ASSESSING CHANGES IN SPATIAL AND TEMPORAL PATTERNS OF CROPPING SEQUENCES IN NORTHEAST MONTANA

by

John Allen Long

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Ecology and Environmental Sciences

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2014

©COPYRIGHT

by

John Allen Long

2014

All Rights Reserved

ACKNOWLEDGEMENTS

I would like to briefly acknowledge a few of the many people and organizations that helped me along the way.

To Rick Lawrence for his guidance, patience, and help in making the conversion from mathematician to scientist.

To Perry Miller for tolerating my agricultural naivety and helping me to see the connections.

To Mark Greenwood and Lucy Marshall for their support and helping me to stay true to the statistical side of my work.

To my wife, Danette, for her support and her ‘positivity’.

And to those who provided financial support, the US Department of Energy and the National Energy Technology Laboratory under Award Number DE-FC26-05NT42587.

TABLE OF CONTENTS

1. INTRODUCTION 1

2. OBJECT-ORIENTED CROP CLASSIFICATION USING MULTITEMPORAL ETM+ SLC-OFF IMAGERY AND RANDOM FOREST 6

Contribution of Authors and Co-Authors 6
Manuscript Information Page 7
Abstract 8
Introduction 8
Methods 15
Study Area 15
Data Acquisition 16
Image and Vector Data Processing 17
Image Segmentation and Data Extraction 19
Classification 20
Results 22
Object-Orientated Approach to Classification 22
Pixel-Based Approach to Classification 24
Discussion 27
Acknowledgements 33

3. CHANGES IN FIELD-LEVEL CROPPING SEQUENCES: INDICATORS OF SHIFTING AGRICULTURAL PRACTICES 41

Contribution of Authors and Co-Authors 41
Manuscript Information Page 42
Abstract 43
Introduction 44
Materials and Methods 47
Study Area 47
Data 47
Satellite Imagery 47
Ground Reference Data 48
The Cropland Data Layer 49
Ancillary Vector Data 50
Data Preparation 51
Creating the ‘Fields’ Layer 51
Creating the CDL Mosaics 51
Image Pre-Processing 52
Determining Field-Level Crop Class (2007–2012) 53

TABLE OF CONTENTS - CONTINUED

Determining Field-Level Crop Class (2001–2006) 54
String Construction and Pattern Determination 55
Results 57
Individual Classes 57
Two-Year Sequences 58
Three-Year Sequences 60
Discussion 61
Total Production and Monoculture Sequences 61
Cereal-Fallow Sequences 63
Cereal-Fallow Sequences Managed by Strip-Cropping 64
Pulse Sequences 65
Conclusion 66
Acknowledgements 67

4. ADOPTION OF CROPPING SEQUENCES IN NORTHEAST MONTANA: A SPATIO-TEMPORAL ANALYSIS 74

Contribution of Authors and Co-Authors 74
Manuscript Information Page 75
Abstract 76
Introduction 77
Material and Methods 83
Study Area 83
Data 84
Analyses 86
Results 91
Point Pattern Analysis 91
Adoption of Cropping Sequences as a Contagion Process 92
Adoption of Cropping Sequences as a Function of the Environment 93
Discussion 95
Key Indicators of Changing Agricultural Practices 95
The adoption of Cereal-Pulse Cropping Sequences 96
The adoption of Block-Managed Cereal-Pulse Sequences in Lieu of Strip-Cropping 100
Conclusion 102
Acknowledgements 103

5. CONCLUSION 110

REFERENCES CITED 116

LIST OF TABLES

Table Page

2.1. Error Matrices for Object-Oriented Classifications 35

2.2. Error Matrices for Pixel-Based Classifications 36

2.3. Tests for Statistically Significant Differences between Kappa Statistics 36

2.4. The Top Variables for Multitemporal Classification as Ranked by Importance 37

2.5. The Top Variables for Mid-Season Classification as Ranked by Importance 37

2.6. The Top Variables for Late-Season Classification as Ranked by Importance 38

3.1. Dates and Sensors for the Landsat Imagery Used in the Study 69

3.2. Definition of Cropping Sequences 69

4.1. Maximum Clustering Range by Cropping Sequence And Time Period 104

4.2. Coefficients for the Top-Ranked Logistic Regression Models 105

4.3. Conversions from Continuous Strip-Cropping to Cereal-Based Sequences as a Percentage of All Conversions 106

LIST OF FIGURES

Figure Page

2.1. Map of the Study Area and the Study Area Subset 38

2.2. The Top 8 Variables in the Object-Oriented Classification....39

2.3. The Top 8 Variables in the Pixel-Based Classification 39

2.4. Detailed Map of the Study Area Subset....40

3.1. Map of the Study Area....70

3.2. Crop Area (2001–2012) 71

3.3. Abundance of 2-Year Monoculture Sequences 71

3.4. Abundance of 2-Year Rotational Sequences 72

3.5. Abundance of 3-Year Monoculture Sequences 72

3.6. Abundance of 3-Year Rotational Sequences 73

4.1. Map of the Study Area....106

4.2. Kernel Density Maps for CP/PC Sequences....107

4.3. Kernel Density Maps for SC Sequences....108

4.4. Mean Minimum Distances between Fields....109

ABSTRACT

Initiatives to mitigate the effects of climate change have focused largely on the reduction of greenhouse gas production and on carbon capture and storage technologies. Changes in agricultural management practices have shown the ability to sequester carbon by increasing soil organic carbon and include reduced tillage intensity, decreased fallow, and changing from monoculture to rotational cropping. All have become more common in portions of the Northern Great Plains; but, despite the growth of these practices, it is unknown to what extent farmers have adopted particular cropping sequences or how they have spread temporally or spatially. My purpose here was to investigate the patterns of changing agricultural practices in northeast Montana during 2001–2012 by focusing on the increased adoption of cereal-pulse sequences and the adoption of block-managed cereal-based sequences in lieu of continuous strip-cropping. A method to identify crops via geospatial data and Landsat imagery was developed, and annual crop maps were created. Crop classifications were extracted from the maps for each field to create a 12-character string for the temporal sequence of crops, and specific 2- and 3-yr sequences were identified with a string-matching algorithm. Finally, I examined the observed spatial patterns of sequence adoption to determine if observed spatial patterns were random or were they consistent with the spread and adoption due to social interaction as described in innovation diffusion theory, adoption based on environmental factors, or neither. The major findings were: (1) cereal-fallow rotations, whether managed in blocks or by strip-cropping, no longer dominate the region; (2) there has been a substantial increase in the adoption of cereal-pulse sequences; (3) producers did not adhere strongly to specific sequences; (4) using 3-yr sequences added no additional information than 2-yr sequences; (5) the adoption of these practices was not randomly located but clustered; and (6) the adoption of these practice are not well-explained by innovation diffusion theory, although social interactions might have played a role in the early stages; the patterns are more consistent with suitability of the physical environment since available water was strongly associated with whether or not a field was managed with either practice.

CHAPTER 1

INTRODUCTION

Carbon dioxide (CO_2) is a natural component of the Earth's atmosphere that selectively absorbs longwave radiation. This absorption results in the retention of thermal energy within the atmosphere, which will remain thermally stable as long as CO_2 concentration remains relatively constant. The data are quite clear that the mean global CO_2 concentration is not constant, but increasing, and has been since the beginning of the industrial age circa 1750 (IPCC, 2007). The most recent data indicates a mean global CO_2 concentration of approximately 396 ppm (NOAA, 2014), a substantial increase over the mean pre-industrial level of 280 ppm (Friedli et al., 1986).

Numerous studies conclude that the increase in atmospheric CO_2 concentration has been accompanied by: (1) an increase in the mean global surface temperature; (2) an increase in the mean global ocean temperature; (3) substantial ablation of snowfields and glaciers; and (4) a slight rise in the mean global sea level (IPCC, 2007). These effects are key indicators of climate change and are largely the result of two anthropogenic activities: increased use of fossil fuels and changes in land use (IPCC, 2007). Climate change is expected to cause a range of deleterious effects such as increases in the frequency and severity of heat waves, changes in ecosystem structure and function, and a redistribution of water resources (IPCC, 2007). Researchers have suggested that mean global concentration of CO_2 has likely already exceeded the maximum level that could be supported without consequences (Hansen et al., 2008). Political response at the national

level to mitigate the effects climate change has been slow at best and antagonistic at worst, despite the concerns of the scientific community.

Regional and state-level initiatives to mitigate the effects of climate change have focused, in part, on greenhouse gas (GHG) reduction and carbon capture and storage (CCS) technologies. One such initiative in Montana is the Big Sky Carbon Sequestration Partnership (BSCSP), a collaboration of regional universities, national laboratories, and private corporations. The BSCSP is supported by the Department of Energy (DOE) as one of seven Regional Carbon Sequestration Partnerships (RCSPs) that are to help develop the technology, infrastructure, and regulations to implement large-scale sequestration efforts (BSCSP, 2012). Here, the goal is to characterize the sequestration potential of regional geologic formations, such as the Kevin Dome in north-central Montana, as well as terrestrial sequestration options.

Terrestrial sequestration is the removal of atmospheric CO_2 by vegetation through photosynthesis and the subsequent storage of carbon as biomass or soil organic carbon (SOC). Improved forestry and agricultural practices can offer relatively rapid reductions in atmospheric CO_2 concentrations and are widely recommended (e.g., Hansen et al., 2008; IPCC, 2007). Changes in agricultural practices have shown the ability to sequester carbon by increasing the accumulation of soil organic carbon (e.g., West & Post, 2002). Potential agricultural practices for sequestering carbon include reduced tillage intensity, decreasing fallow periods, the use of winter cover crops, and changing from monoculture to rotation cropping (West & Post, 2002; IPCC, 2007).

Crop rotation is a planned sequence of different crops, or fallow, in the same field. As an agricultural practice, it has been around for thousands of years. A well designed rotation brings many benefits; however, it is a complicated practice and many farmers opt for the simplicity of monocultures or a crop-fallow rotation, despite the disadvantages. Cereal-pulse sequences have become common in portions of the Northern Great Plains, particularly southern Saskatchewan and northeastern Montana (Tanaka et al., 2010), which has seen a five-fold increase in the area devoted to the practice between 2001 and 2012. Cereal-pulse sequences are rotations of a cereal crop, typically spring or winter wheat (*Triticum aestivum* L.), durum (*Triticum turgidum* L.), or barley (*Hordeum vulgare* L.), with a pulse crop such as dry peas (*Pisum sativum* L.), lentils (*Lens culinaris* Medik.), or chickpea (*Cicer arietinum* L.) (Lemke et al., 2007). Concurrently, the area of fallowed cropland has decreased by one-third.

Incorporating pulses into rotational sequences with the region's dominant crop – small grain cereals – improves the robustness and resilience of local agricultural systems (Zentner et al., 2001, 2002; Burgess et al., 2012). These sequences offer many field-level benefits including: (1) biological nitrogen (N) fixation, making additional N available to the succeeding crop and reducing fertilizer requirements; (2) substantially lower water requirements than cereals or alfalfa; (3) weed control; (4) reduced insect and disease problems; (5) improved soil tilth and stability; (6) reduced soil erosion; and (7) higher protein levels and improved yields for the following cereal crop (Cochran et al., 2006; Lee, 2011, Peel, 1998). Some regional cereal-pulse crop rotations have resulted in smaller

overall CO_2 budgets than non-pulse rotations through reduced energy inputs (e.g., less fuel for farm machinery) as well as increased SOC (Lemke et al., 2007).

The overarching purpose of this study was to investigate temporal and spatial patterns of changing agricultural practices in northeast Montana during 2001–2012 using geospatial techniques. More specifically, the study sought to: (1) develop a remote sensing technique to classify regionally relevant crops at the field-level from imagery with missing data; (2) identify the principal 2- and 3-yr cropping sequences and the changes in their relative proportions; and (3) characterize and explain the spatial patterns of adoption over time. These goals comprised the three main chapters of this dissertation and focused on the increased adoption of cereal-pulse sequences and the adoption of block-managed cereal-based sequences in lieu of continuous strip-cropping. Changes in agricultural practices are important because they indicate how farmers are coping with a changing climate and provide an insight into the general long-term sustainability of regional agriculture. Furthermore, identifying the prevailing cropping sequences can help establish which ones have been successful and, therefore, which sequences might be more likely to succeed in similar regions.

In order to identify cropping sequences, as well as their temporal and spatial patterns of adoption, it was first necessary to have maps of crops at the field-level for each year. This work used geospatial data and Landsat imagery to identify the crops, but the only Landsat satellite available during the 2012 field season was Landsat 7. Unfortunately, Landsat 7's scan-line corrector (SLC) failed in 2003, and imagery since

the failure have data gaps of approximately 22%. Chapter 2 details the development of a method to classify an agricultural landscape despite these gaps.

Once a method to classify crops had been developed, it was necessary to classify the crop in each of the 13,076 fields over a 12-yr period, and then to identify the temporal characteristics of the cropping sequences. Here, multitemporal Landsat imagery and geospatial data were used to develop a 12-character string for each field that represented the temporal sequence of crop classifications. Specific *a priori* 2- and 3-yr crop sequences were identified from the character strings with a string-matching algorithm. Chapter 3 details this work.

The final step was to characterize and explain the spatial patterns of adoption over time. The results from Chapter 3 indicated that the adoption of cereal-pulse sequences and conversion from continuous strip-cropping to block managed cereal-based sequences were the key indicators for changing agricultural practices. The goal in Chapter 4 was to determine if observed spatial patterns were random, and if not, were they consistent with spread and adoption due to social interaction as described in innovation diffusion theory, adoption based on environmental factors, or none of the above.

CHAPTER 2

OBJECT-ORIENTED CROP CLASSIFICATION USING MULTITEMPORAL ETM+ SLC-OFF IMAGERY AND RANDOM FOREST

Contribution of Authors and Co-Authors

Manuscript in Chapter 2

Author: John A. Long

Contributions: Conceived and implemented the study design. Collected and analyzed the data. Wrote manuscript and revisions, and submitted manuscript for publication.

Co-Author: Dr. Rick L. Lawrence

Contributions: Helped conceive the study design. Provided guidance and feedback on manuscript drafts, remote sensing issues, and statistical analysis.

Co-Author: Dr. Perry R. Miller

Contributions: Provided feedback on manuscript drafts and agricultural issues.

Co-Author: Dr. Mark C. Greenwood

Contributions: Provided feedback on manuscript drafts and statistical analysis.

Co-Author: Lucy A. Marshall

Contributions: Provided feedback on manuscript drafts and statistical analysis.

Manuscript Information Page

John A. Long, Rick L. Lawrence, Perry R. Miller, Mark C. Greenwood, Lucy A. Marshall
GIScience and Remote Sensing
Status of Manuscript: (Put an x in one of the options below)
___ Prepared for submission to a peer-reviewed journal
___ Officially submitted to a peer-review journal
___ Accepted by a peer-reviewed journal
_X_ Published in a peer-reviewed journal

Published by Taylor and Francis in 2013, Volume 50, Issue 4, pages 418-436.

Abstract

The utility of Enhanced Thematic Mapper Plus (ETM+) has been diminished since the 2003 scan-line corrector (SLC) failure. Uncorrected images have data gaps of approximately 22% and gap-filling schemes have been developed to improve their usability. We present a method to classify a northeast Montana agricultural landscape using ETM+ SLC-off imagery without gap-filling. We use multitemporal data analysis and employ an object-oriented approach to define objects, agricultural fields, with cadastral data. This approach was assessed by comparison to a pixel-based approach. Results indicate that an ETM+ SLC-off image can be classified with better than 85% overall accuracy without gap-filling.

Introduction

The Landsat series of satellites have provided researchers with a 40+ year archive of multispectral, moderate-resolution observations of the Earth. These observations comprise the world's longest continuously acquired collection and have been used by various educational, governmental, and industrial organizations, both within the United States and the international community (USGS, 2012). Landsat-derived data are essential to researchers in diverse fields such as land cover mapping (e.g., Fuller et al., 1994), ecosystem and vegetation dynamics (e.g., Savage and Lawrence, 2010), biophysical modeling (e.g., Cohen et al., 2003), and classification of urban landscapes (e.g., Lu and Weng, 2005).

The United States Geological Survey (USGS) ceased acquisition of Landsat 5 Thematic Mapper (TM) imagery in November 2011 after detecting a failing amplifier in the satellite's downlink (Clark, 2011). The loss of Landsat 5 left Landsat 7 as the sole operational Landsat satellite until May 2013 when Landsat 8 became operational. Landsat 7 carries the Enhanced Thematic Mapper Plus (ETM+) sensor; however, its scan-line corrector (SLC) failed permanently on 31 May 2003, leaving Landsat 7 with degraded performance (Markham et al., 2004). The SLC compensates for the satellite's forward motion during the cross-track scanning of the ETM+ sensor, thereby keeping the forward and reverse sweeps parallel to each other. The primary consequence of the SLC failure is gaps in coverage that range from none at the center of a scene to 14 pixels at the scene's lateral edges. In total, data gaps comprise approximately 22% of an image (Markham et al., 2004). Images with these data gaps are known as SLC-off images, whereas those acquired prior to the SLC failure are designated SLC-on images (Arvidson et al., 2006). Data gaps are not identical for all spectral bands, and slight shifts in the position of the gaps result in the occasional pixel with valid data for some bands, but not others (Zhang, 2007).

The radiometric performance and geometric characteristics of the ETM+ sensor have not been affected by the SLC failure (Markham et al., 2004). Researchers have continued to rely on SLC-off images, despite the availability of concurrent TM images, because of the superior quality of the ETM+ sensor. Improvements in quality include enhanced radiometric precision because of the addition of a two-gain state (low-gain and high-gain) to optimize sensitivity without detector saturation. Other improvements

include better geodetic accuracy, reduced levels of spatial noise (e.g., banding and striping) due to improved electronics and detector calibration, a more reliable radiometric calibration, the addition of a panchromatic band, and substantial improvements in the spatial resolution of the thermal band (Masek et al., 2001). Research efforts that require Landsat imagery from 2012 must use SLC-off imagery as there is no alternative. We emphasize that complications arising from missing data are not restricted to ETM+ SLC-off imagery and might occur in imagery from any sensor due to a variety of causes such as sensor noise, clouds, or shadows. Consequently, recovering the missing data through various gap-filling schemes has become an increasing important endeavor.

Gap-filling schemes can be divided into three general approaches: (1) direct prediction of the missing data based on interpolation of the surrounding pixels; (2) predictions based on multitemporal Landsat data; and (3) predictions based on data from other sensors (Zeng et al., 2013). Interpolation approaches attempt to infer missing values by using the values from nearby pixels in conjunction with an empirically derived formula or by using a geostatistical interpolation scheme such as kriging (Zhu et al., 2012). These approaches can improve the usability of SLC-off imagery, but under restrictive conditions. Landscapes in which pixel values can potentially change rapidly over a small distance, such as heterogeneous landscapes and landscapes with small or linear features, are difficult to predict when using approaches that rely on surrounding pixel values (Maxwell, 2004; Zhu et al., 2012). Gap-filling is most useful in homogeneous areas with little landscape structural changes (Zeng et al., 2013) or when the missing data represents small spatial gaps (Pringle et al., 2009). It also is unclear if

interpolation approaches based on geostatistical methods meet the intrinsic stationarity assumption (Pringle et al., 2009).

Missing values have also been filled with a multitemporal Landsat approach by using an SLC-on image of the same area to fill the gaps in a SLC-off image (e.g., Storey et al., 2005, Maxwell et al., 2007); however, this method is unsuitable for applications that require spectral data from closely spaced dates. A good example is the classification of specific crops because dates more than two weeks apart might represent different stages in phenological development. An alternative is a multi-sensor approach in which data from a different sensor on the same, or nearly the same, day is used to predict the missing data. The multi-sensor approach to filling ETM+ gaps has been partially successful with a variety of auxiliary sensors. The Advanced Land Imager (ALI) aboard the Earth Observer One (EO-1) satellite (Boloorani et al., 2008), the Moderate Resolution Imaging Spectroradiometer (MODIS) (Roy et al., 2008), the Charge Coupled Device (CCD) camera aboard the China Brazil Earth Resources Satellite (CBERS-02B) (Chen et al., 2010), and the Linear Imaging Self-Scanning System aboard the Indian Remote Sensing Satellite (IRS/1D) (Reza and Ali, 2008) have all been used. Nonetheless, concerns regarding spectral compatibility, spatial resolution, clouds, and financial constraints remain (Zeng et al., 2013).

Gap-filling approaches predict the missing values of individual pixels; however, not all remote sensing tasks require pixel-level data. Classification algorithms, for example, can operate on individual pixels (pixel-based classifiers) or on groups of contiguous pixels (object-oriented classifiers). Classification using an object-oriented (O-

O) approach is most appropriate when the landscape consists of areas that can be delineated into meaningful homogeneous regions, such as agricultural fields (e.g., Forster et al., 2010). The O-O approach classifies at the object-level; the number of pixels in an object is largely irrelevant if the object is reasonably homogeneous. An O-O approach to classification, therefore, does not require estimation of the missing data in SLC-off imagery provided there are a sufficient number of extant pixels in the objects (Turker & Arikan, 2005).

Classification of agricultural regions at the pixel level presents a challenge in that monoculture fields are often classified as multiple classes. The Cropland Data Layer (CDL), a geo-referenced, crop-specific data layer derived from a fusion of moderate resolution satellite imagery (USDA, 2012), contains numerous examples. A number of studies conclude that O-O approaches produce better results (e.g., Pedley and Curran, 1991; Janssen and van Amsterdam, 1996). There are several reasons why O-O methods outperform pixel-based methods. A fundamental reason is that classification accuracy decreases with increasing spectral variability (e.g., Ashish et al., 2009). Variability is high in heterogeneous landscapes and lower in homogeneous landscapes, such as agricultural fields. One compelling reason supporting the O-O approach in the classification of crops is that management decisions are generally made on a per-field basis (Forster et al., 2010). The O-O approach can incorporate ancillary information such as spatial relationships and management practices that, while not unique to the O-O approach, can be advantageous. The use of agricultural field boundaries defines the spatial relationship and identifies a particular object. Good results have resulted from

merging images with a suitable vector layer, such as cadastral data (e.g., Wu et al., 2007; Watts et al., 2009).

The use of multitemporal imagery to classify vegetation is well-documented and has improved crop classification accuracies by 6–9% (e.g., El-Magd and Tanton, 2003; Conrad et al., 2011). The main advantage is that this approach uses images from different dates within a single growing season to better capture the spectral diversity in crops due to differences in phenological stages (Gómez-Casero et al., 2010). The data are fused into a single classification scheme for the growing season. Multitemporal analysis is particularly attractive in crop classification for several reasons: (1) the spectral characteristics of a crop varies with phenological stage; (2) agricultural areas are spatially dynamic in that adjacent fields might be in different phenological stages, thereby increasing heterogeneity; and (3) criteria based on phenological stage have ecological significance and are easier to generalize and more robust than statistical criteria derived from a specific dataset (Simonneaux et al., 2008). The use of multitemporal imagery raises obvious questions regarding quantity and seasonal timing. Research suggests that classification accuracy increases with the number of images and that images acquired at peak growth are more useful than those from low-growth periods (Pax-Lenney and Woodcock, 1997). Others have presented similar findings (e.g., Oetter et al., 2000). Multitemporal classifications have produced accuracies exceeding 84% (e.g., Pax-Lenney and Woodcock, 1997; Wardlow et al., 2007; Simonneaux et al., 2008; Gómez-Casero et al., 2010).

Random forest (RF) is a tree-based ensemble method useful for classification when there are many weak explanatory variables (Breiman, 2001) and is well-established in the remote sensing literature. RF classification is often superior to standard classification approaches, performing well with large numbers of classes (e.g., Rodrigues-Galiano et al., 2012), with difficult to classify data (e.g., Lawrence et al., 2006), and routinely achieving classification accuracies well above 90% (e.g., Chapman et al., 2010). It generates an internal accuracy assessment that has been shown to be equivalent to an independent accuracy assessment (assuming no bias in the data), thereby, eliminating the need for separate validation data (Lawrence et al., 2006). One of the advantages of RF is that it is that it can accommodate high dimensional and redundant data, negating the need to reduce the dimensionality of the data or to transform the data to orthogonal components (Brieman, 2001). This property permits, for example, the use of NDVI as a predictor simultaneously with bands 3 and 4. RF has been used in agricultural studies to classify agricultural practices (e.g., Watts et al., 2009), species-specific crop discrimination (e.g., Adam et al., 2012; Chapman et al., 2010), and multi-seasonal landcover classification (e.g., Rodrigues-Galiano et al., 2012).

We present here a method to classify a northeast Montana agricultural landscape into the region's dominant classes using SLC-off ETM+ imagery from the 2012 growing season – without using gap-filling schemes. Crops in the cereal class (Cereal) are familiar grain crops and include spring wheat (*Triticum aestivum*), durum wheat (*T. durum*), and oat (*Avena sativa*). The pulse crops (Pulse) are leguminous annual crops that include dry pea (*Pisum sativum*) and lentil (*Lens culinaris*). The other class (Other) refers to

agricultural land that is either in fallow or is growing something other than cereal or pulse crops. This method is part of a larger project that seeks to investigate the spatio-temporal patterns of adoption for cereal-pulse crop rotation in northeast Montana during 1994–2012 using geospatial techniques. Two images, one from mid-season and one from late-season, are used to capture phenological variations. We employ an O-O approach and define objects, agricultural fields, with regional cadastral data. Spectral data are extracted from the extant pixels in an object and used in a random forest classification to identify the class of crop in a particular field. The usefulness of the O-O approach was assessed by comparison to a pixel-based approach, each with a random forest classifier.

Methods

Study Area

The study area is in northeast Montana and includes the counties of Daniels, Sheridan, and Roosevelt, and the eastern portion of Valley (Figure 2.1). The region is bounded by Canada to the north, North Dakota to the east, the Missouri River to the south, and public lands such as the Bitter Creek Wilderness Study Area along much of the western border. The area is in the Missouri Plateau (glaciated) region of the Great Plains Physiographic Province and is characterized by low relief. The climate is semi-arid with low humidity. Average total annual precipitation is just over 31 cm, which is primarily rain during April to September (WRCC, 2012). Average daily maximum temperatures in July are 31°C, while January temperatures average -10°C (WRCC, 2012). Agriculture and shortgrass prairie are the dominant land cover types. Agricultural

practices in the study area are primarily dryland cropping systems (Tanaka et al., 2010). Regional agriculture is largely cereal crops, primarily spring wheat, and increasingly substantial acreages of pulse crops (NASS, 2012).

