MAPPING SAGEBRUSH/GRASSLANDS FROM LANDSAT TM-7 IMAGERY: A COMPARISON OF METHODS

Long Draw Creek *BLM Bird Crew Photo*

Prepared for the
USDI, Bureau of Land Management
Montana/Dakota State Office
P.O. Box 36800
Billings, MT 59107-6800

By
Chip Fisher, Will Gustafson, and Roland Redmond
Wildlife Spatial Analysis Lab
Natural Science 205
The University of Montana
Missoula, MT 59812

11 December 2002

ABSTRACT

We compared three different remote sensing techniques for mapping rangeland vegetation, and sagebrush in particular, from Landsat TM imagery. The first two methods involved supervised classifications of regions derived from two different image segmentation methods (SILC3 vs eCognition). The implementation of these two methods differed in two fundamental ways. The SILC3 segmentation came directly from a previous classification of a July 2000 TM image; whereas we used a September 1999 image for the eCognition segmentation. Second, the pixels from the September 1999 image were resampled from 30 m^2 to 15 m^2. Both image segmentations then were classified to land cover type, and canopy closure class for sagebrush, using the same training data and classifier(s). The third method involved a supervised classification of pixels based on a principal component analysis and guided clustering. All three techniques were able to classify four sagebrush canopy classes and five xeric shrub species types (four sagebrush plus greasewood), but at an uncertain level of accuracy. Because the 15m^2 resampled imagery (from September 1999) appeared to better capture fine-scale landscape features, such as woody draws and sagebrush stringers, we expected the results from these two methods (eCognition and PCA) would be better than SILC3 which was based on 30 m input data. Nevertheless, acreage tallies and map overlays indicated a generally high correspondence or similarity among the results. We emphasize therefore that without an independent accuracy assessment of all three datasets it is not possible to objectively determine: 1) if one method is substantially better than another, or 2) how these three methods compare with other available ones.

INTRODUCTION

This report provides the results of a comparison of remote sensing techniques for mapping rangeland vegetation, and sagebrush in particular, for the Clark's Fork valley of the Yellowstone River in southern Carbon County, Montana, and northern Park County, Wyoming (Figure 1). Three techniques were evaluated, and all were applied to Landsat TM-7 imagery. The first two involved supervised classifications of raster polygons or regions, derived from two different image segmentation methods. In the first case, the original imagery (30 m^2 pixels) was segmented into raster polygons using a proprietary object and rule-based merging algorithm (Barsness 1996). In the second case, the TM-7 imagery was resampled.to 15 m^2, and then segmented into raster polygons using eCognition – a commercially available software product. Both resulting image segmentations then were classified to land cover type, and canopy closure class for sagebrush, using the same training data and classifier(s). The third method was developed by Heather McClure and David Prevedel at the USDA Forest Service, Intermountain Regional Office in Ogden, UT (McClure and Prevedel 2002). It involved a supervised classification of pixels based on a principal components analysis and guided clustering (McClure and Prevedel 2002). The primary objective of this study was to compare the results from the different methods, and in the process determine which one(s) could, with reasonable accuracy, map sagebrush and greasewood shrub types according to several different canopy closure classes. A secondary objective was to evaluate the ability of the different methods to accurately classify and map greasewood (*Atriplex* spp.) plus four distinct sagebrush species.

STUDY AREA

The study area comprised a portion of Landsat TM-7 scene Path 37/Row 29 (Figure 1). The full scene was clipped to this smaller area to allow us to focus on sagebrush/grasslands that are most abundant in the Bighorn Basin. Furthermore, because relatively few training data were available from the Wyoming portion of the scene, we opted to reduce data volumes and limit the classifications to the northern third of the scene. The actual comparisons reported and illustrated herein apply to an even smaller area centered on the Clark's Fork valley of the Yellowstone River (Figure 1). This valley lies between the Beartooth Plateau and Pryor Mountains in south-central Montana. Climate conditions in the area range from being cold and wet in the Beartooths to very arid in the Bighorn Basin. Vegetation also varies from coniferous forest at higher elevations to sagebrush steppe and short-grass prairie across the valley floor. The predominate land uses are recreation in the mountains and livestock grazing at lower elevations. More details about the study area can be found in Nesser et al. (1997) and Jones (2001).

Figure 1. Study area map

Part of Carbon County was classified in the USFS Region One Eastside Analysis of Management Situation land cover type mapping project (called SILC3) in 2001. One of the mapping techniques in the comparison was a re-classification of the 30m SILC3 land cover classification. Carbon County was the only SILC3 classification area based on year 2000 Landsat TM-7 imagery (P37/R29) which also had extensive sagebrush communities. To reduce spectral and ecological variation in the other two techniques, which were based on 15m resampled September 1999 Landsat imagery, only part of the P37/R29 scenes were classified. The seven hundred thousand acre comparison area within the clipped image was selected due to limits in the sagebrush training data extent.

MAP LEGEND CLASSES

Land cover types to-be-mapped were based on prior Satellite Image Land cover Classifications (SILC) carried out by this lab. The most recent classification of this TM scene, SILC3, was completed for the Forest Service in June, 2002. But for this project, we were able to classify and map just a single xeric shrub cover type which was predominately sagebrush. For this study, we expanded the SILC3 legend to include five xeric shrub species types, plus four canopy closure classes for these xeric shrubs. Table 1 provides a complete list of the cover types that were classified and mapped using each of the three methods. Note that all land cover types were mapped across the entire study area using both the SILC3 and eCognition methods. But because the PCA method was designed to be applied to rangeland vegetation only (see below), only a subset of the cover types was actually mapped using this method.

