Assessing grassland distributions and spatio-temporal variations in the U.S. Great Plains with MODIS imagery

---

A Thesis

Presented to

The Faculty of the Graduate School

At the University of Missouri

In Partial Fulfillment

Of the Requirements for the Degree

Master of Arts

By

Qing Chang

Dr. Cuizhen Wang, Thesis Supervisor

DECEMBER 2011

The undersigned, appointed by the dean of the Graduate School,

have examined the thesis entitled

Mapping grasslands in the Great Plains and their spatio-temporal variations responding to Climate Change

Presented by Qing Chang

A candidate for the degree of

Master of Arts

And hereby certify that, in their opinion, it is worthy of acceptance.

Dr. Cuizhen Wang

Dr. Hong He

Dr. Grant Elliott

ACKNOWLEDGEMENTS

I extremely wish to express my feeling to be deep graceful for all has helped me all the way that I cannot do without:

This thesis could not have been written without TIMESAT program developed by Per Jönsson in 2002 to finish a part of time-series analysis and image smoothing. And thanks for Parameter-elevation Regressions on Independent Slopes Model Climate Group in Oregon State University, without their comprehensive and detailed climatic information I could not analyzing vegetation variation under different climate situations. I also should thank USDA NASS and USGS for providing me good quality MODIS TERRA satellite imagery. Without the image my research could have never been done.

Thanks to Dr. Cuizhen Wang, who not only served as my supervisor but also encouraged and challenged me throughout my academic program. She always taught me the knowledge as much as she can. She guided me through the thesis process, never accepting less than my best efforts. I still remember the first time I went to her office: she encouraged me and performed a totally new research world to me. After two years, I have successfully transferred from an engineer to geographer. I thank her so much.

I should direct my thanks to Dr. Grant Elliott, who always gave me great support with suggestions together with his perspectives toward my research and provided me for huge knowledge in GIS study. He is the good teacher taught me various spatial data processing methods as the significant parts of my thesis study.

I would also thanks to Dr. Hong He for providing the generous help in my research, especially he has given me important advices in building the objective and correct classification model. I really admire his professionalism and preciseness to science and the passion to share his knowledge with me without asking any reward.

I am also grateful to for Geography Resource Center of Geography department and University of Missouri-Columbia for supporting me my research assistant for my Master's study, for giving me the opportunity to quickly adapt to GIS and Remote Sensing research in a good environment. I would also like to thank the Graduate Professional Council of University of Missouri-Columbia for supplying a chance to present my achievements in the glassland mapping in the Great Plains at the American Society for Photogrammetry and Remote Sensing Annual Conference in Milwaukee, 2011 with favorable reception.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The North American Great Plains is the primary grassland in the United States. Both cool season grasses (CSG) and warm season grasses (WSG) grow in the Great Plains, with their distributions vary with environmental conditions and climate dynamics. As global climate change is being a hot topic today, there is a need of detailed maps of these grass functional types in the Great Plains to enhance our understanding of grassland ecosystems and their responses to climate change. This study delineates CSG and WSG based on their unique phenology features that could be extracted from time-series satellite imagery; and preliminarily examines their inter-annual variation responding to the changing climate conditions in this region.

Several spatial analysis methods are examined in this study. I have 1) built user-defined programs to perform time series analysis of 10 years' MODIS imagery, and extracted useful phenological parameters as innovative inputs of image classification; 2) Developed a sequential multivariate regression algorithm in SAS to classify crop and grass maps in the Great Plains; 3) compared the 10-year climatic data with the classified grass distributions to examine the relationship between climate change and vegetation dynamics. The achievements filled the gap of lacking detailed distribution of grass functional types in the Great Plains, and reported the credible analysis in climate-induced land cover changes.

CHAPTER I: INTRODUCTION

1.1 Research Background

The North America Great Plains covers the central United States and extends to Canada in the North. As the primary land covers in this region, grasslands are composed of two major plant functional types: cool season grasses (CSG) and warm season grasses (WSG), with their distribution varying with spatial and temporal dynamics. These two grass types hold different photosynthetic pathways: the CSG absorbs 3 carbon atoms while WSG holds 4 carbon atoms during photosynthesis.

The alteration of WSG/CSG relative abundances plays an important role in testing regional or global biogeochemical cycle models. Knowing detailed geographic distributions of CSG and WSG is also important in estimating grassland productivity and quantifying land-surface carbon sinks in global climate studies. Therefore, mapping potential distribution of WSG/CSG with remote sensing techniques could be applied for phenology and climate change prediction. Meanwhile, vast of grasslands have been converted to croplands and urban development lands, which increased unknown places and uncertain factors in classification study. How could we find a suitable approach that would be performed well in large-area grassland mapping? It is the focal issue in today's society.

1.2 CSG and WSG Grasses

1.2.1 Geographic distributions

The North American Great Plains covers vast area of prairies expanding from ShortWSG in the west of Rocky Mountain to TallWSG in the east of Mississippi River. Cool season grasses gradual increase from south to north along with higher latitude. Two grass types are commonly observed in the Great Plains grasslands: cool season grass (CSG) and warm season grass (WSG), which dominate the grass covers in the region. Warm season grass produces four carbon molecules in photosynthesis, and cool season grass has three carbon molecules. Therefore they are also named C4 and C3 grasses, respectively, and their relative abundance distribution changes geographically (Wang, 2009). Seasonal distributions of CSG and WSG grasses have been mapped using remote sensing techniques (e.g. Goodin et al., 1997; Peterson et al., 2002; Wardlow et al., 2007). Most of these studies applied medium-resolution Satellite imagery such as Landsat Thematic Mapper (TM) and the Enhanced Thematic Mapper Plus (ETM+). Therefore, these studies are often limited in local areas. Consequently, it is difficult to validate and adopt their findings in regional/global climate studies. Taking advantage of large-coverage, frequent observations of Terra MODIS imagery, and my research covers the vast area of the U.S. Great Plains to examine grassland distributions and dynamics in this region.

1.2.2 Phenology Differences

Each type of grass has its own phenology characteristics. The WSG grasses in the tallWSG prairie appear as tall grass and CSG grasses are shorter. These features are often utilized in past studies to classify CSG and WSG with satellite images (Wang et al. 2010, Chen et al., 1996). In large geographic areas, however, it is often difficult to standardize these phenology characteristics due to the complexity of grass species and growth conditions in different areas. In the west of Great Plains, WSG grasses are much shorter than those in the east of Great Plains when it is extended into the TallWSG Prairie. These grasses are typically observed in the ShortWSG Prairie in western states of the Great Plains. Hereafter these WSG grasses are called ShortWSG. It should be noted that it still belongs to warm season grasses in phenology and holds the same C4 photosynthesis functions as TallWSG in the eastern states.

In theory, CSG mostly grow in high latitudes or elevations, while WSG flourish in warm, tropical to subtropical regions (Winslow et al, 2003). Because of "warm season" and "cool season" grasses refer to different photosynthesis pathways, we could make distinction by tracking carbon atom. WSG pathway evolves in species in the wet and dry tropics (marked in Table 1). The seasons of CSG grasses are sensitive to water availability. In high intensity of heat and light, CSG grasses close their stomates to reduce moisture usage in leaves, and the photosynthesis process could decline up to 50%. WSG grass has its photosynthesis six times more efficient than CSG, so it grows better in the warm, CO_2 enhanced environment (www.Primary Industry Agriculture.com).

In addition, the evidences from the leaf fossil of CSG and WSG prove that they hold different kind of photosynthetic pathway. In previous research, the fossil shape of CSG leaf's epidermal cells always is circular, rectangular, or elliptical (Phytolith Systematics, 1992). Oppositely, the leaf cells in WSG have many intersections inside. Therefore, totally dissimilar types of photosynthetic pathway and photosynthetic efficacy could be caused by distinct vegetation epidermal cells (Phytolith Systematics, 1992).

Table.1 Features of CSG and WSG grasses (www. Industry&investment nsw.com)

	C3	C4
Initial molecule formed during photosynthesis	3 carbon	4 carbon
Growth period	cool season or yearlong	warm season
Light requirements	Lower	higher
Temperature requirements	Lower	higher
Moisture requirements	Higher	lower
Frost sensitivity	Lower	higher
Feed quality	Higher	lower
Production	Lower	higher
Examples	Weeping grass and common wheatgrass	Kangaroo grass, red grass and wire grass

Furthermore, phenological differences of CSG and WSG also widely applied into grass biomass production (McLaughlin et al, 1999), which is useful in animal husbandry, especially in WSG by using its unique photosynthetic pathway (Heaton et al, 2004). The conclusions from past studies are that CSG obtains more biomass production during whole growth season than WSG, and WSG represents its productive advantage mostly in the periods of summer when CSG turns to be senescent in hot environments (Robins, 2010).

1.3 Remote Sensing of grasslands

1.3.1 Satellite Imagery and published maps

With the availability of airborne and spaceborne imagery in past decades, intensive efforts have been made to create vegetation maps at local, state or national levels. The reference data I collected for this research is the so-called Cropland Data Layers (CDL) developed by USDA/National Agricultural Statistics Service (NASS). The NASS have started to collected national agricultural statistics since 1997. The first set of crop maps are North Dakota and part of Arkansas (http://nassgeodata.gmu.edu/CropScape/). It requires huge efforts from individual state governments and USDA as rich information and process are needed from both sides. This product combines remote sensing imagery (MODIS and AwiFS) and NASS survey data in precise ground locations, farm services agency and other accessorial data to estimate major annual crops, other annual plants, and land use &land cover types in every state. It also provides annual state- or county-level crop acreage data. The program published these results after accuracy assessment. It was found that crop types turned to have high interaction with each other (Figure.1).

The CDL data sets are also treated as “census by satellite”. These products offer yearly information about detailed land use and land covers in the United States. In the Great Plains, The CDL data identify more than 120 types of vegetation such as corn, soybean, winter wheat, and nonagricultural areas. However, the different grass types in the Great Plains, the cool season grass, tall grass and short grass, are grouped into one category- “pasture/grassland”. Therefore, further efforts are needed to map these grass types to understand their ecological conditions and to examine their phenological variations and responses to global climate change and human activities.

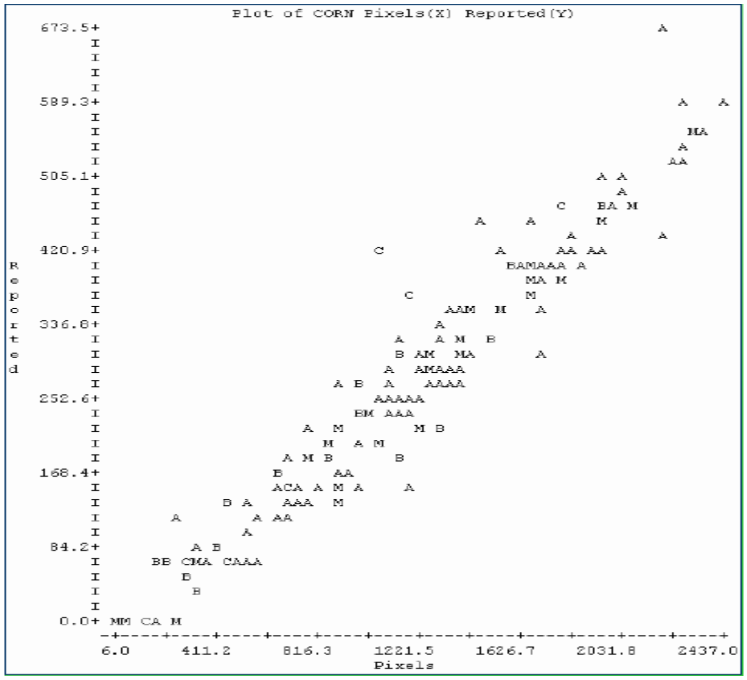

Figure.1. NASS Acreage Estimation & Cropland Data Layer Methodology, (ftp://ftp.iluci.org/GEO_Ag/PowerPoints/Doorn.Cropland%20Data%20Layer%20Program.pdf)

Vegetation's greenness and health are significant factors in annual crop and grass research in my study. The expressive forms are reflected as red and near infrared band. Having these two spectral bands, the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Terra satellite that was launched in 2000 provides coarse-resolution, global coverage, daily observations for regional and global studies. It is the most appropriate data sources in this study for remote sensing of grasslands in the Great Plains. The MODIS mission provides various products to meet the needs of different users. The resolution is from 250 meters to 1000 meters, and the composite could be at 8-day, 16-day, or monthly intervals to reduce cloud noises. It also contains 36 spectral bands to store data information in a range of visible, near infrared, shortwave infrared and thermal infrared spectral regions (https://lpdaac.usgs.gov/). The shortwave infrared bands could be applied in studying vegetation distribution, forest moisture or soil (Xiao et al. 2004). The USGS Data Pool (https://lpdaac.usgs.gov/lpdaac/get_data/data_pool) provides a series of MODIS products that are open to the general public. The resolution and temporal granularity are in a variety of data types and formats to meet the needs of different users.

The satellite product I (MOD09A1) was applied in my research. There are several advantages for using MOD09A1 in vegetation distribution recognition: Firstly, the 500-m resolution imagery contains more information than commonly adopted 1000-m in other products, and it holds much better data quality than raw imagery at 250-m resolution. Secondly, 8-day scan granularity (temporal composite interval) records primary vegetation growth in a growing season to enhance the delineation of WSG and CSG. It also saves computation time, which is of great importance in multi-year MODIS image processing. Another benefit is that the MOD09A1 data highly reduce cloud noises and reach better atmospheric correction (Vermote & Vermeulen, 1999).

Even though the MODIS imagery contains 36 spectral bands, only 7 bands could be recognized in geospatial software that I could access (ERDAS/Imagine and ENVI). These bands, in fact, are most useful in investigating land use and land cover in vegetated environments:

1. Blue band (620–670 nm)
2. near infrared (841–876 nm)
3. Green band (459–479 nm)

4. Red band (545–565 nm)

5. near infrared (1230–1250 nm)

6. Shortwave infrared (1628–1652 nm)

7. Shortwave infrared (2105–2155 nm)

Some other bands, such as Reflectance Band Quality, Solar Zenith Angle, View Zenith Angle, Relative Azimuth Angle, State Flags, and Day of Year, are designed for advanced modeling process and are not explored in this study.

1.3.2 NDVI and Spatio-Temporal Trajectory analysis

As an indicator of vegetation greenness and growing conditions, the Normalized Difference Vegetation Index (NDVI) is a ratio of near infrared band and red band from plant reflectance in satellite imagery (Equation.1), which indicates the vegetation absorptive and reflective capabilities in spectrogram (Jin et al, 2003).

$$NDVI = \frac{(NIR-RED)}{(NIR+RED)} \tag{1}$$

The NDVI value ranges from -1 to 1 (Weiss et al, 2004), and has been widely used in different purposes such as measuring vegetation health and density (Mwangi J. Kinyanjui, 2010), monitoring long-term local or large-scale land cover change (Jin et al, 2003), biomass productivity and energy transformation (Myneni et al, 1995), and temporal variations in grassland growth cycles (Stephen.L et al, 2007). Due to variations in precipitation and temperature, soil moisture, soil PH values, and other elements that affect vegetation dynamics, some past studies found that spring-summer is the best monitoring time to acquire impactful vegetation growth information (Raynolds et al, 2008; Epstein et al, 2003).

More recently, time series analysis bases on multi-day NDVI trajectories provide a unique and advanced way of vegetation mapping. The trends of spatio-temporal NDVI trajectories reveal the phenology cycles of crops and the associated management activities such as harvesting, planting or grazing. This important information can be used in crop mapping. The accuracies could reach 80% and higher (Wang et al, 2010) with reduced spatial and temporal influences. Different Peak NDVI dates and values of different plants optimally reveal their phenological characteristics in peak growing season. Besides, if in the summer, NDVI values are keeping raise that index vegetation is more and more greenness during this period because of the red band decrease and more radiation energy has been absorbed (Kremer and Running, 1993). Therefore, trajectory analysis of NDVI could play a better role in this research with the rich information from frequent MODIS observations.

In spite of huge advantage of MOD09A1 data in CSG and WSG mapping, it still contains lots of noise to interfere final results. The noise occurs owing to various atmospheric effects such as cloud and snow covers, data losing, or hazy conditions. For this reason, time-series NDVI values are often lower than ground observations, sometimes even represent as negative values (Weiss et al, 2004), or data loss. Spectral smoothing is needed to reduce these effects (Macho, 2008).

Lots of smoothing methods have been established for remote sensing spectral analysis, Kalman filter is one of a widely-recognized applied algorithm in spectral smoothing. It is based on hidden markov model and simultaneous linear algebraic equations to finish time-series analysis in noise surroundings (R. E. Kalman, 1960). It has been used in satellite navigation, positioning system, and time-series model (Dan Simon, 2006). More recently, Macho (2008) developed a new algorithm to perform better smoothing based on the maximum likelihood estimation. This method is built on picking stochastic samples and Fourier transform to smooth two-Dimension data by estimating smoothing parameter-the signal to noise ratio and noise variance, and require few calculate time and storage memory.

The smoothing process in my research is a pre-requisite in extracting phenological characteristics from a growth cycle. TIMESAT is a program that performs smoothing and phenology extraction (Jönsson and Eklundh, 2002). The Savitzky-Golay filter is still a mostly popularly used algorithm (Abtaham Savitzky and J.E.Golay, 1964; Chaichoke Vaiphasa, 2006). It reached the best results in removing noise and smoothing the NDVI time-series curves and was thus adopted in TIMESAT (Jönsson, 2002). The most important factor in Savitzky-Golay filter is the filter window size. An oversized window cannot remove a majority of noise, while useful signals are often removed if the window is too small.

After smoothing of NDVI series, an asymmetric Gaussian model in TIMESAT could be used to extract phenological parameters, some of which could serve as inputs of WSG/CSG classification. However, this model only can be applied to normal distribution growth cycle. Some vegetation types, such as winter wheat and cool season grass' growth curves do not match normal distribution. Therefore more process will be tested in my research.

1.4. Classification procedures-Regression Analysis

Land use & land cover mapping and change detection are a core part of digital image processing. While intensive studies have been conducted in vegetation mapping in specific and small areas, currently more focus has been turned to large-area or regional studies that corporate with local ground data. It broadens our scope of mind in global change, and at the same time, increases the complexity and challenge in accuracy (Schwartz et al. 1999). The commonly applied classification approaches are thus limited in large-area studies due to these challenges. More advanced approaches need to be integrated in these studies.

Regression analysis is a statistical method in testing the relationship between independent and dependent variables, and in extracting typical features in pattern recognition and prediction (Dennis Lindley, 1987). Both logistic regression and regular regression can be applied in large-scale studies, and each method has its own strength and weakness in classification process.

1.4.1. Logistic regression classification

Logistic regression is to calculate the probability of event happened by setting logistic determination. This binary regression analysis is popularly used in medical and investment analysis in predicting the incidence rate of independent events by considering its particular marks, such as how much probability that the patient has a definite diagnosis as cancer.

