INTRODUCTION

The rate and spatial extent of biological invasions are increasing in an unprecedented manner across the globe. This invasion trend is recognized as a major component of global environmental change and as an escalating and expensive national problem (Vitousek et al. 1997; Lodge et al. 2005). The overall cost of biological invasions in the United States, in terms of damage, loss, and control, is estimated at $120 billion each year for approximately 50,000 species (Pimentel et al. 2005). Invasive plant species can alter ecosystem functions and cause negative economic impacts in a number of ways, including devaluation of land, reduction of agricultural productivity and rangeland, loss of native habitat, decline of species diversity, and alteration of fire regimes and soil dynamics (Olson 1999). Leafy spurge (Euphorbia esula L.) is an introduced plant listed as a noxious weed in parts of Canada and the north central and western United States. Once established, leafy spurge invasions can spread rapidly, causing particularly serious economic problems on rangelands, where grazing capacity sharply declines (Hein and Miller 1992). Leafy spurge now infests approximately 2 million ha of rangeland, pastures, hillsides, and riparian areas in North America, where the size of infested areas has been doubling nearly every 10 years (Quimby and Wendel 1997).

The invasion mechanisms and reproductive characteristics of leafy spurge are such that complete eradication is unlikely. Land managers report that seeds are persistent and easily dispersed and transported by way of animals, mud, hay, and water. Reproduction can occur through extensive seed production and vegetative reproduction from both the crown and root buds (Hanson and Rudd 1933; Bakke 1936; Bowes and Thomas 1978). As such, infestations of leafy spurge are often
widespread, and cost-effective tools such as remote sensing are needed to monitor changes in leafy spurge distribution and abundance over time (Anderson et al. 2003). Remote sensing has become a useful tool for efficiently mapping the distribution of some invasive plant species, including leafy spurge, over large areas that would otherwise be difficult to survey (Everitt et al. 1995; Lewis et al. 2000; Lamb and Brown 2001; Parker Williams and Hunt 2002, 2004; Everitt and Yang 2004).

One of the challenges with remote sensing of leafy spurge is to leverage the accuracy of detection with the cost of the acquisition and processing of the data. Detecting leafy spurge with coarse spatial and spectral resolution imagery (e.g., at best Landsat 5 scale) is optimal from a cost analysis perspective. Classification accuracy tends to decrease at coarser spatial resolutions because there is an increase in the number of pixels that exhibit a mixed response or overlap between classes. At the same time, classification accuracy may decrease at finer spatial resolutions because of more spectral noise or heterogeneity, which tends to average out at coarser resolutions (Markham and Townshend 1981). This study expands the scope of previous work on remote sensing of leafy spurge by further shifting the focus from establishing finer scale detection limits using high spectral and/or spatial resolution sensors toward investigating coarser scale detection limits. In this study our first objective was to compare the relative importance of spatial and spectral detection components for leafy spurge discrimination. Our second objective was to assess the suitability of using widely available multispectral satellite imagery (i.e., Landsat 5) for surveying core leafy spurge infestations (both current and historic) and monitoring regional distribution and abundance patterns.

**Comparison of Remote Sensors**

Most multispectral remote sensors detect solar radiance and absorption of earth materials at a moderate spatial resolution by way of a few broad bands in the visible and infrared (near, short, and thermal) portions of the electromagnetic spectrum. Imaging spectrometry, or hyperspectral imaging, is a remote sensing technology whereby many narrow bands collect surface radiance information throughout a near-contiguous range of the visible, near-infrared, and shortwave infrared portions of the electromagnetic spectrum (Goetz et al. 1985).

Multispectral sensors such as the Landsat 5 Thematic Mapper, ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer), and SPOT (Satellite pour l’Observation de la Terre) are satellite based and provide global coverage at regular (Landsat 5 and SPOT) or semiregular (ASTER) time intervals. Standard multispectral classification techniques have been developed to classify images into broad categories (Jensen 2005). Remotely sensed data with spatial resolutions of 15–30-m pixels are frequently used for vegetation applications such as land use or land cover classification and rangeland and forestry monitoring (Johnson 1999). In some cases in which a target has spectrally unique characteristics or grows in clusters, multispectral sensors are capable of differentiating individual species (Johnson 1999).

Hyperspectral sensors such as HyMap and AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) are airborne and cover relatively small, narrow geographic areas at irregular time intervals and with spatial resolution typically ranging from 3 m to 20 m. Compared to multispectral sensors, airborne hyperspectral sensors have higher spectral, spatial, and radiometric resolution than multispectral data. Thus, they are more capable of distinguishing subtle spectral responses among species and improving quantitative model estimations of canopy structure and biochemical properties (Aspinall et al. 2002; Parker Williams and Hunt 2002; Root et al. 2004; Ustin et al. 2004; Underwood et al. 2007). High spectral resolution facilitates the use of linear spectral mixture analysis classification techniques that estimate subpixel abundance (Boardman 1998; Aspinall et al. 2002). High spatial resolution airborne data increase the probability of detecting smaller infestations, but such images, when used for repeat monitoring, can present unique challenges in the way of georegistration and geometric errors (e.g., image rotation or nonuniform pixel shifts; Aspinall et al. 2002; Glenn et al. 2005). Additional challenges include successfully requesting and coordinating image acquisition, the need for extensive image processing techniques, and overall costs. As such, high-resolution hyperspectral imagery is less suitable for frequent vegetative monitoring. Satellite-based hyperspectral imagery has the potential to overcome some of these challenges and has been collected by the Hyperion sensor (30-m pixels) onboard NASA’s Earth Observing-1 (EO-1) and by the sensor onboard the Airforce Research Lab’s MightySat II (Otten et al. 1997; Ungar et al. 2003).

