006 to predict the cheatgrass percentage of year 2009. Since we have a range of values for parameter E and K, it is therefore possible to construct a stochastic

model for the cheatgrass percentage prediction. For simplicity, a random value was taken for each of the parameters from the distribution given in Figure 4.3. Starting with the resampled cheatgrass percentage product of the year 2006 (Figure 4.9.c), and running the model for 3 generations with 1,000 replications, allows the behavior of the model to be compared with the resampled cheatgrass percentage product of 2009. The arithmetic average of the 1,000 replications was the final predicted percentage map for 2009 (Figure 4.9.d).

After using the model to predict the cheatgrass percentage of year 2009, the R^2 and MAE are used to evaluate the performance of the model (Table 4.3). It could be easily seen that the predicted cheatgrass percentage of 2009 has higher significant R^2 and lower MAE with the ground truth than compared with the cheatgrass percentage of 2006. Although the MAE value does not decrease significantly, these results demonstrate that the spatial-temporal model could greatly simulate the spatial-temporal dynamics of cheatgrass in simple cheatgrass-dominant environment.

Cheatgrass Management

After the calibration and evaluation, the model could then be used to analyze the effect of cheatgrass management.

Cheatgrass spread rapidly over the past century and is now found in all 50 states, but is most prevalent in Idaho, Nevada, and Utah, where it has turned massive and once flourishing rangelands into arid, desolate fields. Nearly 17.5 million acres in Idaho and Utah are infested by cheatgrass (Haferkamp et al., 2001). Increasing by an estimated 20% each year, cheatgrass also threatens significant areas in Colorado, Kansas, Montana,

North Dakoto, Oregon, South Dakoto, Washington and Wyoming (Diamond, 2009). As its name implies, cheatgrass cheats landowners, and ranchers from earning the full economic benefit of their land by displacing native plants, reducing biodiversity and spreading fires (Haferkamp et al., 2001).

According to Cornell University researchers (Pimentel et al., 1999), invasive species such as cheatgrass cost the nation some $138 billion annually in ecosystem damages, reduced yields, lost forest products, and control efforts. A study conducted in the mid-1990s showed that cheatgrass was costing wheat farmers in the West an estimated $350 to $375 million in lost yield and control costs annually. And in a 2001 study (Haferkamp et al., 2001), cattle that grazed on cheatgrass-infested rangeland gained 2.0 to 2.3 pounds less per head per day than cattle that grazed on non-cheatgrass-infested land.

Managing and controlling cheatgrass can often be difficult and costly, depending on the measures used. In general, the integration of chemical management tools with cultural practices is recommended for successful control of cheatgrass (Mennlled et al., 2008). Because cheatgrass seeds usually die in the seed bank after several years, preventing seed production in the spring reduces the number of seeds in the soil which may improve the outcome of integrated management. Practices like prescribed grazing, prescribed fire, irrigation management, and nutrient management will help maintain the vigor of desirable species and prevent an increase of cheatgrass, but these methods are limited and can be costly and not wholly effective (Diamond, 2009). So far, No biological control agents currently exist for control of cheatgrass, but researchers continue to investigate soil fungi and bacteria as potential agents (Mennlled et al., 2008).

Prescribed grazing is a possible management strategy to partially control cheatgrass, but there are significant costs (Diamond, 2009). First, a great number of animals need to be fenced within a small area for a brief amount of time, requiring labor and fencing supply costs. Then, the animals must be moved often to prevent damage to desired vegetation and provide sufficient feed, requiring additional labor costs. Even if all of this is done correctly, a cheatgrass plant still needs to be clipped off at least 12 times to stop seed head formation. This is nearly impossible, even under intensive grazing.

The long-term ability to control cheatgrass through another alternative — prescribed burning — is uncertain, due to the tendency of cheatgrass to actually proliferate in postburn areas and reduce the prevalence of desirable plants. However, prescribed burning can be appropriate if used to remove thatch and followed by herbicide treatment and reseeding (Diamond, 2009).

In a recent study (Diamond, 2009), one simulation was designed to compare the cost-effectiveness of using cattle grazing and herbicide to create fuel breaks on cheatgrass-dominated landscapes in the northern Great Basin. It shows that targeted grazing and the herbicide Plateau® had similar reductions in flame length and rate of cheatgrass spread. Cattle grazing had high fixed costs (primarily fencing), and was more cost-effective than applications of Plateau® under five fuel loadings.

Since the naturally occurring disturbance activities of animals and fire are essential for the stability of ecosystems (Jentsch et al., 2008), a disturbance function simulating fire that randomly deletes cheatgrass in cells and thus creates free cells (i.e open land) at a given frequency (each year) is implemented into the model as a management strategy of prescribed fire.

Assuming a fixed reasonable dispersal distance for *cheatgrass* of 3 meters, a 100x100 cheatgrass percentage image with value from [0,50] is created in the spatial-temporal model to investigate the effect of prescribed fire.

The rate of the prescribed fire probability is varied from 0, 3%, 10%, and 30%; accordingly, rankings of no fire, low fire rate, medium fire rate, and high fire rate are assigned. In each time step, the same percent of probability of disturbance of the image is randomly selected, and the values of these selected cells are set to 0.

Figure 4.10 shows the impact of different fire rates on the average cheatgrass percentage of the image.

The y-axis shows the average cheatgrass percentage of the image. The x-axis is a time axis for a 200-year simulation time. With a fire rate of above 10% per year (medium fire rate), the average cheatgrass percentage will converge to zero. The high fire rate prevents the invasion of cheatgrass by creating 10-30% of space to open land in each year. With a lower fire rate of 3%, the average cheatgrass percentage converged to a value of 50%, i.e. a cover of about half of the space. The average cheatgrass percentage will approximate to about 80% if no fire is allowed each year. These results demonstrate that the higher the fire rate each year, the lower the cover of cheatgrass in the space, which seems to be common sense.

Conclusions and Discussions

In this research, a spatially explicit population-based model is built upon the knowledge gained from literature review and experimental data to simulate the spatial-temporal dynamic of cheatgrass. Compared with the popular Agent-Based Model (ABM),

which is a bottom-up model, the population-based model cannot approach the true reality of how the population growth is dependent on individual behavior. However, this top-down model could be used to model the growth of the population without understanding the low-level processes and elements which generate the aggregate system behavior. The construction of this spatially explicit population-based model in this research could help us understand how spatial dispersal of invasive species seeds influences the invasion, and answer what-if questions for invasive species management and controlling.