A linear regression is defined (Equation.2 and 3):

$$f(z) = \frac{e^z}{e^z+1} \quad (2)$$

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k \quad (3)$$

Where x_1, x_2, … x_k represent different independent variables, and the β is the coefficient of each variable. The variable z represents the logical contribution of all independent variables, while the output $f(z)$ shows the probability of happening is from 0 to 1.

In general, logistic regression works well when the model has clear criteria and limited influencing factors. In theory it requires plenty of training samples. If the sample number is less than 500, the estimate values could have a considerable range of deviation. Therefore, this “yes” or “no” model does not apply in my study.

1.4.2 Regular regression classification

Most statistical software such as SAS and SPSS could accomplish the regression analysis including regular regression. Sometimes, this approach is also treated as a general linear regression.

In a general regression model, the dependent variable is the vegetation type we need to predict, and independent variable should be the random or discrete sample data in study area. And the independent variables can be linked by a linear equation to generate class values. The equation of a multi-variable regression model is:

$$y_i = \varepsilon_i + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} \quad (4)$$

Where y_i is the dependent variable and is a result of linear combination from all independent variables, x_{ip}. The Y value must be linear with the independent variables in analysis. x_{ip} represents independent variables, which should not have correlative relationship between each other, or contains very low correlative parameter with each other in regular regression process. The ε_i is the intercept.

Same as the logistic regression model, a regular regression still requires large number of training data. But if we lack of samples, the prediction estimation will not have obvious bias. In sum, the regular regression analysis does not have very serious requirements and could be widely applied.

In sum, logistic regression is applied when the dependent variables are divided in dichotomy situation, the predict class type can be either binary or multiple outputs (Rice.J.C, 1994). This method has been applied in land cover mapping combining the GIS system, such as using maximum likelihood ratio and logistic regression coefficient to map the landsides location and the level of hazard ratio (Saro Lee, 2004) or predicting forest fire on the basis of tree healthy level mapping (D.L.Martell et al., 1987). This method will perform well under a small quantity of predicted class types and mapping probabilities of dependent variables.

The regular regression is built on the assumption of samples must be independent and obey the normal distribution. When with multiple dependent outputs in study area, regular regression will be stronger than logistical regression (McCullagh et al., 1989). For example, estimate the distribution of invasive alien grass to evaluate grassland variation and biological cycle (Dark S. J, 2004). If the variables cannot reach the requirements for regular regression, logistic regression is a better choice (NC States University PA 765-766, http://faculty.chass.ncsu.edu/garson/PA765/index.htm).

1.5. Grassland Responses to Climate Change

The temporal and spatial distributions of CSG and WSG indicate their responses to climate change. For example, using C-13 isotopic into grass leaf's fossil, CSG reduces its photosynthesis in warm and high oxygen environment. Therefore, CSG faces the competition from WSG owing to global warming. As WSG's photosynthetic pathway adapts temperature change better, many plants changed their photosynthetic pathway closing to this type (Spicer, 1993). Therefore, it is very useful to study plant evolution and their ecosystem structure responding global climate change.

Various empirical physical bio-geospatial models have been developed to simulate vegetation distributions under different climate scenarios. The Grassland ecosystem model (GEM series) (e.g. Charles F. Rodell, 1977; Nouvellon, Y, 2001; Dev Niyogi et al., 2009;) was developed in early years. It is a hierarchical approach that links vegetation's photosynthetic efficacy, energy exchanges and ecosystem structures to test the interactions between plant growth and ecosystem progress responding to carbon dioxide increase and other climate change (Chen et al, 1996). The next climate change model simulation is the dynamic global vegetation model (e.g. Dieter Gerten et al., 2004; John Hughes et al., 2006; Peter Levy et al., 2010; Peng et al., 2010). This model collects climate change factors in land and atmosphere such as humidity level, precipitation, and energy. Then it analyzes whether and how these changes could influence vegetation dynamics. Previous studies proved this dynamic global vegetation model can provide detailed description about the feedback of climate change to vegetation variation through energy, moisture, and momentum (Bonan et al, 2003).

Another popularly applied model is the seasonal availability of water (SAW) model (Winslow et al, 2003). It was developed to predict the fraction of CSG grasses in 1° grid cells all over the globe. In this model CSG and WSG vegetation was separated based on their phenological differences (Winslow et al, 1994).

The application of SAW model proves that temperature and precipitation change will change the cool season and warm season grass based on their photosynthesis efficiency (Niu et al., 2005). Even though cool season grass needs more carbon dioxide than warm season grass, its photosynthetic advantages could be reduced when the outside environment changed. It is very important to get more accurate information about timing and spatial distribution of WSG and CSG grasses to better monitor their response to future climate change (Winslow et al, 2003).

1.6 Research Goal and Objectives

The proposed research goal is to delineate CSG and WSG form time-series satellite imagery and to examine their inter-annual variation responding to the changing climate in the Great Plains. To fulfill this goal, I will address the following specific objectives:

1) To build 10-year time series of NDVI data sets from MODIS imagery (MOD09A1) covering the Great Plains in 2000-2009, and to extract phenological metrics that better reveal the differences of WSG and CSG grasses;
2) To develop a multi-tier phenology-assisted regression classification algorithm to delineate WSG and CSG as well as major annual crops in the past 10 years;
3) To link spatio-temporal shifts of WSG and CSG grasses to local climate variations in the past 10 years as a preliminary study about grassland ecosystems' responses to climate change in the Great Plains.

CHAPTER II: STUDY AREA AND DATA SETS

2.1 Brief Introduction

The study area of my research is the U.S. Great Plains, which covers 10 states (Montana, Wyoming, Colorado, New Mexico, Texas, Oklahoma, Kansas, Nebraska, South Dakota, North Dakota) in the central and Midwest United States. Distributions of vegetation types and wildlife habitats vary dramatically in this region and therefore, it is regarded as a perfect study area for investigating grassland dynamics, land use and land cover changes, the associated human activities, and climate variation.

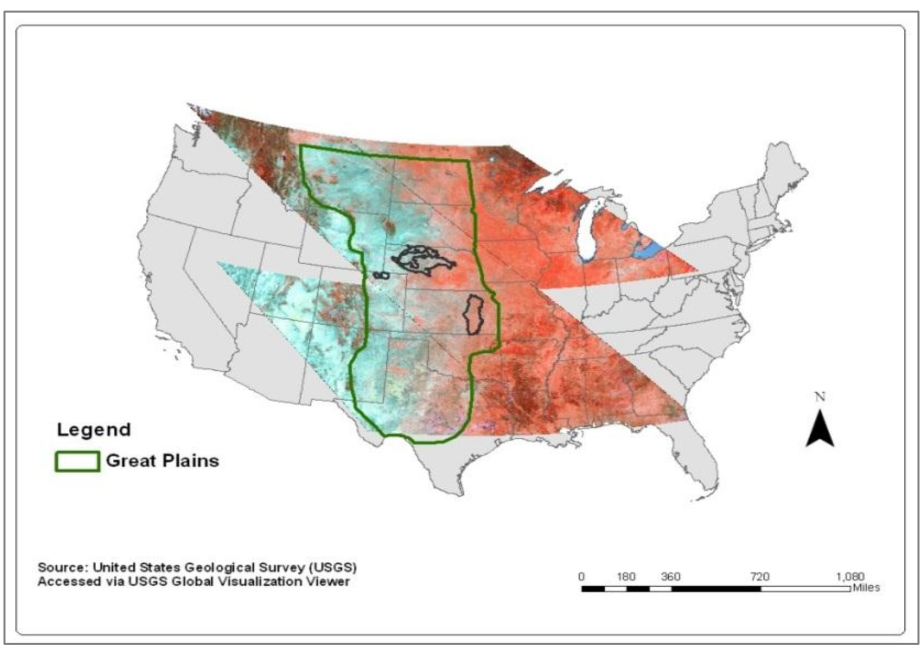

Figure 2. The study area (The example image mosaic is extracted from Terra/MODIS acquired in July 2007).

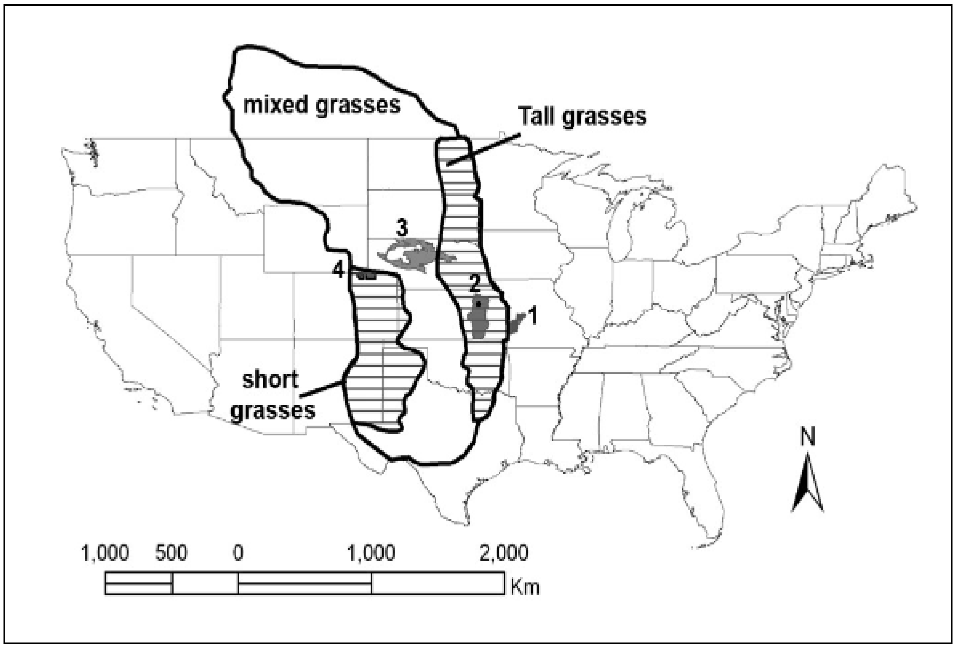

Figure 3. General distributions of CSG and WSG in the Great Plains. (Modified from Risser et al. 1981)

Four focus areas in the Great Plains are selected to represent CSG and WSG grass types and their relative abundances. From east to west of the plains, these focus areas are TallWSG (Flint Hills, KS), ShortWSG (Central Plains Experimental Range, CO), mixed grass (Sandhills, NE), and a special short-lived CSG grass environment (Montana Prairie) in the very north of the plains.

The Flint Hills focus area is marked as #1 in Figure 4. It is located inside the TallWSG Prairie in Kansas and WSG occupies more than 80% in this area. The Sandhills focus area is in Nebraska. It is the largest mixed prairie in North America. The primary grass species in this area include little bluestem (Schizachyriumscoparium (Michx)), sand bluestem (Andropogonhallii Hack), and blue Gramma (Boutelouagracilis).

The Sandhills grassland has been largely used as pasture lands and hay fields (Wang et.al, 2009). From previous researches and ground observations, warm season grasses (both TallWSG and ShortWSG species) still dominate this area. The Central Plains Experimental Range (CPER) in Colorado is operated by the USDA Agricultural Research Service (ARS), and is the smallest of the four focus sites.

The CPER has similar CSG and WSG mixture and growth environment as the Sandhills in Nebraska. However, ShortWSG species, primarily blue Gramma (Boutelouagracilis), are the dominated grass types in this focus area. The CPER is part of the NSF ShortWSG Steppe Long-Term Ecological Research (LTER) program and is covered in the Pawnee National Grassland, a 78,100 hectare of public land managed by US Forest Service (Wang et al. 2009). The 4th focus area is the Montana Prairie located in Montana and North Dakota States. Very limited studies of vegetation distribution change in this area have been published. The CDL products classified this area as grassland. The Montana Prairie contains approximately one quarter of grassland areas in the Great Plains. Thus, it plays an important role in climate change and carbon cycle. Due to its high latitudinal geographic locations, it is reasonable to assume that this area is dominated by CSG species.

2.2 Data Acquisition

The Terra/MODIS time-series imagery is applied in this research. From USGS LP DAAC: Land Processes Distributed Active Archive Center (https://lpdaac.usgs.gov/lpdaac/get_data/data_pool), I downloaded the MOD09A1 surface reflectance images acquired in 2000-2009. To simplify the process, only data in odd years (2001, 2003, 2005, 2007, and 2009) are downloaded in my research. Even though there are only 5-year results, according to information from many channel resources, the years of 2007 and 2009 are very representative in observing climate change and vegetation variation, such as suddenly increased precipitation. Therefore, the relative abundance of CSG and WSG in the four focus areas is useful in climate factors analysis.

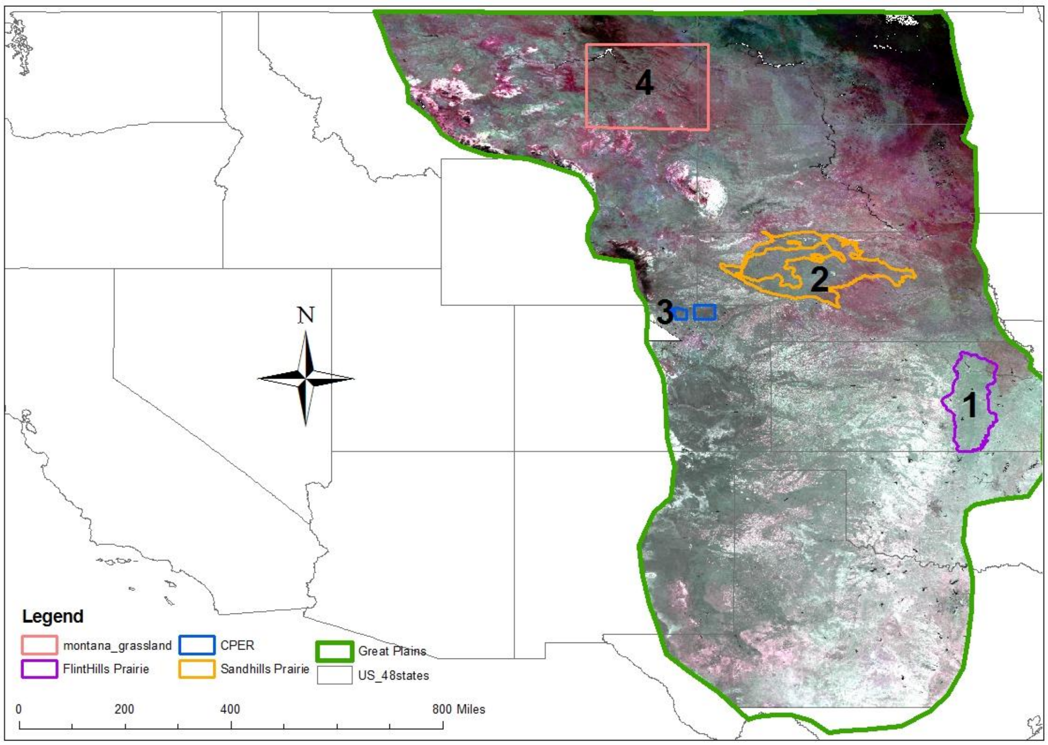

Figure 4. The four study focus areas in the Great Plains: 1-Flint Hills Prairie in Kansas; 2-Sand hills Prairie in Nebraska; 3-the Central Plains Experimental Range (CPER) in Colorado; 4-Montana Prairie.

This 8-day, 500-m resolution composite image contains 7 spectral bands in visible - shortwave infrared region. The original data sets are stored in the sinusoidal projection and are converted to Orthographic Geographic Coordinate System, the North American Datum 1983 (NAD83), for better visualization. Arranged in path and row sequences, four MOD09A1 tiles (H09V05, H10V04, H10V05 and H11V04) are needed to cover the whole area of the Great Plains. These MODIS images are the primary dataset in my studies.

The secondary data is the cropland data layer (CDL) maps developed by the USDA National Agricultural Statistics Services. The CDL products are classified from satellite images for detailed documentation of annual crops in major agricultural regions of the United States. Major annual crops in the Great Plains include soybean, corn, winter wheat, and spring wheat. The resolutions of these layers vary from 30 meters to 60 meters depending on the satellite images used in classification. The availability of state-level CDL products varies in different years for different states. I managed to download the 2007 CDL data sets for most of the states in the Great Plains. For several states in which the CDL products are not available in 2007, I used their 2008 or 2009 products as an alternative. With these data sets I created one CDL mosaic layer for whole coverage of the U.S. Great Plains. At much finer resolution (56-meter) than MODIS imagery, this data set serves as ground truth sources for the primary annual crops. It also serves as a filter to mask out non-herbaceous covers such as forest, wetland, water and urban development. These areas are not the concern of this study. As mentioned earlier, the CDL data groups' cool season grass and warm season grass into one single class, which will be delineated in this research.

CHAPTER III: Phenology-assisted time series analyzing for WSG/CSG delineation

The original MODIS imagery and Cropland Data Layer contain different image resolutions coordinate systems, and other geographic factors. Also, both data types contain noises leading to inaccuracies in classification. Several preprocessing algorithms, including georeferencing and spectral smoothing for noise removal, have to be applied before building classification models.

With the time-series MODIS data sets, phenological features of the two grassland functional types could be extracted, which maximally differentiate WSG and CSG along their growing season. In this research, several user-defined approaches were applied to obtain different kinds of vegetation's phenological parameters, which were adopted in WSG/CSG classification.

3.1 Data Pre-processing

3.1.1 MODIS Imagery

3.1.1MODIS Imagery

The MODIS data contains 4 tiles (H09V05, H10V04, H10V05 and H11V04) and 5 years (2001, 2003, 2005, 2007, and 2009) of individual MOD09A1 images. Each tile has 46 images per year, so the total number of images is 1840 over the whole study area.

These data sets were processed in the following steps. Firstly, each image was re-projected into the Geographic Coordinate System (NAD83) in the ENVI software batch process. Then I calculated NDVI values from red and near infrared bands in each image. The leaves of vegetation absorb luminous energy in visible band from 0.6 μm to 0.7 μm, and reflect in near infrared band (from 0.7 μm to 1.1 μm) that could be captured from sensor. Therefore, the NDVI represents vegetation greenness and healthiness in the study region. In each year, there were 46 NDVI images, which were stacked into one 46-band NDVI composite for the following time-series analysis. The four scenes of these NDVI composite images were mosaicked to cover the whole U.S. Great Plains, and also clipped by Great Plains' boundary that removed nearly 50% redundant data for saving storage and operation processing. The final image sets in this research included five NDVI images (2001, 2003, 2005, 2007, and 2009); each contains 46 bands to represent NDVI series at 8-day interval in each year.