Table 1. Land cover types and canopy closure classes mapped by method

Cover type	Description	SILC3	eCog	PCA
1100	Urban / Developed	X	X	
2010	Agriculture – Dry	X	X	
2020	Agriculture – Irrigated	X	X	
3130	Very Low Cover Grasslands	X	X	X
3150	Low / Moderate Cover Grasslands	X	X	X
3170	Moderate / High Cover Grasslands	X	X	X
3370	Sagebrush / Xeric Shrubs 05-14%	X	X	X
3380	Sagebrush / Xeric Shrubs 15-24%	X	X	X
3390	Sagebrush / Xeric Shrubs 25-34%	X	X	X
3395	Sagebrush / Xeric Shrubs ≥ 35%	X	X	X
3610	Mesic Shrubs / Willow	X	X	X
4101	Aspen	X	X	
4150	Mixed Broadleaf / Cottonwood Forest	X	X	
4203	Lodgepole Pine Forest	X	X	
4204	Whitebark Pine Forest	X	X	
4205	Limber Pine Forest	X	X	X
4206	Ponderosa Pine Forest	X	X	
4212	Douglas-fir Forest	X	X	
4216	Utah Juniper	X	X	X
4223	Douglas-fir / Lodgepole Pine Forest	X	X	
4237	Subalpine Fir / Spruce Forest	X	X	
4241	Mixed Upper Subalpine Forest	X	X	
4242	Mixed Lower Subalpine Forest	X	X	
4244	Mixed Xeric Forest	X	X	
5000	Water	X	X	X
7300	Rock / Barren	X	X	X
9100	Snow	X	X	

Sagebrush and Greasewood Species Types

Species	Description	SILC3	ECog	PCA
3311	Greasewood	X	X	X
3351	Mountain Big Sagebrush	X	X	X
3352	Wyoming Big Sagebrush	X	X	X
3353	Basin Big Sagebrush	X	X	X
3354	Black Sagebrush	X	X	X

LANDSAT TM IMAGERY

Two different Landsat TM-7 images were used in the comparison: one acquired on September 8, 1999 and the other on July 24, 2000. The latter was the same image that we classified previously for the SILC3 project, and for consistency we used it again in this comparison. Because McClure and Prevedel's modified PCA method calls for September TM imagery to best discriminate sagebrush, we purchased the former image from the MRLC Archive (http://edc.usgs.gov/products/satellite/mrlc2000.html) and used it for both the PCA and eCognition classifications. This means that differences between the SILC3 and eCognition (or PCA) classifications could be the result of the different methods, the different imagery, or a combination of both. Whereas any differences between the eCognition and PCA classifications can only be the result of the different methods.

METHODS

Image Resampling

For both the eCognition and PCA methods, the September 8, 1999 TM-7 image was resampled from 30 m^2 to 15 m^2 pixels. Because this resampling quadruples the volume of data to be processed, the full extent of the TM scene was clipped to the study area boundary (Figure 1) prior to resampling. Multispectral data (TM bands 1-5, + 7) were resampled with Erdas Imagine software using the Resolution Merge function, a cubic-convolution analysis window, and the 15 m panchromatic band for contrast. To help identify changes in ground features due to soil reflectance, we also calculated a Normalized Difference Vegetation Index (NDVI) value for each 15 m^2 pixel according to: NDVI = (TM4 - TM3) / (TM4 + TM3).

Image Segmentation

SILC3

A three-step image segmentation process was developed for the SILC3 project (Winne 2000). The first step involved an unsupervised classification of the seven multispectral TM bands using the ISODATA routine in Erdas Imagine. Potential Vegetation (PV) type, obtained from the Forest Service in the form of a 30 m Arc/Info grid, was used to seed the unsupervised classification by extracting statistics from the TM imagery for each PV type, and then using these statistics in a minimum distance classification. The ISODATA routine was run iteratively until each pixel in the image was assigned to one of 130 different spectral classes.

In the second step, fine-scale boundaries and linear features were detected and delineated using a "structure" image derived from a principal component analysis (PCA) of the seven TM bands. PCA was used to maximize spectral variation of the input bands within a single band (the first principal component, PC1). A relative index of spectral uniformity was calculated which identified spectral "hills" and "valleys" and delineated boundaries between ground features. More specifically, mean PC1 values within 3x3 and 7x7 windows were calculated, and the difference between the two means was divided by the standard deviation of the 3x3 window. Pixels that fell in areas relatively uniformly across scales received values near 0, whereas pixels in areas with greater scale differences received higher or lower values. A structure image mask was created to define areas of negative and positive departure from zero and used to guide the first two iterations of the image segmentation process (step 3).

In the third step, pixels from the unsupervised classification (step one) were aggregated into larger spatial units using a rule- and object-based merge process that was run iteratively and with varying threshold values inside and outside the image structure mask (step two). The intent was to capture land cover inclusions and linear features, while at the same time avoiding excessive fragmentation of landscape patches. The minimum map unit varied by iteration and ranged from one to 22 pixels. Finally, any patches or raster polygons larger than 2248 pixels (500 ac or 202.4 ha) were identified and then broken up into smaller units based on spectral class and variation in brightness values of the TM bands associated with its component pixels.

The resulting output file identified 360,043 raster polygons within the study area, ranging in size from 1 to 2248 pixels (0.1 - 500 ac). This file then was imported into ARC/INFO as a 30 m^2 zonal grid. Mean values were calculated from all pixels in each region for all TM bands, including the panchromatic, and NDVI (as calculated above). These values were assigned as separate attributes to each region in the GIS database. Additionally, 7.5 minute DEMs were used to attribute each region according to its mean elevation, mean slope, and majority aspect. Further details may be found in Ma et al. (2001).

eCognition

The 15m^2 Landsat TM and NDVI layers were loaded into eCognition along with a study area boundary layer. The study area boundary layer limited the image segmentation to only areas within it. The eCognition software uses a proprietary algorithm to delineate ground features based on the spectral characteristics of individual pixels and the shape of expanding regions. A synopsis of the segmentation method from the company's website follows below; more details can be found in the User's Manual (Definiens 2002).

"Segmentation in eCognition is a bottom up region-merging technique starting with one-pixel objects. In numerous subsequent steps smaller image objects are merged into bigger ones. The procedure simulates an even and simultaneous growth of segments over a scene in each step. It starts at an arbitrary point in the image with one-pixel objects. The algorithm guarantees a regular spatial distribution of treated image objects. The underlying patented algorithm is essentially a heuristic optimisation procedure, which minimizes the average heterogeneity of image objects for a given resolution over the whole scene. Heterogeneity itself is based not only on the standard deviation of image objects but also on their shape. Weighting between spectral and shape heterogeneity enables an adjusting of segmentation results to the considered application. The stop criterion for the region-merging process is given by the parameter 'scale' and can be edited by the user. It determines the maximum allowed overall heterogeneity of the segments."