Data Acquisition

The study area required three Landsat scenes: Path 36, Row 26 (36/26); Path 36, Row 27 (36/27); and Path 35, Row 26 (35/26). Due to the positions of these three overlapping scenes relative to the study region, data loss due to SLC-off gaps was estimated at 15–18%. SLC-off ETM+ images from the 2012 growing season were obtained from the United States Geological Survey's (USGS) Earth Resources Observation and Science Center (EROS) and were used to create two mosaics. The mid-season mosaic combined scenes from July 11 and July 18, whereas the late-season mosaic used scenes from August 12 and August 19. The disparity in dates for the mosaics was because the study area required imagery from two adjacent paths; thus, an 8-day difference. We chose these dates because they were the only usable images from the 2012 growing season; images from the remaining dates had excessive cloud coverage. All images from Path 36 were cloud-free, whereas the images from Path 35 had occasional cumulus clouds obscuring small portions of the study area.

Ancillary vector data were also used: the Montana cadastral framework and the Montana land cover framework. The cadastral framework is a continuously updated geodatabase of private (taxable) and public (exempt) lands (MSL, 2012). The cadastral data are maintained by the Montana Department of Revenue and are available from the Montana State Library (MSL) Geographic Information Services as a vector layer. The

land cover framework is a geodatabase that consists of land cover classes at three levels of detail. The definitions of the land cover classes follow those in the National Land Cover Database 2001 (NLCD2001) and the Northwest Regional Gap Analysis Project (NWGAP) (MSL, 2010). The land cover data are maintained by the Montana Natural Heritage Program and are available from the MSL as a vector layer. We used the June 2012 version of the cadastral framework and the May 2010 version of the land cover framework.

Ground reference data were collected from 525 randomly chosen sites within the study area during July 2012. Sites were required to be agricultural in nature and viewable from public rights-of-way. We recorded the appropriate class (Cereal, Pulse, or Other) for each site (some producers in the area manage fields by alternating strips of a crop and fallow; these fields were classified as the appropriate crop – cereal or pulse). Some reference sites were not used because they: (1) were not accessible; (2) were not actually agriculture; (3) were duplicate observations of the same land parcel; or (4) had clouds in both the mid-season and the late-season mosaics. The final analysis used data from 434 of the ground reference sites; 278 (64%) were Cereal, 70 (16%) were Other, and Pulse comprised 86 sites (20%). These percentages are similar to those reported by the National Agricultural Statistics Service for the region (NASS, 2012).

Image and Vector Data Processing

All images were level-one terrain-corrected (L1T) products. Invalid pixels, caused by slight shifts in the position of the data gaps per band, were removed by masking out values less than 2. The mid-season mosaic was created in several steps. The two images

from Path 36 (36/26 and 36/27) were directly combined because they were from the same day. This combined image is hereafter referred to as the "western" image. The third image (35/26), hereafter the "eastern" image, could not be combined with the western image as it was from eight days earlier. We combined the eastern and western images as follows: (1) cloud-free and shadow-free regions from the overlapping areas of both images were used to derive a linear regression model with the western image as the reference; (2) the coefficients from the regression model were used to normalize the eastern image to the western image; (3) the western image and the normalized eastern image were combined into a mosaic such that missing data pixels in the western image were automatically filled, if available, by pixel values from the eastern image; (4) any remaining pixels in the final mosaic with invalid data were masked out; and (5) the mosaic was clipped to agriculture-only regions using the land cover vector data. The late-season mosaic was created using the same procedure.

Small clouds and shadows were a persistent problem in both mosaics, primarily in the southeastern corner of the study area. Clouds and shadows in each mosaic were classified with a supervised algorithm and a binary (cloud/shadow or no cloud/shadow) mask created from each mosaic. These binary masks were intersected to identify areas with clouds or shadows in both mosaics and a final binary (clouds or shadows in both mosaics or not) mask created. We used this final binary mask to eliminate pixels that had clouds or shadows in *both* mosaics. Pixels with clouds or shadows in a single mosaic were not removed as long as the corresponding pixels in the other mosaic were clear.

We sought a minimum field size of 16.2 ha (40 acres) to ensure that our analyses focused on the region's major producers – the median agricultural field in the study area is approximately 64.7 ha (160 acres). The cadastral layer was consequently modified by removing all land parcels less than 16.1 ha (39.8 acres) in area; this cutoff allowed for small variations in minimum-sized fields. Finally, we examined the cadastral layer with respect to SLC-off gaps. Some land parcels, particularly the smaller ones, were entirely in a data gap, while others had only a portion of the parcel within a gap. We eliminated land parcels with less than 40 pixels of data to obtain a large enough sample size from a field that any statistical measures (e.g., mean spectral value of the blue band for the field) would closely approximate the entire field. This resulted in 13,143 usable agricultural fields in the study area.

Image Segmentation and Data Extraction

The O-O approach required aggregating the pixels into objects – agricultural fields in this case. Both mosaics were segmented solely on the cadastral layer; land parcels were now objects. Pixels without spectral data (the gaps) were defined as not valid. Finally, we extracted the mean and standard deviation of each spectral band for each object using only the valid pixels (e.g., if an object/field consisted of X pixels, of which Y pixels lacked data because of data gaps, the mean and standard deviation was based on the X–Y valid pixels). These values were exported to a spreadsheet for subsequent analysis. Data for the pixel-based analysis were extracted from the mosaics using a shape file for each class as a spatial filter and exported to a spreadsheet for analysis.

Classification

The default classification procedure used a RF classifier on multitemporal data for regions of the study area that were free of clouds and shadows on both image dates. This was the case for both the O-O and the pixel-based approach. Adjustments were necessary for areas that were free of clouds or shadows only on a single date. We first describe the multitemporal procedure. RF models were developed from data extracted from the reference fields from the mid- and late-season mosaics. The number of potential variables depended on the approach. The multitemporal O-O approach had 45 potential variables: (1) the mean value of each spectral band for each of the two dates (14 variables); (2) the standard deviation of each spectral band for each date (14 variables); (3) the normalized difference vegetation index (NDVI) for each date (2 variables); and (4) the date-wise change in all variables, for example the difference between mid-season NDVI and late-season NDVI (15 variables). Standard deviations are not applicable at the pixel-level. Consequently, a separate RF model was developed for the multitemporal pixel-based approach. This model used only 24 variables: (1) the mean value of each spectral band for each date (14 variables); (2) NDVI for each date (2 variables); and (3) the date-wise changes (8 variables).

Portions of the study area obscured by clouds or shadows on one date were classified with RF models developed from the single clear image; areas obscured in the mid-season mosaic used data from the late-season mosaic and vice versa. Consequently, the RF model for the single-date O-O approach had 15 potential variables: 7 mean spectral values, 7 standard deviations, and NDVI. The single-date pixel-based approach

used only 8 variables: 7 mean spectral values and NDVI. Summarizing, we developed a total of *six* RF models based on approach and the status of clouds. We predicted class membership based on data from both dates when possible; single-date classifications comprised less than 1.4% of the cases.

The data, regardless of approach or number of variables, were classified with a RF model as cereal crops, pulse crops, or other. Adjustable parameters included the number of trees, the number of variables to try at each split, and sample size for each class. We developed RF models with 1000 trees, though error rates were asymptotic with substantially fewer trees. Optimal values for the number of variables per split and class sample sizes were found iteratively with the objective of minimizing the overall classification error rate while maintaining approximately equal class error rates. We used the resultant RF models to predict class membership for the entire study area. This gave predicted classes for complete agricultural fields when using the O-O approach, despite data gaps in the SLC-off imagery, whereas the pixel-based approach can only predict the portion with extant data. Classification using the O-O approach produced vector-based maps, whereas classification based on pixels yielded raster maps.

We computed classification accuracy for all RF models via an error matrix (Congalton and Green, 2009). The RF algorithm uses a random subset of approximately two-thirds of the data (sampling with replacement) to build multiple classification trees. The remaining data, the out-of-bag (OOB) observations, are used to compute a classification accuracy for each tree in the RF model, with the results from all trees combined to form the final OOB accuracy assessment (Breiman, 2001). Overall and class

accuracies based on OOB estimates have been shown to be reliable under the assumption that the reference data are unbiased (Lawrence et al., 2006). We guarded against bias by randomly selecting our reference fields. Cohen's kappa statistic ($\hat{\kappa}$) was calculated for each model as a measure of its predictive ability while accounting for correct classifications due to chance agreement (e.g., Congalton and Green, 2009). We also tested whether the difference in kappa statistics between the O-O approach and the pixel-based approach was statistically significant. Finally, we extracted variable importance from each RF model to identify which variables contributed most to decreasing the overall error rate.

Results

Object-Oriented Approach to Classification

The multitemporal approach was able to classify the object-level data from SLC-off ETM+ imagery, despite data gaps, and had an overall accuracy of 85.5% with a κ statistic of 0.75 (Table 2.1). Producer's accuracies (errors are due to omission) were approximately equal for all classes, ranging from 81.9% to 87.0%. There was more disparity in the user's accuracies (errors are due to commission), which ranged from 67.5% for the Other class to 94.0% for the Cereal class. Confusion between Cereal and Other accounted for slightly more than one-half of the errors, primarily where Cereal fields were classified as Other.

Single-date classifications (which were applied to less than 1.4% of the study area) with the mid-season data and the late-season data were less accurate than the

multitemporal classification. The mid-season model had an overall accuracy of 78.4% with a κ statistic of 0.63. Producer's accuracies ranged from 74.2% to 81.6%, while user's accuracies ranged from 53.9% to 91.8%.The late-season data were slightly less accurate with overall accuracy of 75.6% and a κ statistic of 0.55. Producer's accuracies ranged from 62.7% to 82.6%, and user's accuracies ranged from 55.3% to 85.7%. Confusion between Cereal and Other continued to represent the majority of errors. Testing the significance of the differences in kappa statistics suggests that the multitemporal approach gives a statistically better classification ($\alpha = 0.05$) than either of the single-date classifications (Table 2.3).

The variables that contributed most to decreasing the overall error rate differed according to the specific RF model (Figure 2.2 and Table 2.4). We refer hereafter to the following ETM+ spectral bands by band number: 1 = blue (0.45–0.515 μm), 2 = green (0.525–0.605 μm), 3 = red (0.63–0.69 μm), 4 = near infrared (0.75–0.90 μm), 5 = middle infrared (1.55–1.75 μm), 6 = thermal (10.4–12.5 μm), 7 = middle infrared (2.09–2.35 μm). The multitemporal O-O classification relied most on mid-season values from bands 2 and 4 for overall classification. Differences between mid-season and late-season values were very important as 6 of the top 8 variables as ranked by importance were differenced values, while standard deviation variables appeared only once among the top 8 variables. The single-date O-O classifications, in contrast, did not have differenced variables available and relied primarily on mean spectral band values rather than the standard deviations. NDVI and data from band 6 tended to be less important and were not among the top 8 most important variables.

The variables contributing the most to decreasing the error rate for specific classes also varied according to model, as well as class. The most important variables in the O-O multitemporal model (Table 2.4) for the Cereal and Pulse classes, as in the overall classification, were bands 2 and 4. Discrimination between Cereal and Pulse was based, in part, on differenced values from bands 2 and 7. The Other class relied most on differenced values from bands 5 and 7, while contributions from the mid-season band 2 dropped substantially in importance.

The most important variables for overall accuracy of the mid-season O-O classification (Table 2.5) were bands 2, 4, 5, and 7. The same variables contributed the most to classifying the Cereal and the Other classes, but with slight changes in order. The most important variables for the Pulse class were bands 2 and 4, while the standard deviation of bands 3 and 4 were ranked higher than in either of the other classes or in the overall classification. In contrast, none of the late-season O-O classifications (Table 2.6) were especially dependent on bands 2 and 4. The most important variables in the late-season classifications were instead bands 1, 3, and 7. These were the top three variables for the overall classification and each class with the exception of Pulse, in which band 3 was of lesser importance and band 1 was not among the top 8 variables.

Pixel-Based Approach to Classification

The multitemporal approach was able to classify the pixel-level data, except where data were missing due to SLC-off issues, and had an overall accuracy of 89.4% with a κ statistic of 0.81 (Table 2.2). These accuracies were approximately 5% higher than the corresponding O-O models. Producer's accuracies were nearly equal for all

classes, ranging from 88.3% to 89.7%; the user's accuracies ranged from 74.7% to 95.2%. Confusion between Cereal and Other, as found in models using the object-level data, comprised over one-half of the errors.

Single-date classifications with the mid-season and late-season data were less accurate than the multitemporal classification, the same pattern found with the O-O data. The mid-season model had an overall accuracy of 83.8% with a κ statistic of 0.71. Producer's accuracies ranged from 81.8% to 84.7%, and user's accuracies ranged from 66.8% to 92.2%. The late-season model had an overall accuracy of 81.7% and a κ statistic of 0.68, slightly lower than the mid-season model. Producer's and user's accuracies were similar to mid-season values with the exception that the user's accuracy for Pulse was substantially lower in the late-season model than in the mid-season model. Confusion between Cereal and Other comprised the majority of errors. The pixel-based multitemporal approach produced a statistically better classification ($\alpha = 0.05$) than either of the single-date classifications, and the pixel-based approaches produced statistically better classifications than corresponding O-O approaches (Table 2.3).

The most important variables (Figure 2.3 and Table 2.4) for the overall accuracy of the pixel-based classifications differed by model, and were generally not the same variables important in the O-O classifications. Pixel-based multitemporal classifications relied most on values from bands 4 and 6. This is in direct contrast to the O-O classifications in which band 6 was much less important. Differenced values were less important than they were in the O-O approach, only 2 of the top 8 variables ranked by importance were differenced values. Pixel-based single-date classifications, with only 8

potential variables, relied primarily on the spectral band values, while NDVI ranked low in importance.

Variable importance was also class-dependent. In the pixel-based multitemporal classification (Table 2.5), contributions from the mid-season, late-season, and differenced values of band 6 remained strong for all classes. They accounted for 3 of the top 4 variables for all but the Pulse class, in which bands 4 and 7 were substantially more important. Mid-season band 4 values were especially useful for distinguishing the Other class, while band 4 differenced values were most important for the Pulse class.

The most important variables for the overall accuracy of the pixel-based mid-season classification (Table 2.5) were bands 2, 4, 5, and 6; results similar to the O-O classifications. These bands were the top 4 variables for all but the Other class, in which band 7 was more important and band 5 decidedly less so. The most important variable for overall accuracy of the pixel-based late-season classification (Table 2.6) was band 6; this was the case for all classes as well. Bands 1 and 7 provided the distinctions between classes. Band 1 was of high importance for the Cereal class, intermediate for Other, and low for Pulse; band 7 showed the opposite pattern.

A comparison of the predicted maps with respect to the spatial distribution of the classes would provide a means to assess methodology independent of statistical measures. This comparison, however, requires knowledge of the true spatial distribution, which is not known. Nonetheless, a qualitative assessment was made by conferring with agricultural researchers active in the region (Miller, 2012). We found that our predicted maps, both the object-level and the pixel-level, conformed well to the spatial distribution

of the classes in the study area; pulse crops were predominant in Sheridan County while non-pulse or cereal crops tended to be located near the Missouri River.

Discussion

Data gaps in ETM+ imagery have prompted a variety of approaches to fill the gaps with data interpolated from nearby pixels or to fill the gaps with data derived from a different date or a different sensor. Single-date images in which gaps are filled with data from a different date or sensor are unsatisfactory in crop mapping, which requires spectral data from the same or nearly the same day. Furthermore, interpolation schemes encounter predictive difficulties in heterogeneous landscapes or landscapes linear features (e.g., Maxwell, 2004). Agricultural landscapes are often considered homogeneous, but this is a within-field condition. These landscapes can be heterogeneous at field boundaries in which sharp changes in spectral characteristics might exist; for example, the border between adjoining wheat and fallow fields. Furthermore, linear features are characteristic features in agricultural landscapes. Gap-filling schemes, therefore, are unsuitable for crop mapping.

A number of studies have suggested that O-O approaches yield better classification accuracies for crops than classifications using traditional pixel-based approaches (e.g., Pedley and Curran, 1991). Results from this study, however, indicate that models from the pixel-based approach produced overall classifications that were generally 5% better than equivalent models using the O-O approach. These differences in accuracy are significant at the 95% confidence level. We note that there was a

considerable difference in the sample size of the training data between the O-O and the pixel-based approaches; 182 fields versus 15,000 pixels. The better classification accuracies of the pixel-based approach are obviated, however, by the inability to classify data gaps, potential errors in spatial pattern, location of field boundaries, and area.

A pixel-based model only classifies the extant pixels in an image, limiting its usefulness in a spatial context with ETM+ SLC-off imagery or in other cases where data is missing at the sub-field scale, despite higher accuracies. This was evident when we examined results at the field level (Figure 2.4). The classification based on object-level data is easily interpretable and the field boundaries are clear – even for fields in which data gaps obscure the boundaries. The O-O approach assumes that crop management decisions are made on a per-field basis. This is generally the case, yet some producer's in the study area continue to manage fields by strip cropping (typically alternating strips of cereal and fallow). Fields managed by strip cropping include two classes, yet receive only one classification. This is an obvious disadvantage to the O-O approach; however, the number of fields managed by strip-cropping in the study area is quite small and was estimated at less than 4% of the productive agricultural land. We found no difference in the error rates of strip-cropped fields and monoculture fields. The classification based on the pixel-level data, in contrast, was distinctly different. Data gaps are obvious, and there are numerous fields with mixed classes in both the interior (e.g., fields A, B, and C in Figure 2.4) and along field boundaries (e.g., the southern and eastern boundaries of field B). Furthermore, the locations of many field boundaries are unknown. This is the case when the boundary is in a data gap and when adjoining fields are the same class (e.g., the

western boundary of field C). Classification of the pixel-level data identified fields managed by strip-cropping very well; however, if the goal is to identify the crop growing in a field regardless of fallow or sparsely vegetated areas, as was the case for our study, the pixel-level data essentially failed in fallow areas of strip cropped fields to reflect the crops growing in those fields.

A logical consideration is to use the cadastral layer in a post-classification process, thereby capitalizing on the advantages of both approaches. This is certainly a viable option, but incorporating the cadastral layer post-classification does not allow object-level variables to be used in the classification process itself. Object-level variables, such as various shape attributes or relationships of the object to super- and sub-objects, can improve the reliability of classification if they are included *a priori* (Turker and Arikan, 2005). We used the standard deviation of the spectral bands for each field as a measure of variability in the O-O approach.

Multitemporal models are reported to produce better classification accuracies than single-date models, because they can capture spectral diversity due to phenological differences. We reached the same conclusion. The overall accuracies from the analysis of the multitemporal data were higher than accuracies from the single-date data and were significant at the 95% confidence level. This conclusion is robust with respect to approach, O-O or pixel-based.

The largest source of errors is confusion between the Cereal and the Other classes. This is not surprising given the nature of these classes. As discussed earlier, discriminating between Cereal and Other was a challenging task. This was partially

because each class in the study comprised several different crops and, therefore, had higher class variability than would have been expected in a class consisting of a single crop. Furthermore, the Cereal and Other classes had considerable overlap in spectral characteristics as some of the crops in the Other class are quite similar to the Cereal crops. Two notable examples are hay and fields in the Conservation Reserve Program (CRP) that have largely reverted to shortgrass prairie. We also note that much of the confusion between the Pulse and Other classes is likely due to similarities between alfalfa (Other), which is a member of the pea family, and peas (Pulse). Nonetheless, the RF classifier was largely successful by creating multiple paths, via differing variable choices, to the same class. For example, the spectral characteristics of an object consisting of mid-season wheat, late season wheat, and wheat in a strip-cropped field are very different, but the random forest algorithm creates multiple paths to the correct classification.

Variable importance measures from a RF model are instructive, but a strict interpretation of those measures is not feasible and results should be considered speculative as currently available measures of variable importance have been shown to be problematic (Strobl and Zeileis, 2008). Variable importance is not a particular focus of this paper; nonetheless, we present a brief discussion.

The usefulness of phenological differences, derived from multitemporal data, for crop classification is suggested by noting that in the multitemporal O-O model, 6 of the top 8 most important variables were differenced values. Only 2 of the top 8 variables are differenced values with the pixel-level data. Object-level data are based on a field's mean spectral values, which tends to smooth local (pixel-level) variations. Consequently,

differenced values are more meaningful for fields than for individual pixels. The multitemporal model also relied heavily on mid-season variables, particularly bands 2 and 4, whereas only two of the late season variables, bands 5 and 7, were important. Bands 2 and 4 are highly characteristic of green vegetation and the importance of these bands from the mid-season, when crops are close to maturity but are likely maturing at different rates, is reasonable. Bands 5 and 7 are associated with the amount of water in vegetation and the importance of these bands from the late-season, when crops are potentially water stressed, senescent, or have already been harvested, but again at phenologically different times, is likewise reasonable. Standard deviations of the spectral bands, applicable only to the object-level data, were infrequently among the most important variables, regardless of model or class. Standard deviation captures within-field variability, and to some extent texture. The lack of importance in classification for standard deviation suggests minimal differences in variability among classes.

The reliance of the mid-season O-O classification on bands 2, 4, 5, and 7 suggests that there are some key differences in crops that are not based on phenology (reflected only in multitemporal data) but on the crop themselves. These include differences in: (1) greenness (band 2); (2) leaf structure (band 4); and (3) the amount of plant-held water (bands 5 and 7). The Pulse class, for example, is distinguished by high importance for bands 2 and 4, while bands 5 and 7 are relatively low in importance. In contrast, these bands are the top four most important variables in classifying the Cereals and Other classes. Accordingly, plant-held water is an essential variable in separating the Cereal and Other classes from the Pulse class. We note that the importance of bands 5 and 7 for this

particular separation does not necessarily imply that the Cereal and the Other classes have high amounts of plant-held water, only that the amounts differ in some important way. It is possible, and more likely, that the Pulse class has greater amounts of plant-held water. Bands 2, 4, 5, and 7 remained important for late-season O-O classification, in different relative orders, but bands 1 and 3 were of higher importance. The importance of band 1 is likely associated with harvesting (which does not typically occur for the Other class) since band 1 is associated with soil reflectance. Band 3 is crucial to distinguishing between green vegetation and senescent (or yellow) vegetation; mature cereal crops are yellow.

Patterns in variable importance for the pixel-based classification were generally similar to those found in the O-O classification. There was one key difference; band 6, unimportant with the object-level data, was important at the pixel-level. There were substantial differences among the classes with respect to mid-season band 6 values, suggesting different thermal characteristics among the classes at this time of the year.

The RF classifier is well-known and generally found to be superior to standard classification approaches (e.g., Lawrence et al., 2006). The method performed well in this study and was able to classify the object-level data as well as pixel-level data with acceptable accuracies. One of the advantages of RF models is their ability to handle large datasets and missing values (Breiman, 1996). We used RF to classify nearly 23 million pixels with 24 explanatory variables, and 13,143 fields using 45 variables; neither presented difficulties. The ability of RF to determine variable importance is beneficial, but not unique to RF. Variable importance facilitates an understanding of which variables

contribute the most to classification; but, current measures are problematic and results are speculative.

We sought a method to classify a northeast Montana agricultural landscape into the region's dominant classes: cereal crops, pulse crops, and other, using SLC-off ETM+ imagery – without using gap-filling schemes. Although the pixel-based approach gave higher overall classification accuracies than the O-O approach, it was unable to classify areas within data gaps and produced fields with mixed classifications. The O-O approach, in contrast, enabled the classification of fields partially located within data gaps; we were, of course, unable to classify fields completely within data gaps. The multitemporal approach encapsulated class-wise phenological variations and provided a statistically significant contribution to overall classification accuracies. RF modeling produced acceptable overall classification accuracies with approximately equal class error rates. We did not investigate the minimum number of pixels required to classify an object. We used a minimum of 40 pixels under the supposition that this sample size was sufficiently large, according to the Law of Large Numbers (LLN), to obtain representative means and standard deviations. Research concerning the minimum number of pixels required to classify an object would be an appropriate next step.

Acknowledgements

This article is based upon work supported by the U. S. Department of Energy and the National Energy Technology Laboratory under Award Number DE-FC26-05NT42587. This report was prepared as an account of work sponsored by an agency of

the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. We thank the anonymous reviewer for the thorough and thoughtful comments.

Table 2.1. Error matrices for object-oriented classifications.

Classified Data	Reference Data			
	Cereal	Other	Pulse	User's Accuracy[a]
Multitemporal				
Cereal	187	4	8	94.0%
Other	21	54	5	67.5%
Pulse	7	5	59	81.9%
Producer's Accuracy[b]	87.0%	84.4%	81.9%	
Overall Accuracy	85.5%			
Kappa	0.75			
Mid-season Only				
Cereal	180	10	6	91.8%
Other	34	49	8	53.9%
Pulse	15	7	62	73.8%
Producer's Accuracy[b]	78.6%	74.2%	81.6%	
Overall Accuracy	78.4%			
Kappa	0.63			
Late-season Only				
Cereal	210	16	19	85.7%
Other	25	42	9	55.3%
Pulse	19	9	49	63.6%
Producer's Accuracy[b]	82.7%	62.7%	63.6%	
Overall Accuracy	75.6%			
Kappa	0.55			

[a]User's accuracy is measured with respect to errors of commission. [b]User's accuracy is measured with respect to errors of omission.

Table 2.2. Error matrices for pixel-based classifications.

Classified Data	Reference Data			
	Cereal	Other	Pulse	User's Accuracy[a]
Multitemporal				
Cereal	206615	5286	5183	95.2%
Other	16257	57308	3171	74.7%
Pulse	7540	1267	63232	87.8%
Producer's Accuracy[b]	89.7%	89.7%	88.3%	
Overall Accuracy	89.4%			
Kappa	0.81			
Mid-season Only				
Cereal	193927	7348	9183	92.2%
Other	22991	54114	3877	66.8
Pulse	13494	2399	58526	78.6
Producer's Accuracy[b]	84.2%	84.7%	81.8%	
Overall Accuracy	83.8%			
Kappa	0.71			
Late-season Only				
Cereal	188841	7090	8745	92.3%
Other	20317	52609	5473	67.1%
Pulse	21254	4162	57368	69.3%
Producer's Accuracy[b]	82.0%	82.4%	80.1%	
Overall Accuracy	81.7%			
Kappa	0.68			

[a]User's accuracy is measured with respect to errors of commission. [b]User's accuracy is measured with respect to errors of omission.