Although the estimated values of most parameters in the logistical population model could be found in literature and fixed values for these known parameters are utilized in the model, several uncertainties have not been considered. In this model, all vital rates of the population and its growth rate are fixed and independent, while in reality, those parameters depend not only on density of the population, but on the environments and disturbances (Begon, 1990). Usually, the growth rate decreases with environmental stress or increased population. Therefore, the fixed values of these parameters used in this research only approach the mean values under different environments and disturbances, rather than the true values at each time step. A sensitivity analysis may be useful to investigate the sensitivity of model output to the change of each parameter value.

A constrained Monte-Carlo simulation is used in this paper to calibrate the model and address the issue of unknown parameter estimates. The correlation coefficient and MAE are both used as criteria to find the distribution of two parameters. The correlation coefficient, R, measures the strength of the linear relationship between predictions and observations while the MAE measures how close the predictions are to the observations.

A high-correlation coefficient (0.9) and low MAE (3%) strongly support that the model has a good fit under selected parameter values. The result of the clustered distribution of the two unknown parameters (Figure 4.3) also shows that the Monte-Carlo method is a successful technique in such situations.

Since the complete exponential seed dispersal function is not available for cheatgrass, a simple kernel-based spatial dispersal assuming uniform wind speed and direction is adopted based on descriptions about the seeds dispersal (Hengeveld, 1994; Higgins et al., 2003). The result of simulation is surprising in that it takes almost 70 years for cheatgrass to spread from its origin to a neighborhood of 90,000 square meters. This relatively slow invasion may be due to the fact that no long-distance dispersals by humans and animals have been included in this model.

Examples shown in this research to simulate the spatial dispersal of cheatgrass are limited to a small area of 90,000 square meters with a specific spatial scale of 3x3 meter. This homogeneous cheatgrass area, which includes mostly cheatgrass, soils, rocks, and little sagebrush, is a typical and representative place with cheatgrass invasion in the Western U.S. The experiment at this cheatgrass-dominant place makes it possible to evaluate the constructed model and apply this model to other similar ecosystems.

The evaluation is a measure of degree of how probable the model is to represent the true reality; it usually uses the fit between the model output and the in-situ data to measure the degree; if more in-situ data matches well with the model out-put, we could say that the model constructed is more likely to approximate the reality. In this research, the coefficient of determination (R^2) and MAE are used to evaluate the model. The R^2 is used to determine whether the regression between predictions and observations could

successfully explain a significant portion of the variation. The F-test is performed to test the significance. Table 4.2 shows that the R^2 between predicted and observed cheatgrass of 2009 not only are highly significant ($p<0.01$), but improved significantly from 0.51 to 0.57 compared with the R^2 between observed cheatgrass of 2006 and 2009. The MAE decreases from 7.31% to 6.94%, which means a closer relationship between predicted and observed values. Both of these two measurements demonstrate that the spatial-temporal model could greatly simulate the spatial-temporal dynamics of cheatgrass in simple cheatgrass-dominant environments.

In this model, no physical or human impacts have been included. However, the cheatgrass growth is significantly influenced by many factors like climate, elevation, soils, and topography. For instance, the soil nutrient level plays an especially important role in determining the biomass and competitive ability of cheatgrass. Previous studies show that the increasing soil Nitrogen level leads to greater competitive ability of cheatgrass compared to native bunchgrasses, implying that it may be possible to favor desired plant communities by longer-term reductions in soil N (Sperry et al., 2006; Vasquez et al., 2008; Mazzola et al., 2008). The implications of these studies also explain why cheatgrass growth is favored along roadways. Because nitrogen is a limiting nutrient in arid and semi-arid desert, nitrogen input to the roadside ecosystem has serious implications for species composition. In an arid environment, the nitrogenous gas pollutants released in car exhaust can settle along the roadside through dry deposition and the plot closer to the road usually has a higher Nitrogen level (Trusott, et al., 2005), thus leading cheatgrass to be more invasive along roadways with higher soil N. This explanation is also consistent with the finding that roads are always conduits for exotic

plant invasions in a semiarid landscape (Gelbard et al., 2003). A more complete and accurate model which considers the distance of cheatgrass to the roadside could absolutely capture more details of the potential physical mechanism and approximate the reality with more reasonable results.

In this research, the effect of different disturbance rates to the cheatgrass invasion is investigated. The result (Figure 4.10) shows that higher fire rate prevents the invasion of cheatgrass while lower fire rates have the opposite effect. In consequence, extremely high or medium fire regimes lead to extinction of this invasive species. One interesting finding is that a relatively low or medium fire rate will lead to a state of equilibrium, and the equilibrium value increases when the fire rate decreases. It could be clearly seen that a medium fire rate of 10% can successfully prevent the invasion of cheatgrass. This finding implies that management activities for preventing invasions, such as quarantines or early eradication, can be benefited by such disturbance analysis.

Although the values of parameter e and k in this model are taken randomly from their distribution at each time step, the models are more deterministic than chaotic in that other parameters are all fixed values, as noted above. While plant invasions are influenced by both the biological characteristics of the invasive species and the environmental characteristics of the sites being invaded, these parameters are always dependent on the change of these factors. That is to say, these parameters may change at each time step with a change in temperature, moisture, or other environmental or biological factors. These factors often interact, allowing successful invasions to occur.

Two limiting effects of choosing this small selected field area are acknowledged. First, although the field area has similar ecosystem and land cover with those plots used

for model calibration, it can only represent one specific cheatgrass-dominant landscape, which is not quite pervasive even in the Intermountain West confronting such serious cheatgrass infestation. Second, since this field area occupies a small area, the variance of the cheatgrass percentage may therefore not be large enough compared with the real situation. In other words, this field area may probably not be representative enough to validate and evaluate the model performance; however, this difficulty could be overcome by selecting more homogeneous cheatgrass areas outside of this research area. Due to these two limiting effects, the results of this research could be applicable in other cheatgrass-dominant homogenous areas, most probably in the Intermountain West like Nevada, Wyoming, and Arizona. However, the concept of this spatial-temporal model could be applied in other places with invasive species throughout the world.

Stages of the invasion process are also influenced by interspecific competition. In this model, due to the limitation of information for demographic parameters of different species which coexist with cheatgrass, no competition mechanism is implemented in the present simulation model. That is also another reason that a simple cheatgrass-dominant grassland landscape is selected. In the future, it would be worthwhile to carefully include other plant species in order to extend the analysis towards more complex situations, and intensive field work is required to more accurately simulate the cheatgrass invasion in future research.