We generated a series of segmentations by adjusting the parameters of scale, band weights, color, and shape. The results of each were evaluated by comparing them to digital orthophoto quads and ground reference data. When scale was set to the default value of 10, approximately 569,000 regions were produced, with an average size of 9 acres. Scale values smaller than 10 resulted in segmentations with more than 1,000,000 regions; we felt these were too many to be practical. On the other hand, scale values greater than 10 resulted in the loss of narrow, linear sagebrush features, as well as small sagebrush patches interspersed with grasslands. Thus, we ended up using the default scale value of 10. Because sagebrush patches appeared to be more visible from TM bands 3, 4, 5, & 7 than in bands 1 and 2, we differentially weighted these bands; 3, 4, 5, and 7 were assigned weights of 1.0, and bands 1 and 2 were assigned weights of 0.5. Similarly the NDVI layer was assigned a weight of 2.0 because of the high level of soil influence on reflectance values. Segmentations based on higher color weights and lower shape weights appeared to match imagery and ground features better than ones based on lower color and higher shape weights. Finally, the shape parameter had two sub-weights for smoothness and compactness, and we found that higher smoothness weights and lower compactness weights produced segmentations that best captured narrow, linear sagebrush stands. The final image segmentation was based the following parameter settings:

SCALE	TM1	TM2	TM3	TM4	TM5	TM7	NDVI	COLOR	SHAPE (sm/cp)
10	0.5	0.5	1.0	1.0	1.0	1.0	2.0	0.8	0.2 (0.9/0.1)

The resulting output file identified 549,246 regions within the study area, ranging in size from 2 to 3140 pixels (0.1 – 175 ac). As with the SILC3 process, this file was imported into ARC/INFO as a 15 m^2 zonal grid. Mean values were calculated from all pixels in each region for all TM bands, including the panchromatic, and NDVI (as calculated above). These values were assigned as separate attributes to each region in the GIS database. Additionally, 7.5 minute DEMs were resampled to 15 m and used to attribute each region according to its mean elevation, mean slope, and majority aspect.

Supervised Classification

SILC3 and eCognition

A similar supervised classification process was applied for both the SILC3 and eCognition methods. In each case, the segmented regions were assigned land cover type labels using a combination of manual and supervised classifications. The manual classifications were performed by on screen inspection of the imagery and direct assignment of cover type labels to the following classes: urban or developed lands, agricultural lands (both irrigated and non-irrigated), recent forest burns, water, and mines. The manually assigned labels were filled into the MANLABEL attribute.

Training data files were prepared from essentially the same ground-reference data for the supervised classifications of both the SILC3 and eCognition regions. The ground-reference data came from a variety of sources, including the USDA Forest Service, USDA Natural Resources Conservation Service, USDI Bureau of Land Management, USDI Bureau of Indian Affairs, USDI Fish and Wildlife Service, and Montana Department of Natural Resources and Conservation. In the original SILC3 product completed for the Forest Service, sagebrush and xeric shrublands were classified as a single type and no canopy cover class was assigned. If these training data did not have sufficient information about species or canopy closure to allow them to be used for the more detailed classifications of this study, they were omitted from the training data file.

The training data were analyzed initially for positional accuracy, attribute accuracy, and general life form agreement with the unclassified Landsat TM-7 imagery. To reduce the classification error among the land cover types, the training data were subjected to a sequence of leave-one-out, cross-validation classifications using Dudani's distance-weighted classifier with a nearest-neighbor size (K-NN) between 10 and 15, and the mean inverse distance (MID) spatial adjustment (Steele 2000, Steele and Redmond 2001, Steele and Patterson in press). Outliers were identified for each cover type by examining plots in relation to TM and ancillary data for their respective regions (both visually and in relation to calculated standard deviations). The majority of these outliers were removed from the training dataset. In some cases the x-y location of the training record could be moved one or two pixels to place it in a more representative region. Examples of outliers include training data with: 1) life form confusion, 2) very low cover for a particular life form, and 3) mean Euclidean distances for spectral variables that were much higher than the means for the cover type group.

An initial classification of 20 land cover types was carried out using Dudani's distance-weighted classifier, a nearest-neighbor size of 15, and the mean inverse distance adjustment. The results were mapped and field checked for both methods, and additional training data were collected for the sagebrush and grassland types. These new field data were added to the training data sets, inspected for agreement with the imagery, and checked for outliers via cross-validation. A second supervised classification of the same 20 land cover types was run using the same classifier and parameters. The results of the second classification for both methods were also field checked and additional training data were collected, added to the training data sets, and similarly verified.

The third and final classifications of the 20 land cover types were also carried out using Dudani's distance-weighted classifier, a nearest-neighbor size of 15, and the mean inverse distance adjustment. The three land cover type labels with the highest posterior probabilities for each region were written to the COV_CODE_1, COV_CODE_2, and COV_CODE_3 attribute fields respectively, and the posterior probabilities themselves were written to the database (COV_PROB_1, COV_PROB_2, and COV_PROB_3 attributes). The five sagebrush and xeric shrub species types were assigned in a second classification applied only to regions labeled as a sagebrush/xeric shrub canopy cover type in the land cover classification. Again, this involved several iterations of the Dudani's distance-weighted classifier, with a nearest-neighbor size of 15, and the mean inverse distance adjustment. The species type label with the highest posterior probability was written to the SAGESPP1 attribute for each applicable region.

Principal Component Analysis & Unsupervised Classification

Although only one set of results is reported here, we ran the PCA technique two ways. The first one followed closely the method described by McClure and Prevedel, whereas the second one incorporated the three modifications: 1) as previously noted, the TM imagery was resampled to 15 m^2 pixels; 2) portions of the study area where the National Landcover data indicated the presence of non-rangeland cover types (e.g., forest or agriculture) were masked out prior to analysis; and 3) spectral classes that appeared to represent several different cover types (when overlayed with the SILC3 training data) were split using an iterative ISODATA routine.

