Remote Sensing of Environment 113 (2009) 1843–1852

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Monitoring of cropland practices for carbon sequestration purposes in north central Montana by Landsat remote sensing

Jennifer D. Watts, Rick L. Lawrence *, Perry R. Miller, Cliff Montagne

Department of Land Resources and Environmental Science, Montana State University, P.O. Box 173120 Bozeman, MT 59717, USA

ARTICLE INFO

Article history:
Received 6 October 2008
Received in revised form 30 March 2009
Accepted 11 April 2009

Keywords:
Carbon sequestration validation
Tillage type
Conservation reserve
Crop intensity
Landsat
Random Forest
Object-oriented

ABSTRACT

We used an object-oriented approach in conjunction with the Random Forest algorithm to classify agricultural practices, including tillage (till or no-till (NT)), crop intensity, and grassland-based conservation reserve (CR). The object-oriented approach allowed for per-field classifications and the incorporation of contextual elements in addition to spectral features. Random Forest is a classification tree-based advanced classifier that avoids data over-fitting associated with many tree-based models and incorporates an unbiased internal classification accuracy assessment. Landsat satellite imagery was chosen for its continuous coverage, cost effectiveness, and image accessibility. Classification results for 2007 included producer's accuracies of 91% for NT and 31% for tillage when applying Random Forest to image objects generated from a May Landsat image. Low classification accuracies likely were attributed to the misclassification of conservation-based tillage practices as NT. Results showed that the binary separation of tillage from NT management is likely not appropriate due to surface spectral and textural similarities between NT and conservation-type tillage practices. Crop and CR lands resulted in producer's accuracies of 100% and 90%, respectively. Crop and fallow producer's accuracies were 95% and 82% in the 2007 classification, despite post-senesced vegetation; misclassification within the fallow class was attributed to pixel-mixing problems in areas of narrow (<100 m) strip management. A between-date normalized difference vegetation index approach was successfully used to detect areas having "changed" in vegetation status between the 2007 and prior image dates; classified "changed" objects were then merged with "unchanged" objects to produce crop status maps. Field crop intensity was then determined from the multi-year analysis of generated crop status maps.

1. Introduction

Carbon contract programs are being developed within Montana and across the US. Cropland producers involved in terrestrial carbon sequestration programs are paid to implement practices that might increase soil carbon, including no-till (NT), crop intensity, and conservation reserve (CR). NT systems involve the absence of plowing. Crop intensity is the proportion of years that a field is summer fallowed vs. cropped; increasing crop intensity results in a larger percentage of time that a field is under vegetative cover. CR is the conversion of marginal cropland into diverse perennial plant cover, and is not confined to lands within the Conservation Reserve Program (CRP). Contract validation of these practices is currently limited to on-site field survey (NCOC, 2008), a costly and time-consuming process. Remote sensing might offer a timely and cost-effective option to validate and monitor cropland management for carbon contract purposes, on a per-field basis.

The majority of image-based classifications for NT, field vegetative status (cropped vs. fallow), and grassland-based CR have utilized a per-pixel approach. Pixel-based results are often illogical due to mixed within-field classifications, although fields generally have a single management regime within the field boundaries. The absence of reference field boundaries can also make it difficult to interpret land use patterns. A field-based, object-oriented (O-O), approach offers a practical alternative to pixel-based methods as the classification and analysis process is based on landscape boundaries representing specific spatial patterns. O-O techniques utilize objects, groups of pixels with similar spectral and spatial properties, rather than individual pixels for image classification and analysis (Navulur, 2007). Strengths of using an object-based approach include the partitioning of landscapes into meaningful units (such as fields for agricultural analyses), the generation of shape, texture, and relational features that can be incorporated into the classification process, and an easier integration into GIS systems than traditional pixel-based maps (Hay & Castilla, 2006). Disadvantages might include difficulty in handling large data sets within O-O software and the need to heuristically determine segmentation parameters unless the process is based directly on an existing vector layer, in a "cookie cutter" approach.

* Corresponding author. Tel.: +1 406 994 5409.
E-mail address: rickl@montana.edu (R.L. Lawrence).