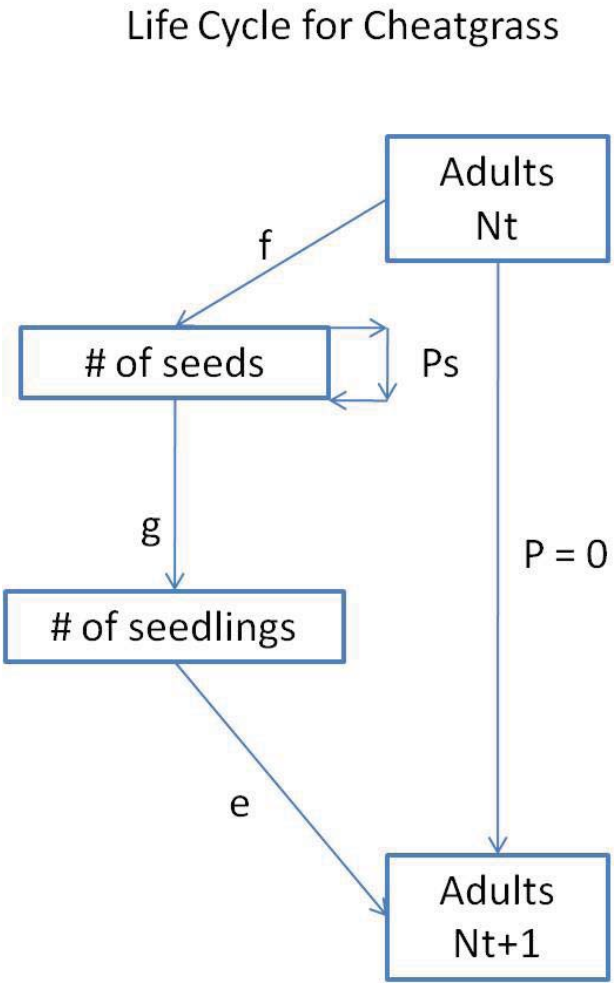


Figure 4.1. The life-cycle for cheatgrass.

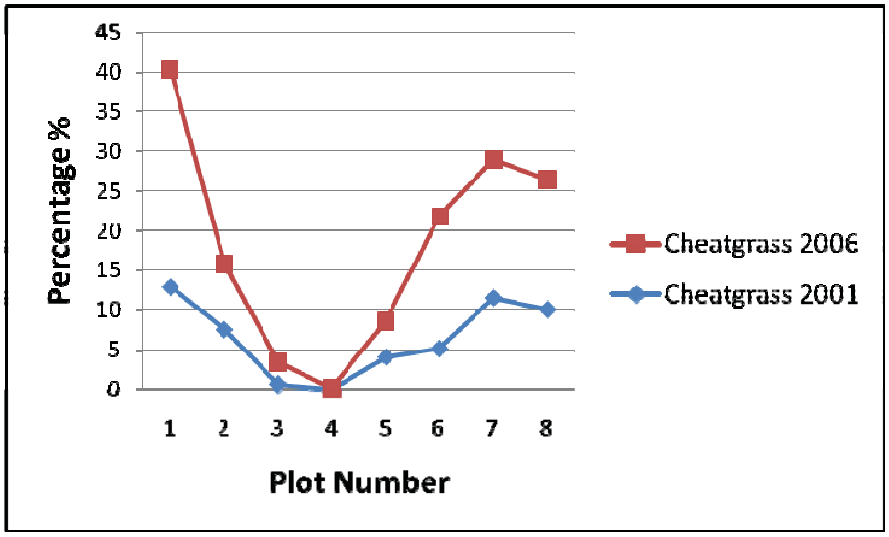


Figure 4.2. Observed cheatgrass percentage in 2001 and 2006.

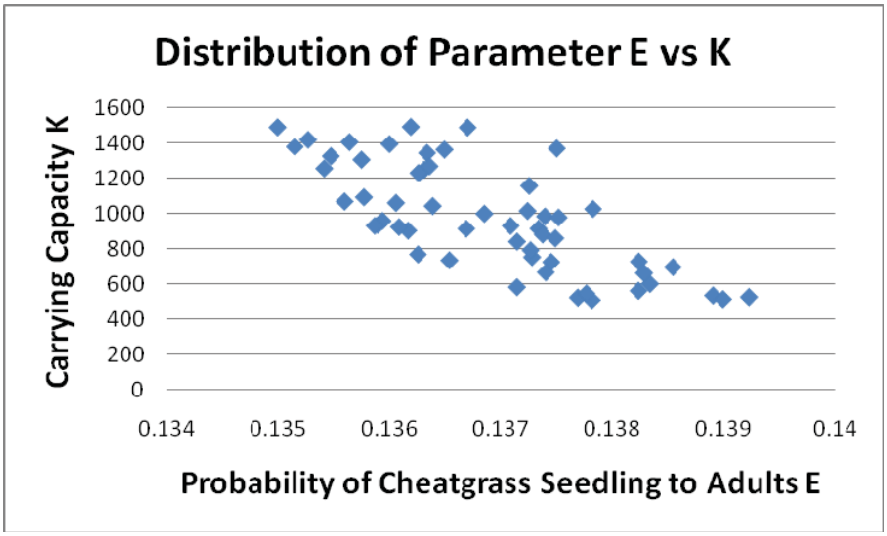


Figure 4.3. The distribution of parameter E vs. K.

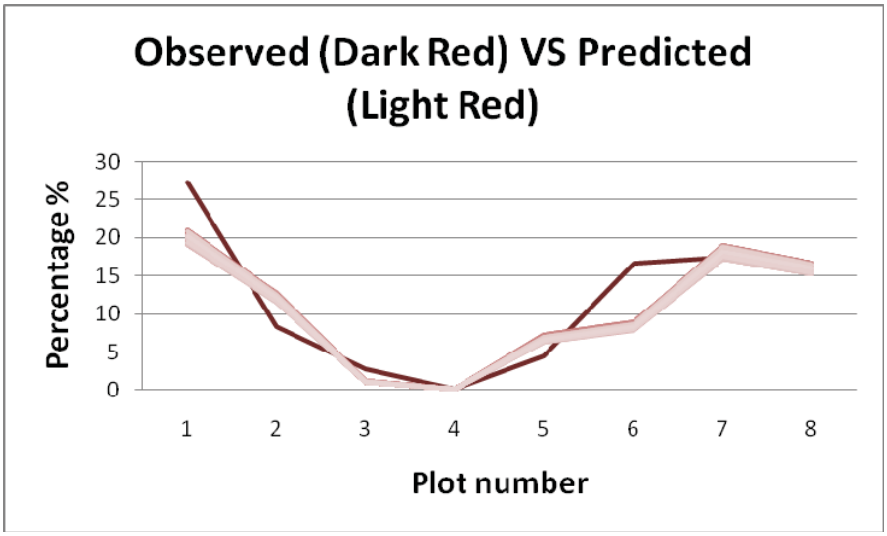


Figure 4.4. Comparisons of predicted with observed percentage (Year 2006).