Table 2.3. Tests for statistically significant differences between Kappa statistics.

Models Compared [a]	Test Statistic (Z)	p-value
O-O Multitemporal vs. O-O Mid-season	2.43	< 0.01
O-O Multitemporal vs. O-O Late-season	4.00	< 0.01
Pixel-based Multitemporal vs. Pixel-based Mid-season	68.79	< 0.01
Pixel-based Multitemporal vs. Pixel-based Late-season	91.04	< 0.01
O-O Multitemporal vs. **Pixel-based Multitemporal**	1.83	0.03
O-O Mid-season vs. **Pixel-based Mid-season**	2.21	0.01
O-O Late-season vs. **Pixel-based Late-season**	3.36	< 0.01

[a] The models in bold are those with the *larger* kappa statistic.

Table 2.4. The top variables for multitemporal classification as ranked by importance [a].

	O-O Multitemporal				Pixel-Based Multitemporal			
Rank	Overall	Cereal	Other	Pulse	Overall	Cereal	Other	Pulse
1	B2 (M)	B2 (M)	B7 (D)	B2 (M)	B6 (M)	B6 (M)	B4 (M)	B4 (D)
2	B4 (M)	B4 (M)	B5 (D)	B4 (M)	B6 (L)	B6 (L)	B6 (L)	B7 (D)
3	B7 (D)	B4 (D)	B4 (M)	B7 (D)	B6 (D)	B6 (D)	B6 (D)	B6 (M)
4	B4 (D)	B2 (D)	B3 (D)	B4 (D)	B4 (L)	B2 (M)	B6 (M)	B6 (L)
5	B5 (D)	B7 (L)	B4 (D) S	B7 (L)	NDVI (L)	B4 (L)	B4 (D)	B7 (M)
6	B4 (D) S	B1 (M)	NDVI (D)	B4 (D) S	B4 (D)	NDVI (L)	B7 (D)	B6 (D)
7	B3 (D)	B5 (D)	B4 (D)	B2 (D)	B7 (M)	B4 (M)	B7 (L)	B7 (L)
8	B2 (D)	B5 (L)	B1 (D)	B5 (L)	B4 (M)	B4 (D)	B5 (M)	B4 (L)

[a] B1–7 refers to the appropriate ETM+ spectral band: 1 = blue (0.45–0.515 μm), 2 = green (0.525–0.605 μm), 3 = red (0.63–0.69 μm), 4 = near infrared (0.75–0.90 μm), 5 = middle infrared (1.55–1.75 μm), 6 = thermal (10.4–12.5 μm), 7 = middle infrared (2.09–2.35 μm). Parenthetical annotations indicate the source: M = values from the mid-season, L = values from the late season mosaic, D = differenced values. S = standard deviation.

Table 2.5**.** The top variables for mid-season classification as ranked by importance [a].

	O-O Mid-Season				PB Mid-Season			
Rank	Overall	Cereal	Other	Pulse	Overall	Cereal	Other	Pulse
1	B2 (M)	B2 (M)	B4 (M)	B4 (M)	B4 (M)	B2 (M)	B6 (M)	B4 (M)
2	B4 (M)	B5 (M)	B7 (M)	B2 (M)	B2 (M)	B6 (M)	B4 (M)	B2 (M)
3	B5 (M)	B7 (M)	B5 (M)	B4 (M) S	B6 (M)	B4 (M)	B7 (M)	B6 (M)
4	B7 (M)	B4 (M)	B2 (M)	B3 (M)	B5 (M)	B5 (M)	B2 (M)	B5 (M)
5	B4 (M) S	B3 (M)	B1 (M)	B3 (M) S	B7 (M)	B7 (M)	B3 (M)	B1 (M)
6	B3 (M)	B1 (M)	B4 (M) S	B7 (M)	B1 (M)	NDVI(M)	B1 (M)	B3 (M)
7	B1 (M)	B4 (M) S	B7 (M) S	NDVI (M)	NDVI (M)	B1 (M)	NDVI (M)	B7 (M)
8	B3 (M) S	B6 (M)	NDVI (M)	B5 (M)	B3 (M)	B3 (M)	B5 (M)	NDVI (M)

[a] B1–7 refers to the appropriate ETM+ spectral band: 1 = blue (0.45–0.515 μm), 2 = green (0.525–0.605 μm), 3 = red (0.63–0.69 μm), 4 = near infrared (0.75–0.90 μm), 5 = middle infrared (1.55–1.75 μm), 6 = thermal (10.4–12.5 μm), 7 = middle infrared (2.09–2.35 μm). Parenthetical annotations indicate the source: M = values from the mid-season, L = values from the late season mosaic, D = differenced values. S = standard deviation.

Table 2.6. The top variables for late-season classification as ranked by importance[a].

	O-O Late-Season				PB Late-Season			
Rank	Overall	Cereal	Other	Pulse	Overall	Cereal	Other	Pulse
1	B7 (L)	B7 (L)	B3 (L)	B7 (L)	B6 (L)	B6 (L)	B6 (L)	B6 (L)
2	B3 (L)	B1 (L)	B7 (L)	B4 (L) S	B7 (L)	B1 (L)	B3 (L)	B7 (L)
3	B1 (L)	B3 (L)	B1 (L)	B5 (L)	B1 (L)	B5 (L)	B1 (L)	B5 (L)
4	B4 (L)	B4 (L)	NDVI (L)	B3 (L)	B5 (L)	B3 (L)	B4 (L)	B3 (L)
5	B5 (L)	B5 (L)	B4 (L)	B4 (L)	B3 (L)	NDVI (L)	B7 (L)	B4 (L)
6	B2 (L)	B2 (L)	B7 (L) S	B7 (L) S	NDVI (L)	B4 (L)	B5 (L)	NDVI (L)
7	B4 (L) S	B4 (L) S	B2 (L)	B4 (L) S	B4 (L)	B7 (L)	NDVI (L)	B1(L)
8	B7 (L) S	B6 (L)	B6 (L) S	B2 (L)	B2 (L)	B2 (L)	B2 (L)	B2 (L)

[a] B1–7 refers to the appropriate ETM+ spectral band: 1 = blue (0.45–0.515 µm), 2 = green (0.525–0.605 µm), 3 = red (0.63–0.69 µm), 4 = near infrared (0.75–0.90 µm), 5 = middle infrared (1.55–1.75 µm), 6 = thermal (10.4–12.5 µm), 7 = middle infrared (2.09–2.35 µm). Parenthetical annotations indicate the source: M = values from the mid-season, L = values from the late season mosaic, D = differenced values. S = standard deviation.

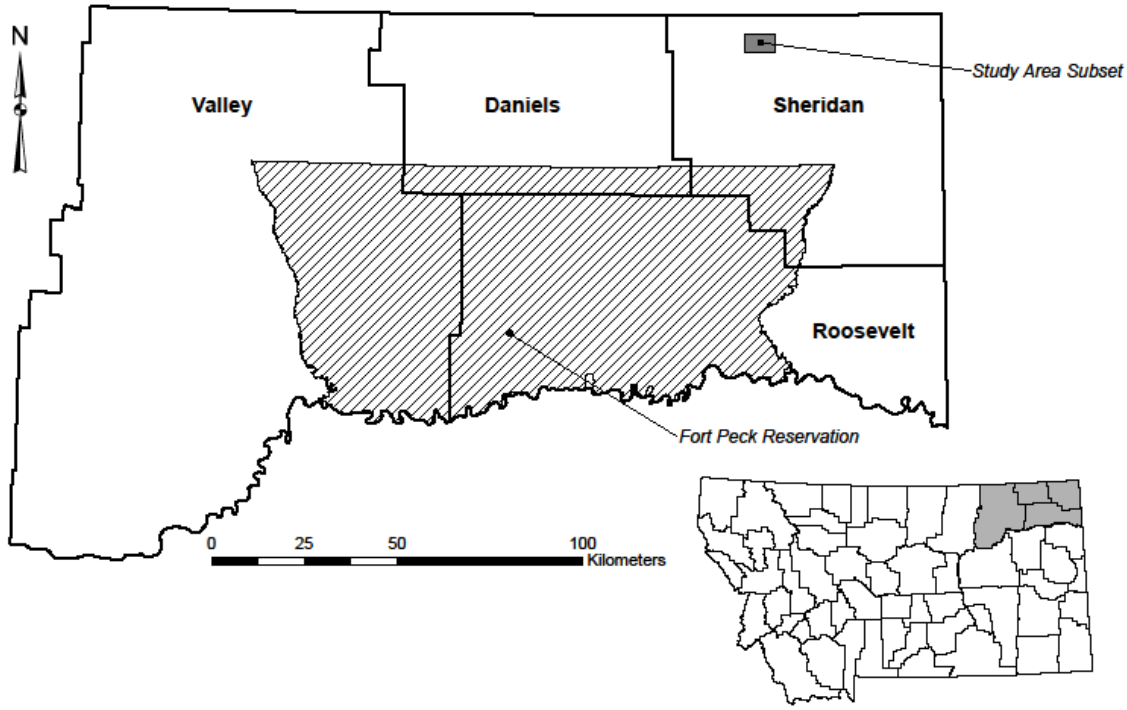

Figure 2.1. The study area is located in northeastern Montana and consists of four counties: Daniels, Roosevelt, Sheridan and Valley; it also encompasses the Fort Peck Reservation. The study area subset is located in northwest Sheridan County.

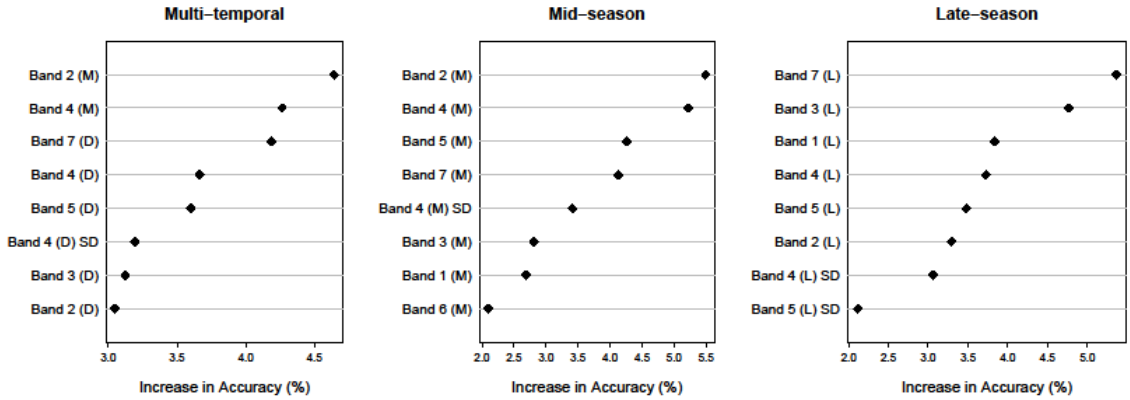

Figure 2.2. The top 8 variables in the O-O classification as ranked by order of importance. Band 1–7 refers to the appropriate ETM+ spectral band: 1 = blue (0.45–0.515 µm), 2 = green (0.525–0.605 µm), 3 = red (0.63–0.69 µm), 4 = near infrared (0.75–0.90 µm), 5 = middle infrared (1.55–1.75 µm), 6 = thermal (10.4–12.5 µm), 7 = middle infrared (2.09–2.35 µm). Parenthetical annotations indicate the source: mid = values from the mid-season mosaic, late = values from the late season mosaic, D = differenced values. SD = standard deviation.

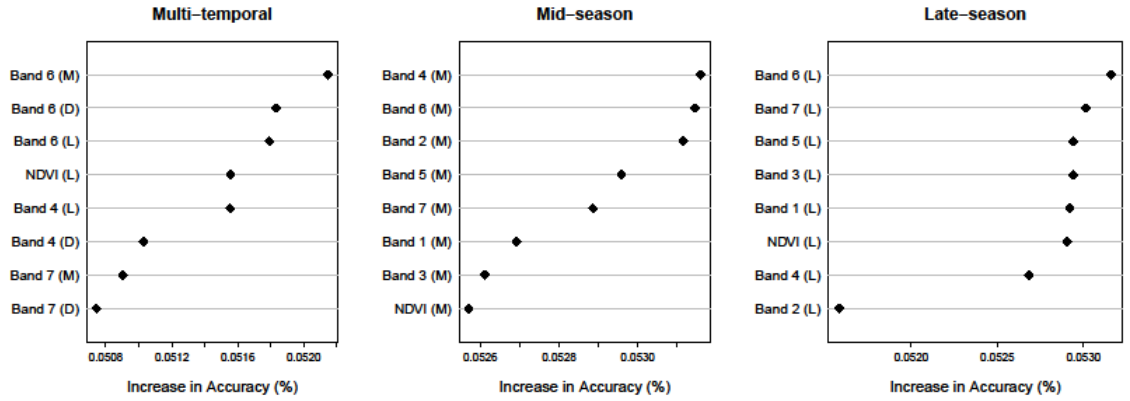

Figure 2.3. The top 8 variables in the pixel-based classification as ranked by order of importance. Band 1–7 refers to the appropriate ETM+ spectral band: 1 = blue (0.45–0.515 µm), 2 = green (0.525–0.605 µm), 3 = red (0.63–0.69 µm), 4 = near infrared (0.75–0.90 µm), 5 = middle infrared (1.55–1.75 µm), 6 = thermal (10.4–12.5 µm), 7 = middle infrared (2.09–2.35 µm). Parenthetical annotations indicate the source: mid = values from the mid-season mosaic, late = values from the late season mosaic, D = differenced values. SD = standard deviation.

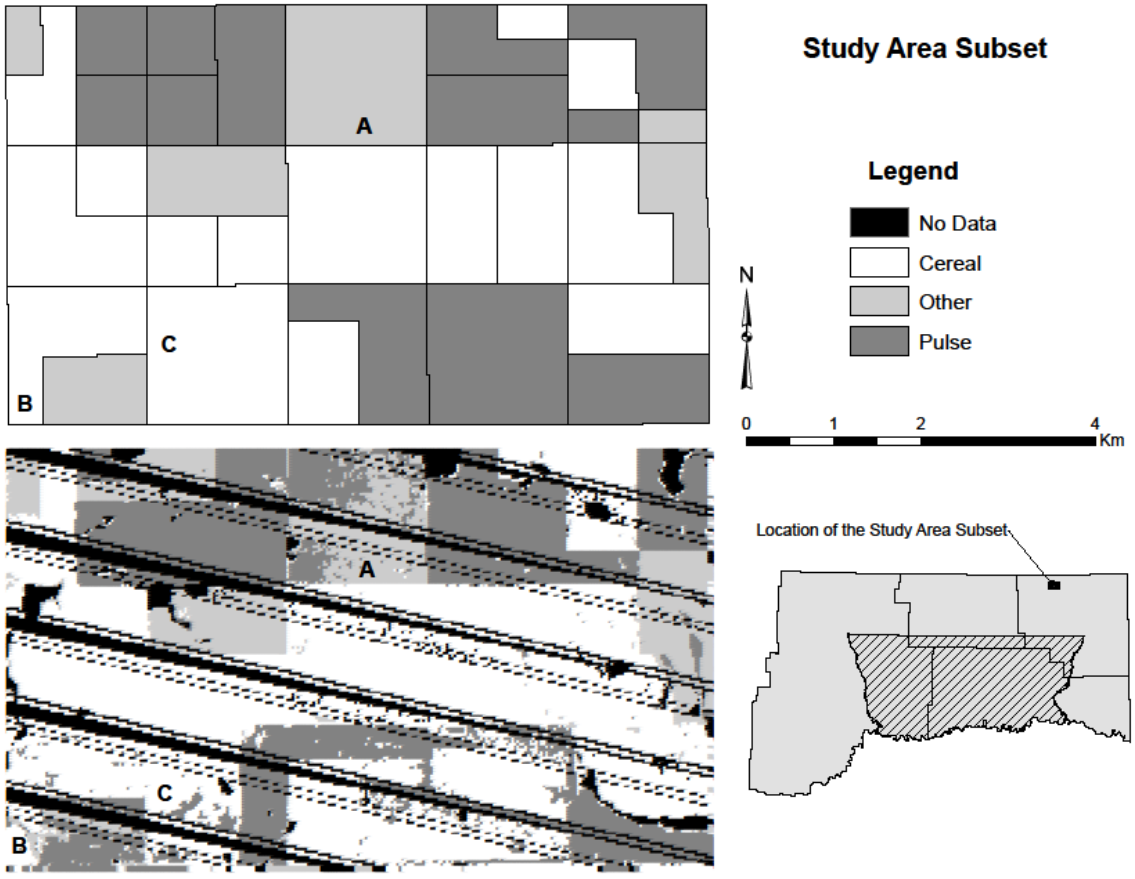

Figure 2.4. The study area subset is located in northwest Daniels County. The top-left portion of the figure is the subset as classified with the object-level data, while the bottom-left is the same area classified with pixel-level data. Locations A, B, and C are discussed in the text.

CHAPTER 3

CHANGES IN FIELD-LEVEL CROPPING SEQUENCES: INDICATORS OF SHIFTING AGRICULTURAL PRACTICES

Contribution of Authors and Co-Authors

Manuscript in Chapter 3

Author: John A. Long

Contributions: Conceived and implemented the study design. Collected and analyzed the data. Wrote manuscript and revisions, and submitted manuscript for publication.

Co-Author: Dr. Rick L. Lawrence

Contributions: Helped conceive the study design. Provided guidance and feedback on manuscript drafts, remote sensing issues, and statistical analysis.

Co-Author: Dr. Perry R. Miller

Contributions: Provided feedback on manuscript drafts and agricultural issues.

Co-Author: Lucy A. Marshall

Contributions: Provided feedback on manuscript drafts and statistical analysis.

Manuscript Information Page

John A. Long, Rick L. Lawrence, Perry R. Miller, Mark C. Greenwood, Lucy A. Marshall
Agriculture, Ecosystems & Environment
Status of Manuscript: (Put an x in one of the options below)
___ Prepared for submission to a peer-reviewed journal
___ Officially submitted to a peer-review journal
___ Accepted by a peer-reviewed journal
_X_ Published in a peer-reviewed journal

Published by Elsevier in 2014, Volume 189, pages 11 – 20.

Abstract

Farmers implement an assortment of management practices to ensure the sustainability, economic viability, and resilience of their operation. Dryland farming practices dominate in the semiarid regions of the US northern Great Plains, where historical practice has been to rotate small-grain cereals with whole-year summer fallow; however, pulse crops (e.g., lentils) have become increasingly common in these regions as an alternative to fallow. The area of fallow in northeastern Montana, for example, has decreased by one-third, while the area of pulse crops has increased more than five-fold. Our objectives were: (1) to characterize the principal cropping sequences in northeast Montana during the period of regional pulse crop adoption (2001–2012); and (2) to identify changes in the relative proportions of these sequences during the same period. We identified crops at the field-level by class (cereal, pulse, fallow, or cereal-fallow strips) for 2001–2012 using multitemporal Landsat imagery in conjunction with the Cropland Data Layer, cadastral data, ground reference data, and local producers' records. The annual crop classifications were combined into a 12-character string for each field that represented the sequence of crop classes for 2001–2012. We then searched these strings for specific 2- and 3-yr crop sequences with a string-matching algorithm. The most abundant sequences involved continuous cereal, block-managed cereal-fallow, and cereal-pulse. We also observed a steady decrease in the abundance of cereal-fallow sequences managed by strip-cropping that were coincident with increases in block-managed cereal-fallow sequences and with increases in pulse production. We conclude that, over the study's time frame, regional producers grew more cereal crops and

fallowed fields less frequently, but did not appear to strongly adhere to specific sequences. Furthermore, strip-cropping as a management practice has declined substantially.

Introduction

The area of fallowed cropland in northeastern Montana has decreased by one-third in the past 15 years, while the area devoted to pulse crops has increased nearly five-fold (Lee, 2011). Pulses, leguminous crops grown for their edible seeds, were grown primarily on land formerly in cereal-fallow sequences, managed by strip-cropping (i.e., alternating 50 or 100 m wide strips) or in block-managed rotations and in idle fields returning to production following expired enrollment in the Conservation Reserve Program (CRP). Incorporating pulses into rotational sequences with the region's dominant crop – small grain cereals – improves the robustness and resilience of local agricultural systems (Zentner et al., 2001, 2002; Burgess et al., 2012). The identification of specific rotational sequences is important because they provide an insight into the general long-term sustainability of regional agriculture. Furthermore, determining the prevailing sequences can help establish which ones have been successful and, therefore, which sequences might be more likely to succeed in similar regions.

Management practices affect the sustainability, economic viability, and resilience of an operation. Management practices at the field-level include logistical (e.g., scheduling planting), crop management (e.g., fertilizer rates, water usage), crop (species and variety), cropping systems (e.g., sequences and rotations), and tillage systems (e.g.,

no till, conservation till, or traditional till) (Meinke and Stone, 2005). We focus on individual fields, because, in most instances, they represent the smallest and most fundamental decision unit for managers.

Climate plays a large role in determining whether a particular management practice is adopted locally and, if adopted, how prevalent the practice becomes. In semiarid regions, the established agricultural practices focus primarily on capturing, conserving, and effectively using water (Cochran et al., 2006). Dryland farming practices dominate the drier portions of the U.S. Northern Great Plains (NGP), such as northeastern Montana. Historical practice in these regions has been to rotate small-grain cereals with summer fallow, a practice that entered widespread use in the late 1930s as a means to manage soil water in regions where rainfall is a limiting resource (Cochran et al., 2006). Cultivating pulses as green manure recently has become an increasingly common alternative to summer fallow in rotations with cereals (Miller et al., 2006; O'Dea et al., 2013; Tanaka et al., 2010).

Cereal-pulse sequences include a cereal, typically spring or winter wheat (*Triticum aestivum* L.), durum (*Triticum turgidum* L.), or barley (*Hordeum vulgare* L.), and a pulse crop such as dry pea (*Pisum sativum* L.), lentil (*Lens culinaris* Medik.), dry bean (*Phaseolus vulgaris* L.), or chickpea (*Cicer arietinum* L.) (Lemke et al., 2007). These sequences provide many field-level benefits including: (1) biological nitrogen (N) fixation, thereby making additional N available to the succeeding crop and reducing fertilizer requirements; (2) substantially lower water requirements than cereals or alfalfa; (3) improved options for controlling weeds; (4) reduced insect and disease problems; (5)

improved soil tilth and stability; (6) reduced soil erosion; and (7) improved yields and higher protein levels for the following cereal crop (Cochran et al., 2006; Lee, 2011, Peel, 1998).

Changes in the relative proportions of cereal-fallow (block-managed and strip-cropped) and cereal-pulse sequences have implications for farm productivity and profitability, as well as broader implications in terms of improved environmental quality and economic impacts on surrounding communities (Zentner et al., 2001, 2002). The pulse industry in the northeast Montana counties of Daniels, Roosevelt, Sheridan, and Valley has recently experienced rapid growth. This region cultivated less than 20,000 ha annually through 2003; however, production increased to more than 125,000 ha during the relatively short span of 2004–2006. The area devoted to pulse crops has stabilized since 2006 with little additional growth; however, other regions such as north-central Montana have experienced recent growth (NASS, 2013). The annual economic benefits to these counties have been substantial with more than $100 million directly attributed to the 2010 pulse crop (Lee, 2011). The lessons learned here can be useful to producers in other regions in which the pulse crop potential has yet to be fully realized. The objectives of this work were: (1) to characterize the principal cropping sequences in northeast Montana during the period of regional pulse crop adoption (2001–2012); and (2) to identify changes in the relative proportions of these sequences during the same period.

Material and Methods

Study Area

The study area comprises the four most northeastern Montana counties of Daniels, Sheridan, Roosevelt, and Valley (Figure 3.1). It is bounded by Saskatchewan to the north, North Dakota to the east, and the Missouri River to the south. Federal and state lands lie along much of the western border and are non-cropland. The region is characterized by low relief and has a semiarid climate (Padbury et al, 2002). Total precipitation averages just over 310 mm annually, occurring primarily as rain between April and September (WRCC, 2013). Maximum daily temperatures in July average 31°C, while January maximum daily temperatures average –10°C (WRCC, 2013). The dominant land cover types are shortgrass prairie and agriculture. Regional agricultural practices are primarily dryland systems (Tanaka et al., 2010), although center-pivot irrigation is not uncommon for producers in close proximity to the Missouri River. Agriculture consists largely of cereal crops, primarily spring wheat, and an increasingly substantial area of pulse crops throughout most of the region; however, relatively small amounts of other crops are grown within the Missouri River corridor (NASS, 2012).

Data

Satellite Imagery. Landsat images for 2001–2006 and 2012 were obtained from the Earth Resources Observation and Science Center (EROS). Three Landsat scenes were required for full coverage of the study area: Path 36, Row 26 (36/26); Path 36, Row 27 (36/27); and Path 35, Row 26 (35/26). We used a mid-season mosaic (~ mid-July) and a

late-season mosaic (~ mid August) from each year to better capture phenological variation and improve classification. Obtaining cloud-free images of the study area required using imagery from two sensors: the Thematic Mapper (TM) aboard Landsat 5 (2001–2006) and the Enhanced Thematic Mapper Plus (ETM+) aboard Landsat 7 (2001–2006, 2012) (Table 3.1). The ETM+ images from 2003 onward have data gaps caused by the permanent failure of its scan-line corrector (SLC) (see e.g., Markham et al., 2004); images with these data gaps are known as SLC-off images. Due to the positions of the three overlapping Landsat scenes relative to the study region, data loss due to SLC-off gaps has been estimated at approximately 15% or less (Long et al., 2013).