0034-4257/$ – see front matter
doi:10.1016/j.rse.2009.04.015

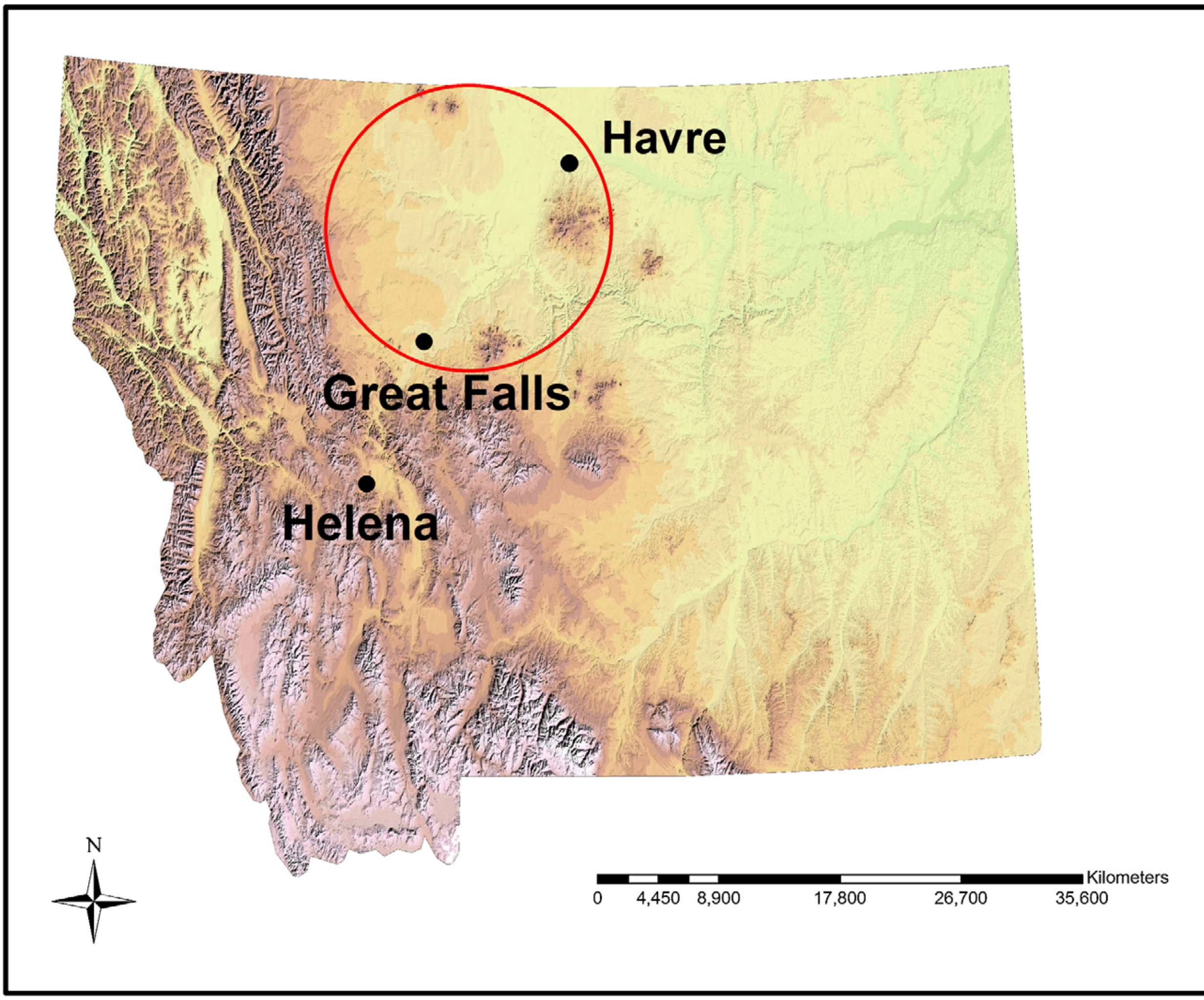

Fig. 1. Geographic location (within the red circle) of the cropland validation study, Montana, USA. Map colors reflect elevation gradients, ranging from gentle lowland prairies in the east, to the more mountainous regions of the west.

Peer-reviewed research concerning the O-O analysis of cropland management practices is lacking. A few pixel-based studies have mapped NT and tillage on surfaces with bare to minimal cover with resulting class accuracies 90% or greater (Bricklemyer et al., 2002; Gowda et al., 2005; South et al., 2004) but have had limited success in the presence of vegetative cover (Bricklemyer et al., 2006). Spectral similarities between grassland-based CR and other management practices have made image classification problematic (Daly, 2001), often requiring the use of multi-temporal techniques (Egbert et al., 2002; Price et al., 1997) or the inclusion of ancillary data (Song et al., 2003). These mentioned studies pertaining to tillage and CR mapping practices were not object-based. We have not identified studies, pixel or object-based, that specifically mapped crop intensity (the multi-year analysis of crop and fallow patterns).

The majority of object-based studies have utilized simple classification techniques rather than advanced ones. Simple classifiers such as the nearest neighbor (NN) (Kamagata et al., 2006; Lucas et al., 2007; Stow et al., 2007) are often used, although continued advancements in image classification algorithms for pixel-based studies have made the use of these methods increasingly obsolete within pixel-based analyses (Gislason et al., 2006). Advanced image classification algorithms that have been used within pixel-based analyses include, but are not limited to, boosting and/or bagging-based classification tree analysis (CTA) (Baker et al., 2006; Lawrence et al., 2004; Lawrence & Wright, 2001) and the CTA-based Random Forest (Ham et al., 2005; Lawrence et al., 2006; Prasad et al., 2006).

The Random Forest (RF) classifier is superior to many tree-based algorithms as it is not sensitive to noise or overtraining, and it is also capable of handling unbalanced data sets (Breiman & Cutler, 2004). RF uses a bagging-based approach, or sampling from the original data set with replacement, to form an ensemble of classification trees (Breiman, 2001; Gislason et al., 2006). Data over-fitting, where the model becomes too statistically conformed to the training data and thus the model performance will not replicate well with future data sets, is avoided as random variable subsets are also used to generate node splits. Node splits occur as the original sample is partitioned into subsamples according to homogeneity; splitting occurs until end nodes result, each representing a particular class, when no more useful splits can be made. The result of the random variable selection at each node split is that each classification tree differs greatly from the next.

Each tree in the resulting forest casts a unit class vote, with the final classification determined by an amongst-tree plurality decision (Breiman, 2001). RF also includes an internal out-of-bag (OOB) accuracy measure that produces results comparable to external accuracy assessments, so long as there is no bias in the reference data (Lawrence et al., 2006). OOB works by withholding a random portion of the original data set from the model-building process, the "withheld" data is then run through the generated classification trees to estimate model accuracy. Studies using RF for O-O classification were not identified, although the remote sensing community has increasingly used RF for pixel-based land cover classifications (Gislason et al., 2006; Pal, 2005).

The objective of this study was to determine if O-O RF classifications using moderate resolution Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) imagery could accurately identify management types specified by National Carbon Offset Coalition (NCOC) carbon contracts, within north central Montana. Three separate classifications types were attempted corresponding to the three types of management practices that might be included in NCOC carbon contracts, namely till from NT, the determination of 4-year crop intensity through an analysis of multi-year crop and fallow patterns, and the separation of grassland-based CR from cropland.

We considered an O-O approach to be advantageous over a pixel-based analysis as each field could be treated as an individual vector object, allowing the determination of management type to occur on a per-unit basis. The object-based approach was also necessary as the validation of management practices defined by carbon contracts correspond to, and thus must be analyzed according to, specific parcel-based areas. The decision to use the RF algorithm to classify object-based data was based on the documentation of prior literature (Breiman, 2001; Lawrence et al., 2006) concerning its advanced capabilities in handling complex data sets, the OOB accuracy measure, and generally higher reported accuracies compared to results by other classifiers. The incorporation of RF with O-O analysis, however, had not been demonstrated previously within remote sensing literature.

2. Methods

2.1. Study area description

We focused our analysis on cropland management practices within north central Montana. This region is roughly bounded by the communities of Great Falls, Fort Benton, Havre, Cut Bank, and Conrad (Fig. 1). This region falls within a semi-arid grassland/shrub (steppe) biotic regime (NRCS, 2007a). Soil type can vary considerably throughout the region (NRCS, 2007b); clay-dominant textures are reportedly common in many cropland fields, with more sandy soils expected within elevated field locations and areas prone to water erosion.

Fig. 2. Data subset locations (outlined in red) used for identifying crop and fallow practices from 2004–2007 as seen within two Landsat images (May 2007) where non-cropland data has been removed ("black" locations represent no data). The imagery is displayed in false color (B2 as blue, B3 as green, and B4 as red).

Precipitation patterns are extremely important as timely spring rain allows crop production to occur in an otherwise semi-arid climate. Temperature and especially precipitation vary strongly within the region. Annual average minimum temperatures have ranged from −0.7 °C Havre to −0.9 °C in Great Falls, with annual average maximum temperatures ranging from 12.7 °C in Havre to 15.5 °C in Fort Benton (NWS, 2007). Annual average precipitation ranges from 265 mm in Chester, 318 mm Cut Bank, and 373 mm in Great Falls (WRCC, 2006).