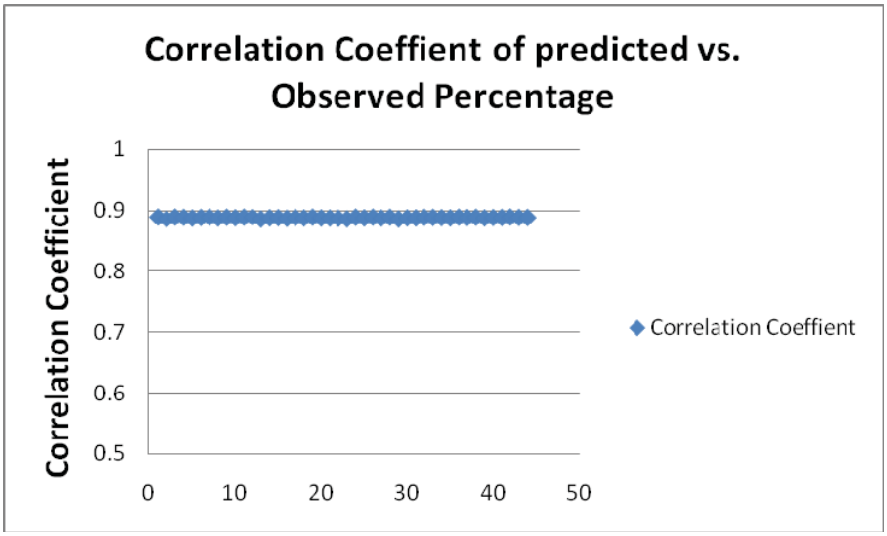


Figure 4.5. The correlation coefficient of predicted vs. observed percentage.

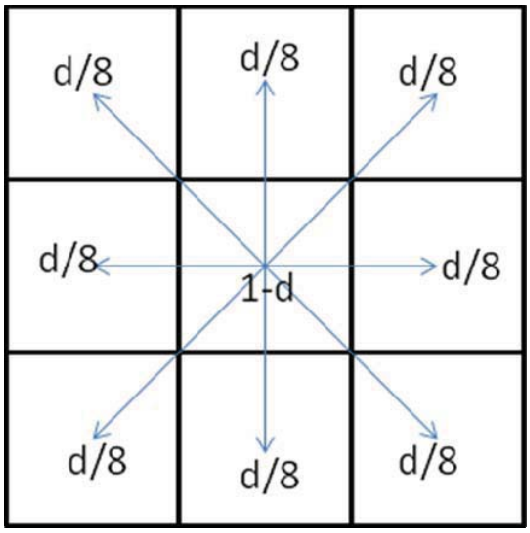


Figure 4.6. The dispersal kernel of cheatgrass.

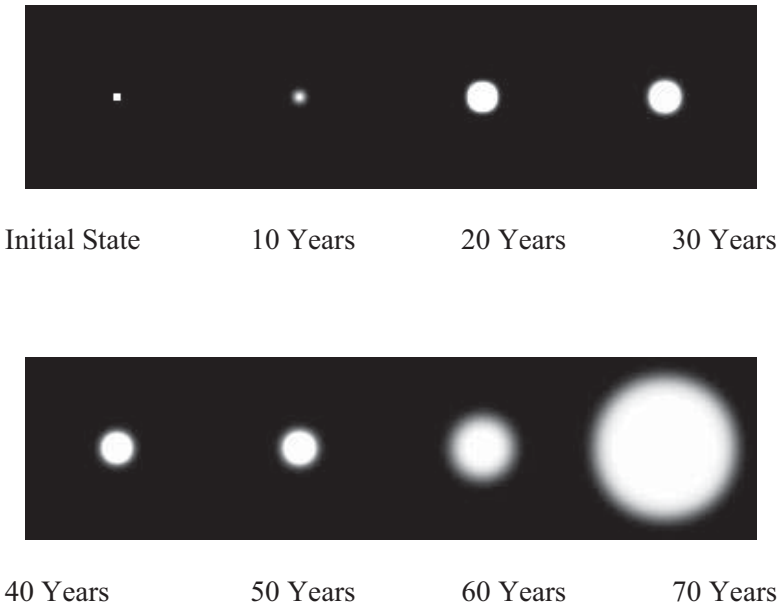


Figure 4.7. The demostration of cheatgrass spatial dispersal.

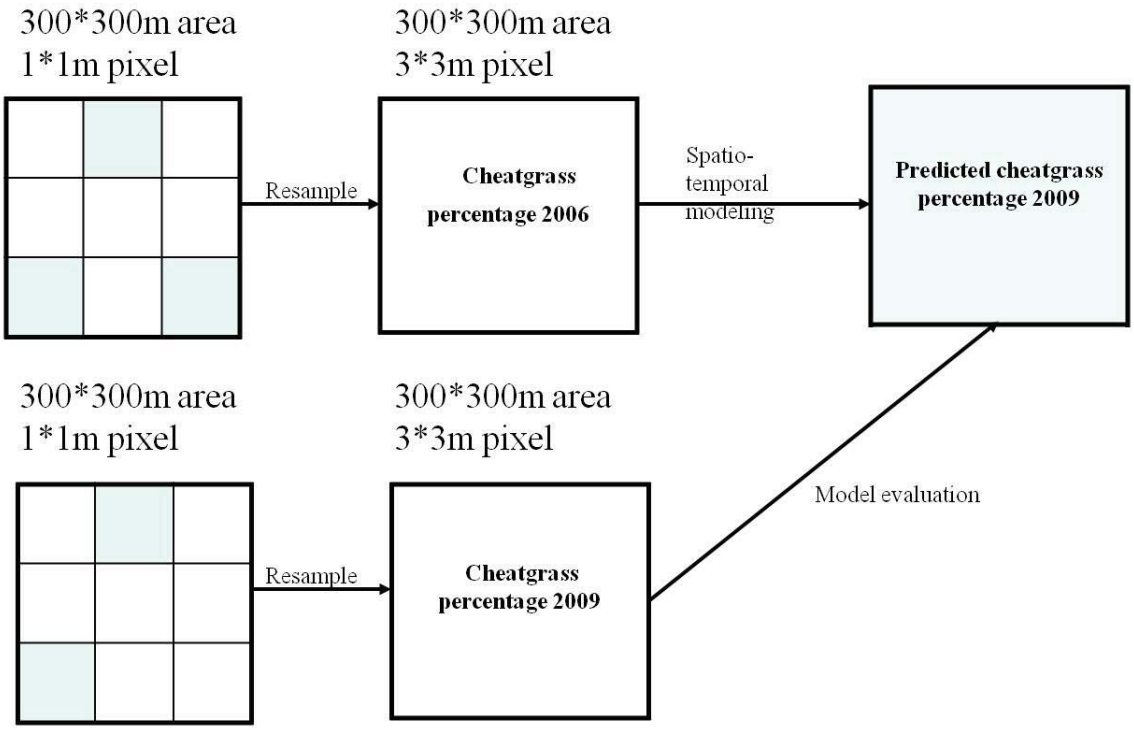


Figure 4.8. Flow chart of spatial-temporal model evaluation.