Ground Reference Data. Ground reference data from 525 locations were collected during July 2012. The locations were randomly chosen from fields identified as agriculture and that were viewable from public rights-of-way. We recorded location and class (Cereal, Pulse, or Other). Cereal-fallow strip-cropping is not an uncommon practice, particularly in the western portion of the study area; these fields were typically classified as ‘cereal’. Some reference sites were unusable because they were: (1) not accessible; (2) not agriculture; (3) duplicate observations of the same field; or (4) cloud covered in both mid-season and late-season images. The final dataset used 434 of the ground reference sites; 278 (64%) were cereal crops, 86 (20%) sites were pulse crops, and 70 (16%) were something else (typically CRP land or, much less frequently, some other crop such as alfalfa). These percentages are similar to regional data from the US Department of Agriculture (USDA) (NASS, 2012).

The Cropland Data Layer. The Cropland Data Layer (CDL) is a geo-referenced raster-based data layer denoting specific agricultural cover types (e.g., wheat, lentil, or corn). It was developed by the USDA's National Agricultural Statistical Service (NASS) primarily to assist in determining seasonal area estimates for major commodity crops (Johnson, 2013; NASS, 2013). The CDL has been used in a wide range of agricultural (e.g., Scheffran et al., 2007; Schultz et al., 2007; Kutz et al., 2012) and environmental (e.g., Linz et al., 2004; Hagen et al., 2005) studies. The CDL extends back to 1997 and is based on medium-resolution satellite data with extensive ground reference data (Boryan et al., 2011; Han et al., 2012). Coverage of the conterminous US for 2008–2012 is complete, while coverage for 1997–2007 depends on the state. Coverage for Montana, for example, dates to 2007, while neighboring North Dakota extends back to 1997.

The CDL is created from a variety of inputs. Imagery inputs are multitemporal and derive from several satellite-based sensors. The list of sensors depends on the year; the CDLs in this study were based on: (1) the Advanced Wide Field Sensor (AWiFS) (2007–2010); (2) TM (2008– 2011); (3) Moderate Resolution Imaging Spectroradiometer (MODIS) 16-day NDVI (Normalized Differenced Vegetation Index) composite (2007, 2009); and (4) the SLIM-6 sensor aboard the DEIMOS-1 satellite (2011–2012) and the UK-DMC-2 satellite (2012) (Boryan et al., 2011; NASS, 2013). Additional inputs to the CDL included the National Elevation Dataset (NED), the National Land Cover Database (NLCD), and the USDA Farm Service Agency (FSA) Common Land Units (CLU), which are the ground reference data (Boryan et al., 2011; Johnson, 2013; NASS, 2013).

Spatial resolution is either 30 m (2008, 2010–2012) or 56 m (2007, 2009) and depends on which sensors were used.

Classification accuracies for the CDL are reported by NASS for each state by year and are typically 80% to 95% (producer's and user's) for the major crops (NASS, 2013). These are state-wide accuracies; however, there are regional variations in accuracy that are not captured by this assessment. The state-wide accuracy report for the 2012 Montana CDL claims an overall accuracy of 73.8%; but when we assessed the accuracy of the study area portion of the CDL using our own ground reference data (434 observations), we found an overall accuracy of 92.6% ($\kappa = 0.86$). This is likely due to high class accuracies of the study area's major crop classes (fallow ~ 93%, cereals ~ 87%, pulses ~ 70%). There is some concern that the CDL tends to underestimate area when calculations are based on pixel counts (Johnson, 2013), but we employ an object-oriented approach (discussed momentarily), which largely avoids this bias. We therefore found the CDL to be a reliable source of ancillary data.

Ancillary Vector Data. We also used ancillary vector data: the Montana cadastral framework and the Montana land cover framework, both available from the Montana State Library's Geographic Information Clearinghouse. The cadastral framework is a geodatabase of private (taxable) and public (tax exempt) land parcels, which generally delineate agricultural fields. The land cover framework is a geodatabase of land cover classes as defined by the 2001 NLCD and identifies, among other things, agricultural land. We used the cadastral framework from June 2012 and the May 2010 land cover framework (MSL, 2012).

Data Preparation

Creating the 'Fields' Layer. We began by intersecting the cadastral layer, which delineates land ownership boundaries (fields), with the NLCD layer. This step eliminated all land parcels from the cadastral layer that lay in regions not defined as agricultural. We sought land parcels of 16.2 ha (40 acres) or larger to define agricultural fields. This minimum field size was chosen for two reasons – it allowed us to easily remove small land parcels (e.g., residential lots) from the cadastral layer, and it helped to focus our analyses on major producers. The median agricultural field in the study area was 65 ha in area, and the 16.2 ha minimum eliminated less than 1% of agricultural land. Consequently, the cadastral layer was modified by deleting all fields less than 16.1 ha; allowing for small variations in minimum-sized fields. Some fields, particularly the smaller ones, were located entirely within the 2012 ETM+ data gaps; these fields were also deleted from the cadastral layer. Fields that were located partially within a data gap were eliminated if they consisted of less than 40 pixels; this was to ensure a large enough sample size of pixels such that any field-based statistical measure (e.g., mean spectral value of the red band) would be representative of the entire field. The modified cadastral layer was designated the 'Fields' layer and comprised 13,145 agricultural fields.

Creating the CDL Mosaics. The CDL data were obtained from the USDA via the CropScape web-based interface. Annual data are provided as TIFF (tagged image file format) images clipped to an area of interest (AOI). We downloaded the CDL data for the study area (Montana counties Daniels, Roosevelt, Sheridan, and Valley) for 2007–2012,

as well as neighboring Divide and Williams counties in North Dakota for 2001–2006 (the North Dakota CDLs were needed to facilitate classification for 2001–2006 (see section 2.5)). The individual county files were combined into regional mosaics by year. CDL data were thematic rasters and comprised 255 classes, the vast majority of which were not present in the study area. We combined these 255 classes into four by recoding the thematic classes for each of the regional mosaics: (1) cereals (barley, durum wheat, spring wheat, winter wheat, rye, oat, and other small grains); (2) fallow (fallow/idle cropland); (3) pulse (chickpea, lentil, and pea); and (4) other (all remaining classes).

Image Pre-Processing. Level-one terrain-corrected (L1T) Landsat images were acquired from the USGS via the EarthExplorer web-based interface. Invalid pixels, caused by line dropout or slight shifts in the position of the data gaps per band, were removed. We employed a multitemporal approach to improve subsequent classification and constructed a mid-season and a late-season mosaic for each year. Complete coverage of the study area required imagery from different dates. Since differences in atmospheric conditions or solar angles affect the amount of reflected radiation, radiometric normalization was required.

Mosaics were produced in several steps. First, we directly combined the two Path 36 images (36/26 and 36/27) into the ‘west’ image without radiometric normalization because these images were from the same day. The third image (35/26), the ‘east’ image, was from a different date since consecutive Landsat orbits are not adjoining, necessitating radiometric normalization. We combined the east and west images by using the overlapping regions to derive a linear regression model with the western image as the

reference, adjusting the values in the east image, and then creating a mosaic using the west and the adjusted east images. Clouds and their shadows were a persistent problem in nearly every year in this study, particularly in the mid-season, and were identified with a supervised classified algorithm and a binary (cloud/shadow or no cloud/shadow) mask created for each mosaic. We used these binary masks to eliminate pixels containing clouds or cloud shadows.

Changes in spectral values from mid- to late-season capture important phenological differences and have proved to be important variables for crop classification (Long et al., 2013). We, therefore, created a differenced mosaic by subtracting the late-season mosaic from the mid-season one. Next, the mid-season, late-season, and the differenced mosaics were stacked and cleaned by removing pixels with a zero in any layer. Zeros represented pixels with no data, due to either data gaps or to clouds and shadows, and this step ensured that the remaining pixels had data for all layers. The final step was to clip the stacked mosaics to the Fields layer.

Determining Field-Level Crop Class (2007–2012)

We used the CDL and the cadastral layer to determine crop class at the field-level for 2007–2012. The CDL mosaics were converted to thematic rasters with four classes during the data preparation phase; however, the mosaics were still pixel-based. Classification at the pixel level is challenging because monoculture fields are often classified as multiple classes; however, classification algorithms can operate on groups of contiguous pixels (object-oriented) as well as individual pixels (pixel-based). Object-oriented classification is most appropriate when the landscape can be delineated into

meaningful homogeneous regions, ideal for agricultural fields (e.g., Forster et al., 2010). Several studies have concluded that object-oriented approaches to classification produce better results than pixel-based approaches (e.g., Pedley and Curran, 1991). Transforming the pixel-based CDL layers to a collection of fields with a single classification required four basic steps: (1) creating crop class-specific versions of the CDL; (2) segmentation of these layers; (3) data extraction; and (4) classification.

We created field-based objects in which objects were assigned to the class corresponding to the majority of the extant pixels from the CDL. The only exception to this scheme is when the proportion of cereal and fallow were nearly equal, in which the fields were classified as cereal-fallow strips. This process resulted in six final classes: cereal (C), fallow (F), pulse (P), other crops or non-agricultural land (O), cereal-fallow strip-cropping (S), and no data or missing (M). The single-letter class identifiers were necessary for subsequent analysis with string matching algorithms; hereafter, we refer to these class identifiers. The accuracy of this procedure was evaluated by comparing the derived field-level classifications to the 2012 ground reference data and to local producers' records for 2007–2012.

Determining Field-Level Crop Class (2001–2006)

Montana CDL layers were not available for 2001–2006; therefore, a separate procedure was developed to determine crop class at the field-level for these years. We took advantage of two key facts: (1) the footprint for our Landsat imagery encompassed Divide and Williams counties in North Dakota, immediately adjacent on the eastern

border of the study area; and (2) CDL data existed for these North Dakota counties for 2001–2006.

We began with an unsupervised classification (36-classes) of the Landsat imagery, since we had no ground reference or CDL data for the study area. We then used the CDL data from the two North Dakota counties as an overlay. This allowed us to identify each of the 36 unsupervised classes, based on pixels, as C, F, O, or P. These classified pixels were aggregated at the field-level. The classified image at this point was a thematic raster with four classes and subsequent processing was as described in the previous section. Accuracy was evaluated by processing the 2012 imagery with this method and comparing the derived field-level classifications to the 2012 ground reference data and to local producers' records for 2012. We made the necessary assumption that, if the method was successful for 2012, it would be successful for other dates, since no reference data was available to test this assumption.

String Construction and Pattern Determination

The annual crop classifications were combined into a 12-character string representing the sequence of classes (e.g., CFCFCFCCPCCP) for each field in the study. There were 106 fields classified as M (missing data) for at least one year in the 2001–2012 sequence. We were able to recover 37 of these by examining the classification for the remaining years. If a field was the same class for all years in which we had data, we assumed that the missing year was the same class. If, for example, a field was missing a class for 2005 and all of the other years were O (OOOOMOOOOOOO), then we felt confident that 2005 was also O and replaced the M with O in the string. We only

recovered missing data in this manner and made no attempts to infer any patterns for fields with missing data. The remaining 69 fields with missing data were removed from the dataset, leaving a total of 13,076 fields with complete sequences.

We then searched these strings for specific *a priori* 2- and 3-yr sequences (e.g., CP, see Table 3.2) using the Biostrings package (Pages et al., 2013) in the statistical program R (R Development Core Team, 2011). Identifying 2- and 3-yr sequences within longer sequences is problematic because every 2-yr sequence is embedded within 3-yr sequences, which are embedded within 4-yr sequences, and so forth (see Hennessy, 2006, for a mathematical discussion). This embedded relationship creates ambiguities when we try to assign a field in a particular year to a specific sequence. The third year, for example, in the 4-yr string CC**F**C is fallow; but, this field might represent the first year of a FC sequence, the second year of a CF sequence, the third year of a CCF, or one of the many other possibilities. In any specified year, a field is simultaneously in several sequences.

A common approach to characterize changes in cropping sequences is to partition the data into non-overlapping blocks (e.g., Plourde et al., 2013). The string CCFCFP, for example, can be considered as three 2-yr sequences (CC, FC, and FP), or as two 3-yr sequences (CCF and CFP) under this scheme. This approach avoids ambiguity by assigning years to specific blocks. If our example, CCFCFP, represented 2001–2006 and we used 2-yr blocks, then 2003 is the first year in a FC sequence. Partitioning the data into non-overlapping blocks ignores the notion that fields are simultaneously in several sequences. An alternative to non-overlapping blocks is to use an *n*-year moving window

– a rotation-conditioned lag operation (Hennessey, 2006). With a 2-yr moving window, the string CCFCFP is considered as five 2-yr sequences (CC, CF, FC, CF, and FP), while a 3-yr moving window yields four 3-yr sequences (CCF, CFC, FCF, and CFP).

We looked at the data in *two* different ways. We used a 2-yr moving window (2001–2002, 2002–2003, …, 2011–2012) to identify and characterize the principal 2-yr sequences, and similarly structured windows to characterize 3-yr sequences. We summarized all sequences in terms of area and the number of fields involved and identified changes in these sequences by identifying increases and decreases in abundance over time. We also identified changes in area over time for the individual classes (C, F, P, and S) and total productive agricultural land.

Results

Individual Classes

Total productive agricultural land in the region increased by approximately 180,000 ha between 2001 and 2012, driven by increases in area devoted to cereal (C) and pulse (P) crops (Figure 3.2a). Cereals increased from 380,000 ha to 580,000, a substantial increase (Figure 3.2b). Pulse crops increased from 20,000 ha to 110,000 ha (Figure 3.2c). Cereal-fallow strip-cropping (S), in contrast, decreased substantially in area, from 210,000 ha to 80,000 ha (Figure 3.2e). Fallowed cropland (F) maintained a constant trend across the study's time frame, but this is due to a doubling of fallow in 2011 during an excessively wet spring that interfered with planting in the region; the area of fallow had been decreasing throughout 2001–2010 (Figure 3.2d). The plots for C and F exhibit a

moderately strong inverse relationship (r = -0.59), which consists of synchronized patterns of alternating highs and lows that typically move in opposition to one another (Figure 3.2f). Fallow reached highs in 2002 and 2011; both years were characterized by above average spring precipitation (WRCC, 2013).

The mean field size increased for all classes between 2001 and 2012. Most of these increases were modest: the average size of cereal fields increased by 3 ha, fallow fields had a 10-ha increase on average, cereal-fallow strip-cropped fields increased an average of 10 ha, and pulse fields increased by an average of 22 ha. These increases were accompanied by overall increases in the number of fields, area, and proportion of land in production. Mean area of the strip-cropped fields increasing while the total area of strip-cropping decreased suggests that smaller fields were preferentially converted to block management, perhaps as a trial prior to full conversion of all fields.

Two-Year Sequences

Two-year monoculture sequences comprised 28–48% of the fields, depending on the year, while multi-class rotational sequences accounted for the remaining 52–72%. Plots of monoculture sequences indicate that the number of fields involved in a specific monoculture sequence (CC, FF, and PP) varied across the time frame. Cereal (CC), for instance, experienced a minimum in 2005–2006, and peaked three years later during 2008–2009 (Figure 3.3a). Continuous fallow (FF) was typically quite low (< 5,000 ha) during the study, but peaked dramatically in 2010–2011 (Figure 3.3b). Continuous pulse sequences (PP) were likewise uncommon (< 10,000 ha) except for a brief period during 2005–2007; yet, these PP sequences comprised less than 1% of all fields (Figure 3.3c).

There was an increase in fallow and a drop in cereal production for 2010–2011 and 2011–2012. These relationships persisted for all sequences, regardless of whether we measured by the number of fields or by area. Hereafter, unless specifically noted, results and comments will apply equally to the number of fields as well as their area. We did not consider fields in continuous cereal-fallow strip-cropping (SS) as monoculture sequences, but rather a form of cereal-fallow rotations, which are discussed shortly.

Plots of multi-class rotational sequences illustrate, as in the monoculture sequences, that the number of fields involved in rotations varied. The most notable trends were the nearly ten-fold increase in the number of fields in cereal-pulse (CP/PC) sequences (Figure 3.4a) and the substantial decline in the number of cereal-fallow sequences (Figure 3.4c) managed by strip-cropping (SS). Cropping sequences involving fallow increased for periods that included 2011. Cereal-fallow (CF/FC) sequences, for instance, were declining, but increased when 2011 was included (Figure 3.4b). Fields in other configurations of cereal-fallow sequences (SC and FS/SF) were less common (Figures 3.4e & 3.4f). The number of SC sequences generally increased over the study's time frame. The SC, and to lesser degree the SF, sequences mark the transition in management practices away from strip-cropping. These transitions were more frequent in the 2003–2004 and the 2006–2007 time frames. Pulse-fallow (PF/FP) sequences were generally not common, but they did exhibit a small rise early in the study (2003–2007) and again in 2011 (Figure 3.4d).

Three-Year Sequences

Monoculture sequences based on a 3-yr moving window comprised 22–42% of the fields in production, while multi-class rotational sequences accounted for the remaining 58–78%. Continuous 3-yr cereal sequences (CCC) were frequent, but less abundant than the 2-yr sequences (CC). The CCC sequences followed the same general pattern of relative abundance (Figure 3.5a) as did the CC sequences. Continuous fallow (FFF) and pulse (PPP) were very uncommon, with less than 15 fields for either sequence (Figs. 5b & 5c). Extended cereal-pulse rotations (CCP, CPC, and PCC) exhibited the same dramatic rise in abundance as in the 2-yr sequences, from 335 fields in the 2001–2003 time frame to 2416 fields in 2010–2012 (Figure 3.6a). Extended cereal-fallow rotations (CCF, CFC, and FCC) also mirrored the 2-yr sequences, slightly declining until 2011 was included in the window (Figure 3.6b). Cereal-fallow rotations managed by continuous strip-cropping (SSS) steadily declined throughout the study's time frame, decreasing by 70% (Figure 3.6d). We also considered 3-yr pulse-cereal-fallow sequences (PCF, PFC, CPF, CFP, FCP, and FPC). These sequences were less common, but not trivial – typically less than 2% of all fields, (Figure 3.6c). PCF sequences typically constituted less than 300 fields; however, there was an increase to 440 fields for the 2009–2011 window (Figure 3.6c), which is likely due to the large number of fallow fields in 2011.

Discussion

Total Production and Monoculture Sequences

An increase in the mean field size for *all* target classes (C, F, P, and S) suggests that additional fields were added to production between 2001 and 2012. This increase in the total area under production is illustrated in Figure 3.2a and is indicative of regional agricultural extensification. Increases in total production are likely primarily due to considerably large areas returning to production following CRP enrollment. Regional reductions of 10.3% in the amount of CRP mirror the decreasing trend at the national level following areal caps on CRP in the 2008 Farm Bill (Fargione et al., 2009, Lee, 2011).

Cereal production varied substantially during 2001–2012, but had an overall increasing trend. Maximum production was in 2012, while minimum production was in 2006 and coincided with a peak in the number of fields enrolled in CRP (Lee, 2011). Cereal abundance is inextricably linked to the abundance of fallow, and the observed increasing trend for regional cereal production is associated with a decreasing trend in fallow. This suggests that either continuous cereal cropping was increasingly common or that block-managed alternate year cereal-fallow sequences were less common. Our data indicate that both situations occurred, but continuous cereal production was the stronger trend. We found that the abundance of the 2-yr continuous cereal sequences (CC) tended to increase over time, but the 3-yr sequences (CCC) remained essentially unchanged.

Montana led the nation in lentil and pea production in 2011 (Lee, 2011). Northeast Montana is the state's premier pulse producing region, and we observed that

pulse production increased six-fold between 2001 and 2012. Pulses accounted for a small amount of cropland in the study area, approximately 22,000 ha per year, during 2000–2003. Pulse crops began a dramatic rise in 2004, which peaked in 2010 at approximately 134,000 ha. Our data closely match those from national and state-level agencies (Lee, 2011; NASS, 2012). Pulse crops reportedly were grown primarily on land that previously had used fallow as a management practice (Lee, 2011). The field-level benefits of pulse crops are mainly beneficial to the successive cereal crop; therefore, continuous pulse sequences (PP or PPP) were uncommon. Two-year pulse sequences (PP) were rare except for the 2005–2006 and 2006–2007 windows, but they never exceeded 100 fields. The 3-yr sequences (PPP) were very infrequent.

Continuous fallow (FF or FFF) is an uneconomical practice, and we found it infrequently throughout the first 10 years of the study. It is possible that some of the fields in 2- or 3-yr fallow sequences were misclassified as fallow when they were more likely strip-cropped fields. Two-year sequences (FF) averaged 3,400 ha during 2001–2010, while 3-yr sequences averaged only 390 ha. We observed a quadrupling of continuous fallow sequences (FF and FFF) above their respective averages for time frames that include 2011. This increase is because many producers were unable to seed their fields in 2011 as they struggled with below average temperatures and precipitation as much as 400% above normal (NOAA, 2011). Many fields were, therefore, unintentionally fallowed; we see this quite clearly in the data as increases in the number of fallow fields (Figure 3.2d) and in the number of FF sequences (Figure 3.3b).

Cereal-Fallow Sequences

The rotation of crops as a beneficial agricultural practice has been known to farmers for thousands of years, yet historical trends in agricultural production have been to increase cropping intensification and simplify rotations. Formal investigations with dryland crop rotations began at least as early as 1908 when the Montana Agricultural Experiment Station initiated a series of experiments with continuous cropping systems and various rotations (Atkinson et al., 1917). Much of this work involved land brought under cultivation for the first time after breaking native prairie sod. Our results indicate that, for northeastern Montana, the trend to increase cropping intensification has continued, but the trend to simplify rotations is reversing.

Fallow decisions in dryland farming conditions are fundamentally issues of soil water management, but they function within the confines of economic considerations. Cereal production is highly responsive to soil moisture at planting time and precipitation during the growing season. Producers, unfortunately, have no control over precipitation, but soil moisture can be influenced by fallowing decisions. Fallowing cereal fields in alternate years is a long-standing practice, which was described in 1703 as "a very ancient piece of husbandry" (Mortimer, 1703). Researchers (e.g., Burt and Allison, 1963; Verburg et al., 2012) have recommended that producers base fallow decisions on soil moisture at planting time, rather than considering continuous cereal or strict CF sequences as the only options. We found cereal-fallow sequences, regardless of form – basic (CF) or extended sequences (CCF/CFC/FCC), to be common, but not consistent in area or in the number of fields. Consequently, rigid adherence to continuous cropping or

rotational sequences by producers in the study area does not seem likely. All cereal-fallow sequences tended to decrease in abundance though 2010, but rebounded in response to the large number of fields fallowed in 2011.

Cereal-Fallow Sequences Managed by Strip-Cropping

There was a substantial decrease, nearly three-fold, in the abundance of cereal-fallow sequences managed by strip-cropping. Producers moved away from strip-cropping, converting those fields to other cereal-fallow sequences or to pulse production. Transitions away from strip-cropping were persistent throughout the study's time frame (Figure 3.2e) and are apparent by noting increases in the frequency of SC and SF sequences. These two sequences mark the transition in management practices away from strip-cropping and we observed spikes in 2003–2004 and in 2006–2007 (Figure 3.4e).

Reductions in strip-cropping practices during the study likely stem from three factors: (1) decreased vulnerability to insect damage; (2) increases in no-till practices; and (3) increases in pulse production. The management of cereal-fallow sequences by blocks (alternating years of cereal and fallow in a field) instead of strips reduces vulnerability to damage from the wheat stem sawfly (*Cephus cinctus* Norton), a major insect pest. Sawfly damage to fields is primarily limited to the edges; transitioning from strips to blocks considerably reduces the extent of edges (e.g., Weiss and Morrill, 1992). Strip-cropping is foremost a management practice to reduce soil erosion, primarily due to wind (e.g., Weiss and Morrill, 1992). No-till practices inject seeds directly into the previous year's stubble, which substantially reduces soil erosion without affecting

productivity. No-till management is the majority tillage practice regionally, with estimates between 50% and 90% (e.g., Watts et al., 2009, Hansen et al., 2012). No-till practices also reduce fallow by efficient use of moisture – crop residues ease the loss of soil moisture due to evaporation and increase infiltration by slowing runoff and trapping snow (e.g., Scholten, 1988). The final factor is dramatic increases in the number of fields in cereal-pulse sequences. Pulse crops replace fallow in traditional cereal-fallow sequences since they have substantially lower water requirements (Cochran et al., 2006). Cereal-pulse sequences can, therefore, supplant cereal-fallow sequences that are managed by strip-cropping.

Pulse Sequences

Formal recommendations in favor of cereal-pulse sequences instead of cereal-fallow can be found more than 300 years ago (e.g., Mortimer, 1703). Our data show a substantial rise in the number of fields involved in cereal-pulse sequences (CP/PC and CCP/CPC/PCC), which peaked in 2010 with Montana's record pulse crop (Lee, 2011). Since 2010, these cereal-pulse sequences have accounted for approximately 17% of all agricultural fields in the study.

Pulse-fallow sequences (PF/FP) were uncommon, since pulse often replaces fallow in cropping sequences. Nonetheless, small numbers of pulse-fallow sequences were observed in 2004–2007 and in 2011. We also considered 3-yr pulse-cereal-fallow sequences (PCF/PFC/CPF/CFP/FCP/FPC). These sequences were likewise uncommon and tended to remain steady at approximately 14,000 ha, except for 2011. These sequences were not abundant for the same reason as pulse-fallow. The 2011 results are

likely an anomaly, as noted previously, where numerous fallow fields represented missed planting windows.

Conclusion

We sought to identify and characterize changes in the principal cropping sequences during the period of pulse crop adoption (2001–2012) in northeast Montana, which we summarize below.

- The amount of cropland in active production increased, largely from declining CRP participation.
- The production of cereals increased overall, but was highly variable. The increases were closely associated with equivalent decreases in fallow. Continuous cereal and cereal-fallow sequences increased, but continuous cereal was the stronger trend – potentially due to improvements in soil water management through the adoption of no-till practices. Producers grew more cereal and fallowed fields less often but did not appear to strongly adhere to specified sequences – it is more likely that producers responded to soil moisture levels when deciding whether to plant a crop or to fallow the field.
- The prevalence of strip-cropped cereal-fallow sequences decreased substantially, particularly in the eastern portion of the study area. Notable decreases occurred in 2004 and 2007, which were coincident with increases in block-managed cereal-fallow sequences and with increases in pulse production. Producers that continued

with cereal-fallow sequences tended to convert to block management rather than strip-cropping.