3.1.2 Cropland Data Layer Imagery

The 10 CDL products in 2007 (in the 10 states in the Great Plains) are downloaded and used as reference in this research. The CDL maps contain more than 120 crop types. As the major concern of this research is grassland, only major annual crops (corn, soybean, winter wheat, and spring wheat) are considered in this grassland study (Table 2). The low-acreage crops could not be further investigated due to the limited spatial resolution of MODIS imagery. Furthermore, the non-agricultural lands, such as urban development, road, lake, and river, need to be masked from the original MODIS images as well. These class types are grouped into “other crops” and “non-crop land” as shown in Table2

Table.2 Percentages of major annual agricultural and non-agricultural lands in the 10-states CDL data (Montana, Wyoming, Colorado, New Mexico, Texas, Oklahoma, Kansas, Nebraska, South Dakota, and North Dakota)

Class Value	**Class Type**	**Percentage**
0	No Data	45.1779 %
1	Non-Crop Land	11.1588%
2	Other Crops	3.5765%
3	Corn	1.7913%
4	Soybean	0.8701%
5	Winter Wheat	2.2974%
6	Spring Wheat	0.9229%
7	Scrubland	12.8683%
8	Grassland	21.3368%

In ENVI software, a decision tree was built to regroup all crop types in the CDL data into 8 classes as listed in Table 2 (Fig.5). . Each class type in the decision tree is designed a unique color represented by a combination of red, green, and blue color values. Only grassland, corn, soybean, winter wheat and spring wheat were used in this research.

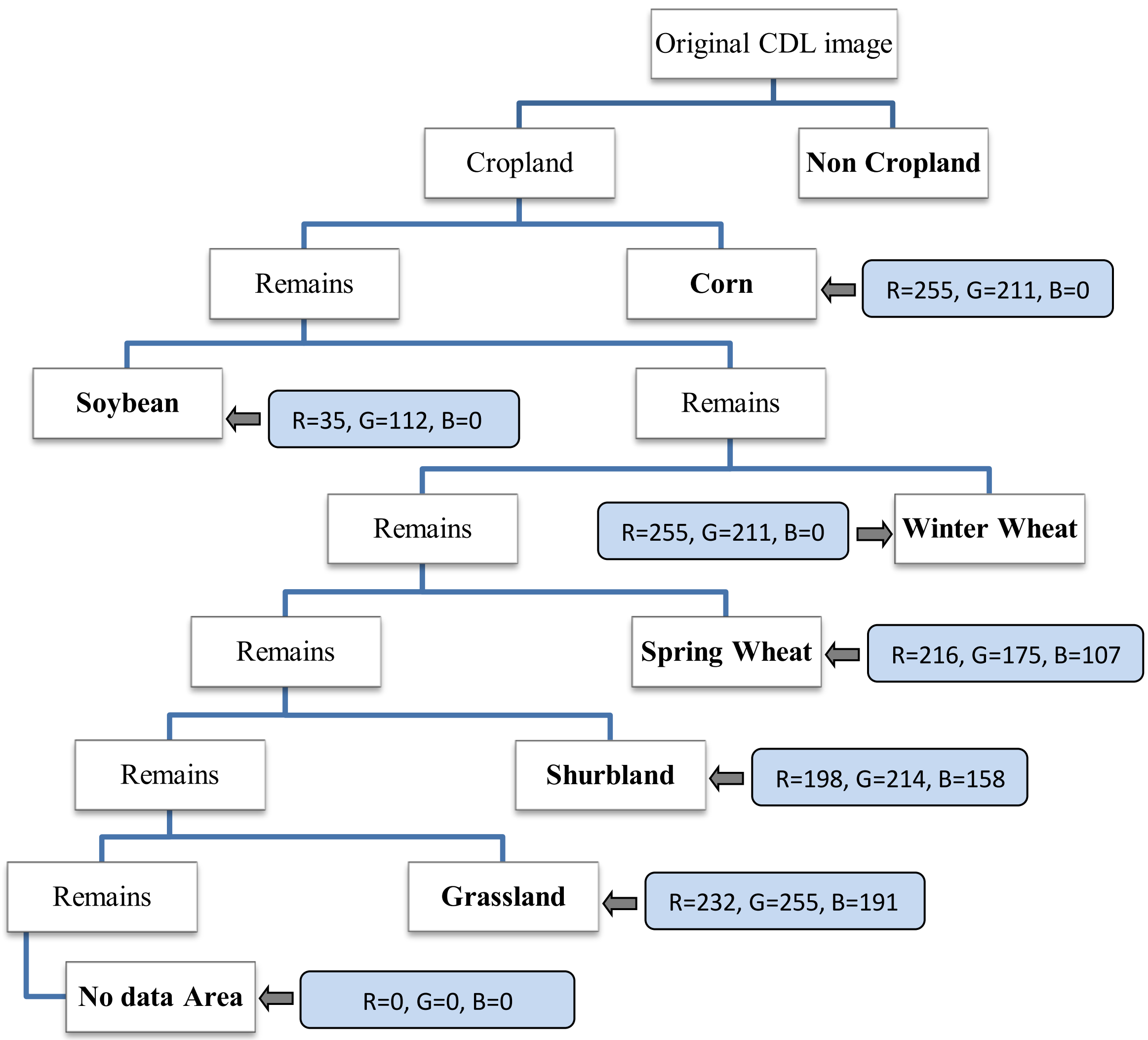

Figure. 5 The decision tree for extracting major class types in the Great Plains. R represents red, G represents green, and B represents blue band value, respectively.

The re-grouped CDL images of the 10 states were then mosaicked and clipped to cover the Great Plains. To match this 56-m CDL mosaic with MODIS imagery, I applied a 7*7 majority filter in ENVI software to resample the CDL mosaic image into 500-m pixel size.

Finally, a mask layer was developed with this resampled CDL image to extract land covers that were not explored in this study (urban, wetland, forest, water, low-acreage crops, etc.). This mask was then applied to the annual 46-band MODIS NDVI series so that the MODIS data only contained major crops and grasslands that were of the primary interest in this research.

3.2 Time-Series Analysis

It has been well know that CSG and WSG grasses hold recognizable phenology differences along a growing season. The CSG grows earlier than WSG, probably started in late February, and WSG usually starts its growth one month later. These differences could be revealed from the MODIS time series of NDVI data. Therefore time-series analysis is one important part of the methodology design in my thesis.

3.2.1 Spectral smoothing

The 8-day NDVI time series are strongly affected by temporal atmospheric variations and cloud noises. Therefore, the 46-point NDVI trajectory of each pixel always faces atmospheric contaminations such as NDVI reduction or data loss due to cloud and snow covers. For this reason, sharp and jagged curves (peaks and troughs in a trajectory) are commonly observed in time-series NDVI trajectories (Dean Fairbanks et al., 2004).

Spectral smoothing process was performed in the MATLAB program to reduce these noises. In previous chapter, I have introduced several image spectral smoothing approaches that were developed in past studies, for example the Kalman filter (R. E. Kalman, 1960) and the Maximum Likelihood Estimate (Macho, 2008). In this research, I adopted the Savitzky-Golay filter, a most commoly used filter because it is an adaptive window size method to optimally reduce cloud noises (Abtaham Savitzky and J.E.Golay, 1964). The equation of Savitzky-Golay smoothing filter is described as:

$$Y = \sum_{j=-n}^{n} c_j y_i + j \tag{5}$$

Where Y is the smoothed value, y_i indicates the original pixel value in images, and n is the filter window size that can be adjusted upon requirements. A 2nd-order polynomial fit in a window size of 4 is commonly accepted in the Savitzky-Golay filter (Jönsson et al., 2002).

I downloaded the TIMESAT program to perform Savitzky-Golay filtering and other simulation process of the NDVI time series (Jönsson et al., 2002). The TIMESAT program requires more than one year data as inputs. Here I doubled one year data and picked $n = 4$ as the first order in Savitzky-Golay smoothing. , the 2nd-order polynomial fit value is little larger than initial filter window size, I chose 5 to avoid important data missing.

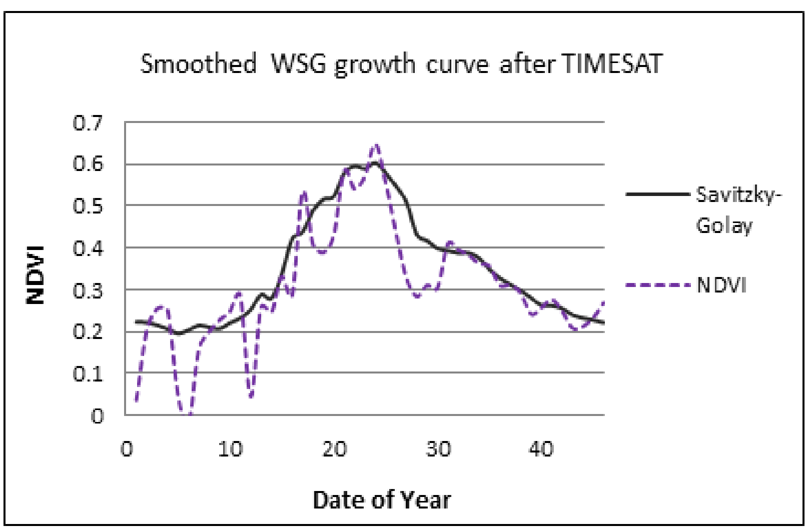

Figure 6. The function of Savitzky-Golay filter.

As shown in Figure 6, the Savitzky-Golay filter smoothed out the time-series trajectory and most of the noises are removed. At the same time, it remains the original trend of the trajectory. With the Savitzky-Golay filter program, the 46-band time-series NDVI image of each year (2001, 2003, 2005, 2007, and 2009) was smoothed for further analysis.

3.2.2 TIMESAT: extracting phenology metrics

In past studies of time series analysis, the smoothed trajectory curves are often simulated with theoretical functions to extract useful phenology metrics (Scott Menard, 1995; Jönsson et al., 2002; Saro Lee, 2004) results. In this research, I applied the TIMESAT program to extract one year growth period's phenological parameters from the NDVI time series. The TIMESAT program could extract 11 phenological parameters below (Jönsson et al., 2002):

1) **Start of the season**: the date in the left that is higher than user-defined starting date value.
2) **Time for the end of the season**: time for right value lower than user-defined starting date value.
3) **Length of the season**: from start date to end date.
4) **Base Level**: left and right minimum value's average in one growth period.
5) **Time for the middle of the season**: the mean of all values in growth period.
6) **Largest data value for the fitted function during the season**: peak value.
7) **Seasonal amplitude**: difference value between peak and minimum.
8) **Rate of increase at the beginning of the season:** calculated as the ratio between the values of start season and at the left 80 % level divided by the corresponding time difference.

9) **Rate of decrease at the end of the season**: calculated as the ratio between the values of end season and at the right 80 % level divided by the corresponding time difference.

10) **Large seasonal integral**: indicates the season from start to end.

11) **Small seasonal integral**: shows the season and base level from start to end.

The simulated trajectory curve with an Asymmetric Gaussian function is demonstrated in Figure 7.

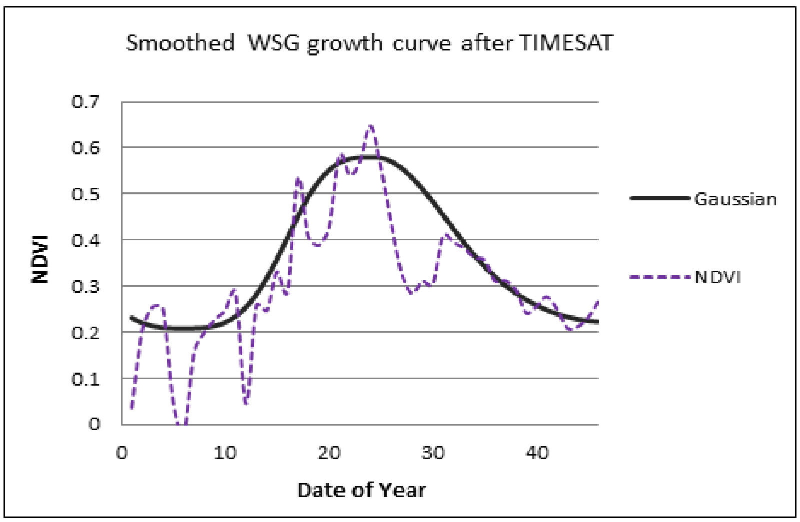

Figure 7. The function of the Asymmetric Gaussian simulation.

3.2.3 User-defined phenology metrics

For a time-series curve with a relatively normal distribution (along a growing season), the Asymmetric Gaussian filter reveals its phenology variation well (as shown in Figure 7). However, no all crops grow in a normally distributed pattern. For example, winter wheat reaches peak greenness in early spring and its NDVI drops down rapidly in summer. The NDVI of winter wheat fields may remain low until re-planting in later fall, or follows a shift planting pattern to be replaced by soybean in fall. More specifically in this research, WSG and CSG follow significantly different growth patterns from annual crops. With much longer growing season, their growing distribution cannot be simply normal distribution. The growth patterns of these major crop types are demonstrated in Figure 8.

Vegetation types shown in Fig.8 are cool season grass, warm season grass, corn, and winter wheat.

Therefore, the Asymmetric Gaussian simulation in the TIMESAT program does not satisfy all land covers in the Great Plains. Here I introduced several additional phenology metrics that I extracted from the NDVI time series in the ENVI Interface Description Language (IDL) programming environment. Applying CDL data as ground truth reference, I extracted a set of time series trajectories for the 4 annual crops (corn, soybean, winter wheat, spring wheat) and the 2 grasses (WSG, CSG). Based on these training data, I calculated three additional phenology metrics that are not extracted from TIMESAT.

1) Peak NDVI value – The maximal NDVI value along a growing season. Higher value represents vegetation grows more luxuriant. It shows the information of vegetation energy absorbance and photosynthetic efficacy that could be applied for evaluating different vegetation. For example, annual crops often reach higher peak NDVI values than grass or shrub.
2) Peak date – the date when NDVI reaches its peak value along a growing season. This value displays the various vegetative growth cycles and rate of growth, and is useful for pattern classification. For instance, corn arrives its peak date in September to October, while grass is in spring or early summer; therefore, corn and grass has different vegetative period and energy absorbance and circulatory system.

3) $\Delta NDVI$ value. It is the difference between the maximum NDVI value in spring and the minimum NDVI value in summer. In general, most of vegetation is assumed to reach peak growth in late spring or summer. But certain special vegetation, especially winter wheat, has separate rhythm that achieves the highest peak value much earlier. Therefore, it has the most NDVI dropdown between spring-early summer. This threshold plays an important role in identifying winter wheat.

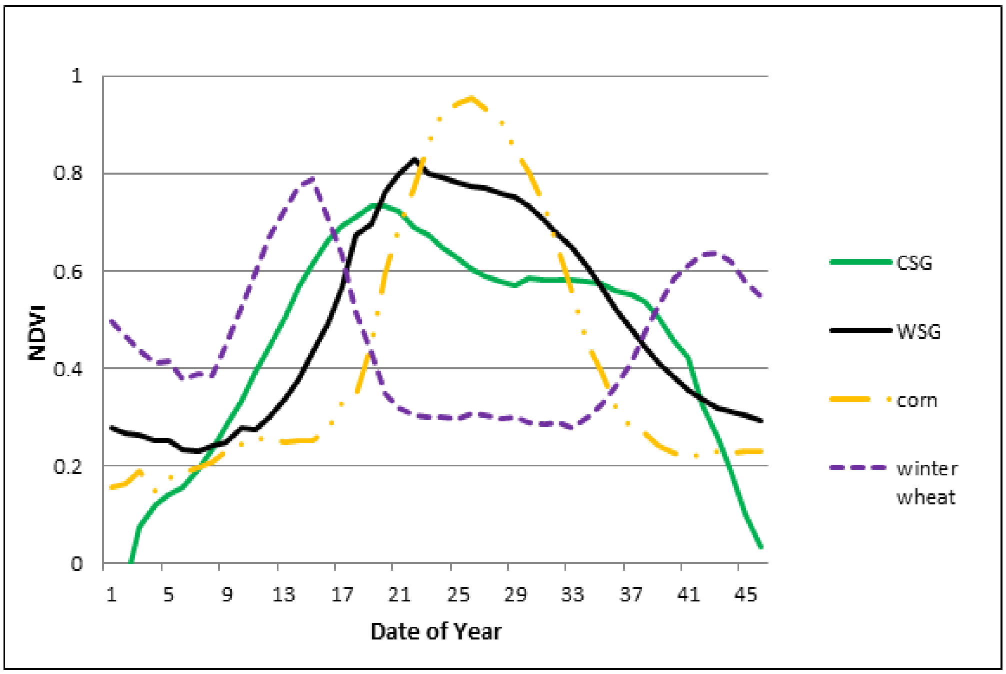

Figure 8. Different vegetation growth patterns in one year (an example of 2007).

3.4 Regular Regression Classification

The phenology matrixes extracted in previous section are the primary inputs for WSG/CSG grassland mapping in my research. Regular classifiers such as maximum likelihood) did not work well because it was not possible for me to collect enough ground truth data, which varied in different geographic locations over the vast area of the Great

Plains. Rather, it was recognized that each of the 6 class types had its unique phenology features, and their geographic variations had been understood well. For example, cool season grasses often grow in northern states while warm season grasses are common in southern and eastern states that were covered in the TallWSG Prairie. Winter wheat was more common in southern states while spring wheat was observed in northern states.

These relationships could be better applied in regression classification methods. In this part, I have tested several different regression methods, such as logistical regression classification and regular regression classification. Objectively and in nominal cases, logistical regression is powerful in estimating relationships between different variables and providing reasonable assessment results for pattern recognition (Rice.J.C, 1994). But my study is unique and demands a quantity of requirements or limitations. Firstly, this is a large-scale classification and is in lack of ground observations that could be used for training data sampling. The amount of my sample points is 216, lower than logistical regression basic demand, so serious deviation of the estimative function occurred in experiment when I implemented logistical regression. Next, I need multiple predicted classes after pattern classification. The logistical regression did not perform well due to the complicated vegetation growth features in large data set. After several tests, I choose the regular regression classification method to classify WSG, CSG and the major annual crops in the Great Plains, considering that these classes hold unique phenology features and common variations in the geographic contexts

3.4.1 Training Data Collection

The CDL products are used as reference to select reasonable sample points for annual crops. The training samples of TallWSG are mostly from the Flint Hills grassland, and the samples of short WSG are randomly picked at the CPER. The method of picking sample points is basing on the information from Cropland Data Layers and predecessor's works (Wang et al. 2010). These training datasets are selected by referring to published resources as well as checking phenology patterns with MODIS time series to remove outliers. For example, past studies display that there are some CSG grasses growing in the west of TallWSG prairie, which locates in central North America and is in the northeast of the Great Plains (Wang et al. 2010). Meanwhile, warm season grasses widely exist in this area.