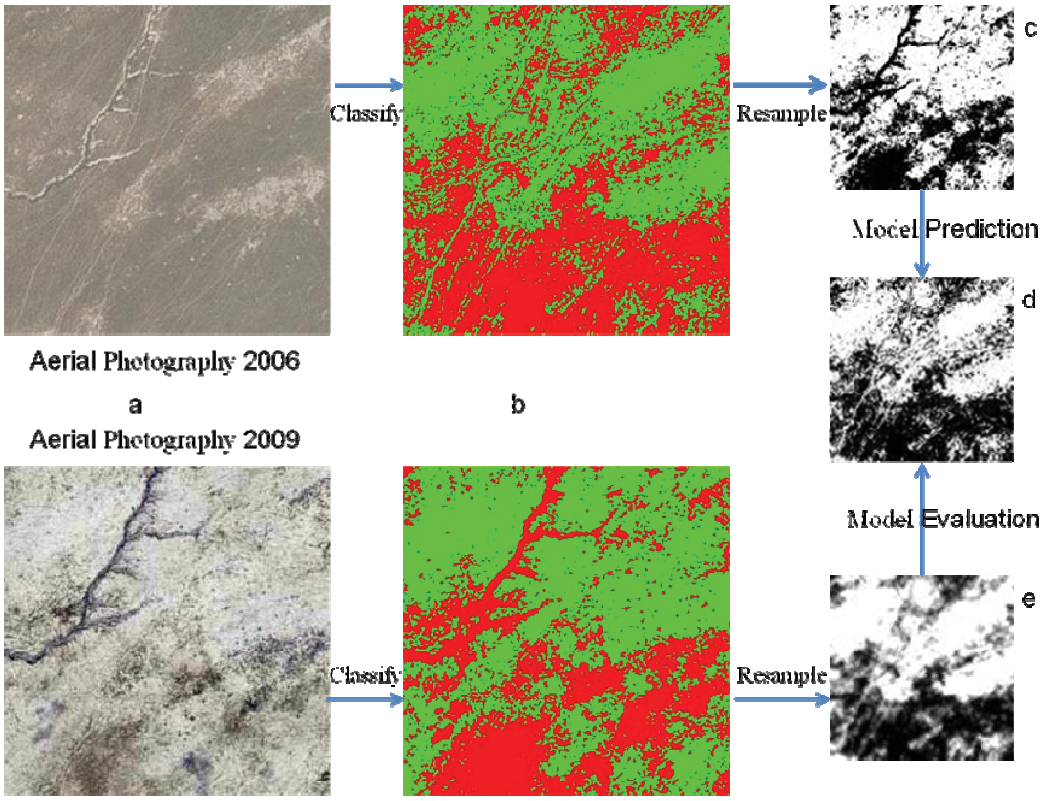


Figure 4.9. The aerial photography in 2006 and 2009.

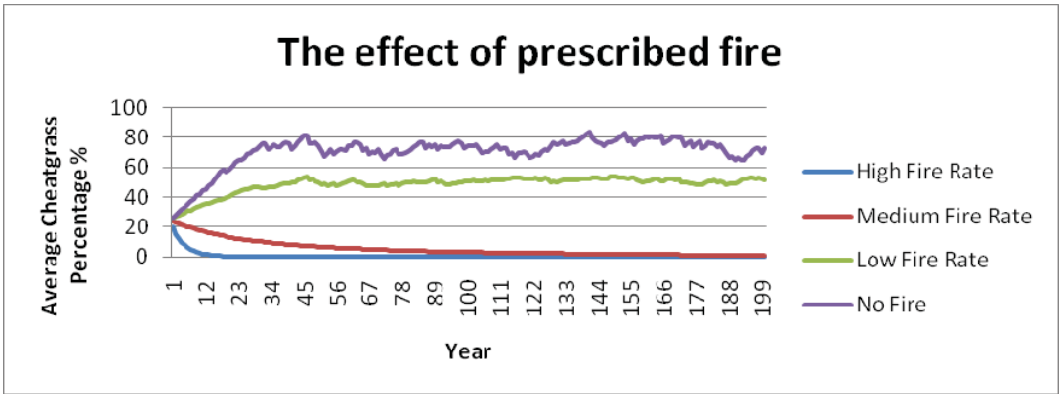


Figure 4.10. The effect of prescribed fire. The y-axis shows the average cheatgrass percentage in the space whereas the x-axis shows the time step in years (up to 200 years).

Table 4.1 Estimates of demographic variables for different stages in the life-cycle of cheatgrass.

Variable	Value	Source
f	10	Young et al., 1969; Hulbert, 1955
Ps	0.95	Young et al., 1969; Hulbert, 1955
g	0.85	Young et al., 1969; Hulbert, 1955
e	Unknown	No
K	Unknown	No

Table 4.2 Utah range trend studies plot number and position with UTM projection.

Utah Range Trend Studies Plot Number	East (Meter)	North (Meter)
12	266042	4597734
51	423322	4596798
52	425426	4593957
56	435053	4573966
58	464007	4538406
61	459759	4544767
67	459327	4552803
86	478254	4496275

Table 4.3 The R^2 and MAE between cheatgrass percentage of 2009 and the cheatgrass percentage of 2006, and the predicted cheatgrass percentage of 2009

	Cheatgrass Percentage of 2006	Predicted cheatgrass percentage of 2009
Cheatgrass Percentage 2009 (ground truth)	MAE = 7.31% R^2= 0.51 *	MAE = 6.94% R^2= 0.57 *

* Significance ($P < 0.01$)

REFERENCES

Adams, J. B., & Smith, M. O. (1986). Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 site. *Journal of Geophysical Research, 91* (88), 8098-8112.