Dryland wheat and barley are the primary crop exports within north central Montana, with market price and fall soil moisture conditions driving the decision to plant spring or winter wheat. Other crops might be planted occasionally, but have contributed to a minute proportion of total cropped ha (CTIC, 2004). A wheat-fallow rotation is common throughout the region; continuous cropping is rarely implemented and is considered by many producers to be too risky due to unreliable late spring and early summer precipitation (Farkell, 2007).

Four data subsets were identified within the general study area for the analysis of crop and fallow patterns. These subsets were located near Dutton (18,500 ha), Chester (11,250 ha), between Big Sandy and Fort Benton (7646 ha), and near Great Falls (13,014 ha) (Fig. 2). Chester was chosen as it represented a drier than average climate within the 2004–2007 period (~250 mm annual precip.), while Great Falls represented a wetter than average climate (~390 mm annual precip.) (HPRCC, 2008). Annual precipitation in the Dutton and Big Sandy/Fort Benton areas was near average (~290–320 mm).

2.2. Reference data collection

Locations for training data were randomly identified throughout the region. Final data collection points were chosen based on their proximity to public roadways to avoid land access issues. Data collection occurred early June 2007, a time by which field tillage type should be apparent for that year, with the exception of chemically managed fallow that might be tilled subsequently during the summer. This practice, however, appears uncommon based on our experience in the region, although there is no data available to quantify this. Reference field information collected through visual, on-site, analyses included vegetative status (cropped or fallow), crop type, and tillage management (tilled or NT). Cropland management class types were based on guidelines from the NCOC (2008) and the Chicago Climate Exchange (CCX, 2008). Tillage management type was classified as either “tilled” or “NT”. The determination of tillage management within a field included a visual examination of stubble position (stubble in a NT field should generally be in a relatively upright position), soil surface disturbance, and the establishment of soil surface crusts. Many of the fields classified as tillage had high levels of surface residue but also showed indication that there had been surface disturbance between the stubble rows. A visual percentage estimate of exposed soil surface was also taken for each field and was based on the relative amount of bare soil seen within the field as opposed to active surface vegetation or plant stubble. Land access restrictions prevented us from obtaining a more quantitative measure of exposed soil surface area.

Verbal communication with farmers throughout the region confirmed that conservation tillage, where only a chisel plough is used once prior to planting for residue management, is common. Although some of these practices can result in minimal levels of soil disturbance, it would still be considered “tillage” under NCOC carbon contract agreement definitions (NCOC, 2008). Rangeland data included vegetation type and the relative percent of soil surface covered by the vegetative canopy. CRP data were provided by the Montana Farm Service Agency (MFSA) to represent grassland management similar to that of the CRP. This step was taken under the assumption that spectral signatures between CRP and CR-based lands did not differ. The CRP sites used within the classification model-building process were randomly selected from the data pool. The resulting 2007 cropland data set included information for 78 NT–fallow, 138 NT–cropped, 48 tilled–fallow, 148 tilled–cropped, and 113 CR field sites. The actual number of field sites utilized within the model-building process was scene dependent due to cloud masking and missing pixel information resulting from the Landsat ETM+ scan-line gaps.

2.3. Satellite data collection

Satellite data consisted of Landsat TM and ETM+ image sets spanning path/row 39-26 and 39-27 (Table 1). Images were selected according to image quality and the degree of atmospheric interference (primarily due to clouds). Landsat data were chosen over other satellite image sources due to its cost effectiveness, relatively large footprint, and temporal coverage. The authors recognize that some error might have been introduced into the model-building process as the 2007 image dates did not match atmospheric interference (primarily due to clouds).

Landsat data were chosen over other satellite image sources due to its cost effectiveness, relatively large footprint, and temporal coverage. The authors recognize that some error might have been introduced into the model-building process as the 2007 image dates did not match that of the June field data collection period, thus there might have been some change in field management status between the time of image data collection and on-site field reference. The extent of this error is unknown, but believed to be relatively minor based on known cropping practices.

2.4. Satellite image pre-processing

Image scenes 39-26 and 39-27 were mosaicked to form one continuous image for each analyzed date. Images were geometrically rectified to a reference TM image used in the creation of the non-agricultural mask template and examined for spectral irregularities, haze, and clouding. Pixels contaminated with cloud, haze, and shadow were identified through a maximum likelihood supervised classification and excluded from image analyses.

Image data were converted to exoatmospheric reflectance to minimize between-image differences due to earth–sun distance and solar angle (Chander et al., 2007; SDH-L7, 2006). Further data correction techniques were used to normalize the 2004–2006 imagery to the 2007 imagery to minimize between-image data differences unrelated to land use change. These techniques included regression-based normalizations (Jensen, 2005) where data values for pseudo-invariant pixel features (~32), such as concrete strips, roadways, and deep water bodies, were identified within 2007 and a prior image date.

A standard linear regression was applied to the pseudo-invariant feature data to obtain the regression equation used to normalize a prior image data set to the 2007 data. A two-sample t-test for means, based on an alpha value of 0.05, was used to compare prior and “corrected” data to determine if improvement through normalization had occurred. The t-test consisted of 22 data points representative of pseudo-invariant features within the same general locations as those used to generate the regression model. Dark object subtraction techniques (Song et al., 2001) were also used to reduce the discrepancy between image data prior to the regression-based normalization, if it was observed that these techniques improved overall normalization

Table 1
Landsat satellite image dates.

Year	2004	2005	2006	2007
Date 1	9 Jul	12 Jul	31 Jul	15 May
Date 2	27 Sep	6 Sep[a]	1 Sep	11 Aug[a]

[a] Denotes that image is Landsat scan-line-corrector-off ETM+ instead of TM.