- Pulse production increased dramatically, primarily on previously fallowed fields. Pulse crops were nearly always in sequences with cereals; continuous pulse and pulse-fallow sequences were rare. Producers replaced a substantial amount of fallow with pulse crops.
- Excessive precipitation in 2011 caused farmers to unintentionally fallow many of their fields. This tripled the amount of fallow and altered sequences.

We examined the 12-yr record of field-level crop classes in two different ways, as 2-yr sequences and as 3-yr sequences. The most prevalent 2-yr sequences were continuous cereal, cereal-fallow, and cereal-pulse. The extended versions of these same sequences dominated when we considered the data as 3-yr sequences: continuous cereal (CCC), cereal-fallow (CCF), and cereal-pulse (CCP), along with their permutations. Continuous strip-cropping, a previously established cereal-fallow sequence, has declined rapidly.

Acknowledgements

This article is based upon work supported by the U. S. Department of Energy and the National Energy Technology Laboratory under Award Number DE-FC26-05NT42587. This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any

information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Table 3.1. Dates and sensors for the Landsat imagery used in the study.

Year	Mid-Season Dates	Sensors[a]	Late-Season Dates	Sensors
2001	5 Aug / 6 Aug	ETM+ / TM	21 Aug / 22 Aug	ETM+ / TM
2002	23 Jul / 24 Jul	ETM+ / TM	23 Aug / 24 Aug	ETM+ / ETM+
2003	18 Jul / 19 Jul	TM / ETM+	11 Aug / 12 Aug	ETM+ / TM
2004[b]	12 Jul / 13 Jul	ETM+ / TM	21 Jul / 14 Aug	ETM+ / TM
			13 Aug / 29 Aug	ETM+ / ETM+
2005	8 Jul / 15 Jul	ETM+ / ETM+	1 Sep / 2 Sep	ETM+ / TM
2006	11 Jul / 18 Jul	ETM+ / ETM+	19 Aug / 28 Aug	ETM+ / ETM+
2012	11 Jul / 18 Jul	ETM+ / ETM+	12 Aug / 19 Aug	ETM+ / ETM+

[a] TM is the Thematic Mapper (Landsat 5); ETM+ is the Enhanced Thematic Mapper (Landsat 7); [b] The 2004 late-season required a different pair for the east (path 35) and west (path 36) halves of the mosaic because of excessive clouds.

Table 3.2. Definition of cropping sequences.

Length	System	Sequences [a]
2-yr Rotations	Monoculture	CC, FF, PP
	2-Class Rotations	CP/PC, CF/ FC, SS, PF/FP, SC, FS/SF
3-yr Rotations	Monoculture	CCC, FFF, SSS, PPP
	2-Class Rotations	CCF/CFC/FCC, CCP/CPC/PCC
	3-Class Rotations	CFP/CPF/FCP/FPC/PCF/PFC

[a] Sequences are shown with all permutations; C = cereal, F = fallow, P = pulse, S = strip-cropped.

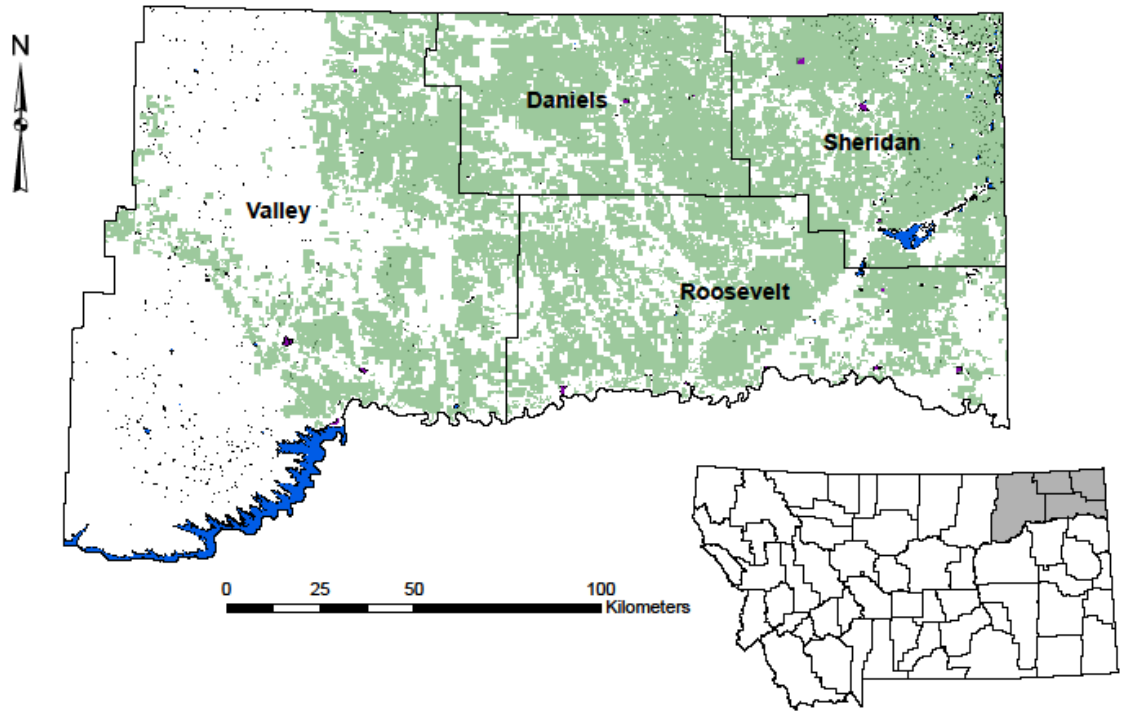

Figure 3.1. The study area encompassed the northeastern Montana counties of Daniels, Roosevelt, Sheridan and Valley. Agricultural areas, as defined in this study, are indicated by shading in the larger map.

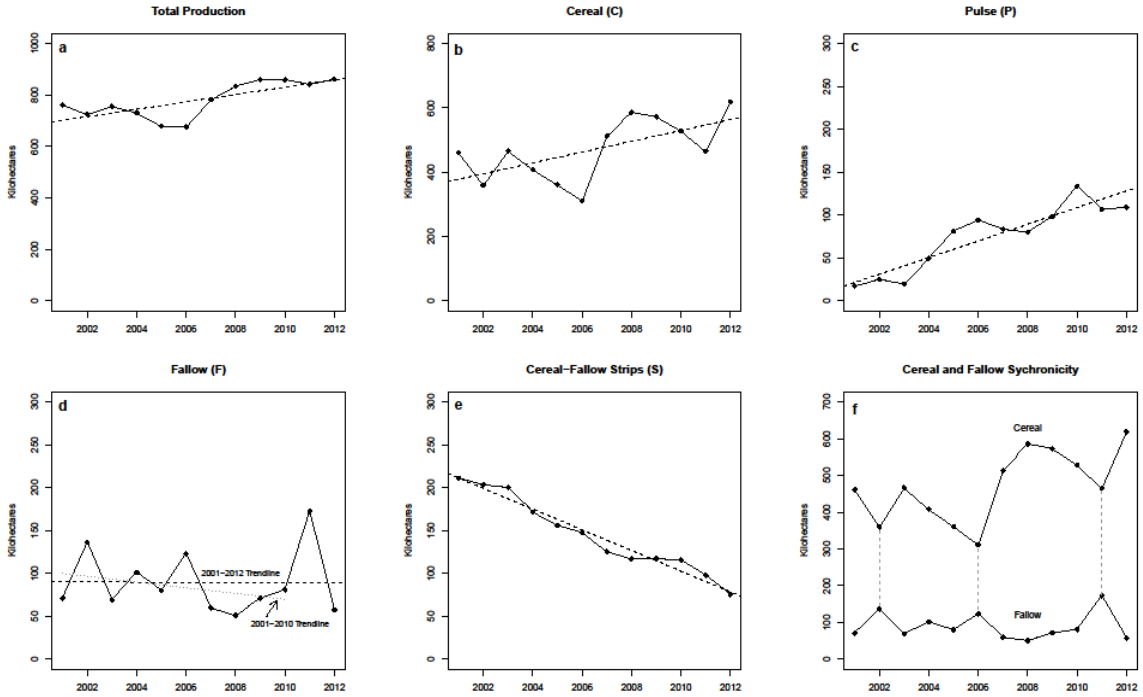

Figure 3.2. Area on an annual basis (2001–2012) devoted to: (a) total agricultural production; (b) cereal (C) crops; (c) pulse (P) crops; (d) fallowed (F) fields; (e) cereal-fallow strip-cropped (S) fields; and (f) cereal and fallow plotted together to illustrate the synchronicity.

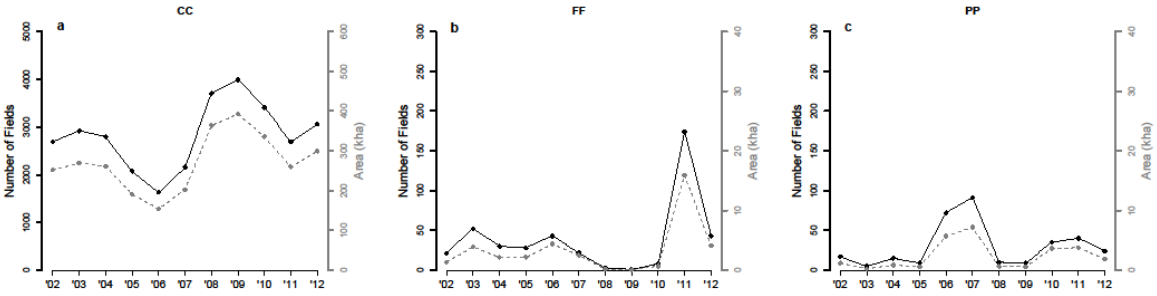

Figure 3.3. Abundance of 2-yr monoculture sequences in number of fields (black solid line) and in area (gray dotted line) for: (a) cereal-cereal (CC); (b) fallow-fallow (FF); and (c) pulse-pulse (PP). The horizontal axis is a 2-yr moving window such that, for example, ’02 = 2001–2002, ’03 = 2002–2003.

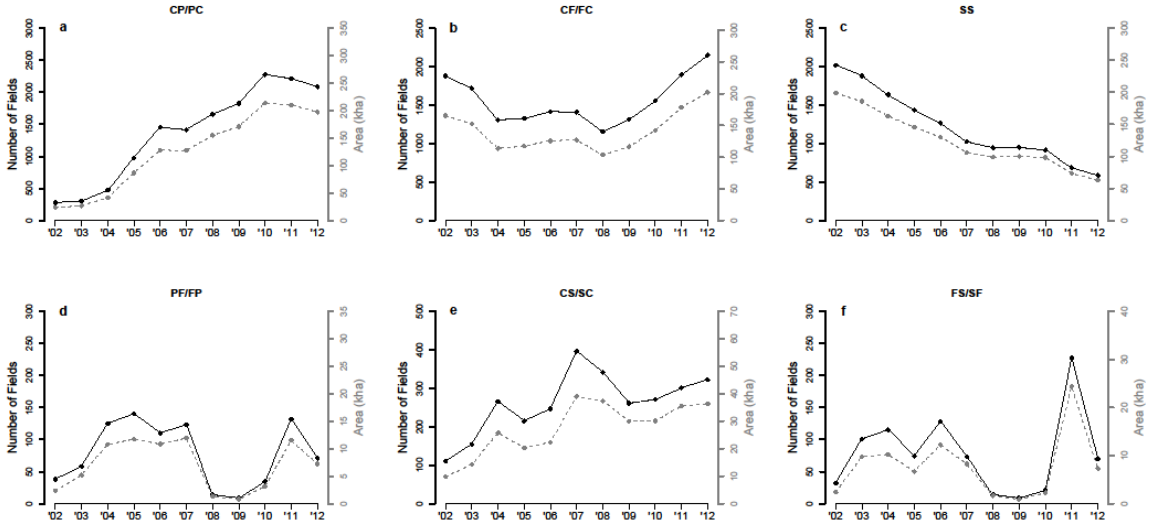

Figure 3.4. Abundance of 2-yr rotational sequences in number of fields (black solid line) and in area (gray dotted line) for: (a) cereal-pulse (CP/PC); (b) cereal-fallow (CF/FC); (c) continuously strip-cropped (SS); (d) pulse-fallow (PF/FP); (e) cereal-strip-cropped (SC); and (f) fallow-strip-cropped (FS/SF). The horizontal axis is a 2-yr moving window such that, for example, ’02 = 2001–2002, ’03 = 2002–2003.

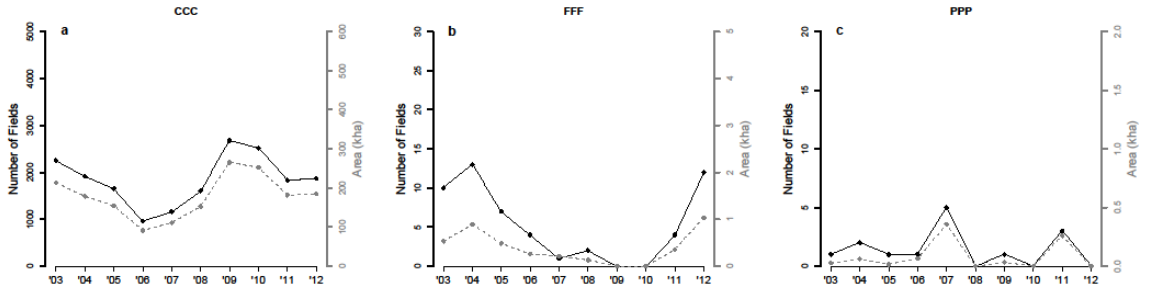

Figure 3.5. Abundance of 3-yr monoculture sequences in number of fields (black solid line) and in area (gray dotted line) for: (a) continuous cereal (CCC); (b) continuous fallow (FFF); and (c) continuous pulse (PPP). The horizontal axis is a 3-yr moving window such that, for example, ’03 = 2001–2003, ’04 = 2002–2004.

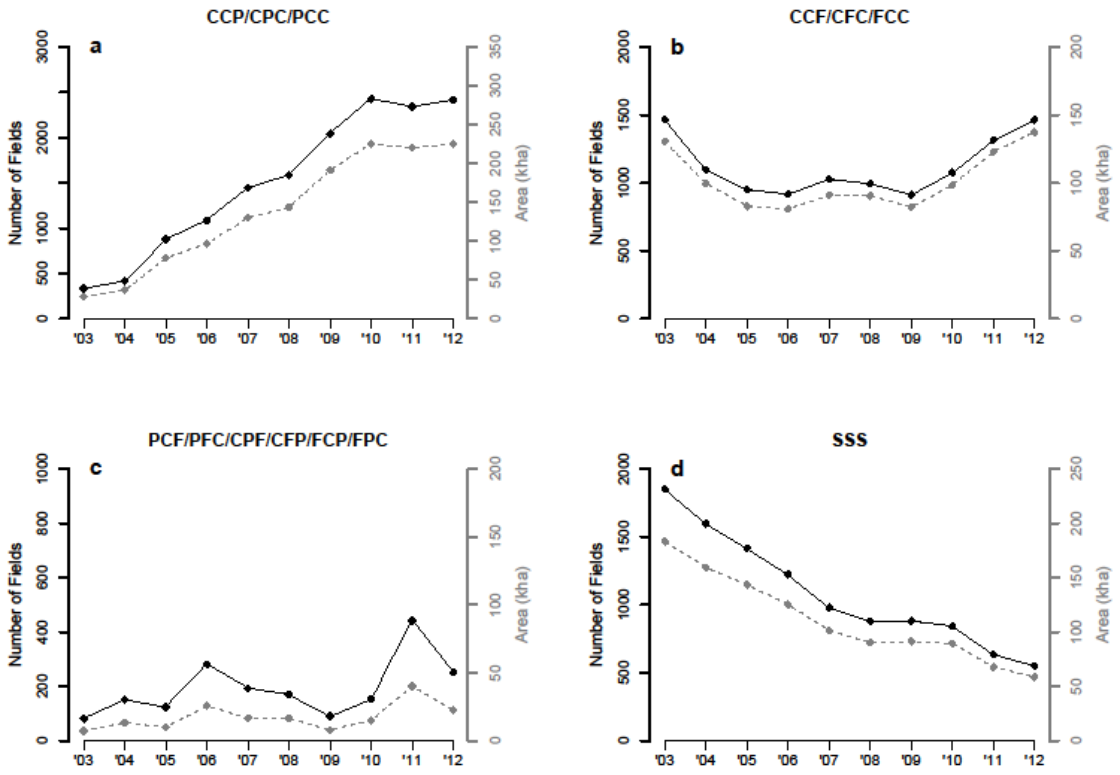

Figure 3.6. Abundance of 3-yr rotational sequences in number of fields (black solid line) and in area (gray dotted line) for: (a) cereal-cereal-pulse (CCP/CPC/PCC); (b) cereal-cereal-fallow (CCF/CFC/FCC); (c) pulse-cereal-fallow (PCF/PFC/CPF/CFP/FCP/FPC); and (d) continuously strip-cropped (SSS). The horizontal axis is a 3-yr moving window such that, for example, ’03 = 2001–2003, ’04 = 2002–2004.

CHAPTER 4

ADOPTION OF CROPPING SEQUENCES IN NORTHEAST MONTANA: A SPATIO-TEMPORAL ANALYSIS

Contribution of Authors and Co-Authors

Manuscript in Chapter 4

Author: John A. Long

Contributions: Conceived and implemented the study design. Collected and analyzed the data. Wrote manuscript and revisions, and submitted manuscript for publication.

Co-Author: Dr. Rick L. Lawrence

Contributions: Helped conceive the study design. Provided guidance and feedback on manuscript drafts, remote sensing issues, and statistical analysis.

Co-Author: Dr. Perry R. Miller

Contributions: Provided feedback on manuscript drafts and agricultural issues.

Co-Author: Dr. Mark C. Greenwood

Contributions: Provided feedback on manuscript drafts and statistical analysis.

Co-Author: Lucy A. Marshall

Contributions: Provided feedback on manuscript drafts and statistical analysis.

Manuscript Information Page

John A. Long, Rick L. Lawrence, Perry R. Miller, Mark C. Greenwood, Lucy A. Marshall
Agriculture, Ecosystems & Environment
Status of Manuscript: (Put an x in one of the options below)
_X_ Prepared for submission to a peer-reviewed journal
___ Officially submitted to a peer-review journal
___ Accepted by a peer-reviewed journal
___ Published in a peer-reviewed journal

Published by Elsevier
Submitted May, 2014

Abstract

Producers make the decision to adopt a particular agricultural practice within a range of social, economic, environmental, and agronomic constraints. The semiarid regions of the US northern Great Plains are dominated by dryland farming practices and the traditional practice has been to rotate small-grain cereals with summer fallow; however, producers are moving away from this practice. The area of fallow in northeastern Montana decreased by one-third and the area of pulse crops increased nearly six-fold during 2001–2012. We previously identified two key practices that are indicative of regionally changing agricultural practices: (1) the broad-scale adoption of cereal-pulse sequences, and (2) the conversion from continuous strip-cropping to block managed cereal-based sequences. Here, we examined the adoption of these two practices from a spatio-temporal perspective to determine if the observed patterns were consistent with those expected from *a priori* processes: random occurrence, spread and adoption of the practices due to social interaction as described in innovation diffusion theory, or adoption based on environmental factors. Our results suggest that the adoption and spread of both practices were likely constrained by the suitability of the physical environment and possibly driven by social processes. Available water, in particular, exerts a fundamental control on the decision whether or not to adopt either of these practices. We also found evidence for the expansion of these practices due, in part, to social factors, particularly during the early period of adoption. We conclude that producers made the decision whether or not to adopt these practices primarily as a function of environmental suitability and, to a lesser extent, within the context of social interactions.

Introduction

Innovations in technology and management practices are vital to the sustainability and increased productivity of agricultural systems; yet, the decisions to adopt them are made by individual farmers faced with incomplete information in terms of cost and price fluctuations, variability in weather conditions, and changes in agricultural policy (Schmit and Rounsevell, 2006). Crop rotation is a well-established 'innovative' practice with many benefits (see, e.g., Plourde et al., 2013), but farmers do not necessarily adhere to strict rotations (Long et al., 2014). The decision to grow crops in a specific sequence is made within environmental, socio-economic, and agronomic constraints (Castellazzi et al., 2007).

Recent research (Long et al., 2014) identified several changes in the frequency of specific cropping sequences between 2001 and 2012 that are indicative of changing agricultural practices in northeast Montana. They concluded that regional producers: (1) increased the total amount of cropland in active production; (2) grew more cereal and reduced fallowed fields; (3) converted a substantial number of fields managed by strip-cropping to block management for cereal-based sequences; and (4) increased the prevalence of cereal-pulse sequences. Adoption of alternative cropping sequences, indicated by the last three findings, is innovative in that they represent new, and presumably more advantageous, practices. This previous work examined cropping sequences from a temporal perspective. Our research reported in this paper entailed the analysis of field-level changes in cropping sequences in a spatio-temporal context by focusing on two of the key changes – the adoption of cereal-pulse sequences and the

conversion from continuous strip-cropping to block managed cereal-based sequences. Our goal was to determine if the observed spatio-temporal patterns of the changing practices were consistent with those expected from specific *a priori* processes: random occurrence, spread and adoption of the practices due to social interaction as described in innovation diffusion theory, or adoption based on environmental factors.

The mapped locations of objects or events can be analyzed as spatial point processes in which the geographical locations are the random element; the actual observed locations represent one of the many possible realizations of a random spatial point process. Spatial point pattern analysis is a common method to assess spatial heterogeneity and detect patterns (e.g., Liu et al., 2007). The primary purpose of spatial point pattern analysis is to determine if the observed locations exhibit some form of systematic patterning over an area and, if so, over what spatial scales does the patterning persist. Spatial point patterns that exhibit complete spatial randomness (CSR) are indicative of a purely stochastic process and imply the absence of causal behavior. Spatial point patterns that do not exhibit CSR possess structure, suggesting spatial dependency and the potential existence of an underlying process, such as a response to environmental heterogeneity (Perry et al., 2002).

CSR assumes that the locations follow a homogeneous Poisson process and have constant intensity over the area, whereas departures from CSR produce regularity or clustering. Regularity exists in patterns that are more evenly distributed (dispersed) than expected, while clustering indicates that the locations are closer together than expected from a purely random process (e.g., Diggle, 2003). These patterns have been identified in

agricultural settings with a variety of established tests. Ripley's K-function has been used, for example, to detect clustering of agricultural areas affected by soil constraints (Dang et al., 2011) and in the yields from cacao stands (Ngo Bieng et al., 2013).

The adoption and spread of farming practices has long been described by innovation diffusion theory. The diffusion of innovation, in its basic form, refers to the spread and adoption of an idea, technology, or behavior (e.g., Wejnert, 2002; Scholnick, 2012). Foundational work (Tarde, 1903) suggested that sociological imitation followed a characteristic S-shaped curve in which the rate of diffusion could be deduced from the slope of the 'S'. Steep slopes indicated a rapid rate of diffusion, while more gradual slopes were indicative of slower rates. The adoption of agricultural innovations, such as new machinery, seed varieties, or management practices, has provided the context for many advances in the theory. It was demonstrated, for example, that the abundance of technological improvements to the sulky plow, determined by the number of patents, followed an S-shaped curve (Chapin, 1928). An investigation of the diffusion of hybrid seed corn – innovative because the new seed substantially improved yields and produced plants with stronger stalks, making mechanical pickers more effective, showed that adoption by Iowa farmers followed an S-shaped curve with three distinct phases: (1) an initial adoption period characterized by slow growth; (2) a period in which the rate of adoption increased rapidly; and (3) a period of declining growth as the resistant farmers adopted the seed (Ryan and Gross, 1943; Ryan and Gross, 1950). Perhaps their most influential claim was that diffusion was a social process in which spread occurred from

early to late adopters, rather than a process of decisions based on economics (Valente, 1995).

The Iowa study (Ryan and Gross, 1950) was framed in a predominantly temporal context and was geographically constrained, as it only considered diffusion in two specific farming regions in Iowa. An analysis of the adoption of various agricultural practices and the acceptance of government subsidies in Sweden more thoroughly considered the diffusion of innovations over space (Hägerstrand, 1953). This research concluded that the adoption of an innovation was a social process and spread outward from an initial cluster of adopters, with declining frequency of adoption as a function of distance and of physical barriers to communication, such as rivers or mountains(Hägerstrand, 1953; Johansen, 1971). Hägerstrand's model is, in many respects, similar to those used to describe the spread of a contagious disease (Clark, 1984). The model incorporates a set of six restrictive assumptions, primarily concerning the nature and means of communication, which diminishes its usefulness in modern applications. The Hägerstrand model, for example, assumes that face-to-face communication between farmers is the only mechanism by which knowledge of the innovation spreads (see, e.g., Hahvey, 1966 for a complete discussion). Later critics have re-analyzed Hägerstrand's data with more sophisticated statistical techniques, arguing that: (1) the initial cluster of adopters was not actually clustered (Cliff, 1968); (2) spatial autocorrelation was inadequately treated (Cliff and Ord, 1974); and (3) the model has little explanatory or predictive power if the socio-economic conditions are different from those experienced in Hägerstrand's study (Clark, 1984).

Models grounded in the idea of 'diffusion of innovation as a social process' persist in the literature, despite their shortcomings. The diminishing role of distance and physical barriers as a spatial constraint to the spread of information was recognized nearly 50 years ago (e.g., Hahvey, 1966). Furthermore, the social network of a twenty-first century farmer is unlikely to be well-characterized by the Euclidian distance between farms as assumed by Hägerstrand (Schmit and Rounsevell, 2006). Others (e.g., Beck, 1992) have argued that in a modern *information* society, traditional social networks are becoming less and less influential. The spread of information, and thus innovation, is no longer confined to a finite geographically-constrained social network of neighbors and colleagues, but includes all forms of modern communication and mass media. Information regarding agricultural innovations is available to all farmers in developed countries nearly simultaneously. Consequently, the strength of social interactions as a function of distance between farms is arguably a tenuous association at best. Research has demonstrated that fields cultivated by farmers in close proximity are statistically no more similar than fields cultivated by farmers who live further away from one another and has concluded that farmer imitation of their neighbors does not affect land use patterns directly (Schmit and Rounsevell, 2006).