Adams, J. B., Sabol, D. E., Kapos, R. A., Roberts, D. A., & Smith M. O. (1995). Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment, 52,* 137-54.

Baret, F., & Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment, *Remote Sensing of Environment, 35,* 161-173.

Bates, J. D., Miller, R. F., & Svejcar, T. J. (1998). Understory dynamics in a cut juniper woodland (1991-1997). *Annual report: Eastern Oregon Agricultural Research Center. Corvallis, OR: Oregon State University, Agricultural Experiment Station,* 24-33.

Begon, M., Harper, J. L., & Townsend, C. R. (1990). *Ecology: Individuals, populations, and communities.* Cambridge, Blackwell Scientific Publisher.

Blank, R. R. (2009). Biogeochemistry of plant invasion: A case study with downy brome (Bromus tectorum). *Invasive Plant Science and Management 1,* 226–238.

Borel, C. C., & Gerstl, S. A. W. (1994). Nonlinear spectral mixing models for vegetative and soil surfaces. *Remote Sensing of Environment, 47*, 403–417.

Bradley, B. A., & Mustard, J. F. (2005). Identifying land cover variability distinct from land cover change: Cheatgrass in the Great Basin. *Remote Sensing of Environment, 94,* 204-213.

Bradley, B. A., & Mustard, J. F. (2006). Characterizing the landscape dynamics of an invasive plant and risk of invasion using remote sensing. *Ecological Applications, 16*, 1132-1147.

Breiman, L. F., Olshen, J. R., & Stone, C. J. (1984). *Classification and regression trees.* New York: Chapman & Hall.

Butera, C. (1986). A correlation and regression analysis of per cent canopy closure versus TMS spectral response for selected forest sites in the San Juan National Forest, Colorado. *IEEE Transactions on Geoscience and Remote Sensing, 24*, 122–129.

Carroll, S. P. (2007). Natives adapting to invasive species: Ecology, genes, and the sustainability of conservation. *Ecological Research, 22,* 892–90.

Chander, G., & Markham, B. (2003). Revised Landsat-5 TM radiometric calibration procedures and postcalibration dynamic ranges. *IEEE Transactions on Geoscience and Remote Sensing, 41*, 2674-2677.

Chavez, P. S. (1996). Image-based atmospheric corrections—revisited and improved. *Photogrammetric Engineering and Remote Sensing, 62,* 1025–1036.

Chen, J. (2004), A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sensing of Environment, 91,* 332-344.

Chen, J. M. (1996). Retrieving leaf area index of boreal conifer forests using Landsat TM images. *Remote Sensing of Environment, 55,* 153-162.

Chen, J. M. (1999). Spatial scaling of a remote sensed surface parameter by contexture. *Remote Sensing of Environment, 69,* 30-42.

Clinton, N. E., Gong, P., Jin, Z., Xu, B., & Zhu, Z. (2009). Meta-prediction of Bromus tectorum invasion in Central Utah, U.S.A. *Photogrammetric Engineering & Remote Sensing. 75,* 689-701.

Clinton, N. E., Christopher, P., Bob, C., Vanessa, G., Peggy, G., & Gong, P. (2010). Remote sensing based time-series analysis of cheatgrass (Bromus tectorum) phenology. *Journal of Environmental Quality, 3,* 955-963.

Cohena, W. B., Thomas, K., Maierspergerb, S. T., & Gowerc, D, P. (2003). An improved strategy for regression of biophysical variables and Landsat ETM+ data. *Remote Sensing of Environment, 84,* 561–571.

Concannon, D. (1978). *Plant succession on burned areas of the Artemisia tridentata/Agropyron spicatum habitat type in southeastern Oregon.* Unpublished manuscript.

Crawley, M. J. (2007). *The R book.* Wiley.

Curran, P. J., & Hay, A. M. (1986). The importance of measurement error for certain procedures in remote sensing at optical wavelengths. *Photogrammetric Engineering and Remote Sensing, 52,* 229–241.

DeFries, R., Hansen, M., Steininger, M., Dubayah, R., Sohlberg, R., & Townshend, J. (1997). Sub-pixel forest cover in central Africa from multisensor, multitemporal data. *Remote Sensing of Environment, 60,* 228-246.

Diamond, J. M. (2009). *Effects of targeted grazing and prescribed burning on fire behavior and community dynamics of a cheatgrass (Bromus tectorum) dominated landscape.* Unpublished manuscript.

Dijik, A., Callis, S.L., Sakamoto, C.M., & Decker, W.L. (1987). Smoothing vegetation index profiles: An alternative method for reducing radiometric disturbance in NOAA/AVHRR data. *Photogrammetry Engineering and Remote Sensing, 53,* 1059-1067.

Ditomaso, J. M. (2000). Invasive weeds in rangelands: Species, impacts, and management. *Weed Science, 48,* 255-265.

Dunning, J. B., Stewart, D. J., Danielson, B. J., Noon, B. R., Root, T .L., Lamberson, R. H., & Stevens, E. E. (1995). Spatially explicit population models: Current forms and future uses. *Ecological Applications, 5,* 3-11.

Eastwood, J. A., Yates, M. G., Thomson, A. G., & Fuller, R. M. (1997). The reliability of vegetation indices for monitoring saltmarsh vegetation cover. *International Journal of Remote Sensing, 18* (18), 3901–3907.

Elmore, A. J., Mustard, J .F., Manning, S .J., & Lobell, D .B. (2000). Quantifying vegetation change in semiarid environments. Precision and accuracy of spectral mixture analysis and the normalized difference vegetation index. *Remote Sensing of Environment, 73*, 87–102.

Ferrero, S. A., Hild, A., & Mealor, B. (2009). Can invasive species enhance competitive ability and restoration potential in native grass populations? *Restoration Ecology, 19* (4), 545-551.

FICMNEW (Federal Interagency Committee for the Management of Noxious and Exotic Weeds). (1997). *Invasive plants – changing the landscape of America.* Washington, D.C.

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, S., Lee, J., & Toom, G. (2011). New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters, 38,* L17706, doi:10.1029/2011GL048738.

Gelbard, J. L., & Belnap, J. (2003). Roads as conduits for exotic plant invasions in a semiarid landscape. *The Journal of the Society for Conservation Biology, 17,* 420-432.

Gurusiddaiah, S., Gea, Y., Kennedy, A. C., & Ogg, A. G. (1994). Isolation and characterization of metabolities from the Pseudomonas fluerescens strain D7 for the control of downy brome (Bromus tectorum L.). *Weed Science, 42,* 492-501.