Because the latter, modified PCA method seemed to produce considerably better results, it is described in further detail below.

A mask of rangeland vegetation was created by selecting pixels representing grass, shrub, and barren cover types from a copy of the National Land Cover Data for the study area (see Table 2). This mask was used to select pixels from the 15 m^2 imagery whose reflectance values (from TM bands 1-5 plus 7) were in turn subjected to a principal components analysis (PCA). We then ran an unsupervised classification of a 3-band image of the first three principal components using the ISODATA clustering algorithm to produce an initial set of 45 classes and their spectral signatures. This output image was examined in relation to ground-reference data (primarily training data used for the supervised classifications described above) to identify spectral classes that were likely to represent multiple rangeland types. These spectral classes were removed from the unsupervised image into separate images that were then subjected to another unsupervised classification to break them into two or more distinct classes. The resulting spectral signatures were added to the original spectral signature set, and a new unsupervised classification was run using the enlarged spectral class set. This classification was then examined in relation to the training data again and additional confused spectral classes were split following the method outlined above. We performed three rounds of splitting spectral classes using the process outlined above. The final unsupervised classification had 69 spectral class signatures which were used to classify the image.

Table 2. NLCD rangeland vegetation types used in mask

NLCD Cover Code	Description
31	Bare Rock / Sand / Clay
33	Transitional
51	Shrubland
71	Grassland / Herbaceous
85	Urban / Recreational Grasses
92	Emergent Herbaceous Wetlands

Training data and visual inspection of the imagery were used to manually assign each of the 69 spectral classes to a single rangeland type (see Table 1). Despite the iterations described above to split spectral classes, many were still associated with training data from more than one rangeland cover type (see Appendix 3). In most of these cases, the rangeland type represented by the majority of associated training data was selected as the label for that spectral class. But if there was confusion between life forms, then the training data were summarized by life form first, and then a rangeland type was chosen from within the majority life form. Several spectral classes associated with water or barren areas were not represented by any training data and consequently were assigned based on visual inspection of the imagery. Field checks were also used to determine the best rangeland type “fit” for a spectral class. Spectral classes assigned a sagebrush canopy cover class were also labeled for sagebrush species using training data and field checks. The spectral class image with the range/sagebrush canopy cover and range/sagebrush species labels was exported from Imagine to an Arc/Info value grid.

Guided clustering was used to remove the “salt and pepper” appearance of the classification and to standardize the minimum mapping unit to 30m x 30m. First, the Arc/Info regiongrid function was used with “eight way” connectivity and “nolink” option to produce a region grid. The “nolink” option was chosen to reduce data volume and processing time, but it also required that the spectral class attribute values be written to the region grid later (see below). Next, to establish a 30 m^2 minimum map unit (MMU), all regions smaller than four pixels were dissolved or merged with surrounding regions by means of the nibble function. Finally, attribute values for spectral class, range/sagebrush canopy cover, and range/sagebrush species labels were written to the final region grid from the spectral class value grid using zonalstats.

Ancillary data were added to the region grid to facilitate manual modifications to sagebrush species types that were not or could not be separated in the unsupervised classification. For example, where it appeared that Wyoming Big Sagebrush was confused with Mountain Big Sagebrush, we used an elevation limit to separate the two. Similarly, to reduce confusion between Wyoming Big Sagebrush and Greasewood along streams, we buffered the streams by 60m and relabeled all Wyoming Sig Sagebrush with a high canopy closure to be Greasewood.

Accuracy Assessment

The users’ accuracy of the land cover type labels was assessed for the eCognition and SILC3 methods (see Table 7 and Appendix 1 & 2) using a leave-one-out cross-validation process with spatial adjustment (Steele and Redmond in press). The actual accuracy assessment involved removing the first training record from the dataset and constructing a new classification rule from the reduced set. The new rule was then used to classify the region represented by left-out training observation. The process was repeated until all training observations had been held out once, and the classification accuracy for each type was estimated by the percentage of the held-out observations that was correctly classified. An accuracy assessment could not performed on the PCA results because of the nature of this classification.

RESULTS

Considering first the acreage subtotals for each life form class (Table 3), all three methods mapped similar amounts of sagebrush/xeric shrub types. But grass types were possibly underclassified by SILC3 in favor of non-rangeland types, and both SILC3 and eCognition methods classified more conifer types (Utah Juniper in particular), than did the modified PCA method. Note also that all three of these newer methods predicted approximately 50% more sagebrush land cover for the study area than did MTGAP (SILC2; Redmond et al. 1998). Not surprisingly, the PCA method mapped relatively little non-rangeland vegetation due to the life form mask applied early in the process.

Within the sagebrush/xeric shrub types, the eCognition and PCA methods produced generally similar results, especially for the more open canopy closure classes (Table 3), whereas SILC3 classified more of the 15-24% canopy class and less of the 5-14% canopy class than the other two methods. Among the xeric shrub species, SILC3 classified the most Wyoming Big Sagebrush, eCognition an intermediate amount, and PCA the least (Table 3); also considerably more Black Sagebrush and less Greasewood were classied by the PCA method than either of the other two methods.