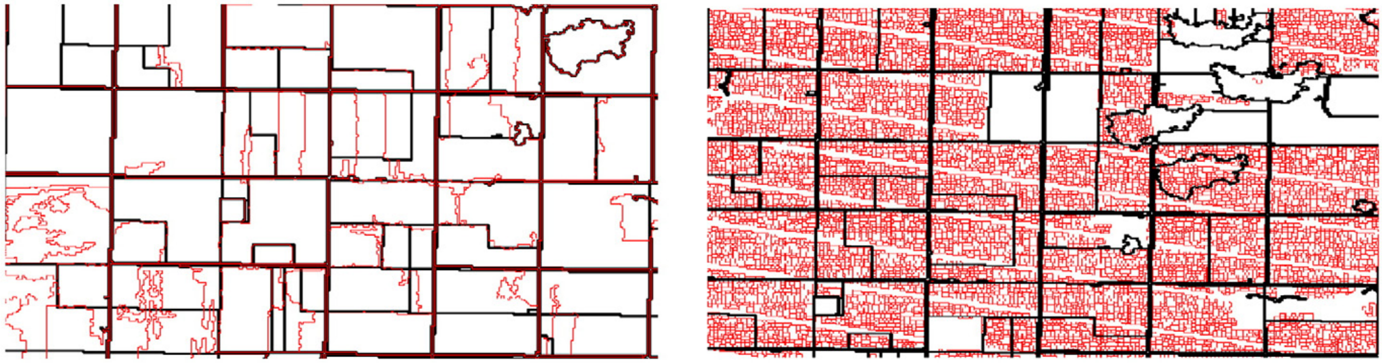

Fig. 3. Object segmentation results (red vector lines) for parcel management strips (left) and within-strip sections (right). Black parameter lines represent taxable field boundaries. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

results. In most cases the means for the normalized and the 2007 base data differed significantly ($\alpha = 0.05$) but showed improvement as the resulting p-value, or the probability of obtaining a test statistic value at least as extreme as the one observed, decreased following the normalization.

An image mask was created to exclude non-agricultural land from the project area. Spatial data pertaining to non-agricultural areas were collected from the Natural Resource Information Service (NRIS, 2007) and included water bodies, wetlands, transportation systems, public lands, and cities. A 15-m radius buffer was applied to point and line features to ensure that the spatial extent was at least 30 m, the Landsat pixel width. The vector-based features were converted into a raster-based template that was used to recode non-agricultural image data to zero.

Spatial data provided by the MFSA were used to exclude rangeland from the study area, while the MFSA-based CRP data were also used to remove known CR land from the cropland tillage and crop/fallow classifications. We recognized the potential for introducing aggregated error into our analysis by including multiple vector layers (each with their own degree of error) but were not able to quantify the extent of this error or how it might have affected the classification results.

Image segmentation was conducted within Definiens Professional Earth LDH O-O software, using the multi-resolution segmentation algorithm (Benz et al., 2004). Definiens Professional was chosen as it provides segmentation, object analysis, and vector export within one package. Two segmentation strategies were used, representing parcel management strips and within-strip sections of spectral and textural similarity (Fig. 3). The within-strip segmentation was used to reduce the inclusion of both crop and bare soil within an image object. A strip-based segmentation was determined to be suitable for tillage and CR classifications, as it was unlikely that these management types would vary within field-based boundaries. Vector information representing taxable field parcels, provided by the Montana Department of Administration, was also included within the segmentation process to ensure that generated objects were constrained within ownership boundaries.

Initial segmentations were applied to a May 2007 Landsat TM image. The image segmentation parameters were determined heuristically. Masked areas within taxable parcels were treated as "no data", thus object-based results were generated using only the "non-masked" data. The resulting image objects were utilized as a vector-based template for the remaining image segmentations. The use of object templates served to produce objects with identical footprints. A spatial joint was used to add the object-based data to the reference management data set for use in the model-building process.

2.5. Image classification

2.5.1. 2007 image classifications

The randomForest package (S-Plus®) was used to generate classification models for NT and till, crop and fallow, and CR and cropland. Object-based spectral, textural, and neighborhood parameters were included within the model-building process (Table 2). The Normalized Difference Vegetation Index (NDVI) (Tucker & Sellers, 1986) and Tasseled Cap components (Crist et al., 1986; Huang et al., 2002) were also included as predictors, as the addition of these indices might better allow for model node splitting. In total there were 121 predictors included into the model-building process, representing object-derived data across the seven Landsat bands, NDVI, and tasseled cap components. Initial forest models were built using 500 generated classification trees, the default number. Adjustments to the

Table 2
Object-based predictive parameters generated through image segmentation within the Definiens Professional software package.

Object data features	Description
Layer mean	The average of all pixels found within an image object. Includes computations for Blue, Green, Red, NIR, MIR, Thermal, MIR2, NDVI, Brightness, Greenness, and Wetness.
Layer standard deviation	The standard deviation calculated from all pixels within an image object.
Min-pixel value	The lowest pixel value within an image object.
Max-pixel value	The highest pixel value within an image object.
Mean difference to brighter neighbors	The layer mean difference computed for each neighboring object, weighted with regard to between-object border length or the area covered by the neighbor objects. A distance of zero is given for direct neighbors.
Contrast to neighborhood pixels	Calculates the mean difference between pixel values and surrounding pixel values.
GLCM homogeneity—all directions	A texture measure concerning the amount of local variation within an image object.
GLCM mean	A texture measure where the pixel values are weighted by the frequency of their occurrence in combination with neighbor pixel values.
GLCM standard deviation	A measure of the dispersion of values around the texture mean.
GLCM dissimilarity	A texture measure of the amount of local variation within the image object; values increase linearly and dissimilarity will be high if there is high contrast within a localized region.
GLCM contrast—all directions	A measure similar to the GLCM mean, differed by all spatial directions (0, 45, 90, 135) being summed prior to inclusion within the texture calculations.

GLCM is the grey level co-occurrence matrix and is a tabulation of how often different combinations of pixel grey levels occur within a given object.

Table 3
Classification (OOB) accuracy for tillage.

Model	Overall accuracy (%)		Classification matrix		Producer's (%)	User's (%)
No-till and tillage			NT	Till	NT	Till
May pixel-based	76	NT	161	13	92	79
		Till	43	13	23	50
May object-based	71	NT	160	14	91	71
		Till	63	29	31	67
August object-based	57	NT	101	35	74	61
		Till	64	31	33	47
May + August object-based	54	NT	93	43	68	60
		Till	63	32	34	43

number of classification trees included within each model were based on an analysis of model error as influenced by the number of trees. This error-based-on-number-of-trees feature is included within RF as a diagnostic measure and allows the user to see at what number of classification trees total model accuracy starts to decrease. Classification matrices and associated class accuracies were determined through the internal OOB accuracy assessment (Breiman, 2001).

The data sets examined for model generation included those from individual image dates (early summer; late summer) or predictive parameters incorporating both image dates. Pixel-based models were also examined, in addition to the object-based models, to ascertain any improvements in classification accuracy that might be attributed to the inclusion of object-based textural and neighborhood parameters into the model. Data sets were then run through the generated RF models to provide object classifications for the 2007 season. Class predictions were exported and joined with the existing vector objects according to field object identification numbers.