The classes I tried to identify in this research include corn/soybean, winter wheat, spring wheat, cool season grass, tall WSG, and short WSG, CSG. The reasons that I combined corn and soybean into one group are: 1) I found that, compared to grass and wheat classes, they have very similar growth rhythm after spectral smoothing, includes peak NDVI value, vegetative cycle, and peak date; 2) Practically soybean and corn are alternative crops and small fields are often mixed in large agricultural lands, which result in mixed corn/soybean pixels in the MODIS image. Therefore corn and soybean are treated as one class in regression analysis.

Training dataset is listed in table 3. Due to the vast coverage of the Great Plains, the geographic coordinate position factors (latitude and longitude) are also considered as input variables in this research. At each sample point, a total of 16 phenological features are used as inputs of the classification process:

1) Time for the start of season
2) Time for the end of the season
3) Length of the season

4) Base level

5) Time for the mid of the season: Peak Value.

6) Largest data value for the fitted function during the season

7) Seasonal amplitude

8) Rate of increase at the beginning of the season

9) Rate of decrease at the end of the season

10) Large seasonal integral

11) Small seasonal integral

12) Peak NDVI value

13) Peak data

14) The different NDVI value between early spring and summer peak value

15) Geographic coordinate: Latitude

16) Geographic coordinate: Longitude

Table 3. Training sampling points for the 9 class types that will be classified in the Great Plains.

Vegetation Type	**The number of samples**
Winter Wheat	44
Spring Wheat	48
Soybean/Corn	34
CSG	28

Tall WSG	32
Short WSG	30

3.4.2 Regular regression classification approach

The regression classification is implemented in the SAS software. The flowchart of the process is shown in Figure 9. The rules of the regression classification are to find the input variables (phenology features) that have the highest R-Square (coefficient of multiple determination for selected variables) and lowest AIC (Akaike Information Criterion) values. The AIC displays a tradeoff between the precision of fit and corresponding number of parameters used. In general, if we add one feature into the regression and the AIC value also increase, then this new feature cannot be used in classification (Hirotugu Akaike, 1974). In this sense, I selected the input parameters that have small AIC and large R^2. Using these selected features as independent variables, the SAS builds the regular regression equation (as shown in table 4) in the Analysis of Variance module. The classes to be identified are the dependent variables in the equation.

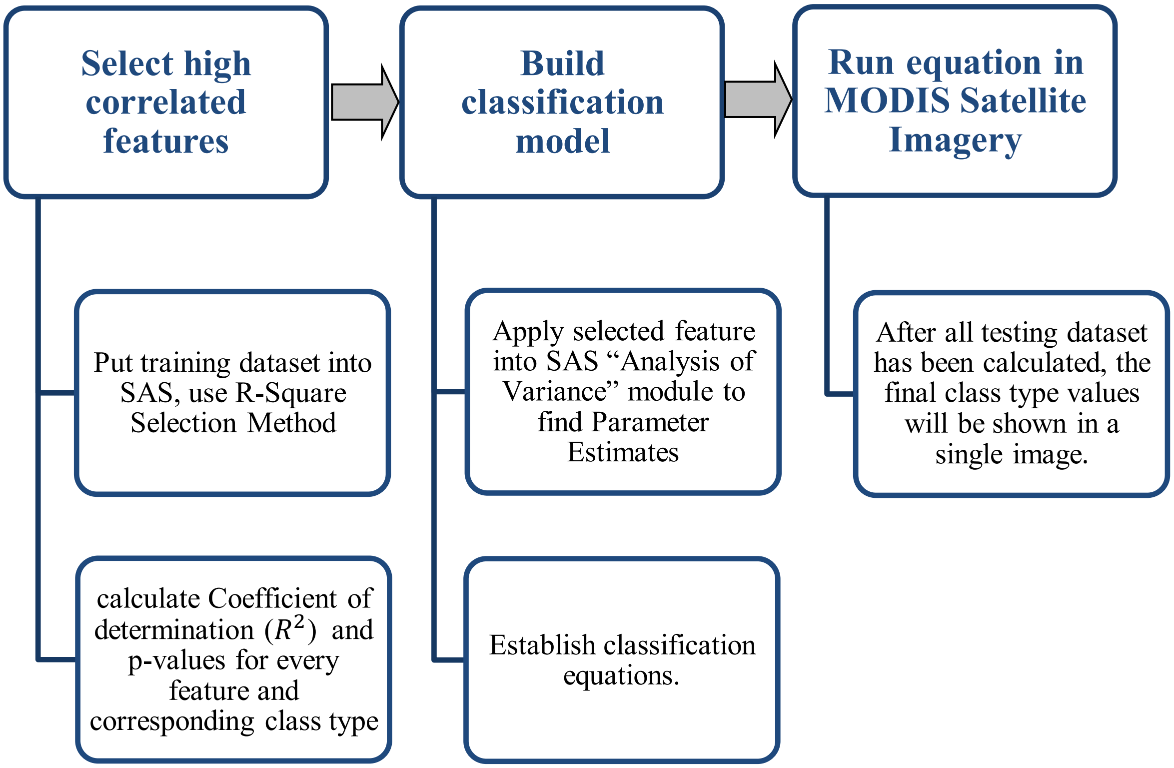

Figure 9. Work flow of the regression classification.

In a multi-linear regression, one important precondition is that all variables must be independent. Without this premise, it cannot be called a rigorous regression classification. Hence, the next step is to test the independency of the selected input parameters.

To test their correlation, I have randomly selected one sample (year in 2009) with 16 phenological parameters from 6 individual class types (winter wheat, spring wheat, corn/soybean, CSG, TallWSG, and ShortWSG). Then put this 6 by 16 matrix into MATLAB to generate cross-correlation matrix. The cross-correlation coefficients of the 16 features are calculated by dividing the phenological parameters' standard deviations (Table4). The coefficients in the table are quite low, indicating that these parameters are independent to each other. As a result, all phenological parameters are valid inputs of the regression classification in this research.

Table 4. The cross-correlation coefficient between 16 phenological parameters, the example of 2009, described their relationship in individual class types. And these phenological parameters are represented by Var. 1 to Var. 16 in following table.

	Var. 1	**Var. 2**	**Var. 3**	**Var. 4**	**Var. 5**	**Var. 6**	**Var. 7**	**Var. 8**	**Var. 9**	**Var. 10**	**Var. 11**	**Var. 12**	**Var. 13**	**Var. 14**	**Var. 15**	**Var. 16**
Var. 1	0.036	0.076	0.040	0.000	0.055	0.001	0.001	0.000	0.000	0.021	0.012	0.001	0.318	0.000	-0.192	0.079
Var. 2	0.076	0.177	0.101	0.001	0.124	0.003	0.002	0.000	0.000	0.056	0.033	0.003	0.771	0.000	-0.449	0.187
Var. 3	0.040	0.101	0.062	0.001	0.069	0.002	0.001	0.000	0.000	0.034	0.021	0.002	0.453	0.000	-0.257	0.108
Var. 4	0.000	0.001	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	-0.002	0.001
Var. 5	0.055	0.124	0.069	0.001	0.088	0.002	0.001	0.000	0.000	0.037	0.022	0.002	0.533	0.000	-0.313	0.130
Var. 6	0.001	0.003	0.002	0.000	0.002	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.014	0.000	-0.008	0.003
Var. 7	0.001	0.002	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.010	0.000	-0.005	0.002
Var. 8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	-0.001	0.000
Var. 9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000
Var. 10	0.021	0.056	0.034	0.000	0.037	0.001	0.001	0.000	0.000	0.020	0.012	0.001	0.256	0.000	-0.141	0.060
Var. 11	0.012	0.033	0.021	0.000	0.022	0.001	0.000	0.000	0.000	0.012	0.008	0.001	0.158	0.000	-0.084	0.036
Var. 12	0.001	0.003	0.002	0.000	0.002	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.014	0.000	-0.008	0.003
Var. 13	0.318	0.771	0.453	0.004	0.533	0.014	0.010	0.001	0.001	0.256	0.158	0.014	3.634	0.002	-1.988	0.843
Var. 14	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	-0.001	0.000
Var. 15	-0.192	-0.449	-0.257	-0.002	-0.313	-0.008	-0.005	-0.001	0.000	-0.141	-0.084	-0.008	-1.988	-0.001	1.150	-0.480
Var. 16	0.079	0.187	0.108	0.001	0.130	0.003	0.002	0.000	0.000	0.060	0.036	0.003	0.843	0.000	-0.480	0.202

3.5 Results

3.5.1 Regression equations

With training data samples for each year, the regression equations were built in SAS. Besides, the location of training samples in 10 years has a slightly change to guarantee reliability of the classification. Table 5 shows the 2005 example of estimating equation's parameters and machine selected features. The reason why I selected the variables in the following equation is the estimated error rates are very low. The regular regression equation is thus represented as:

$$Class\ N = 27.19 + 0.33 * x1 + 11.78 * x2 + (-0.08) * x3 + \cdots + (-0.98) * x12 + (-2.66) * x13 \qquad (6)$$

X1 represents Intercept, X2 represents Length of the season, X3 represents Base Level, X12 represents The Δ NDVI value, and X13 represents Longitude)

Different variables were used in different years due to variations of vegetation's phenological characteristics in the Great Plains in the 10-year period. For instance, the peak date of corn in 2005 was around 217 DOY (day of year), but in 2009 it was delayed to 255. Some parameters may become less important in the regression of certain year.

3.5.2 Class maps

The classified vegetation maps are displayed in Figure 11 to Figure 15 in the 5 years. From classification maps from 2001 to 2009, as shown in these maps, cropland areas have not changed much in these 10 years. Corn and soybean grow in the northeast of the

Great Plains, which is covered in the Corn Belt in the Midwest. Winter wheat keeps the zonal distribution in Oklahoma and Texas, and spring wheat always distributes primarily in the northern area – North Dakota.

Table 5. The 2005 example parameters extracted from SAS

Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr >|t|**
X1: Intercept	1	27.18555	6.00737	4.53	<.0001
X2: Length of the season	1	0.33488	0.05425	6.17	<.0001
X3: Base Level	1	11.78207	4.13940	2.85	0.0049
X4: Time for the mid of the season	1	-0.07721	0.03007	-2.57	0.0109
X5: Largest data value for fitted function during the season	1	-2.72084	1.80019	-1.51	0.1322
X6: Rate of increase at the beginning of the season	1	5.27092	4.85194	1.09	0.2786
X7: Rate of decrease at the end of the season	1	6.77722	5.66073	1.20	0.2326
X8: Large seasonal integral	1	-0.63220	0.14466	-4.37	<.0001
X9: Small seasonal integral	1	0.35121	0.12841	2.74	0.0068
X10: Peak NDVI value	1	0.18770	0.05010	3.75	0.0002

X11: Peak Date	1	-0.11758	0.02341	-5.02	<.0001
X12: The Δ NDVI value	1	-0.97714	1.06436	-0.92	0.3596
X13: Longitude	1	-2.65930	0.52194	-5.10	<.0001

The grass patterns in the Great Plains that is consistent all over the 10 years. For example, TallWSG always dominates in Flint hills; ShortWSG mostly appears in the west.

Temporal shifts of CSG, TallWSG and ShortWSG, however, are obvious in these class maps, especially in the central Plains that belonged to mixed prairie as defined in Chapter2. The CSG grasses dominate the cold areas such as South Dakota and Nebraska in 2001 and 2007, but were less dominant than TallWSG in other years. ShortWSG primarily existed in the western Great Plains and mixes with shrubs in most years, but in 2005, its coverage dramatically declined. These temporal variations will be further examined in climate change analysis next chapter.

3.5.3 Accurate assessment

The CDL products provide fairly reasonable ground truth for annual crops. However, there has not any WSG/CSG product published in the Great Plains. It is also difficult to collect intensive ground observations of grass types in the vast area of the Great Plains. In lack of ground observations for these grass types, I could not run conventional accuracy assessment procedures such as Contingency table approach (Karl Pearson, 1904).

To evaluate the accuracies of the class maps, I used three of the four focus sites of this research. Most of these focus sites are monitored by USDA or other agencies and programs and therefore, ground records are partially available. Class results in these sites are compared with ground observations and records that are published online. The Sandhills grassland focus site was not examined because it is a typical mixed grass prairie covering all types of grasses: CSG, TallWSG, and ShortWSG. Percent covers in this area thus vary in different years, depending on changing climate conditions.

The first site for accuracy assessment is the CPER, the Central Plains Experimental Range in Colorado. Upon the reports from the USDA Agricultural Research Service, blue Gramma, a typical ShortWSG species, is the primary grass type in CPER. As shown in table 6, the results in this research agree with the ARS reports: in any year, the percentage of ShortWSG is much higher than other grass types.

Table 6. The MODIS-derived percent covers of major grass types in CPER: 2001 to 2009.

	2001	**2003**	**2005**	**2007**	**2009**
cool season grass	8.85%	2.74%	12.24%	30.70%	16.95%
shortWSG	46.69%	37.47%	38.31%	59.31%	68.28%
tallWSG	39.48%	52.67%	27.02%	1.54%	11.85%
shrub	4.98%	7.12%	22.43%	8.45%	2.92%

Table 7. The MODIS-derived percent covers of major grass types from 2001 to 2009 in the Flint Hills prairie.

	2001	**2003**	**2005**	**2007**	**2009**
cool season grass	6.19%	0.76%	6.27%	35.88%	14.39%

shortWSG	2.66%	0.24%	12.37%	0.00%	0.97%
tallWSG	90.28%	98.92%	67.30%	64.05%	79.65%
shurb	0.87%	0.08%	14.06%	0.07%	5.00%

The 2nd site is the Flint Hills prairie in Kansas. It is operated by Konza Prairie Biological Station (KPBS). According to their research projects, TallWSG species covers more than 80% of this prairie. In my class maps of the 5 years, the percentage of TallWSG also dominates the prairie (Table 7). The TallWSG coverages in 2005 and 2007 are abnormally lower than the average, which is further examined by linking them to climatic factors in next chapter.

The last focus site is the Montana Prairie in the north of the Great Plains. Without ground observation data in this area, the only way to perform accuracy evaluation of my results is to compare them with CDL data and to lay on basic phenological knowledge in this area. In cropland data layer this region belongs to grassland. In Montana's official state website (http://mt.gov/), the Sandberg Bluegrass (Poa secunda) is a native cool season grass in this grassland. It is also understandable that temperature in Montana is cold enough so that only cool season grass could grow and its growth length is short due to the short growing season.

I also noticed a unique phenomenon about grass growth in this site. In 2005 and 2007, during the classification process, there are clusters of cool-season grasses that displayed extremely short growing season (Figure 10). Compare Figure 10 with CSG and WSG curves, the short-term grass has similar phenological characters as CSG. Both grass types could have two grow periods, one in spring and the other in fall, and have almost peak NDVI values. However, the growth length of short-term grass is much shorter than a typical CSG.

In Table 8, we found that that CSG is the major grass type in the prairie and this special short-term grass only appeared in Montana prairie site and dominated the area in 2005 and 2007. It should be strongly related to climate conditions that are further examined in next chapter.

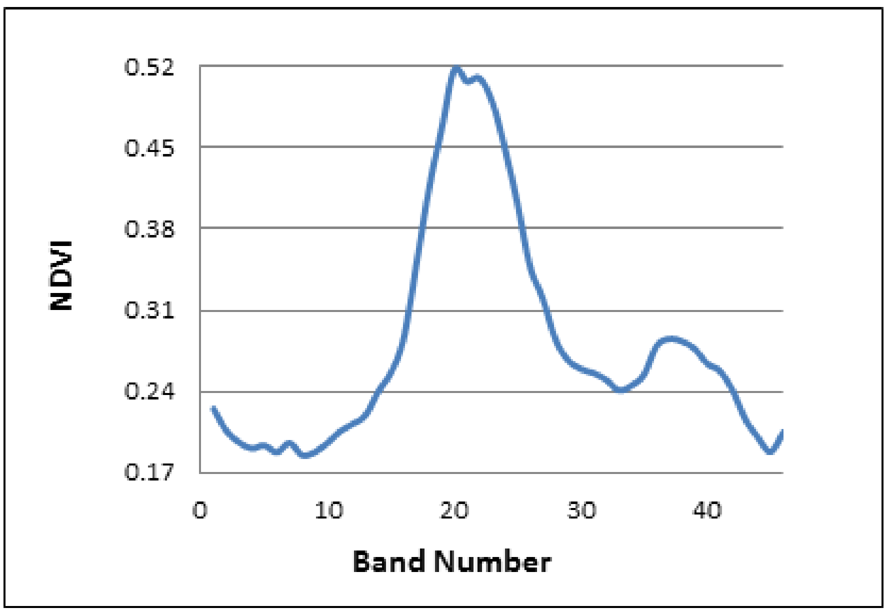

Figure 10. An example growth cycle of a special type of grass (short-term grass) in 2007.

Table 8. The MODIS-derived percent cover of major grass type's variation from 2001 to 2009 in the Montana Prairie.

	2001	2003	2005	2007	2009

cool season grass	86.32%	88.87%	36.26%	30.41%	84.30%
tallWSG	12.72%	11.13%	16.54%	0.00%	14.69%
Shurb	0.96%	0.01%	0.22%	0.01%	1.01%
short-term grass	0.00%	0.00%	46.98%	69.58%	0.00%

3.6 Discussion

In my thesis research, I mapped vegetation distributions of several major vegetation types (corn/soybean, winter wheat, spring wheat, CSG, TallWSG, and ShortWSG). It is very time consuming. I downloaded 10-year MODIS satellite imagery in 4 tiles (H09V05, H10V04, H10V05 and H11V04) and the cropland data layers in 10 states. I re-projected all images to orthophoto maps, calculated MODIS NDVI values, mosaicked and clipped MODIS satellite imagery to match the study area, regrouped attribute data, and resampled all images into 500-m resolution. It is a huge task for image processing. I have processed 3680 MODIS images with total data size over 5 Terabytes.

Via time series analysis, the information I extracted is rich. The MODIS time series reveal growth rhythms of major land cover types. These phenological differences built the foundation for classification process in my research, and it is significant on spectral values used in different classifiers. In this study, phenology information has less dependency of intensive training data, and insensitive to spatial heterogeneity of land cover types. Hence, the regular regression classification method is more adaptive to large coverage area than regular classifier, such as Maximum Likelihood Estimation.

Besides, in lack of ground observation data, the regular regression classification approach performs also better than general method, such as supervised and unsupervised approaches. One of the reasons for this advantage is geographic consideration put into my regular regression approach, although I have limited sample points. The geographic consideration came from: (1) sample points were measured dispersedly, not concentrated in one area; (2) vegetation represents slightly altering owing to distinct elevation, latitude, and other factors. Hence, the measurement standard cannot follow one permanent pattern; (3) the values of latitude and longitude were put into the estimation of regular regression parameters. Then all considerations were combined to reach the good results with limited sample points in this research.

The regular regression classification approach has delineated the clear boundary and few confusion areas among vegetation. They are shown as Figure 11-15. This approach demonstrates the attempt to explore the large area image classification, proofs it could be applied into similar research studies. Given more time to investigate regression methods and parameter estimation the classification equation and results should be further improved.