Table 3. Land cover type with sagebrush/xeric shrub canopy classes

COVER	NAME	SILC3	eCOG	PCA	MTGAP
3130	Very Low Cover Grasslands	111,561	140,911	111,191	156,007
3150	Low / Moderate Cover Grasslands	32,006	22,437	45,231	26,059
3170	Moderate / High Cover Grasslands	5,888	12,859	24,558	2,069
3180	MTGAP Montane / Subalpine Meadows	0	0	0	8,488
	Grass subtotal in acres	149,455	176,206	180,980	192,622
3370	Sagebrush / Xeric Shrubs 05-14% Cover	150,555	193,472	191,480	0
3380	Sagebrush / Xeric Shrubs 15-24% Cover	206,074	160,128	170,683	0
3390	Sagebrush / Xeric Shrubs 25-34% Cover	66,469	41,637	71,028	0
3395	Sagebrush / Xeric Shrubs ≥ 35% Cover	17,500	27,292	8,254	0
3300	MTGAP Sage/Xeric Shrub (3300,3309,3350,3520)	0	0	0	286,583
	Sagebrush/Xeric Shrub subtotal in acres	440,598	422,528	441,444	286,583
3610	Mesic Shrubs	4,578	4,592	17,222	7,041
4205	Limber Pine	3,144	4,588	4,532	10,490
4216	Utah Juniper	56,166	53,230	34,505	18,246
7300	Rock / Barren	32,238	24,249	38,946	8,849
7600	MTGAP Badlands	0	0	0	156,188
9999	Other Non-rangeland Types	31,573	32,360	123	37,734
	Non-rangeland subtotal in acres	127,700	119,019	95,329	238,548
	STUDY AREA TOTAL ACRES	717,753	717,753	717,753	717,753

Species classification of areas classified as a sagebrush canopy class above

COVER	NAME	SILC3	eCOG	PCA	MTGAP
3311	Greasewood	7,296	15,619	2,141	0
3351	Mountain Big Sage	5,080	3,205	2,231	0
3352	Wyoming Big Sage	389,686	368,303	342,633	0
3353	Basin Big Sage	3,233	1,831	0	0
3354	Black Sage	35,303	33,570	94,438	0
3300	MTGAP Sage/Xeric Shrub (3300,3309,3350,3520)	0	0	0	286,583
	Sagebrush/Xeric Shrub subtotal in acres	440,598	422,528	441,444	286,583

At the life form level, there was greater than 80% agreement between the map outputs for three different pair-wise comparisons; PCA and eCognition, PCA and SILC3, and eCognition and SILC3 (Table 4). But when these pair-wise comparisons are made for just the sagebrush canopy cover and other rangeland types (Table 5) or sagebrush species and other rangeland types (Table 6), the level agreement drops to between 20 and 60%. In the former case (sagebrush canopy cover and other rangeland types, Table 5), most areas of disagreement were within grassland canopy cover classes and within sagebrush canopy cover classes. The map results from PCA and eCognition methods showed the highest level of agreement, whereas those from the PCA and SILC3 methods were least alike. For sagebrush species and other rangeland types, most of the disagreement was also within the grass and shrub classes (Table 6), but the map results were most similar between eCognition and SILC3, and least similar between PCA and SILC3.

Table 4. Comparison of sagebrush and rangeland classifications in acres (by life form)

	SILC3		
eCOGNITION	Grassland	Sage/Xeric Shrub	Other
Grassland	105,902	44,384	25,920
Sagebrush/Xeric Shrub	26,958	376,531	19,039
Other Type	16,595	19,683	82,741

	SILC3		
PCA	Grassland	Sage/Xeric Shrub	Other
Grassland	80,746	55,245	44,990
Sagebrush/Xeric Shrub	52,034	356,024	33,386
Other Type	16,675	29,329	49,325

	eCOGNITION		
PCA	Grassland	Sage/Xeric Shrub	Other
Grassland	97,772	39,897	43,312
Sagebrush/Xeric Shrub	54,064	359,839	27,541
Other Type	24,370	22,792	48,166

Table 5. Comparison of sagebrush canopy and rangeland classifications in acres

					SILC3						
eCOG	3130	3150	3170	3370	3380	3390	3395	4205	4216	7300	OTHER
3130	71,060	7,965	117	19,378	18,427	2,145	1,247	234	6,665	12,640	1,034
3150	3,818	13,432	1,065	50	880	154	500	350	489	37	1,664
3170	653	4,673	3,118	66	240	102	1,198	186	23	18	2,583
3370	10,556	28	8	102,942	61,215	8,685	934	2	1,851	6,800	452
3380	12,994	438	15	19,128	96,505	22,020	2,325	42	5,743	609	309
3390	1,148	161	4	1,852	14,214	21,242	1,420	106	1,188	179	122
3395	962	494	151	663	5,004	10,190	8,192	90	986	69	490
4205	738	713	44	1	180	118	111	1,049	572	24	1,040
4216	5,858	610	6	766	6,526	1,238	682	298	36,213	402	632
7300	2,558	83	15	5,353	2,288	401	150	10	1,979	11,226	188
OTHER	1,218	3,409	1,345	357	596	175	743	778	459	234	27,639

					SILC3						
PCA	3130	3150	3170	3370	3380	3390	3395	4205	4216	7300	OTHER
3130	34,970	4,346	110	23,804	17,838	2,402	1,837	228	12,644	10,308	2,705
3150	12,675	17,880	1,855	282	3,139	1,043	881	420	2,123	51	4,883
3170	2,288	4,534	2,088	296	961	425	2,338	556	1,392	85	9,597
3370	27,866	62	22	82,804	54,719	8,544	2,018	7	6,148	6,743	2,547
3380	11,749	94	14	29,113	94,840	26,158	2,290	8	4,323	524	1,572
3390	9,562	817	20	3,890	22,702	22,477	2,964	210	7,591	208	587
3395	1,691	130	8	208	898	1,080	1,320	31	2,514	24	349
4205	374	442	97	7	85	62	138	538	572	12	2,204
4216	4,295	469	65	955	4,284	3,018	2,047	370	16,740	372	1,892
7300	5,164	140	50	9,079	6,252	1,032	415	46	1,379	13,828	1,561
OTHER	928	3,093	1,558	120	356	228	1,252	730	740	85	8,255

					eCOGNITION						
PCA	3130	3150	3170	3370	3380	3390	3395	4205	4216	7300	OTHER
3130	51,990	2,900	189	17,771	11,579	1,533	3,036	345	13,132	6,041	2,675
3150	15,348	13,592	3,459	110	1,145	391	2,092	680	2,716	33	5,665
3170	1,851	2,753	5,690	53	199	224	1,763	776	1,511	66	9,673
3370	32,462	37	15	110,555	32,162	3,829	1,268	6	4,517	4,127	2,502
3380	10,140	48	19	50,820	80,418	17,890	5,162	6	4,245	471	1,465
3390	8,868	540	29	3,812	28,187	14,740	7,784	372	5,778	411	506
3395	1,815	80	12	152	842	545	1,673	64	2,705	20	345
4205	236	453	105	4	19	42	127	829	475	4	2,238
4216	5,713	325	32	1,035	2,954	1,821	3,136	771	16,567	587	1,566
7300	11,851	40	19	9,067	2,512	505	202	44	759	12,409	1,538
OTHER	637	1,670	3,289	92	111	117	1,048	696	825	80	8,780