2.5.2. Multi-year classification to determine crop intensity

Fall images from 2004–2006 also were classified according to crop and fallow practices in order to determine 4-year crop intensity spanning from 2004–2007. The multi-date crop and fallow classifications occurred only within the four study area subsets to reduce the computation time associated with image-object generation (Fig. 2). Resulting vegetation pattern data allowed for the determination of field-based crop intensity. An NDVI threshold method was used for the detection of “unchanged” image objects, or those having the same class type (fallow or cropped) between the 2007 and a prior image date. Between-year vegetative change within the study area is primarily due to crop–fallow rotations; areas of no change, or fields under continuous cropping, are relatively infrequent. “Unchanged” areas, or those cropped in both the 2007 and prior image dates, were determined using object-based data from the May 2007 and July 2004–2006 images as these dates provided the best distinction between photosynthetically active (cropped) and inactive (fallow) field areas. The change analysis consisted of object-based spectral means for the red and near-infrared (NIR) bands. These bands were chosen to detect management changes as they are sensitive to photosynthetic activity. Photosynthetically active vegetation strongly absorbs energy between 0.63 and 0.69 µm (red) and strongly reflects energy between 0.7 and 1.2 µm (NIR) (Jensen, 2005).

Change vectors (CV) were generated according to a distance measure based on the Pythagorean Theorem (Eq. (1)), where Red_1 and Red_2 represent multi-year object-mean values for the red portion of the measured electromagnetic spectrum and NIR_1 and NIR_2 represents multi-year values for the near-infrared portion of the spectrum.

$$\mathrm{CV} = \left((\mathrm{Red}_1 - \mathrm{Red}_2)^2 + (\mathrm{NIR}_1 - \mathrm{NIR}_2)^2\right)^{0.5} \tag{1}$$

Objects representing possible management changes between image dates were visually identified through a comparison of image data and were used as reference points in the determination of appropriate thresholds to separate “changed” from “unchanged”. No clear threshold values were identified through the analysis of CV values from these reference object areas. A more apparent change threshold value was determined within the between-date NDVI difference values, as was also reported by Lyon et al. (1998). Training data for cropped were taken from randomly selected “no-change” objects. “Cropped” class data were also obtained from randomly selected “changed” objects identified based on a visual examination of the data. This step was needed as areas of “no-change” or continuous crop are not often practiced within the general study area, but might be more common in localized areas. Training data for the fallow class were also based on randomly selected “changed” objects because areas of “fallow–fallow” were not identified. These data were used to generate the RF models for the 2004–2006 image-object classifications.

Training object data used to create the 2006 model included 85 cropped and 126 fallow sites, object data used to create the 2005 model included 73 cropped and 70 fallow sites, and 111 cropped and 112 fallow sites were used in the 2004 model. The final number of data points used in the generation of each model were based on results from a random object selection where 230 locations were generated within the study region; object locations that did not fall within areas of masked or “no data” were used to train the models. Areas of “no data” within the Landsat ETM+ scenes also resulted from data gaps due to scan-line corrector failure. The resulting classifications were merged with the “unchanged” object classifications to produce a final class layer. A database analysis was used to determine the four-year cropping patterns.

3. Results

3.1. Tillage type

The highest classification accuracies for tillage type were generated using the May 2007 data set and included information from 113 NT–cropped, 61 NT–fallow, 70 tilled–cropped, and 22 tilled–fallow field locations. This model consisted of 500 trees, with 13 variables examined at each node split. The 13 selected variables were those deemed most important by the RF model, from out of the total number of variables included within the data set. Total model accuracy was 71%, with low user's and producer's accuracies (31% and 67%, respectively) in the tillage class (Table 3). The most important variables for the classification, as observed within the variable importance plots, were the object-mean value for wetness followed by the object standard deviation for greenness. The 2007 August-based model was more likely to misclassify NT–cropped sites as tilled. Lower tillage class accuracies were observed for the 2007 May pixel-based model, although the pixel-based model resulted in the lowest overall model error (~21%). This model included data from 118 NT–cropped, 56 NT–fallow, 34 tilled–cropped, and 22 tilled–fallow sites; 450 trees were included within the forest and 13 variables were examined at each node split.

Table 4
Multi-year classification (OOB) accuracy for crop and fallow, used in determining 4-year crop intensity.

Model	Overall accuracy (%)		Classification matrix		Producer's (%)	User's (%)
Crop and fallow			Crop	Fallow	Crop	Fallow
August object-based 07	91	Crop	178	10	95	93
		Fallow	12	55	82	84
August pixel-based 07	82	Crop	167	21	88	87
		Fallow	24	43	64	67
May + August object-based 07	77	Crop	149	14	91	79
		Fallow	39	29	43	67
May object-based 07	67	Crop	145	18	89	71
		Fallow	58	10	15	36
September object-based 06	96	Crop	80	5	94	95
		Fallow	4	122	96	96
September object-based 05	93	Crop	68	5	93	93
		Fallow	5	65	92	92
July object-based 04	93	Crop	103	8	93	93
		Fallow	7	105	93	93

3.2. Crop and fallow

The August-based RF model was able to distinguish senesced crop from fallow with greater than 82% accuracy (Table 4); the pixel-based model did not yield higher classification accuracy using the August dataset. Data from 188 cropped and 67 fallow sites were used in building the object-based model. The RF variable importance plot indicated that object textural measures such as within-object contrast and homogeneity were often used as model predictive parameters, suggesting that object-derived information allowed for greater predictive ability under certain conditions. Textural measures related to object-based wetness were also found to influence model accuracy, with wetness being greater on cropped surfaces. Misclassification errors within the fallow category were attributed to objects located within landscapes characterized by narrow (<100-m wide) crop and fallow strip management. This was likely due to the within-pixel mixing of crop and fallow spectral signatures. The object-based classification tended to favor the “cropped” class, resulting in a classification bias under these conditions.

3.3. Conservation reserve

The greatest classification accuracies pertaining to the discrimination of CR land from small-grains crop were obtained through the May-based model (Table 5). Data utilized in the building of this model included 304 cropland (95 till and 209 NT) and 127 CR sites. Total model accuracy was 99%, with 100% producer's accuracy in the cropland class (96% user's) and 90% in the CR class (100% user's). Classification error primarily resulted from the misclassification of CR as NT–cropped and tilled–crop. This model consisted of 500 trees, with 13 variables types examined at each node split. Predictive parameters found to be the most important for model generation included the non-directional grey level co-occurrence (GLCM) standard deviation for brightness, followed by the mean object value for thermal, the minimum pixel value for green, the object standard deviation for blue, and the GLCM mean for red. The GLCM is a measure of texture within pixel gray levels within a certain object (Herold et al., 2003).