Classification maps in 2001, 2003, 2005, 2007, and 2009 show reasonable grass and annual crop distributions in comparison with USDA CDL products and ground records. Further discussion: Link variations in your class maps to climate change, which is studies in next Chapter.

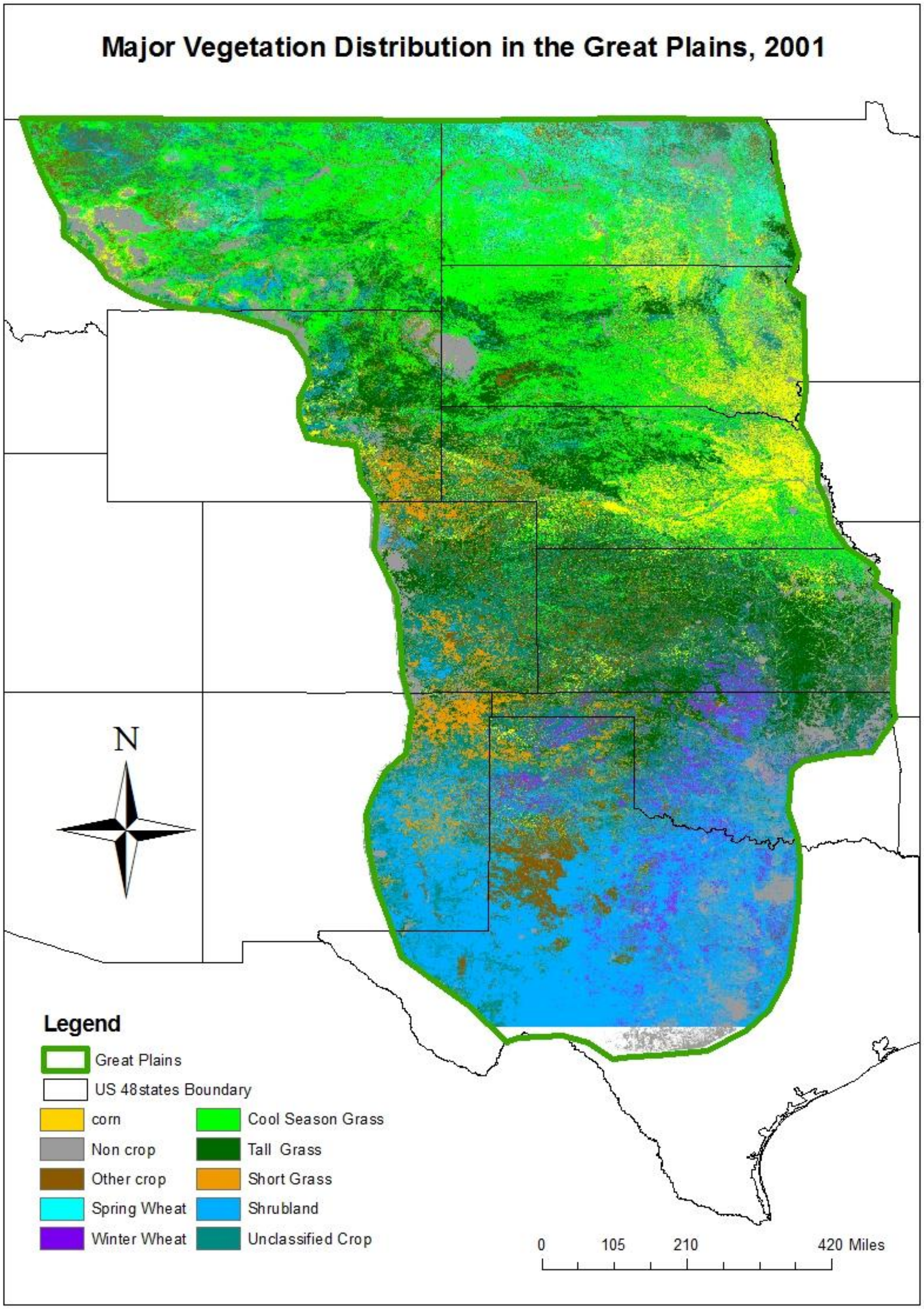

Figure 11. The 2001 vegetation distribution map in the Great Plains.

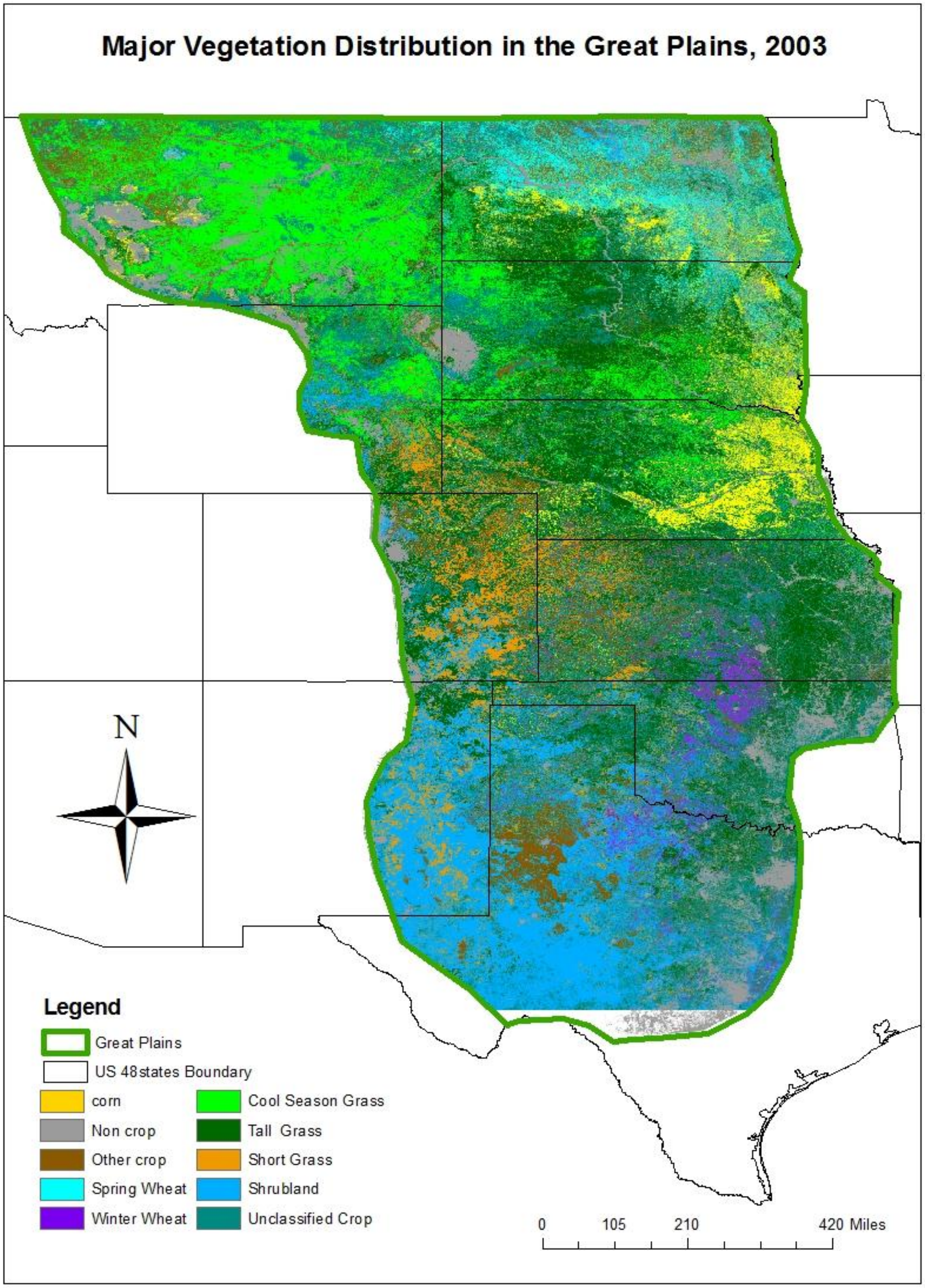

Figure 12. The 2003 vegetation distribution map in the Great Plains.

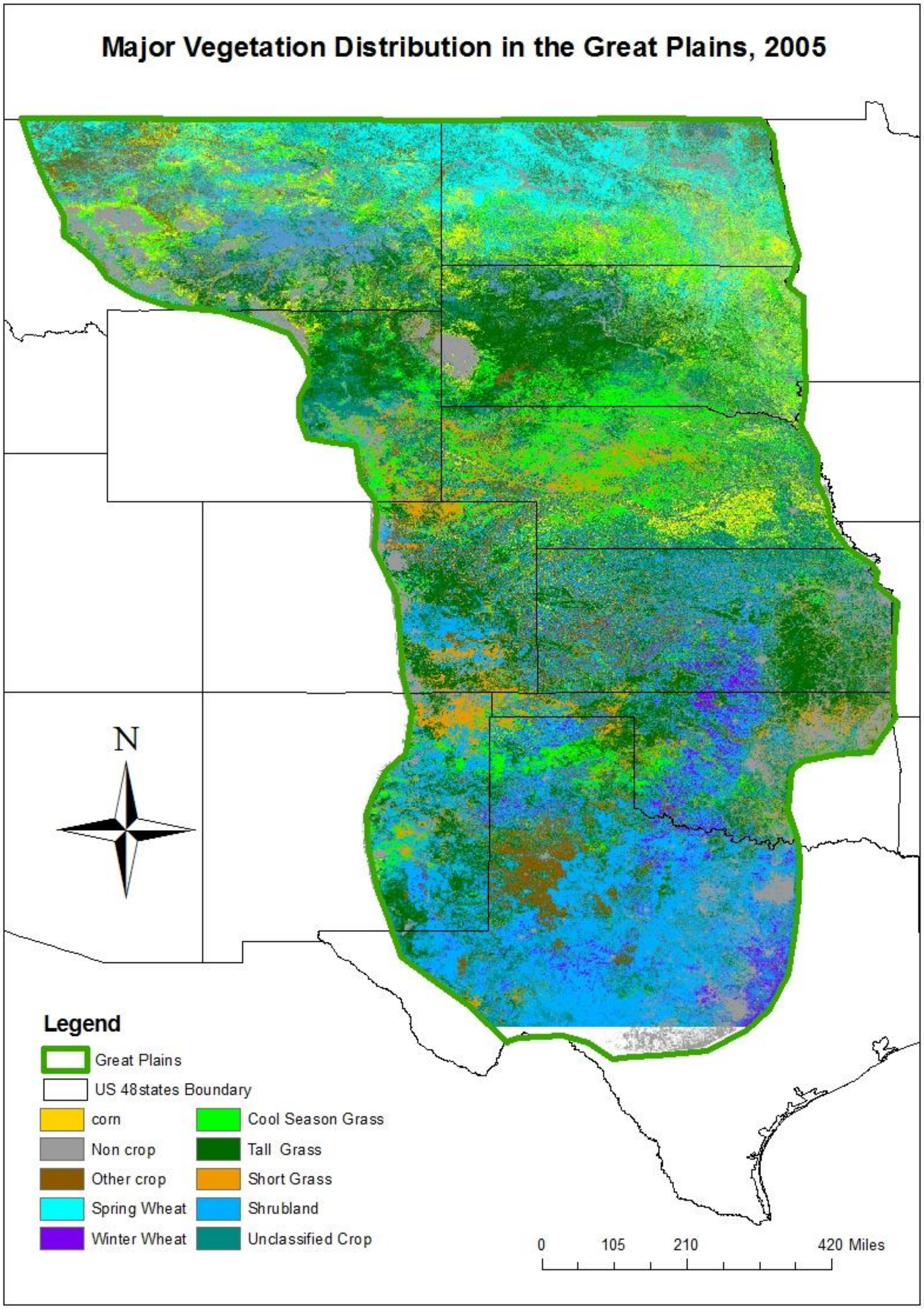

Figure 13. The 2005 vegetation distribution map in the Great Plains.

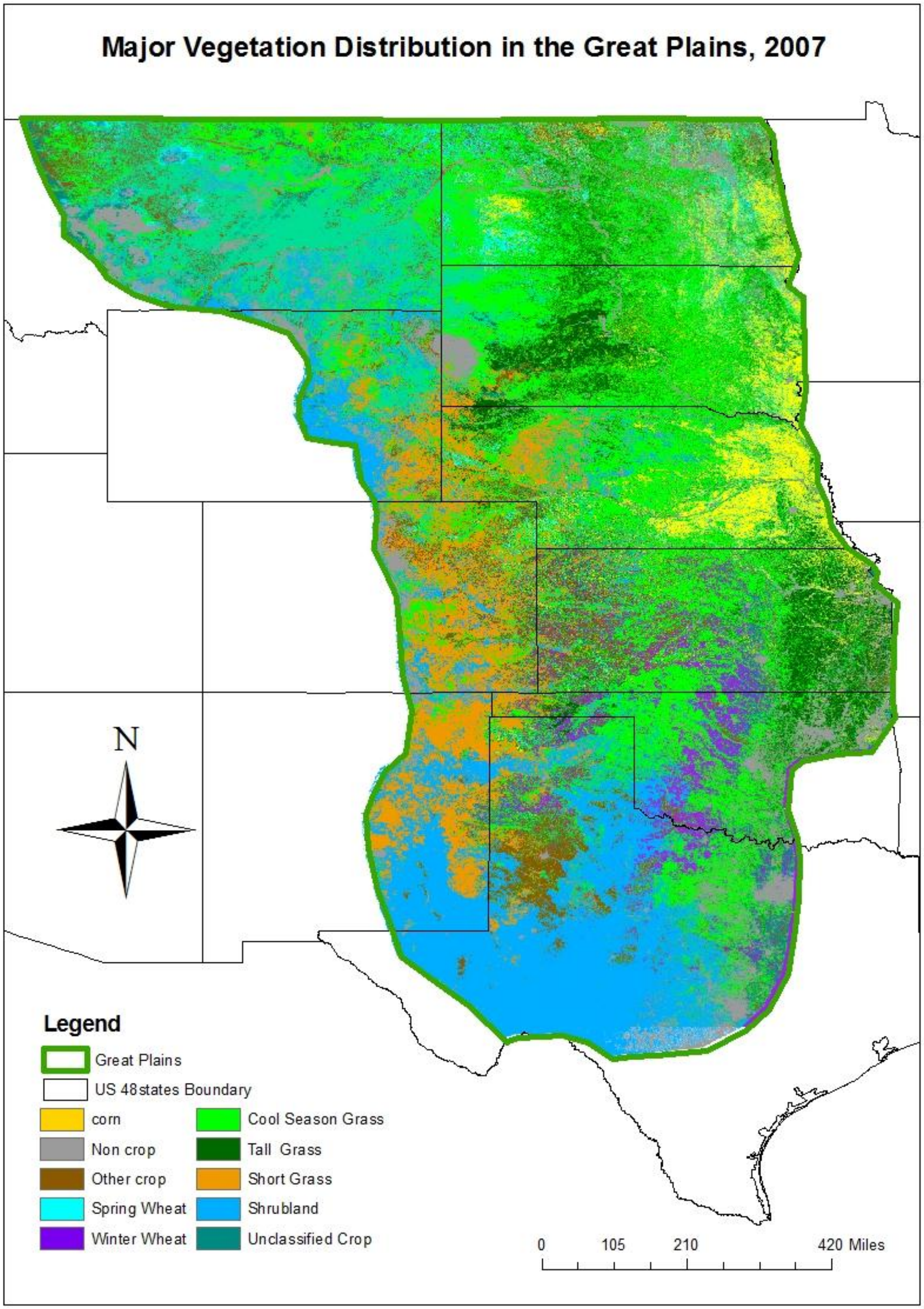

Figure 14. The 2007 vegetation distribution map in the Great Plains.

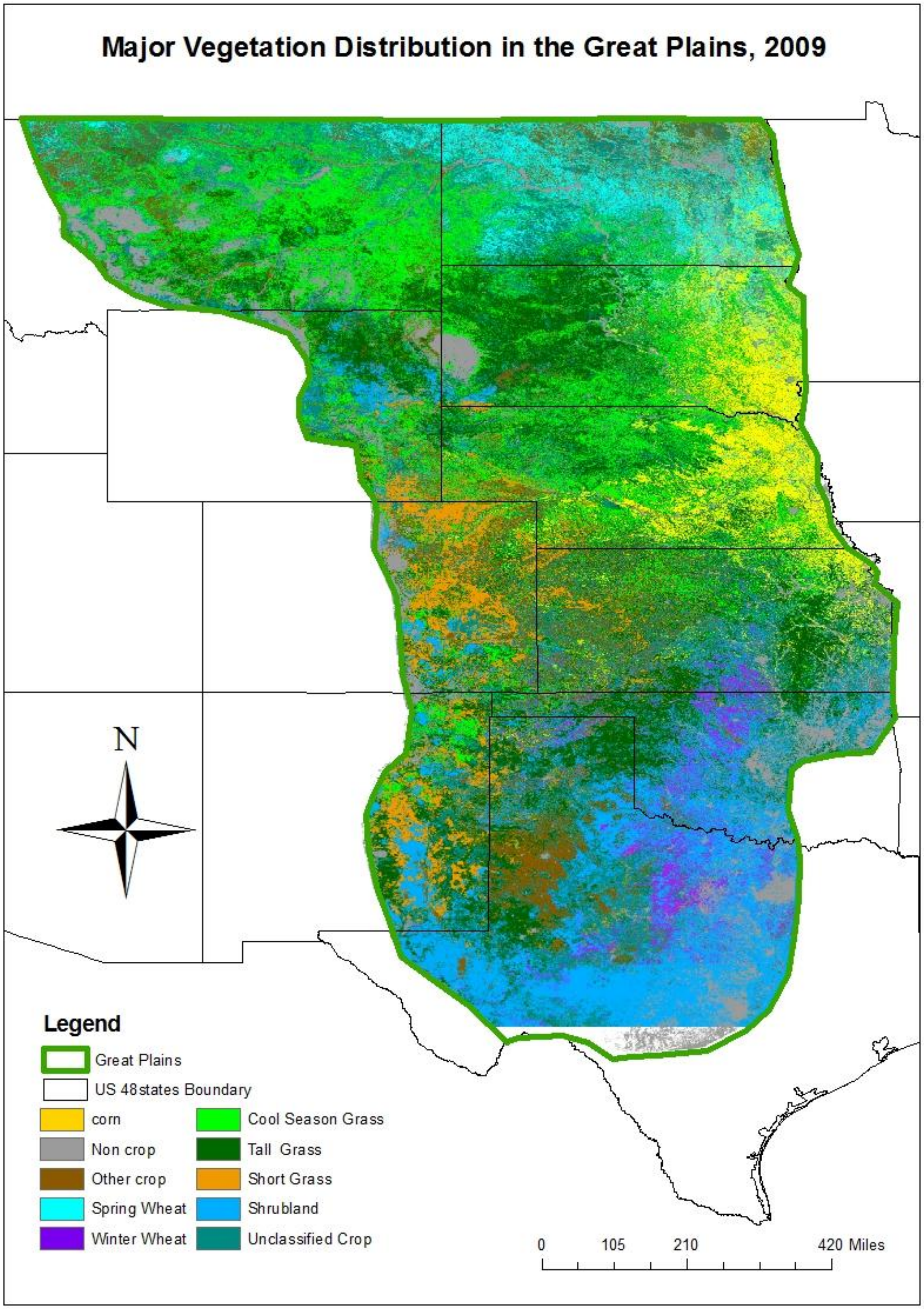

Figure 15. The 2009 vegetation distribution map in the Great Plains.