Alternative viewpoints assert that the diffusion of an innovation over space is not strictly a function of social processes and include environmental context as a fundamental element. The adoption of an agricultural practice, in particular, depends on its suitability to the physical and economic environment (Wejnert, 2002). Management decisions are fundamentally subject to personal preferences, opinions, experiences, and the physical

characteristics of the landscape. Observed differences in agricultural practices are inherently affected by spatial heterogeneity in physical characteristics (Rosenzweig et al., 2002; Rounsevell et al., 2003). Crop yields, for example, are linked to temperature and precipitation (e.g., Rosenzweig et al., 2002; Lobell et al., 2007). Annual precipitation varies and has a substantial effect on the decision to grow a specific crop. Many regional producers were unable to seed fields in 2011, for example, due to excessive precipitation; therefore, many fields were unintentionally fallowed. Precipitation is well-known to vary with northing (latitude), easting (longitude), and elevation. Consequently, northing, easting, elevation, and precipitation are geographic and environmental variables that might be correlated with the adoption of cropping sequences.

The purpose of this work was to investigate changes in specific cropping sequences in northeast Montana within a spatio-temporal context. We focused on the adoption of cereal-pulse sequences and the conversion from continuous strip-cropping to block managed cereal-based sequences. These changes are indicative of shifting agricultural practices and suggest that regional farmers are adopting these 'innovations', because they provide some advantage over the traditional practices. Specifically, we sought to answer the following questions: (1) were the locations of the adopted practices randomly distributed throughout the study area; (2) did the spatio-temporal patterns of adoption resembled those expected under innovation diffusion theory driven by social factors; and (3) were the spatio-temporal patterns of adoption associated with specific environmental factors.

Material and Methods

Study Area

The study area is in northeastern Montana and includes Daniels, Sheridan, Roosevelt, and Valley Counties (Fig. 1). The region borders similar agricultural landscapes to the north and east, Saskatchewan and North Dakota respectively. Lake Fort Peck and the Missouri River form the southern boundary, while the western border consists largely of federal and state land. The region is characterized by low relief (Padbury et al., 2002) with the higher elevations in northern Valley County. Lower elevations are more common in the southeast portion of the study area, as well as along primary drainages such as the Missouri River, Poplar River, and Big Muddy Creek. The area has a semiarid climate with an average of 310 mm of annual precipitation, occurring primarily as rain between April and September (WRCC, 2011). Maximum temperatures in July average 31°C, while January temperatures average -10°C (WRCC, 2011). Shortgrass prairie and agriculture are the dominant land cover types. Dryland farming practices are dominant in the region (Tanaka et al., 2010), although center-pivot irrigation is not uncommon for producers in close proximity to the Missouri River. Agriculture consists largely of cereal crops, primarily spring wheat, and an increasingly substantial area of pulse crops; however, small amounts of other crops are grown within the Missouri River corridor (NASS, 2012; NASS, 2013).

Data

We previously reported cropping sequences for 13,076 agricultural fields in the study area for 2001–2012, which were derived from Landsat imagery in conjunction with the Cropland Data Layer, cadastral data, ground reference data, and local producers' records (Long et al., 2014). Crops were identified by class: cereals (C), fallow (F), pulse crops (P), cereal-fallow strip-cropping (S), or other (O). Cereals included spring or winter wheat (*Triticum aestivum* L.), durum (*Triticum turgidum* L.), or barley (*Hordeum vulgare* L.); pulse crops included dry pea (*Pisum sativum* L.), lentil (*Lens culinaris* Medik.), or chickpea (*Cicer arietinum* L.). Fallow identified uncultivated agricultural land, while cereal-fallow strip-cropping referred to the regionally common practice of alternating 50 or 100 m wide strips of cereal and fallow within a field. Non-agricultural land and crops that were not cereals or pulses comprised the 'other' class. The temporal structure of the data was represented by a 12-character string (e.g., SCPCCPCCFCCC) for each of the 13,076 fields that described the classification of that particular field for each of the years 2001–2012.

These data were then used to identify 2-yr cropping sequences via a 2-yr moving window (2001–2002, 2002–2003, …, 2011–2012), and similarly structured windows to identify 3-yr sequences. A field with SCPCCPCCFCCC for its 12-yr sequence, for example, has eleven 2-yr sequences: SC, CP, PC, CC, CP, PC, CC, CF, FC, CC, and CC. This field also contains ten 3-yr sequences: SCP, CPC, PCC, CCP, CPC, PCC, CCF, CFC, FCC, and CCC. Here we focus only on the 2-yr sequences because the 3-yr sequences reached the same conclusions in our previous work (Long et al., 2014). We did

not examine all 2-yr sequences – only those indicative of the two key changes in agricultural practices: (1) the increase in cereal-pulse sequences, and (2) the conversion from continuous strip-cropping to block managed cereal-based sequences. The sequences of interest, therefore, were CP/PC and SC.

The geographic and environmental factors considered included: northing, easting, elevation, relative effective annual precipitation (REAP), and annual precipitation. These variables were extracted from existing, publicly available datasets. Spatial data consisted of UTM coordinates (northing and easting) for the centroid of each field. Field elevation was defined as the elevation at the field's centroid and was derived from the 1 arc-second (~30 m) National Elevation Dataset (Gesch et al., 2002). Annual precipitation data came from the precipitation distribution model PRISM (Parameter-elevation Regressions on Independent Slopes Model), which uses point data in conjunction with a variety of datasets to produce spatially-gridded estimates of climatological variables (Daly, 2006; Daly et al., 2002, 1994). We used annual precipitation products (PRISM Climate Group, 2014) at a 4-km grid cell resolution to create eleven 2-yr means corresponding to the 2-yr cropping sequences, such that the 2001–2002 cropping sequence, for example, used the mean of the 2001 and 2002 PRISM data. REAP is a gridded (10-m) estimate of available moisture, accounting for precipitation, slope, aspect, and soil properties, based on a 30-yr period (NRCS, 2014). Each field, therefore, had a single REAP value over the study's time frame, unlike the PRISM data. REAP, therefore, better represents the general tendency over time with respect to water availability, compared to precipitation, which

relates to water availability in a specific year. We used field centroids to determine the REAP as fields potentially could have more than one value.

Analyses

All analyses were performed in R, version 3.0.2 (R Core Team, 2013) and with the following packages: 'spatstat' (Baddeley and Turner, 2005), 'gmodels' (Warnes, 2013), 'HH' (Heiberger, 2014), 'MuMIn' (Barton, 2013), and 'effects' (Fox, 2003). Our first objective was to determine if the locations of fields using the adopted practices were randomly distributed throughout the study area. We assessed the spatial dependency among fields in each of the cropping sequences of interest by year to address this objective: (1) visually after applying a smoothing operation to show relative densities across the study area, and (2) quantitatively with point pattern analysis. The location of an agricultural field is not strictly a point, but an area. We used the centroid of each field as 'points' and justify this usage by noting that the centroids are point-like at the scale in this study (e.g., Baddeley, 2008).

We used kernel density estimators (KDE) to map the density $f(s)$ for each cropping sequence and year in order to visually assess the spatial distributions. KDEs are common nonparametric density estimators that estimate the distribution of $f(s)$ from a finite set of points by smoothing. They are weighting functions that smooth the contribution of each observed data point over a local neighborhood such that the weights decrease with distance. The intensity function depends on kernel type and bandwidth; kernel type determines the shape of the weighting function and bandwidth determines the amount of smoothing (e.g., Waller and Gotway, 2004). The general consensus is that the

choice of kernel is largely irrelevant, while the choice of bandwidth is crucial (e.g., Taylor, 1989). We experimented with a variety of kernels and found no discernible differences before settling on a Gaussian kernel. Intensity was, however, responsive to bandwidth. We chose bandwidths that maximized the point process likelihood cross-validation criterion (LCV) (see e.g., Loader, 1999 for a complete discussion). If $s_1, s_2, \ldots$ are locations in the study area, estimated intensity at a given point *s* is given by

$$f(s) = \frac{1}{nh} \sum_{i=1}^{n} \left(\frac{s - s_i}{h} \right) \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{(s-s_i)^2}{h^2}} \right) \quad (1)$$

where *h* is the bandwidth, *n* is the number of locations, $s - s_i$ is the Euclidean distance between *s* and s_i, and the last term on the right hand right is the two-dimensional Gaussian kernel (Cai et al., 2013).

Spatial randomness was assessed quantitatively with the L-function, a normalized version of Ripley's K-function (Besag, 1977; Ripley, 1977). The K-function evaluates the randomness of a point pattern across a spatial domain at a specified distance *r*, by comparing the expected number of points within a local neighborhood of radius *r* of any point against the expected intensity λ, assuming CSR (Ripley, 1977). The discrete form of the K-function (e.g., Streib and Davis, 2011) is defined as

$$\hat{K}(r) = \frac{1}{\lambda N} \sum_{i=1}^{N} \sum_{j=1, i \neq j}^{N} I_r(d_{ij}) \quad (2)$$

where d_{ij} is the Euclidean distance between locations s_i and s_j, λ is the intensity function (total number of points divided by area) and I_r is an indicator function which takes the value of 1 if $d_{ij} \leq r$. Under CSR, the expected value of $K(r)$ is πr^2 and deviations from this value are indicative of clustering ($K(r) > \pi r^2$) or dispersion ($K(r) < \pi r^2$). The L-function is a normalization of the K-function that gives a more interpretable, linear expected value, and is defined as

$$\hat{L}(r) = \sqrt{\hat{K}(r)/\pi} \tag{3}$$

The L-function assumes a homogeneous spatial process (stationary and isotropic) across the entire domain, which is problematic because this is unlikely in many applications (Gabriel and Diggle, 2009). Points near the border of the domain are also problematic because the local neighborhood of radius *r* may lie partially outside of the domain. Consequently, we used the inhomogenous version of the L-function, a generalization applicable to non-stationary point processes (Baddeley et al., 2000), and border-corrected estimates for $K(r)$ to adjust for bias introduced by edge effects (see e.g., Ripley, 1979, 1977). We tested for departures from CSR by creating confidence envelopes for CSR based on Monte Carlo simulations ($n = 500$). We also used chi-squared tests of goodness-of-fit for point process models based on quadrat counts, which determines whether a point pattern is more clustered than would be expected by chance by comparing the observed number of cropping sequences in quadrats of a given size with the frequencies expected from a random process. If O_i is the observed number of

fields in the *i*th quadrat, E_i is the expected number of fields in the *i*th quadrat under CSR, and *k* is the total number quadrats, then the test statistic

$$X^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \quad (4)$$

is asymptotically distributed as a chi-squared distribution with $k-2$ degrees of freedom under the null hypothesis of CSR.

Our second objective was to establish whether or not spatio-temporal patterns of adoption resembled those that would be expected under innovation diffusion theory driven by social factors. Spatial dependency among fields in each of the cropping sequences suggests the possible existence of an underlying process. If the spread of adoption for these cropping sequences was a social process that spread outward from an initial cluster of adopters, as suggested in much of the diffusion of innovation literature, then we would expect clustering around the original adopters, and more generally clustering in any year around adopters in the previous year. The assumptions of the Hägerstrand model and its variants were not met in our data; therefore, we abandoned this method. Instead, we tested for spatio-temporal clustering around the initial locations by considering a general contagion processes. A contagion process is a spatial point process in which the occurrence of an event raises the probability of observing subsequent events nearby (e.g., Waller and Gotway, 2004) – for example, if the locations of fields in a cropping sequence in the first year were geographically closer to fields in the same cropping sequence in the second year than they were to a randomly selected set of field locations from the second year.

For each field in a given cropping sequence at time t_{n+1}, we computed the Euclidean distance between it and the nearest field managed with the same cropping sequence at time t_n. From these values we found the average minimum distance, equivalent to the average nearest neighbor distance, between fields in the same cropping sequence during consecutive years. These were compared to the equivalent metric from Monte Carlo simulations (n = 100) of randomly selected fields. Let A be the set of field centroids for a particular cropping sequence at time t_n such that $A = \{a_1, a_2, \ldots$. Similarly, let $B = \{b_1, b_2, \ldots$ be the set of centroids for fields at time t_{n+1} that are in the same cropping sequence as those in A, and let $C = \{c_1, c_2, \ldots$ be a set of j randomly selected field centroids. We define $\mathbf{D}_{\min}(\mathbf{a}, \mathbf{b})$ as the vector of minimum Euclidean distances between each element of A and each element of B, and $\mathbf{D}_{\min}(\mathbf{a}, \mathbf{c})$ as the vector of minimum Euclidean distances between each element of A and each element of C. A contagion process is suggested when

$$\mathbf{D}_{\min}(\mathbf{a}, \mathbf{b}) < \mathbf{D}_{\min}(\mathbf{a}, \mathbf{c}) \tag{5}$$

We fit a linear model for each cropping sequence with distances as the responses modeled as a function of time (years 2001–2002 / 2002-2003 to 2010– 2011 / 2011–2012) and type (observed or random) and their interaction. Because of severe non-constant variance and skew in the residuals for this model, all analyses were performed on the log-distances, which alleviated concerns about violations of normality and constant variance assumptions. Contrasts were used to test for differences in the mean of the log-distances between types for each year.

Our final objective was to determine if the spatial patterns of cropping sequences were associated with certain environmental factors. We examined potential associations between cropping sequences and the environmental variables (easting, northing, elevation, REAP, and mean annual precipitation) with graphical explorations of the data. Then, we fit two suites of logistic regression models with cropping sequence as the binary response variable (CP/PC or not CP/PC; SC or not SC). All possible models with the predictor variables and an easting-northing interaction were considered; the top-ranked models based on AICc are reported. Here, the log-odds of adopting the practice (CP/PC or SC) relative to a baseline of not adopting the practice were modeled as linear combinations of the environmental predictor variables for each cropping sequence. The baseline model gives the log-odds of a success,

$$\log\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots \qquad (6)$$

where $\hat{\pi}$ is the odds of adopting the practice, $1-\hat{\pi}$ is the odds of not adopting the practice, β_0 is the model intercept, and $\beta_1, \ldots$ are the coefficients for the $X_1, \ldots$ predictor variables.

Results

Point Pattern Analysis

We used point pattern analysis to assess the spatial dependency among fields by cropping sequence and year to determine if the locations were randomly distributed throughout the study area. None of the cropping sequences were randomly distributed during any of the 2-yr periods. This was evident visually in the density maps (Figs. 2 and

3) and statistically from both the analysis of quadrat counts and the border-corrected inhomogeneous L-functions. Assessment of CSR based on quadrat counts consistently indicated clustering for both cropping sequences during 2001–2012 ($p < 0.01$ in all cases). These results were robust to quadrat size and clustering was indicated with quadrats as large as 25 x 25 km and as small as 2.5 x 2.5 km. Assessment of CSR based on L-functions also indicated clustering for both sequences at similar ranges, typically 10–20 km (Table 1).

Adoption of Cropping Sequences as a Contagion Process

Spatio-temporal patterns of cropping sequence adoption driven by social factors suggests that fields in the adopted practice in a given year should be closer on average to locations that previously adopted the practice than would be expected from random locations. This was the case for CP/PC sequences from 2001–2004 (Fig. 4a); they were farther than expected, however, for the remainder of the study. There was strong evidence of different changes over time between the mean minimum log-distance between fields in CP/PC cropping sequences and the mean minimum log-distance between fields under random field selection (F (9, 16820) = 3.9, $p < 0.0001$). The main effects in the model with an interaction also showed strong evidence of differences (Type: $F(1, 16820) = 59.4$, $p < 0.0001$; Time periods: $F(9, 16820) = 180.2$, $p < 0.0001$). Pairwise contrasts indicated that the observed and random fields were significantly different in six of the ten time periods (Fig. 4a).

We also found that SC fields were closer on average to locations that previously adopted the practice than expected from random locations for all time periods except one (Fig. 4b). There was evidence of changes over time between the mean minimum log-distance between fields in SC sequences and the mean minimum log-distance between fields under random field selection (F (9, 4623) = 2.15, p = 0.0023). The main effects in the model with an interaction showed strong evidence of differences (Type: F(1, 4623) = 10.48, p = 0.0001; Time periods: F(9, 4623) = 7.27, p <0.0001). Pairwise contrasts indicated that the observed and random fields were significantly different in only two of the ten time periods (Fig. 4a).

Adoption of Cropping Sequences as a Function of the Environment

Inferences are limited to the observations at hand, because the entire population of interest was available in the dataset for each year. The preliminary exploration of the data suggested that fields in CP/PC sequences were potentially associated with locations that: (1) were farther north and east; (2) were at lower elevations; and (3) had higher REAP. The adoption of block-managed cereal-based sequences, in lieu of continuous strip-cropping, showed the same tendencies. The northing and easting trends are visually evident in the density plots (Figs. 2 and 3), particularly for the CP/PC sequences.

Logistic regression models were initially fit with all environmental variables and an interaction between easting and northing for each cropping sequence by time period; final models were selected based on AICc (Table 2). The logistic models were fit such that the baseline category was fields that had not adopted the cropping sequence;

therefore, positive coefficients indicated variables that were associated with an increase in the likelihood that a field used the adopted practice and negative coefficients indicated variables that decreased the likelihood. We focused primarily on the direction of the association for the selected models, because we were more interested in identifying associations rather than predictive modeling.

The environmental variables that were associated with an increased likelihood that a field had adopted CP/PC sequences varied temporally (Table 2). Positive coefficients for easting and northing in the first half of the study gave way to negative ones by 2007–2008. The coefficient of the easting-northing interaction was always opposite in sign from those of the main effects and therefore served to mediate the main effects. The effect of northing, because of the interaction, ranged from positive to negative depending on the easting value and the time frame; it was generally positive with the exception of the most western eastings from 2007–2008 onwards. The coefficients for elevation were negative early on, but changed to positive in 2005. Coefficients for precipitation and REAP were typically negative, when present in the final model, with few exceptions. This seemingly incongruous result is examined in the Discussion. When interpreting the coefficients in Table 2, we note that we used UTM coordinates for northing and easting; the units are in meters.

The environmental variables associated with an increased likelihood that a field had converted from strip-cropping to block management also varied temporally (Table 2). Northing and easting, when retained in the model, tended to be negative during 2001–2007 and positive afterwards. The easting-northing interaction was generally not retained

prior to 2007, but had a negative sign in the succeeding years. Coefficients for elevation had mixed directions of estimated effects in the models, but they tended to be positive in the first half of the study. The coefficients for REAP and annual precipitation were typically negative, but were seldom selected in the final models.

Discussion

Key Indicators of Changing Agricultural Practices

Agricultural practices in northeast Montana are changing with respect to the management of cropping sequences. Two practices in particular have emerged as key indicators of changing agricultural practices in the region. First, producers have adopted cereal-pulse sequences at a substantial rate – a six-fold increase within a decade. Second, producers have simultaneously adopted block-managed cereal-based sequences in lieu of continuous strip-cropping such that the number of fields managed by continuous strip-cropping decreased four-fold (Long et al., 2014). Producers that converted from strip-cropping to cereal-based sequences adopted continuous cereal cropping sequences 49% of the time on average, block-managed cereal-fallow sequences 35% of the time, while cereal-pulse sequences were adopted 16% of the time (see Table 3 for details). We examined the spatial distribution of the adoption of these practices to see if they were non-random and, if so, were they consistent with patterns expected from the spread and adoption of the practices due to social interaction or adoption based on environmental factors.

The Adoption of Cereal-Pulse Cropping Sequences

The locations of fields that adopted cereal-pulse sequences were not randomly distributed at any time during the study, which suggests an underlying rationale for the adoption. All tests of spatial randomness indicated that these sequences were clustered on the landscape within a local neighborhood, typically 15–20 km. This is visually evident in the density plots (Fig. 2), where we can distinguish regions with higher concentrations of CP/PC sequences, particularly for the early portion of the study. We caution the reader that the density plots are not scaled equally – i.e., the colors do not reflect the same density across all plots.

We tested whether the observed patterns of adoption tended to cluster geographically around the fields that had originally adopted the practice. We defined 'original' to be those fields using the practice in 2001–2002, the first 2-yr sequence in our study. We see in Fig. 2 that the original adopters were in the eastern portion of the region, primarily in the vicinity of Medicine Lake. Clustering around the original adopters would suggest a contagion process – one in which the occurrence of a field in a CP/PC sequence raises the probability of observing nearby fields in CP/PC sequences in subsequent years. There was some evidence that the adoption of cereal-pulse sequences in the region initially spread through a contagion process as fields in successive years were closer together than expected from a random process during 2001–2004; however, these differences were small, less than 600 m, and only showed strong evidence of a difference once. The results for 2005–2012, on the other hand, indicated that successive CP/PC sequences were farther apart than would be expected from a random process; there was

evidence of differences for five of the seven time periods. Collectively, we take these results to suggest that a weak contagion process might have operated during the early stages of cereal-pulse adoption, potentially driven by social processes, and within a radius of no more than 20 km. Several additional clusters formed during 2005–2012 at distances farther than the local clustering neighborhoods that arose via nucleation (e.g., Yarranton and Morrison, 1974), and not contagion processes.

Adoption of a cropping sequence depends on its suitability to the physical environment; therefore, we looked at the association between a suite of environmental variables and whether or not a field had adopted CP/PC cropping sequences. Modeling of the environmental variables (northing, easting, elevation, REAP, and annual precipitation) generally indicated that geography was strongly associated with the likelihood that a field had adopted CP/PC sequences. The collective effect of easting, northing, and their interaction over time suggest that the adoption of CP/PC sequences grew primarily in the northeastern portion of the study area between 2001 and 2007, and then expanded toward the southwest.

The models indicate that fields in CP/PC sequences were initially associated with lower elevations, with the exception of the Missouri River corridor, but had begun to move upslope by 2005–2006. Some of the lowest elevations in the study area are in the vicinity of Medicine Lake – the location of the original clusters. The upslope expansion in the latter portion of the study was congruent with results from easting, as the terrain generally increases in elevation from east to west.

REAP was retained in the models only occasionally and, when present, had negative coefficients; precipitation was likewise negative. Thus, locations with higher REAP or precipitation were less likely to have adopted CP/PC sequences, after controlling for easting and northing effects in the models. These findings contradict the exploratory data analysis, which indicated that adoption of these sequences was associated with locations with high REAP and precipitation. REAP, by definition, does not vary over time and it has a conspicuous northeast-southwest trend such that higher values are in the northeast. REAP also had a very strong positive relationship with easting ($r = \sim 0.79$ depending on time frame) and a strong positive relationship with northing ($r = \sim 0.57$). Similarly, annual precipitation had strong positive relationships with easting ($r = 0.39–0.89$ depending on time frame). This suggests that multi-collinearity was present in the predictor variables in these models and that all interpretations must be made with caution because of this shared information.

The importance of the geographic variables is requires interpretation, as geographic position does not inherently explain anything ecological, but functions as a surrogate for other environmental variables (e.g., Pereira and Fonseca, 2003). We conclude that the strong effect of geographic position is actually driven by water availability. This suggests that the adoption of the CP/PC sequences began in locations with high REAP values, typically greater than 330 mm, and gradually spread to locations with REAP values as low as 270 mm. The same general pattern of spread from higher to lower water availability was true for annual precipitation. The environmental variables, collectively, implied that the adoption of CP/PC sequences were initially located in the

eastern portion of the study area and expanded, first to the northwest and then towards the southwest (see also Fig. 2). This relationship is consistent with a pattern in which adoption occurred in more favorable locations initially and expanded into marginal environments.

The more favorable locations for the adoption of CP/PC sequences were those with more available water, i.e., locations with higher annual precipitation and higher REAP. We see this association with water play out as the original adopters were in locations with higher REAP (north and east) and expansion was progressively into locations with lower REAP. Fields using center-pivot irrigation to grow pulse crops were not common in the study region during 2001–2012; however, they were occasionally present in the west (Valley County) where REAP and annual precipitation were lowest. The evidence suggests that the adoption of CP/PC sequences possibly expanded via a contagion process during 2001–2003, but this is inconclusive. On the other hand, expansion via nucleation during 2004–2012 was evident as several new centers for the adoption of the practice arose nearly simultaneously across the study area. This expansion is likely related to the Farm Security and Rural Investment Act of 2002 (i.e., the 2002 ‘Farm Bill’), which provided meaningful protection in the form of marketing loans and payments if market prices fall too low for producers of pulse crops (Hauser, 2002).

The Adoption of Block-Managed Cereal-Based Sequences in Lieu of Strip-Cropping

Fields that abandoned continuous strip-cropping and adopted block-managed cereal-based sequences were, like the CP/PC sequences, not randomly distributed at any time. Tests for spatial randomness indicated clustering within local neighborhoods of 5–17 km, approximately one-third smaller than the neighborhoods for the CP/PC sequences. This is visually evident in the density plots (Fig. 3), where regions with higher concentrations of SC sequences are readily visible. We caution the reader again with respect to scaling in the density plots.

As before, we tested whether the observed patterns of adoption of block-managed cereal-based sequences tended to cluster around the fields that had adopted the practice. This was true for the majority of the study and there was evidence of differences in two of the ten comparisons (see Fig. 4b). We see in Fig. 3 that the original adopters were primarily located in two clusters: one near the center of the study region and another just north of Medicine Lake. The cluster north of Medicine Lake was coincident with an original cluster of CP/PC adoption. These data suggest that adoption of block-managed cereal-based sequences might have propagated as a contagion process; however, the distances between fields were not consistently significant. We concluded that a weak socially-driven contagion process for the conversion from continuous strip-cropping to block-managed cereal-based sequences potentially operated throughout the duration of the study. New clusters formed at distances farther than the local clustering neighborhoods, not necessarily via contagion processes.

Modeling the environmental variables indicated that geographic position was associated with the likelihood that a field had adopted block-managed sequences in lieu of continuous strip cropping, but not as strongly nor as consistently as in the adoption of CP/PC sequences. Locations farther east and north were more likely to have adopted block-management. The northern and eastern locations were more likely to have adopted either continuous cropping of cereals or cereal-pulse sequences, while locations farther south and west favored the adoption of block-managed cereal-fallow sequences. Elevation generally indicated that the adoption of block-managed cereal-based sequences, like the adoption of CP/PC sequences, tended to move upslope – toward the west. REAP and annual precipitation were rarely retained in the models, but were strongly correlated with easting and northing as we found in the CP/PC sequences. Therefore, we believe it is likely that, similar to the SC sequences, geographic position acted as a proxy for water. Together, the environmental variables suggest that adoption of block-managed cereal-based sequences initially occurred in the central and eastern portions of the study area and expanded first to the north and then toward the southwest (see also Fig. 3).