Table 6. Comparison of sagebrush species and rangeland classifications in acres

	SILC3										
eCOG	3130	3150	3170	3311	3351	3352	3354	4205	4216	7300	OTHER
3130	71,060	7,965	117	745	2,328	34,280	3,844	234	6,665	12,640	1,034
3150	3,818	13,432	1,065	383	819	310	71	350	489	37	1,664
3170	653	4,673	3,118	1,161	121	315	8	186	23	18	2,583
3311	665	370	143	5,301	0	9,456	355	30	624	56	449
3351	1,092	643	18	0	942	15	49	178	154	27	88
3352	20,904	87	17	1,731	5	325,406	5,954	15	5,829	7,570	788
3354	2,999	22	0	58	39	6,651	20,570	17	3,161	5	49
4205	738	713	44	37	292	58	22	1,049	572	24	1,040
4216	5,858	610	6	347	276	4,208	4,380	298	36,213	402	632
7300	2,558	83	15	83	11	8,077	20	10	1,979	11,226	188
OTHER	1,218	3,409	1,345	685	247	909	31	778	459	234	27,639

	SILC3										
PCA	3130	3150	3170	3311	3351	3352	3354	4205	4216	7300	OTHER
3130	34,970	4,346	110	1,186	1,226	40,307	3,161	228	12,644	10,308	2,705
3150	12,675	17,880	1,855	633	1,509	2,274	929	420	2,123	51	4,883
3170	2,288	4,534	2,088	2,125	529	1,269	97	556	1,392	85	9,597
3311	89	18	5	569	1	819	51	4	279	21	286
3351	960	264	3	0	200	20	64	52	535	29	104
3352	40,207	372	41	2,489	529	256,552	13,838	167	17,095	7,396	3,947
3354	9,612	448	16	688	521	64,221	15,460	34	2,668	53	718
4205	374	442	97	107	99	73	13	538	572	12	2,204
4216	4,295	469	65	1,276	205	7,262	1,560	370	16,740	372	1,892
7300	5,164	140	50	308	6	16,351	112	46	1,379	13,828	1,561
OTHER	928	3,093	1,558	1,147	254	538	17	730	740	85	8,255

	eCOGNITION										
PCA	3130	3150	3170	3311	3351	3352	3354	4205	4216	7300	OTHER
3130	51,990	2,900	189	2,414	701	28,639	2,165	345	13,132	6,041	2,675
3150	15,348	13,592	3,459	1,786	1,303	472	177	680	2,716	33	5,665
3170	1,851	2,753	5,690	1,625	363	230	22	776	1,511	66	9,673
3311	149	19	8	719	0	547	48	3	333	16	299
3351	1,480	155	1	0	85	0	96	150	197	13	54
3352	42,582	191	35	4,508	93	256,896	15,732	248	13,560	4,936	3,853
3354	9,074	340	31	2,749	319	63,725	14,322	47	3,156	64	611
4205	236	453	105	103	62	23	4	829	475	4	2,238
4216	5,713	325	32	2,403	143	5,473	926	771	16,567	587	1,566
7300	11,851	40	19	142	6	12,071	67	44	759	12,409	1,538
OTHER	637	1,670	3,289	1,002	129	228	11	696	825	80	8,780

Finally, the users' accuracies were very similar between the SILC3 and eCognition results, not just overall, but also among both the sagebrush canopy closure and also the sagebrush species classes (Table 7). These findings, coupled with the lack of any accuracy assessment for the PCA results, make it difficult to unequivocally select one method over another.

Table 7. Sagebrush canopy cover and land cover type training data cross validation users accuracy results

		SILC3 Classification			**eCognition Classification**		
Cover type	*Name*	*Total Points*	*Number Correct*	*Percent Accuracy*	*Total Points*	*Number Correct*	*Percent Accuracy*
3130	Very Low Cover Grasslands	229	199	86.9	223	184	82.5
3150	Low/Mod Cover Grasslands	199	164	82.4	164	134	81.7
3170	Mod/High Cover Grasslands	43	35	81.4	76	61	80.3
3370	Sage/Xeric Shrubs 05-14%	102	63	61.8	128	72	56.3
3380	Sage/Xeric Shrubs 15-24%	173	91	52.6	163	86	52.8
3390	Sage/Xeric Shrubs 25-34%	87	44	50.6	89	45	50.6
3395	Sage/Xeric Shrubs ≥ 35%	49	41	83.7	50	42	84.0
3610	Mesic Shrubs/Willow	132	100	75.8	125	93	74.4
4101	Aspen	68	47	69.1	72	49	68.1
4150	Mixed Broadleaf/Cottonwood	73	58	79.5	78	58	74.4
4203	Lodgepole Pine	40	25	62.5	37	26	70.3
4204	Whitebark Pine	24	20	83.3	29	22	75.9
4205	Limber Pine	49	37	75.5	46	31	67.4
4206	Ponderosa Pine	66	48	72.7	62	47	75.8
4212	Douglas-fir	404	326	80.7	421	329	78.1
4216	Utah Juniper	117	78	66.7	103	78	75.7
4223	Douglas-fir/Lodgepole Pine	41	19	46.3	39	15	38.5
4237	Subalpine Fir/Spruce	41	24	58.5	41	23	56.1
4240	Mixed Conifer Forest	159	104	65.4	154	101	65.6
7300	Rock/Barren	11	107	972.7	106	105	99.1
	Total	2,207	1,630	73.9	2,229	1,623	72.8