CHAPTER IV: Linkage of grass distributions and climate variation

It is found in previous chapter that the spatial distributions of major grass types (CSG, TallWSG, ShortWSG) in the Great Plains vary in 2001, 2003, 2005, 2007, and 2009. This variation may be related to the changing climate conditions in these years. This chapter explores climate data in the Great Plains and links it to the grassland distributions in the region. It could be treated as preliminary research in investigating grassland responses to climate change at a regional scale.

4.1 Temperature and Precipitation Data in 2000-2009

Climate data in the Great Plains in past years are downloaded from the PRISM (Parameter-elevation Regressions on Independent Slopes Model) program at the Oregon State University (http://www.prism.oregonstate.edu/). PRISM is the parameter-elevation regressions on independent slopes model (http://www.prism.oregonstate.edu/). The goal of this special knowledge-based system is to produce continuous monthly and yearly climatic parameters. The PRISM database provides various climate data records in the recent 20 years, primarily monthly products in United States from 1985 to present that include precipitation, average maximum temperature, average minimum temperature, average dewpoint, and anomalies. PRISM data also includes other yearly world-wide climate metadata, such as the 800-m climatology normal metadata in Pacific islands from

1971 to 2000. The PRISM products' spatial format is ACSII GRID. This continuous data could be easily converted to raster image for climatic analysis.

Climate data off the growing season is not analyzed in this research. To fit in the growing season in the Great Plains, I downloaded monthly precipitation and temperature data from March to October for each year (2000-2009). Further, to simplify the process and to reduce data redundancy, I re-grouped each climate data type into three seasons: spring (March to May), summer (June to August), and fall (September to October). In each season, four climate variables are re-calculated and regrouped after downloading the metadata: total rainfall, average precipitation, maximum temperature, and average temperature. Therefore, in total there are 12 climate factors are tested in the correlation analysis: Table 4 gives all parameters here (22 in total).

4.2 Correlation analysis

To link the variation of grass percent covers in the Great Plains to climate change in the past 10 years, I performed correlation analysis to find the climate factors that are highly correlated to the variation of grass distributions. While climate data are available each year in 2000-2009, percent covers of grasses are extracted for 5 years (2001, 2003, 2005, 2007, and 2009). In correlation analysis, both percent cover of specific grass types and climate factors are analyzed in 5-point series (5 years). The correlation coefficient represents the normalized measure of the strength of linear relationship between two variables. Given variables X and Y, the correlation coefficient could be calculated as (Rodgers and Nicewander, 1988):

$$r_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E((X-\mu X)(Y-\mu Y))}{\sigma_X \sigma_Y} \tag{8}$$

I firstly collected 12 climatic factors in three seasons from 2000 to 2009, and then calculated the percent cover of grasses and annual crops (CSG, TallWSG, ShortWSG, corn/soybean, winter wheat, and spring wheat) in the four study areas. For each year, I grouped these data into a 22 by 5 matrix and put it into MATLAB to run correlation coefficients with the module “corrcoef”. This module generates correlation coefficients matrix and returns a p-value matrix for the no-correlation hypothesis. At 95% confidence interval, when p-value is less than 0.05, the correlation is statistically significant. Therefore, if a climate factor corresponds to the highest r value and p-value less than 0.05, it is determined to be significantly related to the variation of grassland distributions in the Great Plains.

4.3 Results

The Great Plains covers different geographic divisions that vary from the Rocky Mountains in the west to the TallWSG Prairie in the east. The climate ranges from the cold, continental climate in the north to the sub-arid, hot climate in the south. Therefore, the geographical variations of climate conditions could be more dramatic than temporal variations in 5 years. For this reason, I did not run correlations analysis of the average climate factors all over the Great Plains. Rather, I performed detailed analysis in the selected 4 focus sites, assuming that climate conditions in these small sites are relatively stable in its geospatial context.

Table 9. Correlation coefficients between climatic factors and grasses variations in four focus area (Montana Prairie, CPER, Central Prairie, and Flint hill grassland). The highly linked climatic factors will be selected for later analysis.

	Climatic Factors	CSG	ShortWSG	TallWSG	short –term grass
Montana Prairie	spring total rainfall	-0.836	0.922	-0.89	0.925
	summer total rainfall	0.107	-0.296	0.612	-0.357
	fall total rainfall	-0.319	-0.224	0.519	-0.296
	spring average rainfall	-0.728	0.87	-0.921	0.888
	summer average rainfall	0.232	-0.426	0.707	-0.481
	fall average rainfall	-0.081	-0.395	0.466	-0.426
	spring max temperature	0.431	-0.044	-0.012	-0.021
	summer max temperature	-0.386	0.549	-0.611	0.559
	fall max temperature	-0.138	-0.11	0.38	-0.152
	spring average temperature	-0.638	0.503	-0.209	0.447
	summer average temperature	-0.137	0.382	-0.551	0.41
	fall average temperature	-0.208	0.257	-0.332	0.26
Central plains experimental range in Colorado	spring total rainfall	-0.604	-0.515	0.606	N/A
	summer total rainfall	0.098	0.539	-0.357	N/A
	fall total rainfall	-0.596	-0.521	0.616	N/A
	spring average rainfall	0.128	0.577	-0.398	N/A
	summer average rainfall	0.809	0.492	-0.597	N/A
	fall average rainfall	-0.234	-0.655	0.544	N/A
	spring max temperature	0.557	0.201	-0.29	N/A
	summer max temperature	0.081	-0.424	0.211	N/A
	fall max temperature	0.084	0.433	-0.356	N/A

	spring average temperature	0.149	0.465	-0.404	N/A
	summer average temperature	0.644	0.758	-0.623	N/A
	fall average temperature	-0.086	-0.619	0.412	N/A
Central Prairie	spring total rainfall	-0.142	0.484	0.254	N/A
	summer total rainfall	0.443	0.276	-0.07	N/A
	fall total rainfall	0.016	-0.074	0.155	N/A
	spring average rainfall	-0.121	0.556	0.229	N/A
	summer average rainfall	0.353	0.156	-0.002	N/A
	fall average rainfall	0.016	-0.074	0.155	N/A
	spring max temperature	0.547	0.567	-0.311	N/A
	summer max temperature	0.253	0.636	-0.317	N/A
	fall max temperature	-0.27	-0.307	0.287	N/A
	spring average temperature	0.921	0.483	-0.936	N/A
	summer average temperature	0.272	0.533	-0.403	N/A
	fall average temperature	0.069	0.412	-0.196	N/A
Flint-Hills Prairie	spring total rainfall	-0.239	-0.45	-0.303	N/A
	summer total rainfall	0.803	-0.435	-0.141	N/A
	fall total rainfall	0.057	-0.588	0.516	N/A
	spring average rainfall	0.789	-0.448	-0.274	N/A
	summer average rainfall	-0.314	-0.468	-0.032	N/A
	fall average rainfall	0.177	-0.633	0.393	N/A

	spring max temperature	-0.097	0.894	-0.35	N/A
	summer max temperature	0.361	0.869	-0.559	N/A
	fall max temperature	0.016	-0.131	-0.55	N/A
	spring average temperature	0.675	0.318	-0.986	N/A
	summer average temperature	0.125	0.956	-0.43	N/A
	fall average temperature	0.491	0.365	-0.929	N/A

4.3.1 Montana Prairie

As I described in the previous chapter, the Montana Prairie in the north of the Great Plains are mostly composed of cool-season grasses. The example of one year grass distribution shows in Figure. 16. From this picture, the majority of the area is covered by grassland, but the percentage changed by temporal factors. The percent covers of grass types in the 5 years are summarized in Figure 17. It confirms that cool-season grasses in this site have short growing seasons in 2005 and 2007. This abnormally could be explained by examining climate variations in the past 10 years.

Figure 18 shows the 2000-2009 variations of all climate factors in this area. With correlation analysis, I found that the climate factors that are highly correlated to grass covers were total precipitation and the average rainfall in spring. The R values of both factors are larger than 0.92. The linear trends of these two factors are show plotted in Figure 18.

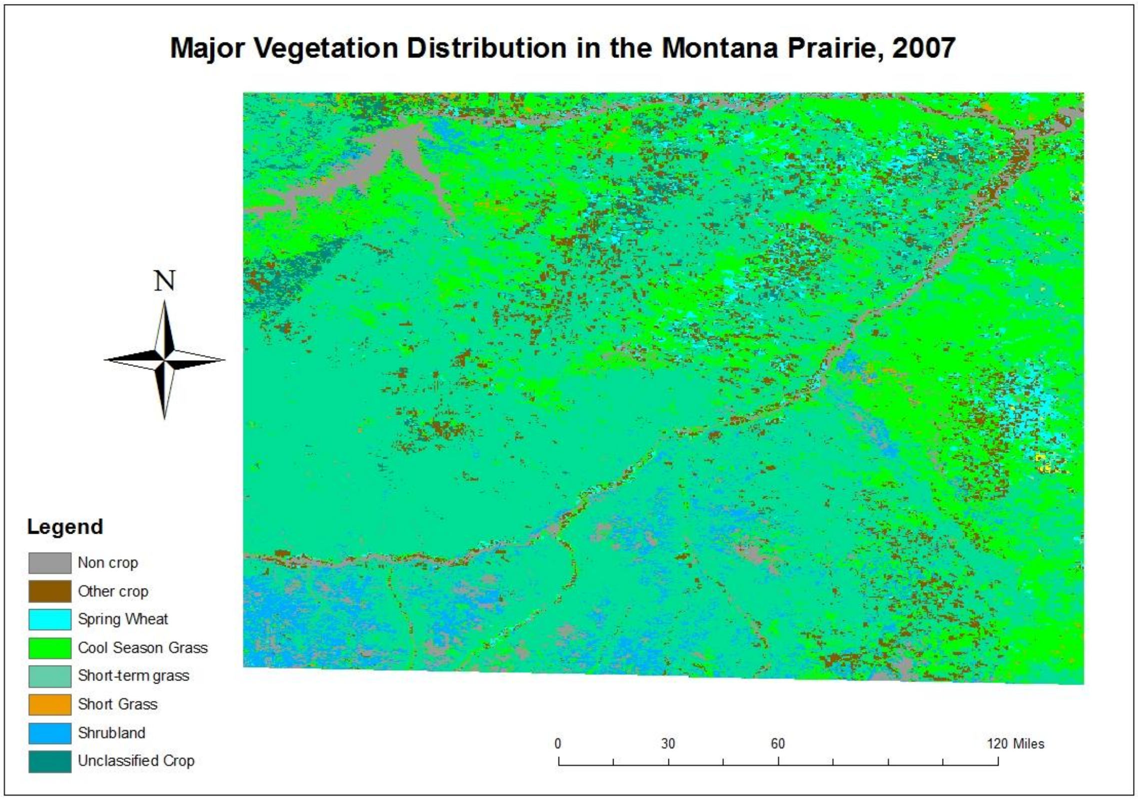

Figure 16. The 2007 vegetation distribution map in the Montana Prairie.

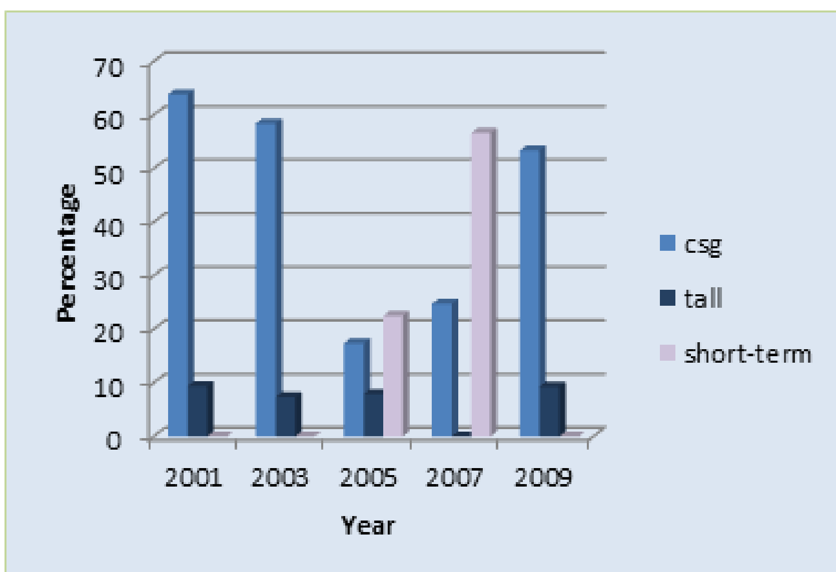

Figure 17. Column chart of grass covers in the Montana Prairie in the 5 years.

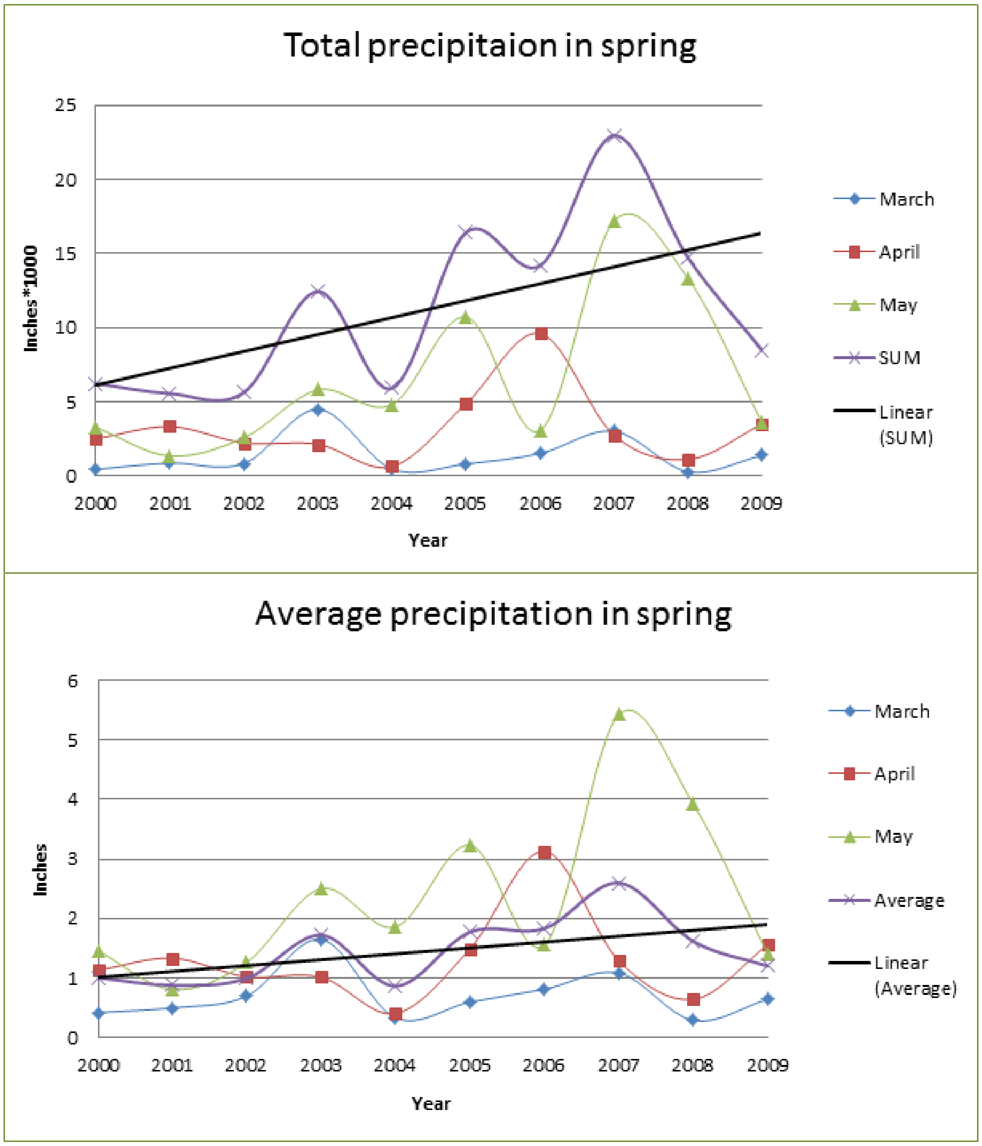

Figure 18. Climate conditions in 2000-2009 in the Montana Prairie.

As shown in the top scatterplot in Figure 18, the total precipitation in spring (March-May) of 2005 and 2007 are much higher than other years. The average precipitation in May is also extremely high in these two years (the bottom scatterplot). The short-term grass growth cycle indicates that this type of grass reaches peak date around May. Therefore, we chould reasonably relate increased precipitation in spring to the expanded distribution of the cool season grass in this grassland.

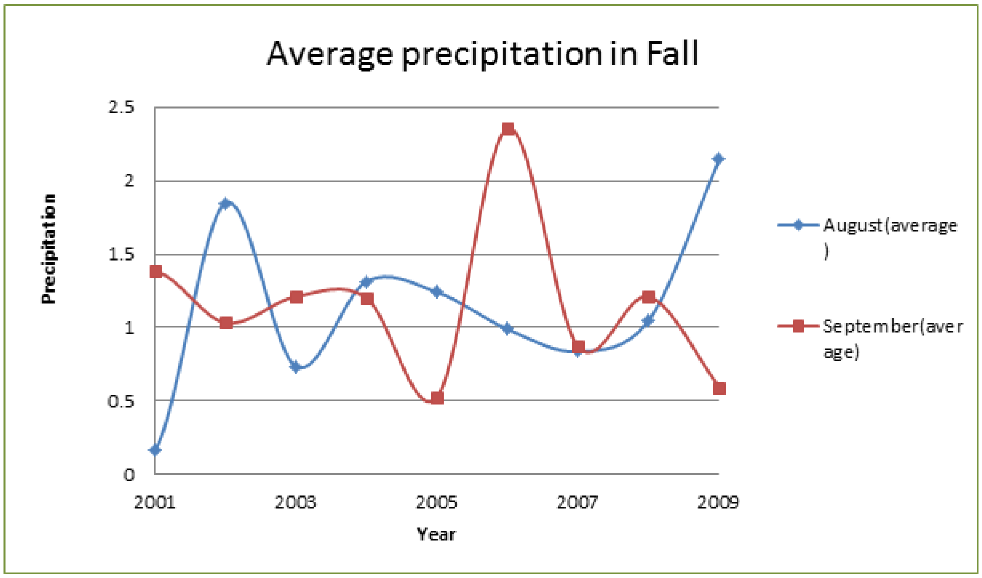

Figure 19. Fall average precipitation variations from 2000 to 2009 in the Montana Prairie.