Haining, R. (1990). The use of added variable plots in regression modeling with spatial data. *The Professional Geographer 42,* 36-45.

Haferkamp, M. R., Grings, E. E., Heitschmidt, R. K., Macneil, M. D., & Karl, M.G. (2001). Suppression of annual bromes impact rangeland: Animal responses. *Journal of Range Management, 54,* 663-668.

Hengeveld, R. (1994). Small-step invasion research. *Trends in Ecology and Evolution, 9,* 339-342.

Higgins, S. I., Nathan, R., & Cain, M. L. (2003). Are long-distance dispersal events in plants usually caused by non-standard means of dispersal? *Ecology, 84,* 1945-1956.

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

Holben, B. N. (1986). Characteristic of maximum value composite images for temporal AVHRR data. *International Journal of Remote Sensing, 7,* 1417-1434.

Huang, C., & Townshend, J. R. (2003). A stepwise regression tree for nonlinear approximation: Applications to estimating sub-pixel land cover. *International Journal of Remote Sensing, 24,* 75-90.

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment, 25,* 295-309.

Huete, A. R., Liu, H. Q., Batchily, K., & Leeuwen, W. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment, 59,* 440-451.

Hui, C. (2006). Carrying capacity, population equilibrium, and envrionment's maximal load. *Ecological Modelling, 192,* 317-320.

Hulbert, L. C. (1955). Ecological studies of Bromus tectorum and other annual bromegrasses. *Ecological Monographs, 25,* 181-213.

Jentsch, A., Friedrich, S., Steinlein, T., Beyschlag, W., & Nezadal, W. (2008). Assessing conservation action for substitution of missing dynamics on former military training areas in central Europe. *Restoration Ecology 17,* 107–116.

Jonsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to time series of satellite sensor data. *IEEE Transactions on Geosciences and Remote Sensing, 40* (8), 1824-1832.

Jordan, C. F. (1969). Derivation of leaf area index from quality of light on the forest floor. *Ecology, 50,* 663-666.

Kimes, D. S., Nelson, R. F., Manry, M. T., & Fung, A. K. (1998). Attributes of neural networks for extracting continuous vegetation variables from optical and radar measurements. *International Journal of Remote Sensing, 19,* 2639–2663.

Larsson, H. (1993). Linear regressions for canopy cover estimation in acacia woodlands using Landsat-TM, MSS and SPOT HRV XS data. *International Journal of Remote Sensing, 14,* 2129–2136.

Laura, C. V., & Larson, D. (2009). Role of invasive Melilotus officinalis in two native plant communities. *Plant Ecology, 200,* 129-139.

Leger, E. A. (2008). The adaptive value of remnant native plants in invaded communities: An example from the Great Basin. *Ecological Applications 18,* 1226–1235.

Lieth, H. H. (1974). *Phenology and seasonality modeling.* New York: Springer.

Li, X., & Strahler, A. H. (1992). Geometric optical bidirectional reflectance modeling of the discrete crown vegetation canopy: Effects of crown shape and mutual shadowing. *IEEE Transactions on Geoscience and Remote Sensing, 30,* 276–292.

López, G. F., Jurado, M., Peña, J. M., & García, T. L. (2006). Using remote sensing for identification of late season grass weed patches in wheat. *Weed Science, 54,* 346–353.

Ma, M. F. (2006). Reconstructing pathfinder AVHRR land NDVI time-series data for the Northwest of China. *Advances in Space Research, 37,* 835-840.

Machault, V., Vignolles C., & Borchi F. (2011). The use of remotely sensed environmental data in the study of malaria. *Geospatial Health, 5,* 151-168.

Mack, R. (1981). Invasion of Bromus tectorum L. into western North America; an ecological chronical. *Agro-ecosystems, 7,* 145-165.

Mazzola, M. B., Kimberly G., Allcock, J. C., Chambers, R. R., Blank, E. W., Schupp, P. S., & Robert S. N. (2008). Effects of nitrogen availability and cheatgrass competition on the establishment of Vavilov Siberian wheatgrass. *Rangeland Ecology & Management, 61,* 475-484.

Mcmorrow, J. (2001). Linear regression modeling for the estimation of oil palm age from Landsat TM. *International Journal of Remote Sensing, 22,* 2243-2264.

Menalled, F., Mangold, J., & Davis, E. (2008). *Cheatgrass: Identification, biology and intergrated management.* Montana State University Extension.

Mitich, L.W. (1999). Downy brome, Bromus tectorum L. *Weed Technology, 13,* 665-668.

Monsen, S. B. (1994). The competitive influences of cheatgrass (Bromus tectorum) on site restoration. S. B. Monsen and S. G. Stanley (compilers), *Proceedings--Ecology and Management of Annual Rangelands.* 1992 May 18- 22; Boise, ID. Gen. Tech. Rep. INT-GTR-313. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station: 43-50.

Noujdina, N. V., & Susan, L. U. (2008). Mapping Downy Brome (Bromus tectorum) using multidate AVIRIS data. *Weed Science, 56,* 173-179.

Novak, S. J., & Mack, R. N. (2001). Tracing plant introduction and spread: Genetic evidence from Bromus tectorum (cheatgrass). *BioScience, 51,* 114-122.

Pallant, J. (2011). *SPSS survival manual.* Open university press.

Peterson, E. B. (2005). Estimating cover of an invasive grass (Bromus tectorum) using tobit regression and phenology derived from two dates of Landsat ETM+ data. *International Journal of Remote Sensing, 26* (12), 2491-2507.

Piemeisel, R. L. (1951). Causes affecting change and rate of change in a vegetation of annuals in Idaho. *Ecology, 32,* 53-72.

Pierson E.A., & Mack, R. N. (1990a). The population biology of Bromus tectorum in forests: Distinguishing the opportunity for dispersal from environmental restriction. *Oecologia, 84,* 519-525.

Pierson E.A., & Mack, R. N. (1990b). The population biology of Bromus tectorum in forests: Effect of disturbance, grazing, and litter on seeding establishment and reproduction. *Oecologia, 84,* 526-533.

Pimentel, D., Lori, L., Rodolfo, Z.,, & Doug, M. (1999). *Environmental and economic cost: association with non-indigenous species in the United States.* College of Agriculture and Life Sciences, Cornell University, Ithaca, NY.

Pinty, B., & Verstraete, M. M. (1992). GEMI: A non-linear index to monitor global vegetation from satellites. *Vegetatio, 101,* 15-20.