Sagebrush species training data cross validation users accuracy results

		SILC3 Classification			**eCognition Classification**		
Cover type	*Name*	*Total Points*	*Number Correct*	*Percent Accuracy*	*Total Points*	*Number Correct*	*Percent Accuracy*
3311	Greasewood	31	23	74.2	34	26	76.5
3351	Mountain Big Sagebrush	47	42	89.4	49	44	89.8
3352	Wyoming Big Sagebrush	295	285	96.6	283	276	97.5
3353	Basin Big Sagebrush	4	2	50.0	3	1	33.3
3354	Black Sagebrush	48	36	75.0	56	38	67.9
	Total	425	388	91.3	425	385	90.6

DISCUSSION

Small or narrow landscape features like woody draws, sagebrush stringers, patches of mesic shrubs or willows or aspen appeared to be better captured by the 15 m imagery used by both the eCognition and PCA methods, than by the 30 m data used by SILC3. We therefore might expect the results from these two methods to be more accurate than those from SILC3. Unfortunately though, the use of different image dates (July for SILC3 vs September for eCognition and PCA) will confound the interpretation of an independent assessment of the three map outputs. In hindsight, we should have used the same TM imagery as inputs for all three methods.

The PCA technique was the fastest method to produce an initial labeled map, albeit for sagebrush/rangeland types only. It was also relatively easy to change the labels through manual modifications. The PCA technique had only 69 spectral classes to label, 20 of which were water or barren (easily discernable in the imagery). Thus, only 49 classes actually required training data or user knowledge to label. The problem was spectral classes where some areas were clearly sagebrush and others were clearly different life forms. Tough choices had to be made on some of the spectral class labels. Some of these problem spectral class areas were corrected after region grouping the classification and applying an ancillary data rule set. Another limitation of the modified PCA method was the rangeland mask used to mask out non-range land pixels in the imagery. Rangelands misclassified as non-rangelands could be masked out of the classification. The modified PCA technique does not generate labeling accuracies in the process. However, due to its fast production cycle and limited costs, a post-classification accuracy assessment could be built into project.

For their supervised classifications, both the methods based on image segmentation (SILC3 and eCognition) require large training data sets that are time consuming and expensive to acquire, especially for large geographic areas. It might be possible to reduce these costs by means of an iterative classification system where a small training data set is used for an initial classification, and then additional training data are collected in a series of field checks and re-classifications. Yet despite the greater time and cost associated with the supervised classifications, one distinct advantage they have over the unsupervised PCA approach is the ability to estimate thematic accuracy iteratively, if need be, throughout the labeling process. Other advantages of the SILC3 and eCognition methods are 1) the segmented regions were usually easier to identify and field check than single pixels, and 2) region grids derived from them were substantially smaller in total size than the region grid derived from the unsupervised classification of 15 m pixels. But this latter issue may not be a concern for users who do not need to convert a pixel-based land cover grid to a region grid or a polygon coverage.

Considering just the two image segmentation methods, eCognition was much faster to run than the SILC3 process (hours to run versus days), but it also tended to over-segment some ground features associated with agriculture and water.

CONCLUSIONS

All three techniques were able to classify sagebrush canopy classes and sagebrush species types, but at an uncertain level of accuracy. Although we believe that they represent substantial improvements over existing land cover datasets, such as MTGAP or NLCD, we strongly advocate the need for an independent accuracy assessment of the three map outputs. This should be relatively easy to accomplish and will help objectively determine if one method is substantially better than the others or how these three methods compare with other available ones.

LITERATURE CITED

Barsness, S. 1996. Aggregating raster polygons derived from large remotely sensed images. MS Thesis. The University of Montana, Missoula, MT.

Definiens Imaging Corp. ECognition User's Manual. (http://www.definiens-imaging.com/documents/index.htm)

Jones, W.M. 2001. Ecologically significant wetlands in the Upper Shields, and Stillwater River drainages. Report to the Montana Department of Environmental Quality. Montana Natural Heritage Program, Helena, MT. 37pp + appendices. (http://nhp.nris.state.mt.us//plants/reports/upper_Yellow_wet.pdf).

Ma, Z., Hart, M.M., and R.L. Redmond. 2001. Mapping vegetation across large geographic areas: Integration of remote sensing and GIS to classify multisource data. Photogrammetric Engineering & Remote Sensing 67:295-307.

McClure, H., and D. Prevedel. 2002. Unpublished report. USDA Forest Service, Intermountain Region, Ogden, UT.

Nesser, J.A., Ford, G.L., Maynard, C.L., and D.S. Page-Dumroese. 1997. Ecological Units of the Northern Region: Subsections. USDA Forest Service General Technical Report INT-GTR-369.

Redmond, R.L., Hart, M.M., Winne, J.C., Williams, W.A., Thornton, P.C., Ma, Z., Tobalske, C.M., Thornton, M.M., McLaughlin, K.P., Tady, T.P., Fisher, F.B., and S.W. Running. 1998. Montana Gap Analysis: Final Report. Unpublished report. Montana Cooperative Wildlife Research Unit, The University of Montana, Missoula. xii + 136 pp. + appendices.

Steele, B. M. and D.A. Patterson. In press. Land cover mapping using combination and ensemble classifiers. *Proceedings of the 33rd Symposium on the Interface* Interface Foundation of North America, Fairfax Station VA.

Steele, B.M., and R.L. Redmond. 2001. A method of exploiting spatial information for improving classification rules: Application to the construction of polygon-based land cover maps. *International Journal of Remote Sensing*, 22, 3143-66.

Steele, B.M., Patterson, D.A., and R.L. Redmond. In press. Estimating thematic map accuracy without a probability test sample.