However, the results indicates that the growth cycle of grasses in 2005 and 2007 are extremely short when comparing with cool season grasses in other areas, even though their phenological characteristics are similar. The only reason to explain this abnormity is that climate after spring must change abnormally, because cool season grass usually has a long growth length from early spring to middle fall. In Figure 19, it shows that the average precipitation in August and September is lower than the medium value. Owing to lack of water, cool season grass cannot have a second growth period in fall. Oppositely, it was quickly dry down. As a result of this short growth, this area has short-term grass arisen in 2005 and 2007.

4.3.2 Central Plains Experimental Range in Colorado (CPER)

This site is a ShortWSG-dominated grassland in Colorado. As shown in the column chart in Figure 20, ShortWSG possesses the largest percentage of the site each year (except

2003), and it's percent cover has an apparent increasing trend in the past 10 years. The increase of ShortWSG cover corresponds to the decline of TallWSG.

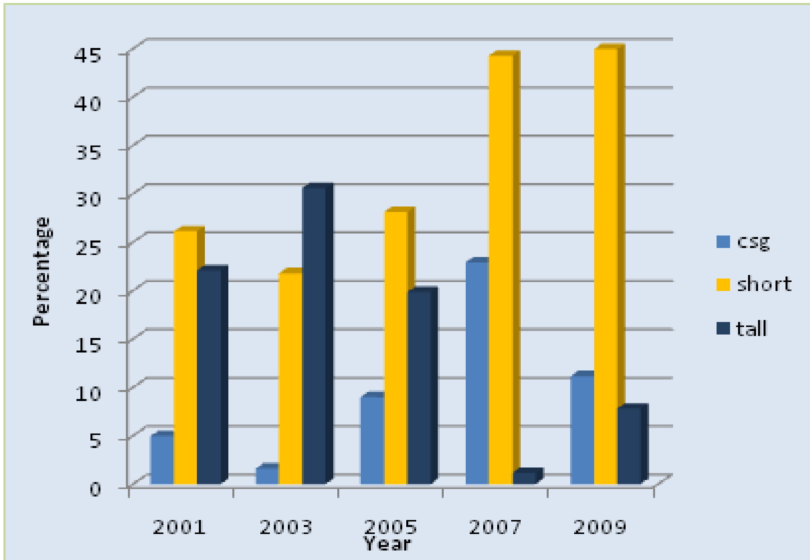

Figure 20. Column chart of grass covers in the CPER in the 5 years.

Upon correlation analysis, grass covers in this site highly correlate with average rainfall and temperature in summer. The R values of these two factors are higher than 0.88. In the scatterplots of these two climate factors (Figure 21), there is a slight decrease of average summer temperature from May to July in the past 10 years, which could be responsible for the decrease of TallWSG covers as shown in Figure 20. This result agrees with physiological development of warm-season grasses. WSG grows better in warmer circumstances, and the findings in my research demonstrate that TallWSG species relies on temperature during its growth season. It is also noticeable in Figure 21 that precipitation in the past 10 years increases, which could be related to increase ShortWSG distribution. It indicates that ShortWSG, another type of warm-season grasses in the west of Great Plains, is more sensitive to precipitation than temperature.

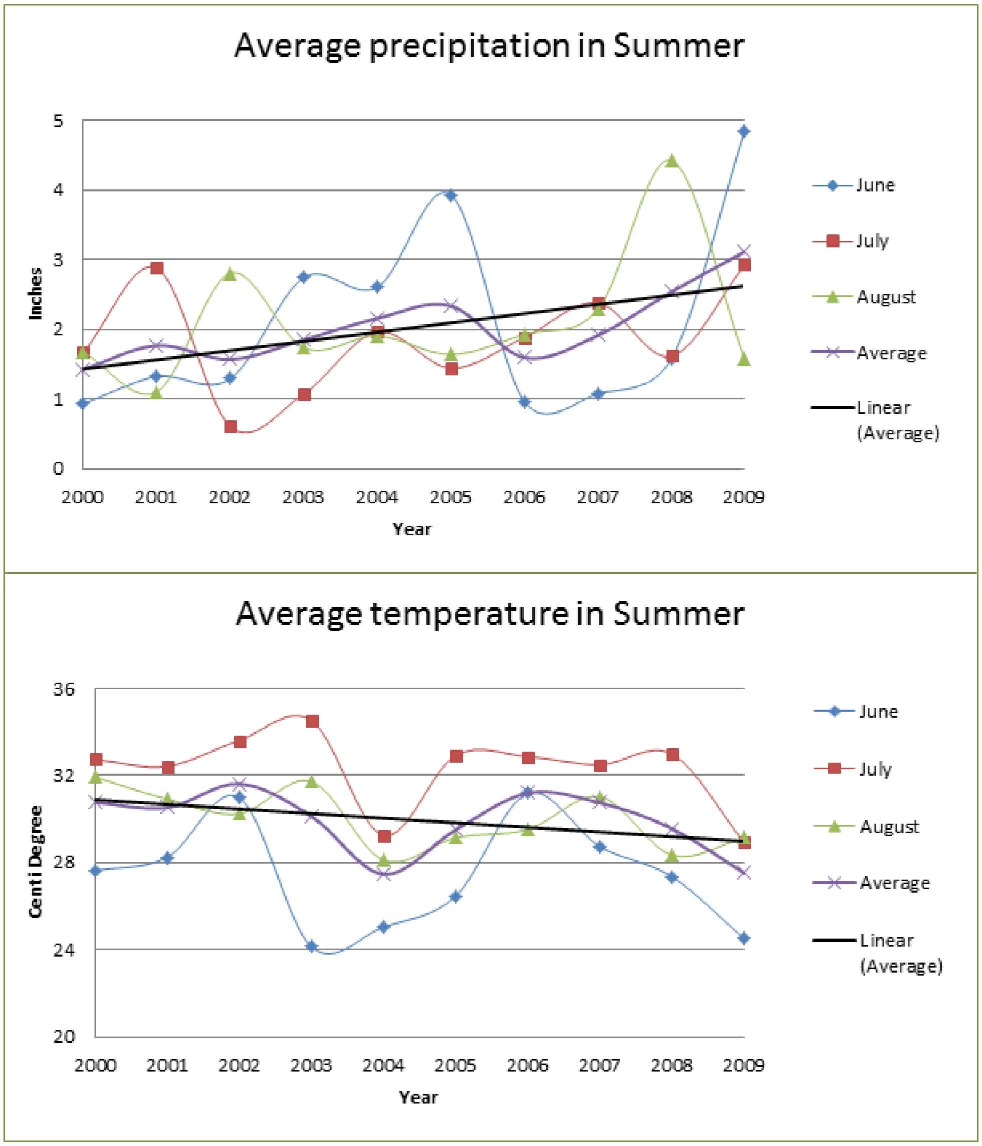

Figure 21. Climate conditions in 2000-2009 in CPER.

4.3.3 Central Prairie - The Sand Hills Grassland

Central prairie is typically mixed grassland. The Sandhills Grassland is located in Nebraska. In its column chart of grass covers (Figure 22), both CSG and TallWSG dominates the grassland, but CSG turns to decrease while TallWSG has an increasing trend in the past 10 years. Not much CSG and short warm season grass growing in this mixed grassland that both TallWSG and ShortWSG species are observed in this area.

The variations of TallWSG and CSG types from 2001 to 2009 are examined with climate data. Upon correlation analysis, I found that the average temperature in spring and total rainfall in summer are significantly correlated to grass covers. The R values of these two factors are higher than 0.921.

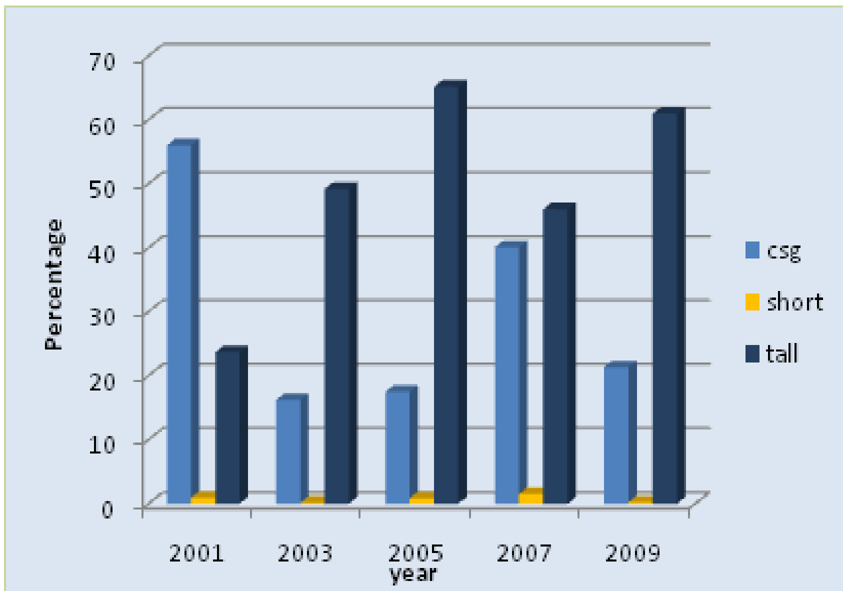

Figure 22. Column chart of grass covers in the Sandhill grassland in the 5 years

In the scatterplots of these two climate factors (Figure 23), spring temperature in this site slightly decreases in the past years. As CSG grasses green up early and reach their peak dates in early May, the low temperature in spring greatly influences its growth. It also shows in the figure that rainfall rises from 2000 to 2009 in this grassland. Meanwhile, although spring temperature decreases, the average temperature in May remains stable. As TallWSG species green up later and reaches peak growth in June-July, low average temperature in spring may not affect TallWSG growth as much than CSG. Abundant rainfall also favors the growth of TallWSG. All these climate influences explain the increased percent cover of TallWSG species in the Sandhills Grassland.

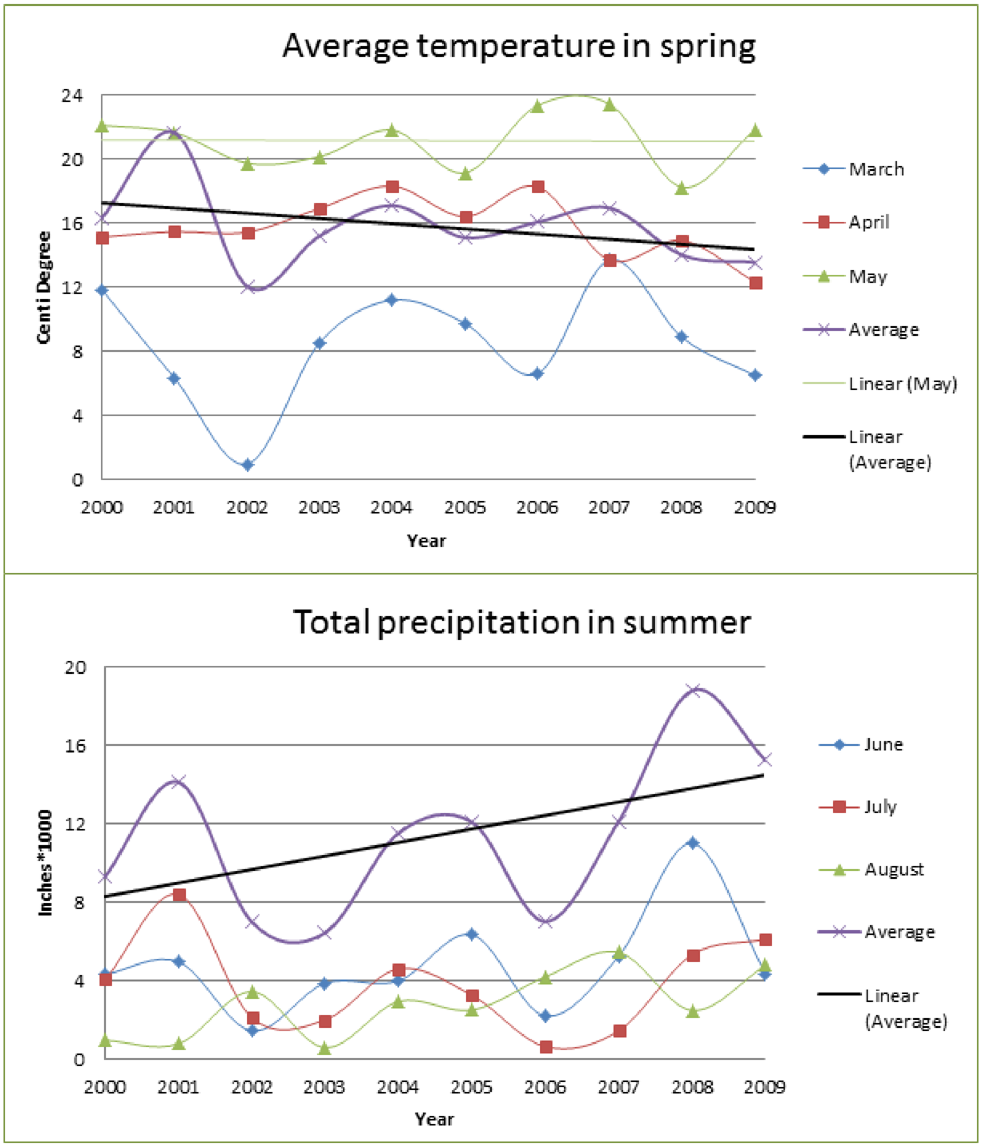

Figure 23. Climate conditions in 2000-2009 in the Sandhills Grassland.

4.3.4 Flint Hills Prairie in Kansas

Basing on ground survey data records from the Konza Prairie Biological Station (KPBS), the Flint Hills grassland in Kansas is extremely dominated with tallWSG species. In my results, this dominance is obvious in the column chart in Figure 24. The WSG percent cover remains dominant although a slight decreasing trend is observed in the 5 years.

CSG grass covers remain low expect 2007, in which CSG dramatically increases and possesses more than 15% of the whole site.

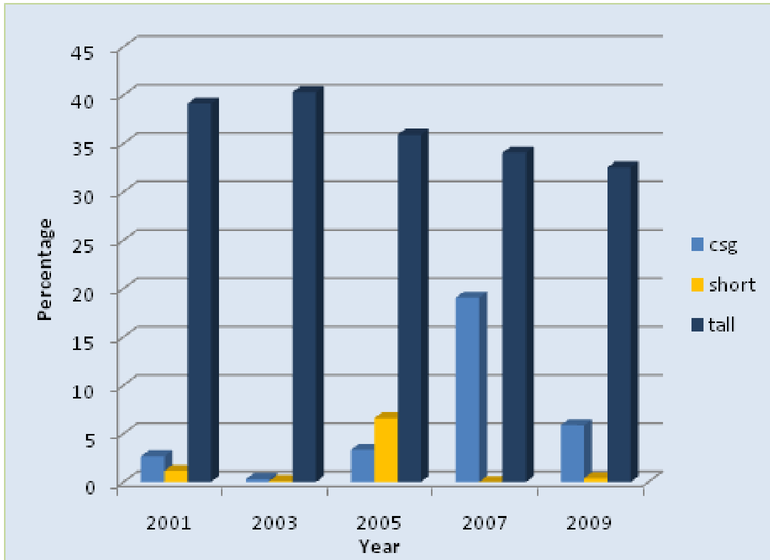

Figure 24. Column chart of grass covers in the Flint Hills Prairie in the 5 years.

Upon correlation analysis, I found that the average temperature in spring and total rainfall in spring are most sensitive to grass covers in this site. The R values of these two factors are higher than 0.893.

We have discussed earlier that CSG is more depended on water availability, while WSG favors higher temperature. In Figure 25, precipitation reached a highest value in 2007, but summer temperature in this year was lower than the average. These climate conditions result in the increased CSG percentage cover in 2007 in this site. On the other hand, the descending trend of the average summer temperature (as shown in the bottom scatterplot in Figure 27) is responsible for the slight decrease of TallWSG cover in 2003-2009.

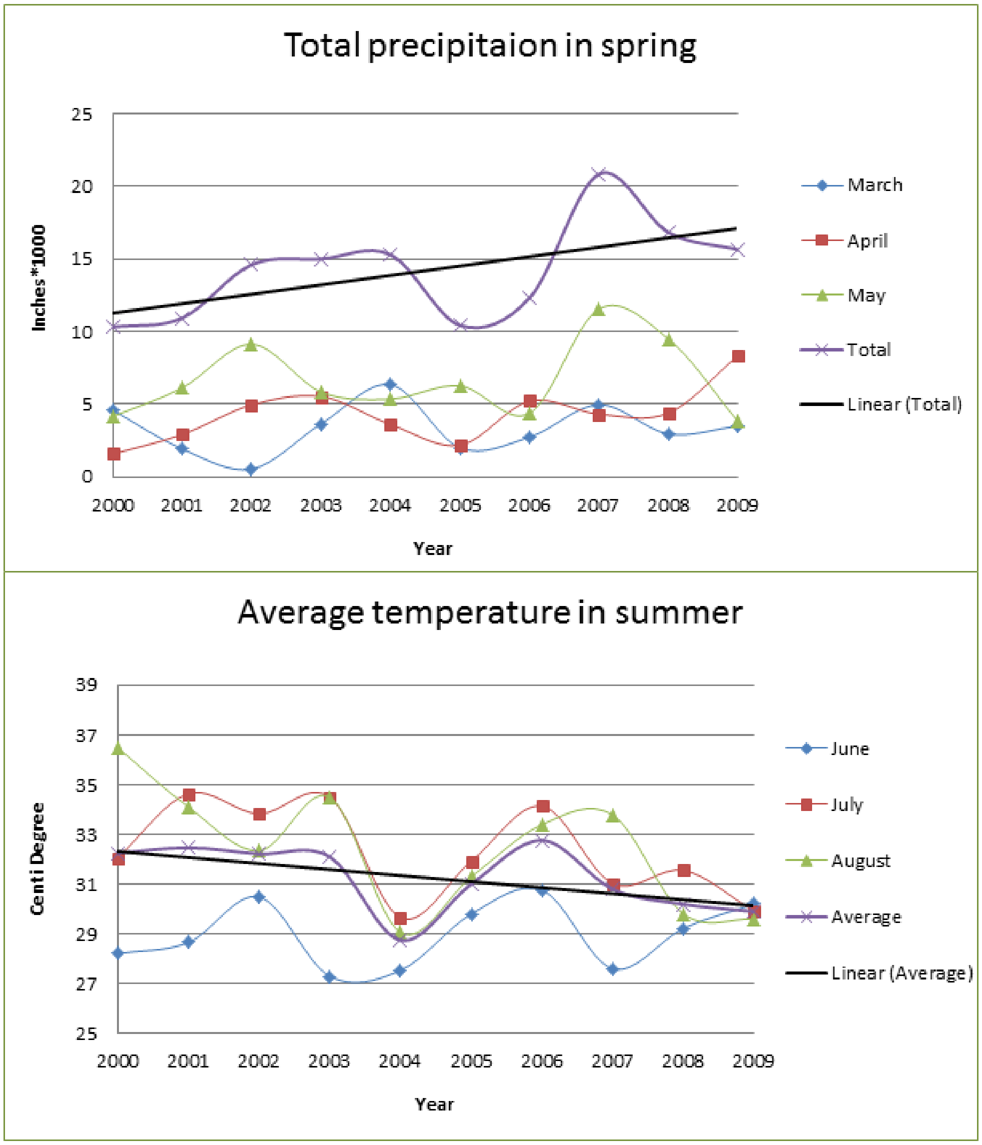

Figure 25. Climate conditions in 2000-2009 in the Flint Hills.