4. Discussion

The objective of this study was to identify cropland management types defined, in part, by anthropogenic boundaries stemming from land ownership. A pixel, when examined as a single unit, does not inherently provide information concerning landscape patterns and spatial relationships. Instead, when landscape objects of interest are distinct and at a scale where they consist of a large number of pixels, it makes more sense to examine an image based on localized spatio-temporal characteristics (Burnett & Blaschke, 2003).

Image segmentation allowed us to incorporate landscape patterns into our analysis by deriving spatial, morphological, and contextual information according to individual cropland field locations (Navulur, 2007). Object-based information representing the field units of interest could then be incorporated into the classification process, in addition to spectral information. Landscape patterns resulting from varied cropping management within ownership parcels could also be used to define individual sub-object units for image analysis and classification purposes.

It was expected that the object-based classification of tillage type, CR, and crop and fallow would yield high class accuracies through the addition of object-based parameters and use of the RF algorithm. Classification success, however, was varied.

4.1. Tillage type

Past studies have reported difficulty in the distinction of tillage from NT in situations where the soil surface is covered by established crop canopy and plant residues. One pixel-based study applying logistic regression to late June TM data successfully classified NT under crop canopy with a 99% producer's accuracy, but failed to adequately classify tillage (29% producer's) (Bricklemyer et al., 2006). It was expected that the object-based approach utilized within this study, aided by the RF algorithm, would yield higher classification accuracies than those generated through the logistic regression approach. The O-O classification, however, produced results very similar to those reported by Bricklemyer et al. (2006). The tillage producer's accuracy was unacceptable for mapping purposes (31%), although the NT producer's accuracy was acceptable (91%).

The RF algorithm theoretically should outperform logistic regression, as it is more robust to noise and outliers (Guo et al., 2004). Studies have demonstrated the superiority of RF to logistic regression

Table 5
Classification (OOB) accuracy for crop and CR.

Model	Overall accuracy (%)		Classification matrix		Producer's (%)	User's (%)
Crop and CR			Crop	CR	Crop	CR
May object-based	97	Crop	304	0	100	96
		CR	12	115	90	100
May + August object-based	91	Crop	87	19	82	88
		CR	12	219	95	92
August object-based	88	Crop	99	27	79	85
		CR	17	214	93	89
May pixel-based	86	Crop	159	22	88	84
		CR	30	158	84	87

and other popular algorithms in classifying a variety of data sets (Hothorn & Lausen, 2005; Lee et al., 2005; Qi et al., 2006). The pixel-based RF model yielded a NT producer's accuracy of 92%, with only a 23% user's accuracy for tillage. This suggests that the lower classification accuracies generated by the RF approach compared to the Bricklemyer et al. (2006) study might be attributed to differences between study data sets. The May data used within this study should have allowed for the increased detection of surface disturbances as vegetative canopy was lower than those reported within the late June data. The study site information incorporated into the RF-based models might also have reflected a greater degree of variability within regional cropland management as it reflected a larger number of study site locations than those used by Bricklemyer et al. (2006). It has also been noted that the survey-based data collection methods utilized by the Bricklemyer et al. (2006) study might have introduced a slight bias towards NT sites, possibly resulting in a bias within the classification model.

The inability of RF to separate NT from conservation-based tillage given May data-based spectral, textural, and object neighborhood information suggests that there is substantial overlap between class feature characteristics. The May data should have minimized vegetation-based spectral similarities between NT and conservation tillage, as senesced crop canopy was avoided. We believe that a large part of the class confusion was a function of the wide use of conservation tillage within the region. Similarities between NT and other forms of conservation tillage likely stem from little difference in crop surface residue amounts between class types. Conservation tillage, by definition, leaves at least 30% of the soil surface covered by crop residue (CTIC, 2004). NT is an extreme form of conservation tillage. Alternative management types might include the use of V-blades (Kuepper, 2001), chisel ploughs, or light duty tandem disks to prepare the soil prior to planting or for weed management on fallow land (Uri, 2000). These tillage types result in a greater amount of soil disturbance compared to NT but do not fully mix surface residues into the soil as would moldboard ploughs or heavy duty tandem disks, or repeated operations of chisel ploughs and/or light duty tandem disks, thus preserving surface residue amounts.

Managers might decide to incorporate minimal tillage methods, as opposed to strictly adhering to a NT system, in the event that surface residue density begins to impede the proper function of planting equipment or to control chemically tolerant weeds in the hot droughty summer and fall periods. This observation was confirmed through verbal communication with area farmers during the 2007 field data collection period. Many managers had reported that contracted combine crews had set cutting lengths too tall in previous years, resulting in the need to reduce surface residue amounts through tillage or burning prior to planting crops for the 2007 season.

4.2. Crop and fallow

It has been shown that crop and fallow can be accurately classified within Landsat ETM+ imagery using simple NDVI differencing techniques (Xie et al., 2007), as were used to detect objects continuously cropped between 2004 and 2007; however, the incorporation of multiple spectral and object-based data parameters into a RF classification approach allowed the identification of crop and fallow in a post-photosynthetic state. This is of value for years when cloud or smoke cover limits the availability of image scene dates. Imagery collected too early might not allow for the detection of early-emergence spring crops, while those collected late summer or early fall would not provide spectral signatures from active crop vegetation, especially in a dryland setting.

RF model accuracies for the classification of crop vs. fallow ranged from 92 to 96%, which were similar to those reported by Xie et al. (2007) (~93% total accuracy) using a spectral angle mapping algorithm and a NDVI threshold approach. That the RF model was able to produce results similar to the Xie et al. (2007) study which used pre-senesced vegetation data, despite the use of data representing post-photosynthetic vegetation, suggests that the incorporation of textural and neighborhood parameters into the classification model increases the classifier's predictive abilities. Classification error was greater in the pixel-based classification, especially for the fallow class (18% greater producer's, 17% greater user's). This was in agreement with expectations that a model based solely on spectral parameters, without added object-based data such as texture, might have greater difficulty in distinguishing between fallow lands characterized by surface stubble (in NT fields) and recently harvested (cropped) fields in a post-senescence setting. The presence of weeds might have contributed to some degree of classification error, wherein a field might have been misclassified as cropped as opposed to fallow, but we were unable to quantify this error as we were not able to obtain these data. Using an object-based approach, however, should have been advantageous as the majority of pixels within an object-area would have to have been weed infested for the resulting object NDVI value to be characteristic of a highly vegetated (cropped) area.

The use of a threshold to separate "unchanged" (crop–crop) from "changed" (crop–fallow or fallow–crop) objects within NDVI difference values provided a direct way to determine change in crop-vegetation status between the May 2007 and prior image dates, despite a temporal difference of more than a month between the reference and preceding image dates. The use of May data, however, restricted the ability to detect late planted spring wheat crop that was still in periods of early emergence. It is recommended that late June or early July imagery be used when possible to avoid this situation.