Appendix 1. SILC3 sagebrush canopy and land cover cross validation producers error matrix

	3130	3150	3170	3370	3380	3390	3395	3610	4101	4150	4203	4204	4205	4206	4212	4216	4223	4237	4240	7300	Total
3130	199	11	0	5	8	1	1	0	0	0	0	0	0	0	0	6	0	0	0	4	235
3150	6	164	1	0	5	2	0	8	0	0	0	0	3	0	0	0	0	0	0	0	189
3170	0	3	35	0	0	0	0	4	1	2	0	0	0	0	0	0	0	0	0	0	45
3370	1	1	0	63	46	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	114
3380	9	3	0	24	91	31	3	0	0	0	0	0	1	0	0	9	0	0	0	0	171
3390	6	3	0	2	16	44	3	0	0	0	0	0	0	1	0	9	0	0	0	0	84
3395	3	0	0	0	3	6	41	0	0	0	0	0	1	0	0	2	0	0	0	0	56
3610	0	13	4	0	0	0	0	100	9	7	0	0	0	0	1	0	0	0	0	0	134
4101	0	0	0	0	0	0	0	2	47	5	0	0	0	0	0	0	0	0	0	0	54
4150	0	0	3	0	0	0	0	16	11	58	0	0	0	1	1	0	0	0	0	0	90
4203	0	0	0	0	0	0	0	0	0	0	25	0	0	1	5	0	7	3	8	0	49
4204	0	0	0	0	0	0	0	0	0	0	0	20	0	0	1	0	0	2	0	0	23
4205	0	1	0	0	1	0	1	1	0	0	0	0	37	1	0	3	0	1	3	0	49
4206	0	0	0	0	0	0	0	0	0	0	1	0	0	48	9	1	0	0	11	0	70
4212	0	0	0	0	0	0	0	0	0	1	5	0	0	2	326	0	8	5	28	0	375
4216	1	0	0	0	3	0	0	0	0	0	0	0	1	0	0	78	0	0	0	0	83
4223	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	0	19	0	1	0	23
4237	0	0	0	0	0	0	0	0	0	0	1	3	1	0	1	0	0	24	4	0	34
4240	0	0	0	0	0	0	0	1	0	0	7	1	4	12	58	0	7	6	104	0	200
7300	4	0	0	8	0	0	0	0	0	0	0	0	1	0	0	9	0	0	0	107	129
Total	229	199	43	102	173	87	49	132	68	73	40	24	49	66	404	117	41	41	159	111	2207

Appendix 2. eCognition sagebrush canopy and land cover type cross validation accuracy assessment producers error matrix

	3130	3150	3170	3370	3380	3390	3395	3610	4101	4150	4203	4204	4205	4206	4212	4216	4223	4237	4240	7300	9100	Total
3130	184	15	0	8	14	0	1	0	0	0	0	0	0	0	0	7	0	0	0	1	0	230
3150	7	134	3	1	1	2	0	1	0	0	0	0	4	0	0	0	0	0	0	0	0	153
3170	0	7	61	0	0	0	0	13	3	1	0	0	0	1	0	0	0	0	0	0	0	86
3370	6	1	0	72	33	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	114
3380	9	2	0	35	86	32	3	0	0	0	0	0	1	0	0	3	0	0	0	0	0	171
3390	4	1	0	3	25	45	2	0	0	0	0	0	1	0	0	3	0	0	0	0	0	84
3395	5	0	0	0	2	7	42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	56
3610	1	3	10	0	0	0	0	93	10	14	0	1	1	1	0	0	0	0	0	0	0	134
4101	0	0	1	0	0	0	0	0	49	3	0	0	0	0	1	0	0	0	0	0	0	54
4150	0	0	0	0	0	0	0	18	8	58	0	0	0	1	5	0	0	0	0	0	0	90
4203	0	0	0	0	0	0	0	0	0	0	26	0	0	1	5	0	12	2	3	0	0	49
4204	0	0	0	0	0	0	0	0	0	0	0	22	0	0	0	0	0	1	0	0	0	23
4205	0	0	1	0	1	0	1	0	0	0	0	1	31	1	1	3	0	4	5	0	0	49
4206	0	0	0	0	0	0	0	0	2	0	2	0	1	47	8	0	0	0	10	0	0	70
4212	0	0	0	0	0	0	0	0	0	2	3	0	0	4	329	0	6	3	28	0	0	375
4216	2	0	0	0	0	1	1	0	0	0	0	0	0	0	0	78	0	0	1	0	0	83
4223	0	0	0	0	0	0	0	0	0	0	2	0	0	0	5	0	15	0	1	0	0	23
4237	0	0	0	0	0	0	0	0	0	0	1	3	2	0	0	0	0	23	5	0	0	34
4240	0	1	0	0	0	0	0	0	0	0	3	2	5	6	67	2	6	7	101	0	0	200
7300	5	0	0	9	1	0	0	0	0	0	0	0	0	0	0	7	0	1	0	105	1	129
9100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	22	22
Total	223	164	76	128	163	89	50	125	72	78	37	29	46	62	421	103	39	41	154	106	23	2229

Appendix 3. Spectral class assignment to land cover type based on available ground reference data for PCA method

Spectral Class[1]	Assigned Cover Type	Number of Ground-Referenced Data		
		Total[2]	Assigned Cover Type[3]	Other Cover Types[4]
1	3170	4	1	3
2	3170	51	5	46
3	3610	55	20	35
4	3610	15	8	7
5	3150	12	6	6
6	3170	7	3	4
7	3380	71	29	42
8	3610	32	12	20
9	3150	65	38	27
10	3370	56	27	29
11	4205	23	5	18
12	3150	49	29	20
13	7300	13	10	3
14	3610	29	20	9
15	3370	19	8	11
16	3370	5	1	4
17	3130	42	30	12
18	4216	26	7	19
19	7300	4	4	0
22	7300	14	12	2
23	7300	3	3	0
29	7300	1	1	0
45	3170	38	7	31
46	3170	17	3	14
47	3150	80	18	62
48	3390	58	9	49
49	3390	47	10	37
50	3390	57	15	42
51	7300	15	12	3
52	3395	33	8	25
53	4216	19	12	7
54	4216	23	10	13
55	3380	36	17	19
56	3150	42	9	33
57	3130	24	19	5
58	3130	24	15	9
59	3130	22	12	10
60	3370	58	25	33
61	3130	17	7	10
62	3130	38	20	18
63	4216	33	4	29
64	3380	27	6	21
65	3370	20	5	15
66	3150	21	6	15
67	4216	21	13	8
68	3130	35	18	17
69	3380	55	29	26

1 - spectral classes (n = 46) for which ground reference data were available
2 - total number of ground reference data
3 - number of ground reference data representing assigned cover type
4 - number of ground reference data representing other cover types