4.4 Discussion and Conclusion

In this chapter, I divide one year into three parts (spring, summer, and fall). Winter data is not applied in climate analysis. I downloaded the PRISM climate data and selected suitable climatic factors in every season, such as average temperature, and total precipitation. Then I chose the four focus areas to calculate percent covers of cool season grass, TallWSG, and ShortWSG in 10 years. With these data I have found several abnormal phenomena existing in these regions. In Montana Prairie, a specific kind of short-term grass appeared and had a large amount in 2005 to 2007, and the percentage of cool season grass suddenly increased in 2007 at flint hill grassland.

The abnormalities of these observations are examined with climate data in each focus area. In order to explain these phenomena and also predict the future change, I linked all percentage data with climatic factors to find their relationship. Water availability is the driving factor for grass variation in Montana Prairie and Flint hill grassland. The Montana's rich spring rainfall during 2005 to 2007 leaded cool season grass growth quickly in the first half year, but the fall arid climate made the grass wilt soon. Therefore, the final classification map appears this short-term type grasses. In a similar way, abundant spring precipitation in Flint hill grassland also caused cool season grass increasing its occupancy. Similarly, ShortWSG appears flourishing with increased precipitation, which explains the continuous increase of ShortWSG distribution percentage in Central Plains Experimental Range in Colorado (CPER). TallWSG fits in warm environment and relies on temperature change more than soil moisture. This ecological feature leads to slightly downtrend of TallWSG occupancy in Flint hill grassland.

My findings in the central prairie agree with the predictions from predictions of ecological models. The Winslow's SAW model, put the result into bio-geospatial model to estimate the climate change, Seasonal Availability of Water Model is one kind of bio-geospatial models. The SAW model predicts that cool season grass is now taking placed by TallWSG (J.C. Winslow et al. 2003). It also shows in the 5-year trend of grassland distributions in my research.

Finally, our findings indicate that precipitation in spring and summer play the most significant role in examining the responses of different grass functional types to climate change. With limited data source, other climate factors such as maximum temperature are not significant in most of the focus sites. More advanced climate data and analysis methods should be explored to enhance our understanding of the linkage between grassland growth and climate variations in the Great Plains.

CHAPTER V: CONCLUSIONS AND FUTURE WORKS

The major goal of this study is to examine the large-scale vegetation distributions in the Great Plains and their sensitivities with climate change. This research was finished via the following analysis: 1) collecting 10- year MODIS TERRA Satellite imagery and Cropland Data Layers. 2) Image processing on all metadata. 3) Spatial and time-series analysis to extract 16 phenological parameters. 4) Established regular regression classification method to generate major annual energy crops and grasses distribution. 5) Correlation analysis to link grassland covers to climate variations in the four focus sites.

Remote sensing data and techniques provide the fundamental basis of this research. The MOD09A1 surface reflectance product has 500-meter pixel resolution and 8-day scan frequency. For the purpose of delineating warm-season and cool-season grasses in the vast area of Great Plains, this product provides proper spatial resolution and temporal composite interval and therefore, it loses less data during cloud removal process and acquires important vegetation growth information along a growing season in each year.

Time-series analysis is highly effective in extracting vegetation's phenological features. Aside from the TIMESAT program that is publically available, time series analysis could be processed with high flexibility using thruster-interactive oriented objective language, such as IDL (Interfaced Description Language) in ENVI. Cool season grass is demonstrated containing up to twice the growth period than warm season grass. Winter wheat could be identified with its unique early peak date when comparing with other vegetation types. It is also revealed via time series analysis that annual crops in general have higher photosynthesis activities in radiation absorption and reflectance, which result in much higher NDVI peak values than grasses in the Great Plains.

The Regression Classification Analysis performs better than logistical regression as multiple classes are examined and more quantitative analysis is conducted. With limited sample points, I found that my phenological characteristics, which serve as independent variables in regression classification, have low cross correlation with vegetation types. It indicates that the requirement of training samples of the regression classification approach is less rigid than machine learning classifiers. This makes it superior to the commonly applied supervised and unsupervised classification, and decision tree methods, which require intensive ground observations as training data. Therefore, the regression classification method could serve as an optimal approach in large-area, multi-class applications. Moreover, the ability of coupling geographic information (latitude and longitude) in the regression classification process takes the temporal variations of phenology into consideration and thus improves the robustness its application in large geographic contents.

The results of my research describe the general and creditable vegetation distributions and variations in the past 10 years. From the class maps, grass percent covers in the selected focus sites agree with ground survey data. Grass changes gradually from the east to the west of the Great Plains. Grasslands in the east (e.g. the Flint Hills) are primarily covered by TallWSG while those in the west (e.g. CPER) are by ShortWSG types. Grass covers in the north of the Great Plains are mostly cool-season grasses. This information could improve the capability of grassland delineation in the CDL crop mapping products.

In a preliminary analysis of climate data, my research also examines the abnormalities and change tendencies of grasslands in the Great Plains. It was found that water availability is the key factor leading to cool season grass increase and decline under other climate factor influences. It also affects ShortWSG species in the west prairie that appear flourishing in rich precipitation surroundings. Temperature change is more related to TallWSG variation because this grass is more suitable for warm environment. In addition, my prediction agrees with ecological models in past studies (Winslow et al. 2003), in which cool season grass is expected to constantly decreases if global temperature maintains rising. Warm-season and cool-season grasses are of the major concern in my thesis research. As these two grass functional types have distinct productivity, the variation of their distributions in central prairie is important for pasture management. Their different photosynthetic pathways results in different roles in carbon sequestration, which is also of great importance in climate change studies. In the future, the results of my research could be applied in these different research activities in regional grassland management and climate change studies. The approaches explored in my result could also be integrated into ecological and climate models to answer large-scale geographic questions in global environment changes.

REFERENCE

[1] Akaike, Hirotugu (1974). "A new look at the statistical model identification". IEEE Transactions on Automatic Control 19 (6): 716–723.

[2] Betts, Richard; Hughes, John K; Valdes, Paul J. "Dynamics of a global-scale vegetation model". Ecological Modelling. Vol: 198 Issue: 3-4 ISSN: 0304-3800 Date: 2006 Pages: 452 – 462.

[3] Brian D. Wardlow, Stephen L. Egbert. "Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains". Remote Sensing of Environment, ISSN 0034-4257, 2007, Volume 112, Issue 3, pp. 1096 – 1116.

[4] Chaichoke Vaiphasa. "Consideration of smoothing techniques for hyperspectral remote sensing". ISPRS Journal of Photogrammetry and Remote Sensing, ISSN 0924-2716, 2005, Volume 60, Issue 2, pp. 91 – 99.

[5] Charles F. Rodell. "A Grasshopper Model for a Grassland Ecosystem". Ecology Vol: 58 Issue: 2 ISSN: 0012-9658 Date: 03/1977 Pages: 227 – 245.

[6] Cuizhen Wang, Brent E. Jamison, Anthony A. Spicci. "Trajectory-based warm season grassland mapping in Missouri prairies with multi-temporal ASTER imagery". Remote Sensing of Environment, ISSN 0034-4257, 03/2010, Volume 114, Issue 3, pp. 531 – 539.

[7] Cuizhen Wang; Felix B Fritschi. "Phenology-based Assessment of Perennial Energy Crops in North American TallWSG Prairie". Association of American Geographers.

Annals of the Association of American Geographers, ISSN 0004-5608, 12/2011, Volume 101, Issue 4, page: 742.

[8] D. L. Martell, S. Otukol, B. J. Stocks. “A logistic model for predicting daily people-caused forest fire occurrence in Ontario”. Canadian Journal of Forest Research, 1987, 17:(5) 394-401, 10.1139/x87-068.

[9] D. Schwartz. “Advancing to full bloom: planning phenological research for the 21st century Mark”. International Journal of Biometeorology, ISSN 0020-7128, 02/1999, Volume 42, Issue 3, pp. 113 – 118.

[10] Dan Simon, 2003. “Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches”.

[11] Dark S.J. “The Biogeography of Invasive Alien Plants in California: An Application of GIS and Spatial Regression Analysis”. Diversity and Distributions Vol: 10 Issue: 1 ISSN: 1366-9516 Date: 01/2004 Pages: 1 – 9.

[12] DeFries, R. S., J. R. G. “Townsend, and M. C. Hansen, 1999. Continuous fields of vegetation characteristics at the global scale at 1-km resolution”. J. Geophys. Res. 104(D14):16,911-16,923.

[13] Epstein, H. E., W. K. Lauenroth, I. C. Burke, and D. P. Coffin, 1997. “Productivity patterns of C3 and C4 functional types in the US Great Plains”. Ecology, 78:722-731.

[14] Goodin, D. G., and G. M. Henebry, 1997. “A technique for monitoring ecological disturbance in tallgrass prairie using seasonal NDVI trajectories and a discriminant function mixture model”. Remote Sensing of Environment, 61:270-278.

[15] De-Xing Chen, H. W. Hunta, and J. A. Morgan. “Responses of a C3 and C4 perennial grass to CO2 enrichment and climate change: Comparison between model predictions and experimental data”. Ecological Modelling. Volume 87, Issues 1-3, June 1996, Pages 11-27.

[16] ENVI + IDL, the Environment for Visualizing Images (ENVI): ENVI Tutorials, Version 4.7, ITT Visual Information Solution, 2009.

[17] Haberlandt, U; Gerten, D; Lucht, W; Sitch, S. “Terrestrial vegetation and water balance—hydrological evaluation of a dynamic global vegetation model”. Journal of Hydrology Vol: 286 Issue: 1-4 ISSN: 0022-1694 Date: 01/2004 Pages: 249 – 270.

[18] J. L. Rodgers and W. A. Nicewander. “Thirteen ways to look at the correlation coefficient”. The American Statistician, 42(1):59–66, February 1988.

[19] Jin Chen, Per. Jonssonc. “A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter”. Remote Sensing of Environment, ISSN 0034-4257, 06/2004, Volume 91, Issue 3-4, pp. 332 – 344.

[20] Kalman, R.E,1960. "A new approach to linear filtering and prediction problems". Journal of Basic Engineering 82 (1): 35–45. Retrieved 2008-05-03.

[21] Karl Pearson “On the Theory of Contingency and Its Relation to Association and Normal Correlation". part of the Drapers' Company Research Memoirs Biometric Series I published in 1904.

[22] Kinyanjui, Mwangi J. “NDVI- based vegetation monitoring in Mau forest complex, Kenya”. African Journal of Ecology, ISSN 0141-6707, 06/2011, Volume 49, Issue 2, pp. 165 – 174.

[23] Lee, Jae-Seok. "Combined effect of elevated CO2 and temperature on the growth and phenology of two annual C3 and C4 weedy species". Agriculture, Ecosystems & Environment, ISSN 0167-8809, 03/2011, Volume 140, Issue 3-4, pp. 484 – 491.

[24] Levy, Peter E; Williams, Mathew; Cox, Peter. "Multiple mechanisms of Amazonian forest biomass losses in three dynamic global vegetation models under climate change". New Phytologist Vol: 187 Issue: 3 ISSN: 0028-646X Date: 08/2010 Pages: 647 – 665.

[25] Lindley, D.V,1987. "Regression and correlation analysis". New Palgrave: A Dictionary of Economics, v. 4, pp. 120–23.

[26] Martha K. Raynolds, Josefino C. Comiso. "Relationship between satellite-derived land surface temperatures, arctic vegetation types, and NDVI". Remote Sensing of Environment, ISSN 0034-4257, 2007, Volume 112, Issue 4, pp. 1884 – 1894.

[27] McCullagh, P. & Nelder, J. A,1989. "Generalized Linear Models, 2nd ed. London: Chapman & Hall. Recommended by the SPSS multinomial logistic tutorial".

[28] McGwire, Kenneth C and Fairbanks, Dean H. K. "Patterns of floristic richness in vegetation communities of California: regional scale analysis with multi- temporal NDVI". Global Ecology and Biogeography Vol: 13 Issue: 3 ISSN: 1466-822X Date: 05/2004 Pages: 221 – 235.

[29] Menard, Scott W. "Applied logistic regression analysis". Vol: no. 07-106. Call Num: QA 278.2 .M461 1995 ISBN: 9780803957572 Date: 1995.

[30] Nippert, Jesse B; Fay, Philip A; Knapp, Alan K. "Photosynthetic traits in C3 and C4 grassland species in mesocosm and field environments". Environmental and Experimental Botany, ISSN 0098-8472, 2007, Volume 60, Issue 3, pp. 412 – 420.

[31] Niu, Shuli; Yuan, Zhiyou; Zhang, Yanfang, “Photosynthetic responses of C3 and C4 species to seasonal water variability and competition”. Journal of experimental botany Vol: 56 Issue: 421 ISSN: 0022-0957 Date: 11/2005 Start Page: 2867.

[32] Niu, Shuli; Zhang, Yanfang; Yuan, Zhiyou. “Effects of interspecific competition and nitrogen seasonality on the photosynthetic characteristics of C3 and C4 grasses”. Environmental and Experimental Botany, ISSN 0098-8472, 2005, Volume 57, Issue 3, pp. 270 – 277.

[33] Niyogi, Dev; Alapaty, Kiran; Chen, Fei; Raman, Sethu. “Development and Evaluation of a Coupled Photosynthesis-Based Gas Exchange Evapotranspiration Model (GEM) for Mesoscale Weather Forecasting Applications”. Journal of Applied Meteorology and Climatology Vol: 48 Issue: 2 ISSN: 1558-8424 Date: 02/2009 Pages: 349 – 368.

[34] Nouvellon, Y. “Coupling a grassland ecosystem model with Landsat imagery for a 10-year simulation of carbon and water budgets”. Remote Sensing of Environment Vol: 78 Issue: 1-2 ISSN: 0034-4257 Date: 10/2001 Pages: 131 – 149.

[35] P. C.Twiss, 1992. “Predicted world distribution of C3 and C4 grass phytoliths”.

[36] Peng, Changhui; Garneau, Michelle; Quillet, Anne, “Toward dynamic global vegetation models for simulating vegetation–climate interactions and feedbacks: recent developments, limitations, and future challenges”. Environmental Reviews Vol: 18 Issue: NA ISSN: 1181-8700 Date: 12/2010 Pages: 333 – 353.

[37] Per Jönsson and Lars Eklundh. “Seasonality extraction from time-series of satellite sensor data”. IEEE Transactions on Geoscience and Remote Sensing, ISSN 0196-2892, 2002, Volume 40, Issue 8, pp. 1824 – 1832.

[38] Per Jönsson and Lars Eklundh, 2006. TIMESATGUI - Users guide for version 2.3.

[39] Peterson, D. L., K. P. Price, and E. A. Martinko, 2002. “Discriminating between cool season and warm season grassland cover types in northeaster Kansas”. International Journal of Remote Sensing, 23:5015-5030.

[40] PRISM (Parameter-elevation Regressions on Independent Slopes Model) climate data. PRISM Climate Group, Oregon State University.

[41] Rice, J. C,1994. "Logistic regression: An introduction". In B. Thompson, ed., Advances in social science methodology, Vol. 3: 191-245. Greenwich, CT: JAI Press. Popular introduction.

[42] Robert A. Spicer. “Palaeoecology, past climate systems, and C3/C4 photosynthesis”. Chemosphere. Volume 27, Issue 6, September 1993, Pages 947-978.

[43] Running, S. W., T. R. Loveland, L. L. Pierce, R. Nemani, and E. R. Hunt, Jr., 1995. “A remote sensing based vegetation classification logic for global land cover analysis”. Remote Sens. Environ. 51 (1):39-48.

[44] Saro, Lee. “Application of likelihood ratio and logistic regression models to landslide susceptibility mapping using GIS”. Environmental Management Vol: 34 Issue: 2 ISSN: 0364-152X Date: 08/2004 Pages: 223 – 232.

[45] Savitzky, A.; Golay, M.J.E,1964. "Smoothing and Differentiation of Data by Simplified Least Squares Procedures". Analytical Chemistry 36 (8): 1627–1639.

[46] Susan C. Mulholland, George Rapp Jr. 1992. Phytolith Systematics: Emerging Issues (Advances in Archaeological and Museum Science) (ISBN: 0306442086 / 0-306-44208-6).

[47] TIMESAT - a Program for Analyzing Time-Series of Satellite Sensor Data. Version 2.3. Published by Per Jönsson and Lars Eklundh, 2002.

[48] USDA National Agricultural Statistics Service – Cropland Data Layer. http://www.nass.usda.gov/research/Cropland/SARS1a.htm.

[49] USGS Land Processes Distributed Archive Center. https://lpdaac.usgs.gov/get_data

[50] WEISS, J. "Long-term vegetation monitoring with NDVI in a diverse semi-arid setting, central New Mexico, USA". Journal of Arid Environments, ISSN 0140-1963, 07/2004, Volume 58, Issue 2, pp. 249 – 272.

[51] Winslow, Jerome C; Hunt Jr, E.Raymond; Piper, Stephen C. "The influence of seasonal water availability on global C3 versus C4 grassland biomass and its implications for climate change research". Ecological Modelling, ISSN 0304-3800, 05/2003, Volume 163, Issue 1-2, pp. 153 – 173.

[52] Wolf, J. "Model simulation of effects of changes in climate and atmospheric CO2 and O3 on tuber yield potential of potato (cv. Bintje) in the European Union". Agriculture, Ecosystems & Environment, ISSN 0167-8809, 02/2003, Volume 94, Issue 2, pp. 141 – 157.

[53] Y. Wang; M. Notaro; Z. Liu; R. Gallimore; S. Levis; J. E. Kutzbach. "Detecting vegetation-precipitation feedbacks in mid-Holocene North Africa from two climate models". Climate of the Past, ISSN 1814-9324, 03/2008, Volume 4, Issue 1, pp. 59 – 67.