4.3. Conservation reserve

A RF O-O classification based on May TM data was able to successfully separate CR from cropland with producer's accuracies of 90% and 100% (Table 5). Previous pixel-based studies had relied on more elaborate multi-year techniques to achieve similar accuracies (Egbert et al., 2002; Price et al., 1997); the pixel-based model accuracy achieved in this study was the lowest of all generated models (86%). Classification error within the May-based model primarily resulted from the misclassification of CR as NT–cropped and tilled–crop. The misclassified sites were often those under recent conversion from cropland to CR, as was determined by an examination of data supplied through the MFSA. It was also observed that the misclassification rate was greater at late vegetative maturity (late August), suggesting that there are greater spectral and textural differences between the management classes during early stages of growth.

An analysis of model predictive parameters used in the data splitting process suggests that object-based texture played a role in the discrimination of crop from CR. Both the textural standard deviation for brightness and the textural mean for red demonstrated some degree of importance in the predictive variable plot, a measure by RF showing the relative degree to which predictor variables influenced classification accuracy. The high standard deviation in the texture for brightness likely resulted from greater variation within cropland surface albedo due to patches of soil exposure, straw stubble, and vegetation than might be expected at CR locations, which generally feature a more uniform grassland surface.

Increased object texture exhibited within the red portion of the spectrum might have resulted from unevenness in soil exposure throughout site locations, possibly driven by the incorporation of spectral signatures from both cropped and fallowed strips within some objects as well as patterns resulting from crop row spacing. CR sites were also characterized, on average, by a higher mean object thermal value than cropped sites. It is suspected that this might have resulted from higher evapotranspiration at the CR sites due to higher photosynthetically active plant biomass densities. This observation is given further support as the average minimum pixel value within each object was found to be greater for CR sites than for cropland sites, also indicative of greater vegetative coverage. The average standard deviation for object reflectance in the blue band was also slightly higher for CR sites than cropland fields, possibly due to differences in surface moisture and evapotranspiration.

5. Conclusion

Our results indicated that the incorporation of object-based parameters into a RF model has the ability to distinguish cropland from grassland-based CR using image data collected at early stages of vegetative growth (~May). Study results also showed the ability of an object-based RF model to separate crop from fallow within a dryland, post-photosynthesis landscape. We recommend that remote sensing might be used successfully for the validation and monitoring of grassland-based CR and crop intensity within north central Montana for carbon contract purposes. Future studies should be aware of possible problems resulting from the misclassification of CR as cropland in fields under recent conversion to CR. Future studies might also avoid the use of moderate resolution imagery (~30 m) in agricultural landscapes where narrow (<100 m) crop strip management patterns are used, as misclassification was often common within these areas due to spectral mixing.

We were unable to adequately separate NT from conservation tillage management using Landsat-based O-O analyses, in conjunction with RF classifications. We foresee difficulty in discovering a contracted farmer who has agreed to follow NT practices but is in fact practicing a less extreme form of conservation tillage (reduced tillage as opposed to NT). Physical validation of tillage management within carbon contract sites will likely be necessary if the current classification scheme (NT vs. all levels of tillage) continues to be used for validation purposes, complimented by satellite-based classifications to detect heavy tillage disturbances.

We suggest that future research continue to incorporate object-based approaches to classify cropland practices for carbon contract validation purposes as these methods allowed for the analyses to occur on a per-field basis. Future research must also continue to investigate the classification strategies that are currently used to separate tillage management types, determining if a binary "NT vs. till" approach is appropriate from a spectral and textural standpoint. This suggestion is made as study results showed great similarity in surface spectral characteristics (specifically surface residues) between NT and sites thought to be under conservation-based tillage management. Alternatively, greater emphasis in future studies might be made on using relative surface residue amounts to indicate degrees of tillage usage and field disturbance. The incorporation of higher temporal resolution data, should it become available at the requisite spatial resolution, could be beneficial as the timing of practices in NT and conservation tillage scenarios differs.

Acknowledgements

This project was funded by the U.S. Department of Energy, 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.

References

Baker, C., Lawrence, R., Montagne, C., & Patten, D. (2006). Mapping wetlands and riparian areas using Landsat ETM+ imagery and decision-tree-based models. *Wetlands*, *26*, 465−474.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, *58*, 239−258.

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

Breiman, L., & Cutler, A. (2004). Random forests. The Random Forests User's Manual. Last Accessed 29 March 2009 <http://www.stat.berkeley.edu/~breiman/RandomForests.htm>

Bricklemyer, R. S., Lawrence, R. L., & Miller, P. R. (2002). Documenting no-till and conventional till practices using Landsat ETM+ imagery and logistic regression. *Journal of Soil and Water Conservation*, *57*, 267−272.

Bricklemyer, R. S., Lawrence, R. L., Miller, P. R., & Battogtokh, N. (2006). Predicting tillage practices and agricultural soil disturbance in north central Montana with Landsat imagery. *Agriculture, Ecosystems and Environment*, *114*, 210−216.

Burnett, C., & Blaschke, T. (2003). A multi-scale segmentation/object relationship modeling methodology for landscape analysis. *Ecological Modeling*, *168*, 233−249.

Chander, G., Markham, B. L., & Barsi, J. A. (2007). Revised Landsat 5 Thematic Mapper radiometric calibration. *IEEE Geoscience and Remote Sensing Letters*, *4*, 490−494.

Chicago Climate Exchange (CCX) (2008). *Soil carbon management offsets. CCX Guidelines: Agricultural soil carbon.* Chicago, Illinois, USA: CCX 2 pp.

Conservation Technology Information Center (CTIC) (2004). Conservation tillage and other tillage types in the United States (1990–2004). *National crop residue management survey* West Lafayette, Indiana, USA: Conservation Technology Information Center.

Crist, E. P., Laurin, R., & Cicone, R. C. (1986). Vegetation and soils information contained in transformed Thematic Mapper data. *Proceedings of IGARSS '86 symposium, 8–11 September 1986, Zurich, Switzerland* (pp. 1465−1470).

Daly, P. (2001). Remote sensing for crop assessment. *A report to the Franklin Conservation District, Pasco WA* 4 pp.

Egbert, S. L., Park, S., Price, K. P., Lee, R., Wu, J., & Duane Nellis, M. (2002). Using conservation reserve program maps derived from satellite imagery to characterize landscape structure. *Computers and Electronics in Agriculture*, *37*, 41−156.

Farkell, J. (2007). Personal Interview. Control Agricultural Consultant. Tillage and cropping systems near Conrad, Montana.

Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random forests for land cover classification. *Pattern Recognition Letters*, *27*, 294−300.

Gowda, P. H., Peters, R. T., & Howell, T. A. (2005). Mapping tillage practices in the Texas Panhandle with Landsat Thematic Mapper (TM) data. USDA-ARS Conservation and Production Research Lab. Wetting Front News, 7, 5–7.

Guo, L., Ma, Y., Cukie, B., & Singh, H. (2004). Robust prediction of fault-proneness by random forests. *Proceedings of the 15th international symposium on software reliability engineering, 2–5 November 2004, Saint-Malo, France* (pp. 417−428).

Ham, J., Chen, Y., Crawford, M. M., & Ghosh, J. (2005). Investigation of the random forest framework for classification of hyperspectral data. *International Journal of Remote Sensing*, *243*, 492−500.

Hay, G. J., & Castilla, G. (2006). Object-based image analysis: strengths, weaknesses, opportunities and threats (SWOT). *International archives of photogrammetryRemote Sensing and Spatial Information Sciences*, *vol. 36*. 3 pp.

Herold, M., Liu, X., & Clarke, K. C. (2003). Spatial metrics and image texture for mapping urban land use. *Photogrammetric Engineering and Remote Sensing*, *69*, 991−1001.

High Plains Regional Climate Center (HPRCC). 2008. Climate summary maps and archived station data. Last Accessed 22 April 2008 < http://www.hprcc.unl.edu>

Hothorn, T., & Lausen, B. (2005). Bundling classifiers by bagging trees. *Computational Statistics and Data Analysis*, *49*, 1068−1078.

Huang, C., Wylie, B., Yang, L., Homer, C., & Zylstra, G. (2002). Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance. *International Journal of Remote Sensing*, *23*, 1741−1748.

Jensen, J. R. (2005). *Introductory digital image processing: A remote sensing perspective.* Upper Saddle River, New Jersey, USA: Pearson Prentice Hall 526 pp.

Kamagata, N., Hara, K., Mori, M., Akamatsu, Y., Li, Y., & Hoshino, Y. (2006). A new method of vegetation mapping by object-based classification using high resolution satellite data. *Journal of the Japan Society of Photogrammetry and Remote Sensing*, *45*, 43−49.

Kuepper, G. (2001). *Pursuing conservation tillage systems for organic crop production.* Butte, Montana, USA: National Center for Appropriate Technology 28 pp.

Lawrence, R., Bunn, A., Powell, S., & Zambon, M. (2004). Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote Sensing of Environment*, *90*, 331−336.

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.

Lawrence, R. L., & Wright, A. (2001). Rule-based classification systems using classification and regression tree (CART) analysis. *Photogrammetric Engineering and Remote Sensing*, *67*, 1137−1142.

Lee, J., Lee, J., Park, M., & Song, S. (2005). An extensive comparison of recent classification tools applied to microarray data. *Computational Statistics and Analysis*, *48*, 869−885.

Lucas, R., Rowlands, A., Brown, A., Keyworth, S., & Bunting, P. (2007). Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping. *ISPRS Journal of Photogrammetry and Remote Sensing*, *62*, 165−185.

Lyon, J. G., Yuan, D., Lunetta, R. S., & Elvidge, C. D. (1998). A change detection experiment using vegetation indices. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*, 143−150.

National Carbon Offset Coalition (NCOC). (2008). Exchange soil offset eligible practices and offset issuance rates by zone. Cropland Protocol Guidelines. 23 pp.

National Weather Service (NWS) (2007). *Past weather local climate information.* National Oceanic and Atmospheric Administration.

Natural Resources Information Service (NRCS). (2007a). Major biomes map. Last Accessed 10 March 2009 <http://soils.usda.gov/use/worldsoils/mapindex/biomes.html>

Natural Resources Information Service (NRCS). (2007b). Interactive soil survey map. Last Accessed 15 March 2009 <http://websoilsurvey.nrcs.usda.gov/app/WebSoilSurvey.aspx>

Natural Resources Information Service (NRIS). (2007). Collective spatial information clearinghouse for Montana. Last Accessed 11 February 2009 < http://nris.mt.gov/gis/default.asp>

Navulur, K. (2007). *Multispectral image analysis using the object-oriented paradigm.* Boca Raton, FL: CRC Press 165 pp.

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, *26*, 217−222.

Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, *9*, 181−199.

Price, K. P., Egbert, S. L., Nellis, M. D., Lee, R. Y., & Boyce, R. (1997). Mapping land cover in a High Plains agro-ecosystem using a multi-date Landsat Thematic Mapper modeling approach. *Transactions of the Kansas Academy of Science*, *100*, 21−33.

Qi, Y., Bar-Joseph, Z., & Klein-Seetharaman, J. (2006). Classification methods for use in protein interaction prediction. *Proteins, Structure, Function, and Bioinformatics*, *63*, 490−500.

Science Data Users Handbook—Landsat 7 (SDH-L7). (2006). National Aeronautics and Space Administration. Last Accessed 10 March 2009 <http://landsathandbook.gsfc.nasa.gov/handbook/handbook_toc.html>

Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P., & Macomber, S. A. (2001). Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sensing of Environment*, *75*, 230−244.

Song, X., Fan, G., & Rao, M. N. (2003). Machine learning approaches to multisource geospatial data classification: application to CRP mapping in Texas County, Oklahoma. *IEEE workshop on advances in techniques for analysis of remotely sensed data, 27–28 October 2003, Greenbelt, MD* (pp. 202−211).

South, S., Qi, J., & Lusch, D. P. (2004). Optimal classification methods for mapping agricultural tillage practices. *Remote Sensing of Environment*, *91*, 90−97.

Stow, D., Lopez, A., Lippitt, C., Hinton, S., & Weeks, J. (2007). Object-based classification of residential land use within Accra, Ghana based on Quickbird satellite data. *International Journal of Remote Sensing*, *28*, 5167−5173.

Tucker, C. J., & Sellers, P. J. (1986). Satellite remote sensing of primary production. *International Journal of Remote Sensing*, *7*, 1395−1416.

Uri, N. D. (2000). Perceptions on the use of no-till farming in production agriculture in the United States: an analysis of survey results. *Agriculture, Ecosystems and Environment*, *77*, 263−266.

Western Regional Climate Center (WRCC). (2006). Historical climate information. Last Accessed 05 April 2007 <http://www.wrcc.dri.edu/CLIMATEDATA.html>

Xie, H., Tian, Y. Q., Granillo, J. A., & Keller, G. R. (2007). Suitable remote sensing method and data for mapping and measuring active crop fields. *International Journal of Remote Sensing*, *28*, 395−411.