The principal driving factor for the adoption of block-managed cereal-based sequences, or alternatively the abandonment of continuous strip-cropping, during 2001–2012 was water. These conversions happened in locations with consistently higher levels of moisture, while the decision to continue managing fields via strip-cropping was confined to locations where the availability of water was less consistent. This pattern is logical given that strip-cropping entered widespread use in regions where rainfall is a limiting resource as a means to better manage soil water. While environmental variables

explained the locations of conversions, we cannot rule out the possibility that the decision to adopt block-management or to continue strip-cropping also occurred within a social context. Adoption of block-management, in regions where it is feasible, seemed to happen more frequently than not amongst neighbors within 20 km of one another.

Conclusion

We investigated changes in two common cropping sequences in northeast Montana: (1) adoption of cereal-pulse cropping sequences; and (2) adoption of block managed cereal-based sequences in lieu of continuous strip-cropping. These sequences are key indicators of shifting agricultural practices in the region. Dryland farming has a history of developing and adopting management practices to capture and conserve rainfall, and to better manage soil moisture. Water efficiency is essential in the semiarid portions of the Northern Great Plains and, not surprisingly, water exerts a fundamental control on the decision whether or not to adopt either of these practices. Traditional cereal-fallow sequences are inefficient and producers in northeast Montana are moving away from strip-cropping, a form of cereal-fallow sequences, to block-management of cereal-based sequences, including those with pulse crops in locations where the environment is suitable. It is possible that the expansion of these practices was due, in part, to social factors, particularly during the early periods of adoption. This suggests a process in which the ‘neighbors’ of the original adopters emulated nearby practices.

Acknowledgements

This manuscript is based upon work supported by the US Department of Energy and the National Energy Technology Laboratory under Award Number DE-FC26-05NT42587. It was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Table 4.1. Maximum clustering range by cropping sequence and time period [a].

	Time Period										
	2001-2002	2002-2003	2003-2004	2004-2005	2005-2006	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011	2011-2012
CP/PC	3.0	18.5	20.6	20.5	25.5	14.7	15.4	15.8	14.9	17.4	16.8
SC	11.9	9.8	10.3	16.8	11.3	12.5	16.1	12.0	17.1	5.0	5.2

[a] Approximate maximum range (km) for clustering based on the 95% confidence envelope from a border-corrected inhomogeneous L-Function. CP/PC = cereal-pulse, SC = conversion from strip-cropping to cereal-based sequences.

Table 4.2. Coefficients for the top-ranked logistic regression models [a].

	Time Period	*df*	*n*	Easting	Northing	Easting:Northing	Elevation	REAP	Precipitation
CP/PC	2001–2002	5	6338	**7.05e-04**	**6.41e-05**	**-1.29e-10**	**-4.24e-03**	---	---
	2002–2003	4	6445	**7.64e-04**	**7.53e-05**	**-1.39e-10**	---	---	---
	2003–2004	7	6097	**7.94e-04**	**7.99e-05**	**-1.46e-10**	**-3.93e-03**	-7.60e-02	**-1.79e-02**
	2004–2005	6	5681	**5.51e-04**	**5.72e-05**	**-1.01e-10**	---	-5.78e-02	**-1.09e-02**
	2005–2006	6	5763	**5.43e-04**	**5.93e-05**	**-9.61e-11**	**4.50e-03**	---	**-6.44e-03**
	2006–2007	6	6044	2.95e-04	**3.76e-05**	-5.02e-11	**3.19e-03**	---	**-8.23e-03**
	2007–2008	7	7037	**-8.34e-04**	**-6.00e-05**	**1.57e-10**	**4.87e-03**	**-1.56e-01**	**8.60e-03**
	2008–2009	7	7509	**-1.33e-03**	**-1.05e-04**	**2.49e-10**	**4.95e-03**	**-1.13e-01**	-4.77e-03
	2009–2010	6	7659	**-1.41e-03**	**-1.13e-04**	**2.65e-10**	**3.41e-03**	---	**-1.26e-02**
	2010–2011	6	7516	**-1.52e-03**	**-1.19e-04**	**2.85e-10**	**3.27e-03**	---	**-1.58e-02**
	2011–2012	6	7542	**-1.07e-03**	**-8.12e-05**	**2.01e-10**	**2.36e-03**	-6.64e-02	---
SC	2001–2002	6	6338	-8.28e-04	**-8.49e-05**	1.55e-10	**7.40e-03**	---	-2.29e-02
	2002–2003	4	6445	3.58e-06	---	---	**3.19e-03**	---	-1.92e-02
	2003–2004	3	6097	---	-7.19e-06	---	1.65e-03	---	---
	2004–2005	3	5681	---	---	---	1.43e-03	---	8.71e-03
	2005–2006	3	5763	-2.46e-06	---	---	3.33e-03	---	---
	2006–2007	4	6044	-7.04e-06	-6.36e-06	---	1.50e-03	---	---
	2007–2008	5	7037	**5.81e-04**	5.19e-05	**-1.09e-10**	-2.94e-03	---	---
	2008–2009	5	7509	**1.27e-03**	**1.14e-04**	**-2.37e-10**	**-5.01e-03**	---	---
	2009–2010	6	7659	**9.48e-04**	**9.15e-05**	**-1.76e-10**	---	**-2.20e-01**	**1.26e-02**
	2010–2011	5	7516	**1.09e-03**	**9.52e-05**	**-2.01e-10**	---	**-2.03e-01**	---
	2011–2012	4	7542	**8.62e-04**	**6.86e-05**	**-1.61e-10**	---	---	---

[a] Coefficients for the top-ranked logistic regression models by sequence and time period. Coefficients with *p*-values < 0.05 are in bold; *df* = degrees of freedom, *n* = sample size.

Table 4.3. Conversions from continuous strip-cropping to cereal-based sequences as a percentage of all conversions [a].

	Time Period									
	2001-2003	2002-2004	2003-2005	2004-2006	2005-2007	2006-2008	2007-2009	2008-2010	2009-2011	2010-2012
SCC	65.4	72.2	35.4	31.1	36.4	55.7	60.0	54.1	50.0	28.3
SCF/SFC	25.3	23.1	45.1	32.9	36.7	29.3	27.7	32.9	36.4	57.1
SCP/SPC	9.3	4.7	19.5	36.0	26.9	15.0	12.3	13.0	13.6	14.6

[a] The percentage of conversions from strip-cropping to cereal-based rotations by the type of rotation. SCC = conversion to continuous cereal, SCF/SFC = conversion to cereal-fallow, SCP/SPC = conversion to cereal-pulse.

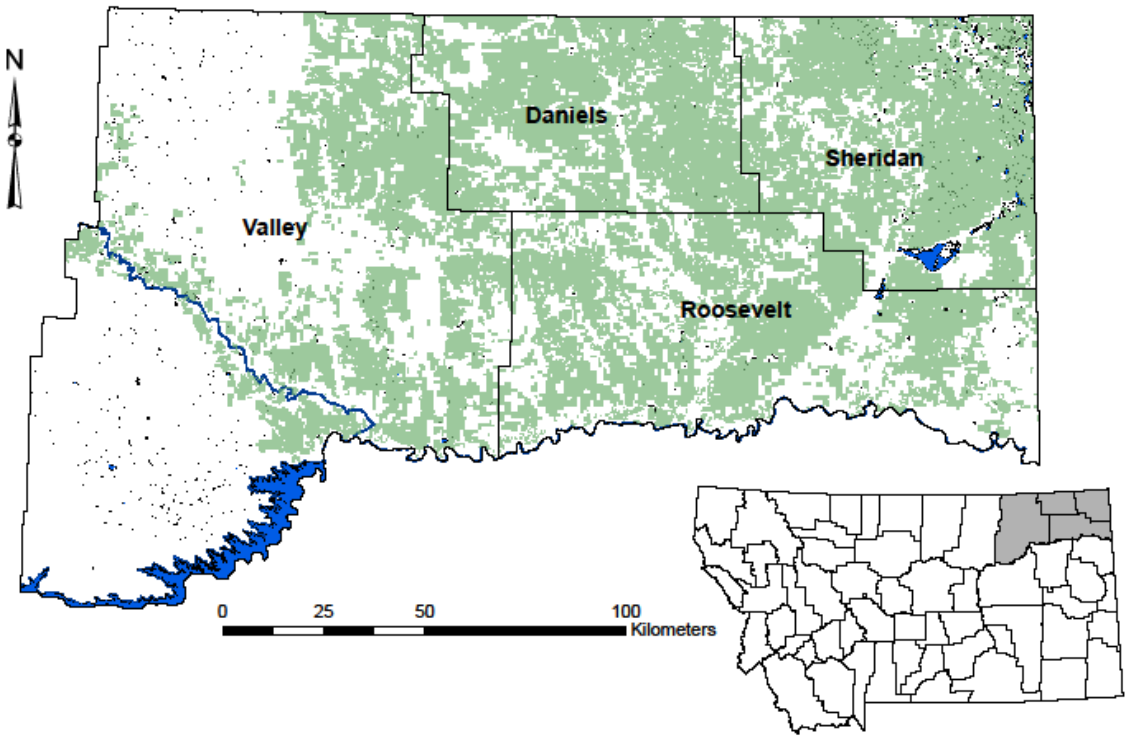

Figure 4.1. The study area in northeastern Montana included the counties of Daniels, Roosevelt, Sheridan and Valley. The agricultural regions are green.

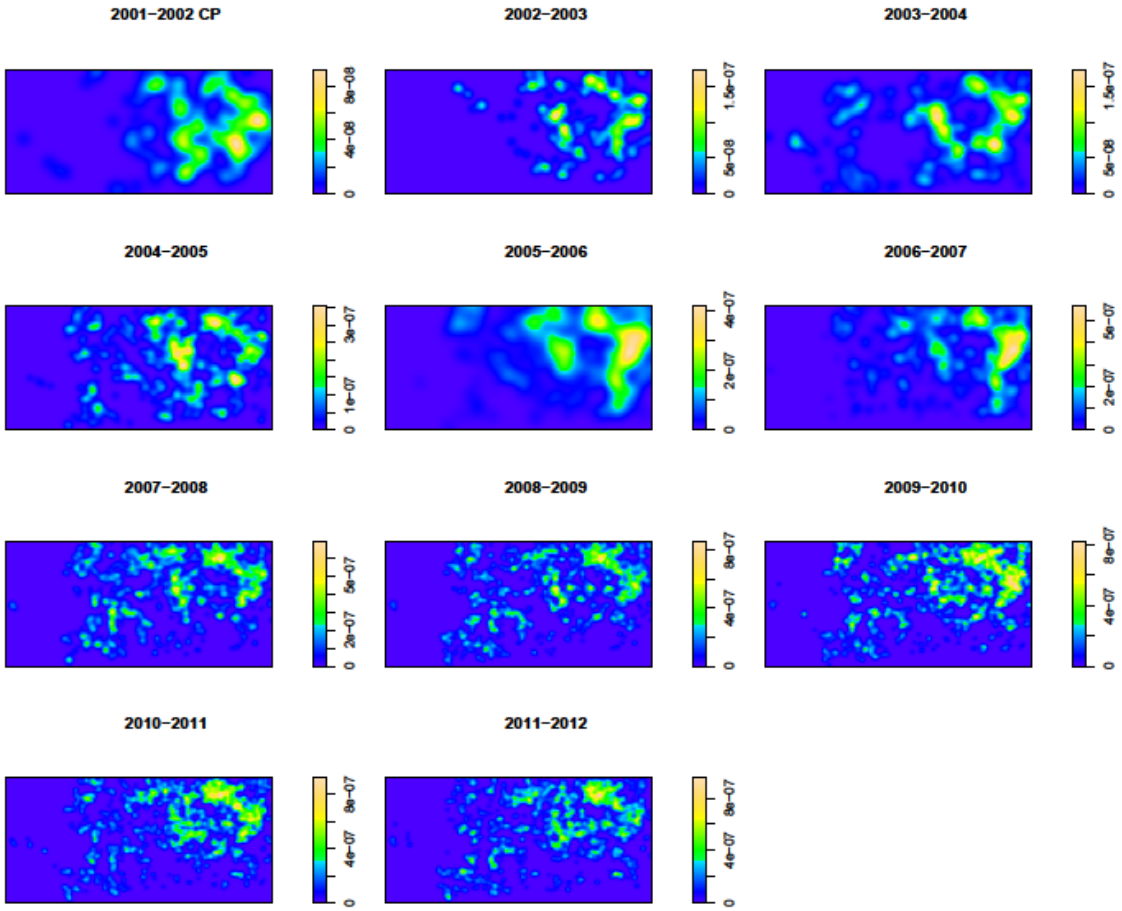

Figure 4.2. Kernel density maps for the locations of fields that had adopted CP/PC sequences by year. Note that the probability scale next to each map is scaled independently for each year and cannot be used to compare maps from year to year.

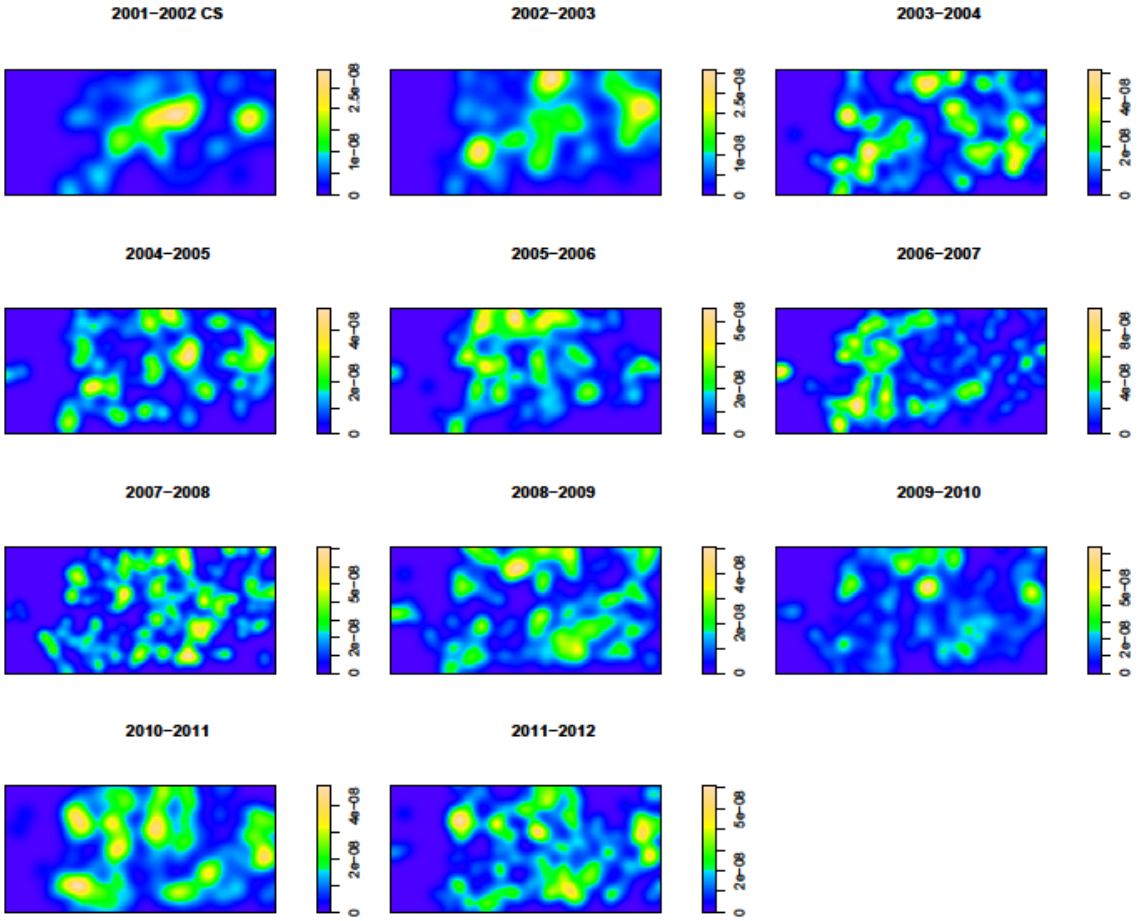

Figure 4.3. Kernel density maps for the locations of fields that had adopted block-managed cereal-based sequences in lieu of continuous strip-cropping by year. Note that the probability scale next to each map is scaled independently for each year and cannot be used to compare maps from year to year.

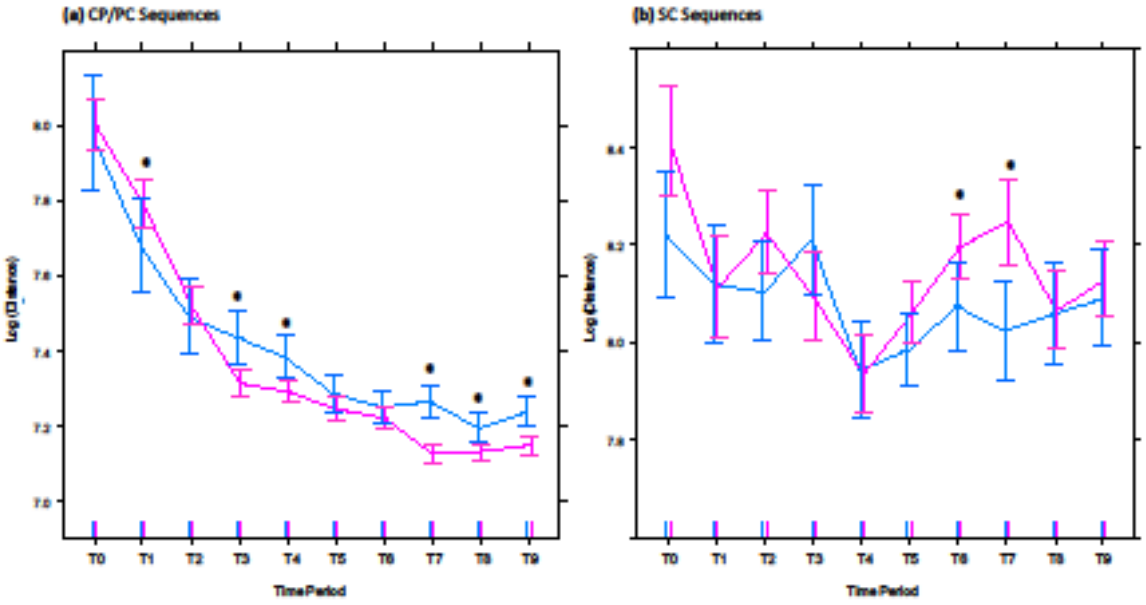

Figure 4.4. Plots of the mean minimum log-distance between fields (+/- 1 SE) in (a) CP/PC sequences and (b) SC sequences, and the mean minimum log-distance between fields under random field selection. Blue lines are between observed sequences and pink lines are between observed and random selection. Time periods with asterisks indicate time periods where the difference in mean log-distances between "Type" is detectably different (p-values < 0.05). Time periods represent the comparison of two 2-yr sequences: T0 = 2001–2002 / 2002–2003; T1 = 2002–2003 / 2003–2004; T2 = 2003–2004 / 2004–2005; T3 = 2004–2005 / 2005–2006; T4 = 2005–2006 / 2006–2007; T5 = 2006–2007 / 2007–2008; T6 = 2007–2008 / 2008–2009; T7 = 2008–2009 / 2009–2010; T8 = 2009–2010 / 2010–2011; T9 = 2010–2011 / 2011–2012.

CHAPTER 5

CONCLUSION

Changing agricultural practices in northeast Montana during 2001–2012 presented the opportunity to study the adoption of cropping sequences from temporal and spatial perspectives using geospatial techniques. My purpose in this dissertation was threefold: (1) develop a remote sensing technique to classify regionally relevant crops at the field-level from satellite imagery with missing data; (2) identify the principal 2- and 3-yr cropping sequences and the changes in their relative proportions; and (3) characterize and explain the spatial patterns for the adoption of cropping sequences over time. Identification of the prevailing cropping sequences in a region can establish which ones have been successful and, therefore, which sequences might be more likely to succeed in similar regions. The adoption of cereal-pulse sequences and the adoption of block-managed cereal-based sequences in lieu of continuous strip-cropping were particularly important sequences to providing an insight into the long-term sustainability of regional agriculture.

Satellite imagery provides a convenient means to study large areas such the nearly 40,000 km^2 in the study area. The Landsat series of satellites has given researchers a 40+ year archive of multispectral, moderate-resolution observations of the Earth; however, the 19 months between November 2011 and May 2013 was a time in which Landsat 7 was the sole operational Landsat satellite. The onboard sensor, ETM+, experienced an unfortunate permanent SLC failure in May 2003, leaving SLC-off ETM+ imagery with

gaps in coverage of approximately 22% per image. Therefore, the first task was to devise a means to classify crops despite missing pixels in the satellite imagery that coincided with my ground reference data.

The classification of regionally dominant crops by class (cereal, pulse, etc.) using SLC-off ETM+ imagery from the 2012 growing season was successfully accomplished by using a multitemporal, object-oriented approach. Two images, a mid- and a late-season one, were used to capture phenological variations. The object-oriented approach allowed agricultural fields to be defined as objects by delineating them with regional cadastral data. Spectral data were then extracted from the extant pixels in an object and used in a random forest classification to identify the class of crop in the fields under the assumptions that the extant pixels were representative of the entire field. The results indicated that the method was able to classify fields with an overall accuracy of 85.5% and with a κ statistic of 0.75.

Several studies have suggested that object-oriented approaches yield better classification accuracies for crops than classifications using traditional pixel-based approaches. I tested this claim and found that the pixel-based approach yielded overall classifications that were generally 5% better; however, the better classification accuracies of the pixel-based approach are obviated by its inability to classify pixels with missing data. Multitemporal models have been reported to produce better classification accuracies than single-date models in agricultural settings due to their ability to capture phenological variations. This study supported this claim, as multitemporal classification accuracies were 5–7% higher than any single date image could manage.

The importance of this section of the dissertation is not specifically related to agriculture. Complications arising from missing data are not specific to ETM+ SLC-off imagery – an image from any sensor might be missing data due to a variety of causes such as sensor noise, clouds, or shadows. Consequently, the devised method is applicable to any situation in which the landscape can be delineated into meaningful homogeneous areas that have missing data.

Once classification was possible, the focus turned to my primary goals, characterizing the temporal and spatial patterns. The original impetus for my work stemmed from the substantial growth of cereal-pulse cropping sequences in the region; however, it became apparent in the second phase of the dissertation that two other practices were changing as well: (1) reductions in fallow and (2) conversions from continuous strip-cropping of cereals to block-managed cereal-based sequences. Collectively, these three agricultural practices were the focus of the remainder of the dissertation. Defining the temporal sequence of crops in a given field required classifying 13,076 fields for each year during 2001–2012, which I was able to do using multitemporal Landsat imagery in conjunction with the Cropland Data Layer, cadastral data, ground reference data, and local producers' records. The 12-yr temporal sequences were searched for specific 2- and 3-yr crop sequences with a string-matching algorithm.

Historical practice in dryland farming regions has been to rotate cereal with fallow, primarily as a means to better manage soil water. These rotations are either managed in blocks or by strip-cropping where wind erosion was a concern. These practices are beneficial because of their simplicity; but, they are less resilient in the long-

term than a more diverse system. Cereal monocultures, for example, are more susceptible to disease and insect damage, while pulse crops provide a variety of beneficial field-level effects. The results from this phase of my dissertation suggest that historical practice no longer dominates the region. Several changes in agricultural practices have happened, somewhat simultaneously.

The amount of cropland in active production, for example, increased as did the overall production of cereals. Both of these increases were closely associated with equivalent decreases in fallow made possible due to improvements in soil water management through the adoption of practices such as no-till. The main findings were the characterization of the temporal dimension for the substantial increase in the adoption of cereal-pulse sequences and the discovery that producers were adopting block-managed cereal-based sequences in lieu of continuous strip-cropping. There were two other important findings: (1) producers did not strongly adhere to specified sequences, but likely responded to soil moisture levels to decide whether to plant a crop or to fallow the field and (2) using 3-yr sequences added no additional information than using 2-yr sequences, the temporal characteristics were the same.

The spatial analysis of the broad-scale adoption of cereal-pulse sequences and the conversion from continuous strip-cropping to block managed cereal-based sequences provided insight into why producers might adopt either of these sequences. The initial thrust of the analysis was to determine if the locations of fields in these cropping sequences were randomly distributed across the landscape. The results indicated that the

fields were not randomly located but clustered out to approximately 20 km – an important finding, as it suggested the presence of an underlying process.

One of the leading concepts with respect to the geographic spread of practices, such as improved agricultural systems, is innovation diffusion theory and is rooted in the adoption of innovations due to social influence from neighbors. The theory is well-established in agriculture applications; however, my results imply that the adoption of cropping sequences in northeastern Montana largely is not explained by this theory, although early adoption patterns indicated that, at that stage, diffusion through social interactions might have played a role. Instead, the adoption and spread of both practices were more consistent with suitability of the physical environment. Available water, in particular, was strongly associated with whether or not a field was managed with either practice.

There were, of course, intriguing aspects in this research that I did not pursue. In Chapter 2, for example, I was able to classify fields using imagery with missing data by considering them to be homogeneous objects and using the extant pixels. I did not explore the limits of this process in terms of just how few pixels would be needed, but simply set a minimum of 40. Another opportunity for the expansion of this work is the spatial analysis. Potential factors that might also explain the observed patterns of adoption were left unexplored. The distance to delivery points for pulse crops, for example, affects transportation costs and might influence the decision to adopt cereal-pulse cropping sequences.

Another useful line of inquiry would be to examine the generality of my conclusions in terms of location and crop type. The adoption of pulses in north-central Montana, for example, has grown substantially in the past few years, which would provide a prime opportunity to test whether the same temporal and spatial patterns hold for an analogous region. Similarly, the rotation of cereals with oilseed crops offer potential benefits, and an analysis of the adoption of these sequences would be instructive with respect to crop type.