Purevdorj, T., Tateishi, R., Ishiyama, T., & Honda, Y. (1998). Relationships between percent vegetation cover and vegetation indices. *International Journal of Remote Sensing, 19* (18), 3519-3535.

Purkis, S. J., Graham, N. A. J., Riegl, B. M. (2008). Predictability of reef fish diversity and abundance using remote sensing data in Diego Garcia (Chagos Archipelago). *Coral Reefs, 27,* 167-178.

Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., & Sorooshian, S. (1994). Modified Soil Adjusted Vegetation Index (MSAVI). *Remote Sensing of Environment, 48,*119-126.

Ray, T. W., & Murray, B. C. (1996). Nonlinear spectral mixing in desert vegetation. *Remote Sensing of Environment, 55,* 59–64.

Reed, B.C., Brown, J. F., Vanderzee, D., Loveland, T.R., Merchant, J.W., & Ohlen, D.O. (1994). Monitoring phonological variability from satellite imagery. *Journal of Vegetation Science, 5,* 703-714.

Riano, D., Emilio, C., Javier S., & Inmaculada A. (2003). Assessment of different topographic corrections in Landsat-TM data for mapping in vegetation types, *IEEE Transactions on Geoscience and Remote Sensing, 41* (5), 1056-1061.

Rice K. J., Black, R. A., Radamaker, G., & Evans, R .D. (1992). Photosynthesis, growth, and biomass allocation in habitat ecotype of cheatgrass (Bromus tectorum). *Functional Ecology, 6,* 32-40.

Richardson, A.J., & Wiegand, C.L. (1977). Distinguishing vegetation from soil background information. *Photogrammetric Engineering and Remote Sensing, 43* (2), 1541-1552.

Robert, D. A., Smith, M. O., & Adams, J. B. (1993). Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data. *Remote Sensing of Environment, 44,* 255–269.

Rogerson, P. (2001). *Statistical methods for geography.* London: SAGE Publications.

Savitzky, A., & Marcel J.E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36,* 1627–1639.

Sellers, P.J., Tucker, C.J., Collatz, G.J., Los, S.O., Justice, C.O., Dazlich, D.A. & Randall D.A. (1994). A global 1 by 1 NDVI data set for climate studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI. *International Journal of Remote Sensing, 15 (17),* 3519-3545.

Sheley R. L., & Petroff, J. K. (1999). *Biology and management of noxious rangeland weeds.* Corvallis: Oregon State University Press.

Skellam, J.G. (1951). Random dispersal in theoretical populations. *Biometrika, 38,* 196 – 218.

Sperry, L. J., Belnap, J., Evans, R. D. (2006). Bromus tectorum invasion alters nitrogen dynamics in an undisturbed arid grassland ecosystem, *Ecology, 87,* 603-615.

Swets, B. C., Reed, J. R., & Rowland, S. E. (1999). *A weighted least-squares approach to temporal smoothing of NDVI.* ASPRS annual conference, Portland, Oregon.

Teillet, P. M., Guindon, B., & Goodenough, D.G. (1982). On the slope-aspect correction of multispectral scanner data, *Canadian Journal of Remote Sensing, 8,* 84-106.

Truscott A., Palmer S., McGowan M., Cape J., & Smart, S. (2005). Vegetation composition of roadside verges in Scotland: The effects of nitrogen deposition. *Disturbance and Management, 136,* 109-118.

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

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, Mass: Addison-Wesley Publishing Co.

Upadhyaya, M. K., Turkington, R., & Mcilvride, D. (1986). The biology of Canadian weeds. *Canadian Journal of Plant Science, 66,* 689-709.

Van, E.J., Van, D. V., & Berdowski, H. (1995). Estimation of ground cover composition per pixel after matching image and ground data with subpixel accuracy. *International Journal of Remote Sensing* 16 (1), 97-111.

Van, J. W., & Root, R. R. (2001). The use of multi-temporal Landsat Normalized Difference Vegetation Index (NDVI) data for mapping fuels in Yosemite National Park, USA. *3rd Int. Workshop on Remote Sensing, GIS, and Fire Management, Paris, France, May 17-18,* 2001.

Vasquez, E.A., Sheley, R.L., & Svejcar, A.J. (2008). Nitrogen enhances the competitive ability of cheatgrass (bromus tectorum) relative to native grasses. *Journal of Invasive Plant Science and Management, 1,* 287-295.

Velleman, P. F., & Hoanglin, Da. C. (1981). *Applications, basics, and computing of exploratory data analysis.* North Scituate, Mass.: Duxbury Press.

Viovy, O., & Arino, A. S. (1992). The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series. *International Journal of Remote Sensing, 12* (8), 1585-1590.

Vitousek, P., Loope, L., Rejmanek, M., & Westbrooks, R. (1997). Introduced species: A significant component of human-caused global change. *New Zealand Journal of Ecology, 21,* 1-16.

Wadsworth, R. A., Collingham, Y. C., Willis, S. G., Huntley, B., & Hulme, P. E. (2000). Simulating the spread and management of alien riparian weeds: Are they out of control? *Journal of Applied Ecology, 37,* 28-38.

Weber, E. (2003). *Invasive plant species of the world.* Wallingford: CAB International Publishing.

Whitson, T. D. (1991). *Weeds of the west.* Laramie: University of Wyoming.

White, M. A., Thornton, P. E., & Running, S. W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles, 11,* 217– 234.

With, K. A. (2002). The landscape ecology of invasive spread. *Conservation Biology, 16,* 1192-1203.

Xiao, J. F., & Moody, A. (2005). A comparison of methods for estimating fractional green vegetation cover within a desert-to-upland transition zone in central New Mexico, USA. *Remote Sensing of Environment. 98,* 237-250.

Xu, B., Gong, P., & Pu, R. (2003). Crown closure estimation of oak savannah in a dry season with Landsat TM imagery: Comparison of various indices through correlation analysis. *International Journal of Remote Sensing, 24* (9), 1811-1822.

Young, J. A., Evans, R A., & Eckert, R. E. (1969). Population dynamics of downy brome. *Weed Science, 17,* 20-26.

Young, J. A., & Clements, C. D. (2000). Cheatgrass control and seeding. *Rangelands, 22,* 3-7.

Zamora, D. L., & Thill, D. C. (1999). Early detection and eradication of new weed infestations. *Biology and management of noxious rangeland weeds.* Corvallis, OR: Oregon State University Press: 73-84

Zhang, X., Friedl, M. A., Schaaf, C. B., Gao, F., Reed, B. C., & Huete, A. (2003). Monitoring vegetation phenology using MODIS. *Remote sensing of Environment, 84,* 471-475