REFERENCES CITED

Adam, E.M., Mutanga, O., Rugege, D. & Ismail, R. (2012). Discriminating the papyrus Vegetation (Cyperus papyrus L.) and its co-existent species using random forest and hyperspectral data resampled to HYMAP. *International Journal of Remote Sensing*, *33*(2), 552–569.

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

Ashish, D., McClendon, R.W. & Hoogenboom, G. (2009). Land-use classification of multispectral aerial images using artificial neural networks. *International Journal of Remote Sensing*, *30*(8), 1989–2004.

Atkinson, A., Stephens, J.M. & Morgan, G.W. (1917). *Dry farm crop rotations and cultural methods (Agricultural Experiment Station Bulletin No. 116)*. Bozeman, MT: University of Montana.

Baddeley, A. & Turner, R. (2013). Spatstat (R package version 2.20.4) [Software]. Available from http://www.R-project.org/.

Baddeley, A. (2008). *Analysing spatial point patterns in R* [WWW document]. Retrieved from http://www.csiro.au/Portals/Publications/Research--Reports/Spatial-Point-Patterns-in-R.aspx.

Baddeley, A., Moller, J. & Waagepetersen, R. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, *54*, 329–350.

Barton, K. (2013). MuMIn: Multimodel Inference. (R package version 1.9.13) [Software]. Available from http://www.R-project.org/.

Beck, U. (1992). *Risk society: Towards a new modernity*. London, UK: Sage.

Besag, J. (1977). Contribution to the discussion of Dr. Ripley's paper. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *39*, 193–195.

Boloorani, A. D., Erasmi, S. & Kappas, M. (2008). Multi-source image reconstruction: exploitation of EO-1/ALI in Landsat 7/ETM+ SLC-off gap filling. In J.T. Astola, K.O. Egiazarian, & E. R. Dougherty (Eds.), *Image processing: Algorithms and systems VI, Proceedings of SPIE-IS&T, 6812* (pp. 681219-1–681219-12). San Jose, CA: SPIE.

Boryan, C., Yang, S., Mueller, R. & Craig, M. (2011). Monitoring US agriculture: The US Department of Agriculture, National Statistics Service, cropland data layer program. *Geocarto International*, *26*(5), 341–358.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*, 123–140.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

BSCSP (Big Sky Carbon Sequestration Project). (2012). *About BSCSP* [WWW Document]. Retrieved from http://www.bigskyco2.org/about.

Burgess, M., Miller, P. & Jones, C. (2012). Pulse crops improve energy intensity and productivity of cereal production in Montana, U.S.A. *Journal of Sustainable Agriculture*, *36*, 699–718.

Burt,O.R. & Allison, J.R. (1963). Farm management decisions with dynamic programming. *Journal of Farm Economics*, *45*(1), 121–136.

Cai, X., Wu, Z. & Cheng, J. (2013). Using kernel density estimation to assess the spatial pattern of road density and its impact on landscape fragmentation. *International Journal of Geographic Information Science*, *27*, 222–230.

Castellazzi, M.S., Perry, J.N., Colbach, N., Monod, H., Adamczyk, K., Viaud, V. & Conrad, K.F. (2007). New measures and tests of temporal and spatial pattern of crops in agricultural landscapes. *Agriculture, Ecosystems, and Environment*, *118*, 339–349.

Chapin, F. (1928). *Cultural change*. New York, NY: The Century Co.

Chapman, D.S., Bonn, A., Kunin, W.E. & Cornell, S.J. (2010). Random forest characterization of upland vegetation and management burning from aerial imagery. *Journal of Biogeography*, *37*, 37–46.

Chen, F., Tang, L. & Qiu, Q. (2010). Exploitation of CBERS-02B as auxiliary data in recovering the Landsat7 ETM+ SLC-off image. In Y. Liu & A. Chen (Eds.), *18th International Conference on Geoinformatics*. Beijing, China: Institute of Electrical and Electronics Engineers.

Clark, G. (1984). *Innovation diffusion: Contemporary geographical approaches*. Norwhich, UK: Geo Books.

Clark, S. (2011). *Veteran Landsat 5 satellite on the brink of failure* [WWW Document]. Retrieved from http://www.spaceflightnow.com/news/n1111/18landsat5/.

Cliff, A. & Ord, K. (1974). *Spatial autocorrelation*. London, UK: Pion.

Cliff, A. (1968). The neighbourhood effect in the diffusion of innovations. *Transactions of the Institute of British Geographers*, *44*, 75–84.

Cochran, V., Danielson, J., Kolberg, R. & Miller, P. (2006). Dryland cropping in the Canadian Prairies and the U.S. Northern Great Plains. In G.A. Peterson, P.W. Unger, & W.A. Payne (Eds.), *Dryland agriculture* (2nd ed.) (pp. 293–339). Madison, WI: American Society of Agronomy, Inc., Crop Science Society of America, Inc., & Soil Science Society of America, Inc.

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

Congalton, R.G. & Green, K. (2009), *Assessing the accuracy of remotely sensed data: Principles and practices* (2nd ed.). Boca Raton, FL: CRC Press.

Conrad, C., Colditz, R.R., Dech, S., Klein, D. Vlek, P.L.G. (2011). Temporal segmentation of MODIS time series for improving crop classification in central Asian irrigation systems. *International Journal of Remote Sensing*, *32*(23), 8763–8778.

Daly, C. (2006). Guidelines for assessing the suitability of spatial climate data sets. *International Journal of Climatology*, *26*, 707–721.

Daly, C., Gibson, W., Taylor, G., Johnson, G. & Pasteris, P. (2002). A knowledge-based approach to the statistical mapping of climate. *Climate Research*, *22*, 99–113.

Daly, C., Neilson, R. & Phillips, D. (1994). A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology*, *33*, 140–158.

Dang, Y.P., Pringle, M.J., Schmidt, M., Dalal, R.C. & Apan, A. (2011). Identifying the spatial variability of soil constraints using multi-year remote sensing. *Field Crops Research*, *123*, 248–258.

Diggle, P. (2003). *Statistical analysis of spatial point patterns*. Oxford, UK: Oxford University Press.

El-Magd, I.A. & Tanton, T.W. (2003). Improvements in land use mapping for irrigated agriculture from satellite sensor data using a multi-stage maximum likelihood classification. *International Journal of Remote Sensing*, *24*(21), 4197–4206.

Fargione, J.E., Cooper, T.R., Flaspohler, D.J., Hill, J., Lehman, C., McCoy, T., McLeod, S., Nelson, E.J., Oberhauser, K.S. & Tilman, D. (2009). Bioenergy and wildlife: Threats and opportunities for grassland conservation. *Science*, *59*(9), 767–777.

Forster, D., Kellenberger, T.W., Buehler, Y. & Lennartz, B. (2010). Mapping diversified peri-urban agriculture – potential of object-based versus per-field land cover/land use classification. *Geocarto International*, *25*(3), 171–186.

Fox, J. (2013). Effects (R package version 3.0-3) [Software]. Available from http://www.R-project.org/.

Friedli, H., Lötscher, H., Oeschger. H., Siegenthaler, U. & Stauffer, B. (1986). Ice core record of the $^{13}C/^{12}C$ ratio of atmospheric CO_2 in the past two centuries. *Nature, 324*, 237-238.

Fuller, R.M., Groom, G.B. & Jones, A.R. (1994). The land-cover map of Great-Britain – An automated classification of Landsat Thematic Mapper data. *Photogrammetric Engineering and Remote Sensing*, *60*, 553–562.

Gabriel, E. & Diggle, P.J. (2009). Second-order analysis of inhomogeneous spatio-temporal point process data. *Statistica Neerlandica*, *63*, 43–51.

Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M. & Tyler, D. (2002). The national elevation dataset. *Photogrammetric Engineering and Remote Sensing*, *68*, 5–11.

Gómez-Casero, M.T., Castillejo-González, I.L., García-Ferrer, A., Peña-Barragán, J.M., Jurado-Expósito, M., Garciá-Torres, L. & López-Granados, F. (2010). Spectral discrimination of wild oat and canary grass in wheat fields for less herbicide application. *Agronomy for Sustainable Development*, *30*, 689–699.

Hagen, S.K., Isakson, P.T. & Dyke, S.R. (2005). *North Dakota comprehensive wildlife conservation strategy* [WWW Document]. Retrieved from http://www.gf.nd.gov/conservation/cwcs.html.

Hägerstrand, T. (1953). *Innovation diffusion as a spatial process*. Chicago, IL:The University of Chicago Press.

Hahvey, D. (1966). Theoretical concepts and the analysis of Agricultural land-use patterns in geography. *Annals of the Association of American Geographers*, *56*, 361–374.

Han, W., Yang, Z., Di, L. & Mueller, R. (2012). Cropscape: A web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support. *Computers and Electronics in Agriculture*, *84*, 111–123.

Hansen, J., Sato, M., Kharecha, P., Beerling, D., Berner, R., Masson-Delmotte, V., Pagani, M., Raymo, M., Royer, D. & Zachos, J. (2008). Target atmospheric CO_2: Where should humanity aim? *The Open Atmospheric Science Journal*, *2*, 217-231.

Hansen, N.C., Allen, B.L., Baumhardt, R.L. & Lyon, D.J. (2012). Research achievements and adoption of no-till, dryland cropping in the semi-arid U.S. Great Plains. *Field Crops Research*, *132*, 196–203.

Hauser, B. (2002). Brief highlights of the 2002 Farm Bill. [WWW Document]. Available from http://www.farmdoc.illinois.edu/policy/farmbill/policy_farmbill_2002.html.

Heiberger, R.M. (2014). HH (R package version 3.0-3) [Software]. Available from http://www.R-project.org/.

Hennessy, D.A. (2006). On monoculture and the structure of crop rotations. *American Journal of Agricultural Economics*, *88*(4), 900–914.

IPCC (Intergovernmental Panel on Climate Change). (2007). *Climate change 2007: Synthesis report*. *Contribution of working groups I, II and III to the fourth assessment report of the intergovernmental panel on climate change* [Core Writing Team, Pachauri, R. K. and Reisinger, A. (eds.)]. Geneva, Switzerland: IPCC.

Janssen, L.L.F. & van Amsterdam, J.D. (1996). An object-based approach to the classification of remotely sensed images. In *Proceedings of the International Geoscience and Remote Sensing Symposium, IGARSS '91, Remote sensing: Global monitoring for earth management, Vol. 4* (pp. 2192–2195).

Johansen, H. (1971). Diffusion of strip cropping in southwestern Wisconsin. *Annals of the Association of American Geographers*, *13*, 671–684.

Johnson, D.M. (2013). A 2010 map estimate of annually tilled cropland within conterminous United States. *Agricultural Systems*, *114*, 95–105.

Kutz, F.W., Morgan, J.M. III, Monn, J. & Petrey, C.P. (2012). Geospatial approaches to characterizing agriculture in the Chincoteague Bay subbasin. *Environmental Monitoring and Assessment*, *184*, 679–692.

Lawrence, R.L., Wood, S.D. & Sheley, R.L. (2006). Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest)," *Remote Sensing of Environment*, *100*, 356–362.

Lee, C. (2011). *Montana's pulse industry: How it has developed, economic impact & potential for growth* [WWW Document]. Retrieved from http://agr.mt.gov/agr/Programs/Commodities/PulseCrops/Pulse_Potential_PDFs/Montana_Pulse_Potential.pdf.

Lemke, R.L., Zhong, Z., Campbell, C.A. & Zentner, R. (2007). Can pulse crops play a role in mitigating greenhouse gases from North American agriculture? *Agronomy Journal*, *99*, 1719–1725.

Linz, G.M., Schaaf, D.A., Mastrangelo, P., Homan, H.J., Penry, L.B. & Bleier, W.J. (2004). *Wildlife conservation sunf lower plots as a dual-purpose wildlife management strategy*. Fort Collins, CO: USDA National Wildlife Research Center.

Liu, D., Kelly, M., Gong, P. & Guo, Q. (2007). Characterizing spatial–temporal tree mortality patterns associated with a new forest disease. *Forest Ecology and Management*, *253*, 220–231.

Loader, C. (1999). *Local regression and liklihood*. New York, NY: Springer.

Lobell, D.B., Cahill, K.N. & Field, C.B. (2007). Historical effects of temperature and precipitation on California crop yields. *Climate Change*, *81*, 187–203.

Long, J.A., Lawrence, R.L., Greenwood, M.C., Marshall, L. & Miller, P.R. (2013). Object-oriented crop classification using multitemporal ETM+ SLC-off imagery and Random Forest. *GIScience and Remote Sensing*, *50*(4), 418–436.

Long, J.A., Lawrence, R.L., Miller, P.R. & Marshall, L.A. (In Press). Changes in field-level cropping sequences: Indicators of shifting agricultural practices. *Agriculture, Ecosystems, and Environment*.

Lu, D. & Weng, Q. (2005). Urban classification using full spectral information of Landsat ETM+ imagery in Marion County, Indiana. *Photogrammetric Engineering and Remote Sensing*, *71*(11), 1275–1284.

Markham, B.L., Storey, J.C., Williams, D.L. & Irons, J.R. (2004). Landsat sensor performance: History and current status. *IEEE Transactions on Geoscience and Remote Sensing*, *42*(12), 2691–2694.

Masek, J.G., Honzak, M., Goward, S.N., Liu, P. & Pak, E. (2001). Landsat ETM+ as an observatory for land cover: Initial radiometric and geometric comparisons with Landsat 5 Thematic Mapper. *Remote Sensing of Environment*, *78*, 118–130.

Maxwell, S. (2004). Filling Landsat ETM+ SLC-off gaps using a segmentation model approach. *Photogrammetric Engineering and Remote Sensing*, *70*(10), 1109–1111.

Maxwell, S.K., Schmidt, G.L. & Storey, J.S. (2007). A multi-scale segmentation approach to filling gaps in Landsat ETM+ SLC-off images. *International Journal of Remote Sensing*, *28*(10), 5339–5356.

McVay, K., Burrows, M., Menalled, F., Jones, C., Wanner, K. & O'Neill, R. (2013). *Montana cool-season pulse production guide (Montana State University Extension Guide EB0210)*. Bozeman, MT: Montana State University Extension.

Meinke, H. & Stone, R.C. (2005). Seasonal and inter-annual climate forecasting: The new tool for increasing preparedness to climate variability and change in agricultural planning and operations. *Climate Change*, *70*, 221–252.

Miller, P.R., Engel, R.E. & Holmes, J.A. (2006). Cropping sequence effect of pea and pea management on spring wheat in the northern Great Plains. *Agronomy Journal*, 98, 1610–1619.

Mortimer, J. (1703). *The whole art of husbandry; or, the way of managing and improving of land* (2nd ed.). London, UK: H. Mortlock.

MSL (Montana State Library). (2012). *Montana geographic information clearinghouse* [Data files]. Retrieved from http://nris.mt.gov/gis/default.asp.

NASS (National Agricultural Statistics Service) (2012). *Montana agricultural statistics, volume XLIX*. Washington, DC: United States Department of Agriculture.

NASS (National Agricultural Statistics Service). (2013). *Data and statistics: Crops and plants* [Data files]. Retrieved from http://www.nass.usda.gov/Statistics_by_Subject.

Ngo Bieng, M.A., Gidoin, C., Avelino, J., Cilas, C., Deheuvels, O. & Wery, J. (2013). Diversity and spatial clustering of shade trees affect cacao yield and pathogen pressure in Costa Rican agroforests. *Basic and Applied Ecology*, 14, 329–336.

NOAA (National Ocean and Atmospheric Administration). (2011). *Montana weather/ precipitation summary, May 2011* [WWW Document]. Retrieved from http://www1.ncdc.noaa.gov/pub/data/cmb/regional-reports/wrccsumm1105.pdf.

NOAA (National Oceanic and Atmospheric Administration). (2012). *Trends in atmospheric carbon dioxide: Recent global CO_2*. Retrieved from http://www.esrl.noaa.gov/gmd/ccgg/trends/global.html.

NRCS (Natural Resources Conservation Service). (2014). *Relative effective annual precipitation data* [Data file]. Retrieved from http://nris.mt.gov/nrcs/reap/.

O'Dea, J.K., Miller, P.R. & Jones, C.A. (2013). Greening summer fallow with legume green manures: On-farm assessment in north-central Montana. *Journal of Soil and Water Conservation*, *68*(4), 270–282.

Oetter, D.R., Cohen, W.B., Berterretche, M., Maiersperger, T.K. & Kennedy, R.E. (2000). Land cover mapping in an agricultural setting using multi-seasonal Thematic Mapper data. *Remote Sensing of Environment*, *76*(2), 139-155.

Padbury, G. Waltman, S., Caprio, J., Coen, G. ,McGinn, S., Mortensen, D., Nielsen, G. & Sinclair, R. (2002). Agroecosystems and land resources of the northern Great Plains. *Agronomy Journal*, *94*, 251–261.

Pages, H., Aboyoun, P., Gentleman, R., & DebRoy, S. (2013). Biostrings (R package version 2.20.4) [Software]. Available from http://www.R-project.org/.

Pax-Lenney, M. & Woodcock, C.E. (1997). Monitoring agricultural lands in Egypt with multitemporal Landsat TM imagery: How many images are needed? *Remote Sensing of Environment*, *59*, 522–529.

Pedley, M.I. & Curran, P.J. (1991). Per-field classification: An example using SPOT HRV imagery. *International Journal of Remote Sensing*, *12*(11), 2182–2192.

Peel, M., 1998. *Crop rotations for increased productivity (NDSU Agriculture and University Extension Publication No. EB-48 (Revised))*. Fargo, ND: North Dakota State University Extension Service.

Perry, J., Liebhold, A., Rosenberg, M., Dungan, J., Miriti, M., Jakomulska, A. & Citron-Pousty, S. (2002). Illustrations and guidelines for selecting statistical methods for quantifying spatial pattern in ecological data. *Ecography (Cop.)*, *25*, 578–600.

Plourde, J.D., Pijanowski, B.C. & Pekin, B.K. (2013). Evidence for increased monoculture cropping in the Central United States. *Agriculture, Ecosystems, and Environment*, *165*, 50–59.

Pringle, M.J., Schmidt, M. & Muir, J.S. (2009). Geostatistical interpolation of SLC-Off Landsat ETM+ images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*, 654–664.

PRISM (PRISM Climate Group). (2014). *Prism climate data* [Data file]. Retrieved from http://prism.oregonstate.edu.

R Development Core Team. (2011). R: A Language and Environment for Statistical Computing (version 2.13.1) [Software]. Available from http://www.R-project.org/.

Reza, M.M. & Ali, S.N. (2008). Using IRS products to recover ETM+ defective images. *American Journal of Applied Sciences*, *5*, 618–625.

Ripley, B. (1977). Modelling spatial patterns. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *39*, 172–192.

Ripley, B. (1979). Tests of "randomness" for spatial point patterns. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *41*, 368–374.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M. & Rigol-Sanchez, J.P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *67*, 93–104.

Rosenzweig, C., Tubiello, F.N., Goldberg, R., Mills, E. & Bloomfield, J. (2002). Increased crop damage in the US from excess precipitation under climate change. *Global Environmental Change*, *12*, 197–202.

Rounsevell, M.D.., Annetts, J., Audsley, E., Mayr, T. & Reginster, I. (2003). Modelling the spatial distribution of agricultural land use at the regional scale. *Agriculture, Ecosystems, and Environment*, *95*, 465–479.

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

Ryan, B. & Gross, N. (1943). The diffusion of hybrid seed corn in two Iowa communities. *Rural Sociology*, *8*, 15–24.

Ryan, B. & Gross, N. (1950). *Acceptance and diffusion of hybrid corn seed in two Iowa communities (Research Bulletin 372)*. Ames, IA: Agriculture Experiment Station, Iowa State College of Agriculture and Mechanical Arts.

Savage, S.L. & Lawrence, R.L. (2010). Vegetation dynamics in Yellowstone's northern range: 1985–1999. *Photogrammetric Engineering and Remote Sensing*, *76*(5), 547–556.

Scheffran, J., BenDor, T., Wang, Y. & Hannon, B. (2007). A spatial-dynamic model of bioenergy crop introduction in Illinois. In *The 25th international conference of the system dynamics society*. Boston MA.

Schmit, C. & Rounsevell, M.D. (2006). Are agricultural land use patterns influenced by farmer imitation? *Agriculture, Ecosystems, and Environment*, *115*, 113–127.

Scholnick, J.B. (2012). The Spatial and Temporal Diffusion of Stylistic Innovations in Material Culture. *Advances in Complex Systems*, *15*, 1150010–1 – 1150010–18.

Scholten, H. (1988). Snow distribution on crop fields. *Agriculture, Ecosystems, and Environment*, *22/23*, 363–380.

Schultz, S., Schmitz, N. & Leitch, J. (2007). A spatial evaluation of agricultural property tax inequity associated with productivity-based assessments. *Journal of Property Tax Assessment and Administration*, *3*(3), 53–65.

Simonneaux, V., Duchemin, B., Helson, D., Er-Raki, S., Olioso, A. & Chehbouni, A.G. (2008). The use of high-resolution image time series for crop classification and evapotranspiration estimate over an irrigated area in central Morocco. *International Journal of Remote Sensing*, *29*(1), 95–116.

Storey, J., Scaramuzza, P. Schmidt, G. & Barsi, J. (2005). Landsat 7 scan line corrector-off gap-filled product development. In *ASPRS 2005 - Pecora 16 Conference Proceedings*, Sioux Falls, SD: American Society of Photogrammetry and Remote Sensing.

Streib, K. & Davis, J.W. (2011). Using Ripley's K-function to improve graph-based clustering techniques. In *IEEE conference on computer vision and pattern recognition 2011* (pp. 2305–2312). Colorado Springs, CO: IEEE.

Strobl, C. & Zeileis, A. (2008). Danger: High power! – Exploring the statistical properties of a test for random forest variable importance. In *Proceedings of the 18th International conference on computational statistics*. Porto, Portugal.

Strobl, C., Boulesteix, A-L., Zeileis, A. & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, *8*, 1–25.

Tanaka, D., Lyon, D., Miller, P., Merrill, S. & McConkey, B. (2010). Soil and water conservation advances in the semiarid Northern Great Plains. In T. Zobeck & W. Schillinger (Eds.). *Soil and water conservation in the United States (SSSA Special Publication 60)* (pp. 81–102). Madison, WI, American Society of Agronomy.

Tarde, G. (1903). *The laws of imitation*. New York, NY: Henry Holt and Company.

Taylor, C. (1989). Bootstrap choice of the smoothing parameter in kernel density estimation. *Biometrika*, *76*, 705–712.

Turker, M. & Arikan, M. (2005). Sequential masking classification of multitemporal Landsat ETM+ images for field-based crop mapping in Karacabey, Turkey. *International Journal of Remote Sensing*, *26*(17), 3813–3830.

USDA (United States Department of Agriculture) (2012*). USDA National Agricultural Statistics Service cropland data layer* [Data file]. Retrieved from http://nassgeodata.gmu.edu/CropScape/], last accessed on June 11, 2013.

USGS (United States Geological Survey). (2012). *Landsat: A global land-imaging mission*., Reston, VA: U.S. Geological Survey.

Valente, T. (1995). *Network models of the diffusion of innovations*. New York, NY: Hampton Press.

Verburg, K., Bond, W.J. & Hunt, J.R. (2012). Fallow management in dryland agriculture: Explaining soil water accumulation using a pulse paradigm. *Field Crops Research*, 130, 68–79.

Waller, L. & Gotway, C. (2004). *Applied spatial statistics for public health data*. Hoboken, NJ: John Wiley & Sons, Inc.

Wardlow, B.D., Egbert, S.L. & Kastens, J.H. (2007). Analysis of time-series MODIS 250m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sensing of Environment*, *108*(3), 290–310.

Warnes, G.R. (2013). Gmodels (R package version 2.14.4.1) [Software]. Available from http://www.R-project.org/.

Watts, J.D., Lawrence, R.L., Miller, P.R. & Montagne, C. (2009). Monitoring of cropland practices for carbon sequestration purposes in north central Montana using Landsat imagery. *Remote Sensing of Environment*, *113*, 1843–1852.

Weiss, M.J. & Morrill, W.L. (1992). Wheat stem sawfly (Hymenoptera: Cephidae) revisited. *American Entomologist*, *38*(4), 241–245.

Wejnert, B. (2002). Integrating models of diffusion: A conceptual framework. *Annual Review of Sociology*, *28*, 297–326.

West, T. & Post, W. (2002). Soil organic carbon sequestration rates by tillage and crop rotation: A global data analysis. *Soil Science Society of America Journal, 66*, 1930-1946.

WRCC (Western Regional Climate Center). (2011). *Local climate summary for Glasgow, Montana* [WWW Document]. Retrieved from http://www.wrcc.dri.edu/summary/ggw.mt.html.

Wu, S., Silván-Cárdenas, J. &Wang, L. (2007). Per-field urban land use classification based on tax parcel boundaries. *International Journal of Remote Sensing*, *28*(12), 2777–2800.

Yarranton, A.G.A. & Morrison, R.G. (1974). Spatial dynamics of a primary succession : Nuceation. *Journal of Ecology*, *62*, 417–428.

Zeng, C., Shen, H. & Zhang, L. (2013). Recovering missing pixels for Landsat ETM+ SLC-off imagery using multitemporal regression analysis and a regularization method. *Remote Sensing of Environment*, *131*, 182–194.

Zentner, R.P., Campbell, C.A., Biederbeck, V.O., Miller, P.R., Selles, F., & Fernandez, M.R. (2001). In search of a sustainable cropping system for the semiarid Canadian prairies. *Journal of Sustainable Agriculture*, *18*, 117–136.

Zentner, R.P., Wall, D.D., Smith, D.G., Young, D.L., Miller, P. R., Campbell, C.A., Lafond, G.P., Brandt, S.A. & Johnston, A.M. (2002). Economics of crop diversification opportunities for the northern Great Plains. *Agronomy Journal*, *94*, 216–230.

Zhang, C., Li., W. & Travis, D. (2007). Gaps-fill of SLC-off Landsat ETM+ satellite image using a geostatistical approach. *International Journal of Remote Sensing*, *28*(22), 5103–5122.

Zhu, X., Liu, D. & Chen, J. (2012). A new geostatistical approach for filling gaps in Landsat ETM+ SLC-off images. *Remote sensing of Environment*, *124*, 49–60.