**Previous Work**

During peak phenology, the yellow-green flower bracts of leafy spurge are spectrally unique and can be distinguished from surrounding vegetation using remote sensors because of higher reflectance in the visible region (0.5–0.7 μm) and higher reflectance values and different spectral signatures in the chlorophyll absorption region (0.55–0.69 μm; Everitt et al. 1995; Anderson et al. 1996, 1999; Parker Williams and Hunt 2002, 2004; Hunt et al. 2004).

Both hyperspectral and multispectral sensors have been used, with varying degrees of success, to identify leafy spurge (Everitt et al. 1995; O’Neill et al. 2000; Parker Williams and Hunt 2002, 2004; Root et al. 2002; Dudek et al. 2004; Hunt and Parker Williams 2006; Glenn et al. 2005; Stitt et al. 2006). Root et al. (2002) and Dudek et al. (2004) used AVIRIS (20-m pixels, 224 bands [0.4–2.5 μm]) and Hyperion (30-m pixels, 220 bands [0.4–2.5 μm]) with classification accuracies ranging from 39% to 63%. Classification methods in these studies included mixture-tuned matched filtering (MTMF; Harsanyi 1993; Harsanyi and Chang 1994; Boardman 1998) and spectral angle mapper (SAM; Kruse et al. 1993). Parker Williams and Hunt (2002, 2004) used AVIRIS imagery to map leafy spurge in northeastern Wyoming with classification accuracies of 75–95% for large, high-density leafy spurge infestations. Glenn et al. (2005) found similar results in Idaho using HyMap imagery (3.5-m pixels, 126 bands [0.45–2.48 μm]), with overall classification accuracies above 84%. In this study leafy spurge infestations at 10% cover could be detected within a 3.5-m pixel, and infestations at 40% cover could be repeatedly detected over the same area.

Mladinich et al. (2006) used Landsat 7 imagery to classify leafy spurge with overall classifications of approximately 63%. The authors concluded that although the imagery was
inappropriate for small-scale detection, the study did demonstrate the potential for regional distribution mapping. In addition, the authors suggested that advanced image-processing techniques such as MTMF may increase leafy spurge detection with Landsat. Stitt et al. (2006) used the Advanced Land Imager (ALI) to produce conservative accuracy assessments in the range of 59–66%. The ALI sensor has seven bands that are spectrally and spatially comparable to Landsat (30-m pixels), as well as three panchromatic bands with a 10-m spatial resolution. It should be noted that the signal to noise ratio (SNR) of ALI is 2.5 times higher than that of Landsat 7 (Kutser et al. 2003).

A limited number of studies have compared the use of hyperspectral and multispectral imagery for leafy spurge detection (Hunt and Parker Williams 2006; Root et al. 2004; Stitt et al. 2006). To detect the distribution and abundance of leafy spurge in prairie, riparian, and woodland cover types, Hunt and Parker Williams (2006) found similar classification accuracies between AVIRIS, Landsat 7, and SPOT (approximately 63–68%). Root et al.’s (2004) review of leafy spurge research from 1998 to 2003 determined that hyperspectral data yielded slightly higher overall classification accuracies (63–78%) than multispectral classification accuracies (60–70%), although multispectral classification techniques and accuracy assessment details are unpublished. This study included cost-benefit analyses of satellite and aircraft-based sensors, which indicated that the use of multispectral sensors, possibly combined with predictive modeling, is the most efficient means of mapping leafy spurge infestations at the regional scale.

METHODS

Study Site

Research was conducted on approximately 7 700 ha of sagebrush steppe on and in the vicinity of Medicine Lodge (lat 44°19'N, long −112°30'W), and Spencer (lat 44°21'N, long −112°10'W), Idaho, USA (Fig. 1). Both sites are located just south of the Continental Divide, in the Centennial Mountains of Clark County, within 20 km of the town of Dubois. The Spencer area has a long history of leafy spurge invasion. The weed was likely first introduced into the towns of Dubois and Spencer by way of the Union Pacific Railroad, which was built in the late 1800s and is now located alongside Interstate 15, both of which span the length of the Spencer study site. The Medicine Lodge and surrounding drainages (Rocky Creek, Middle Creek, and Indian Creek) have a somewhat shorter invasion history than Spencer because the railroad is farther away, although Medicine Lodge infestations were exacerbated by fire in 2003. Similar general distribution patterns that were observed at both the Medicine Lodge and Spencer sites include well-established colonies of leafy spurge associated with rock outcrops and areas of concentrated livestock use, and absence or low concentrations of leafy spurge associated with xeric knolls along hill slopes. Despite similar general distribution patterns, leafy spurge is present at high densities throughout the Medicine Lodge site, with ground cover estimates averaging 60%. In contrast, infestations at the Spencer site are characterized by a single expansive, core infestation (~0.75 km²), and infrequent occurrences of low-density infestations throughout the remainder of the site.

Figure 1. Location of hyperspectral flightlines and ground reference sites. Three overlapping flightlines were acquired over the Spencer study area, and two perpendicular flightlines were acquired over the Medicine Lodge study area.

Image Acquisition

In this study the high spectral and spatial resolutions of hyperspectral imagery were necessary for obtaining baseline data, from which spectrally and/or spatially degraded images could be derived for comparative purposes. The relatively coarser spectral and spatial resolutions of Landsat imagery were necessary to assess the validity of the degraded image results and the suitability of using widely available multispectral satellite imagery for regional distribution mapping.

Hyperspectral imagery was collected over the study area using the HyMap sensor (operated by HyVista, Inc.) mounted on an aircraft flying about 1 000 m above the ground to obtain 3.2 × 3.2 m pixel resolution. The HyMap sensor collected five flightlines of data on 28 June 2006, which was an optimal date for capturing leafy spurge in peak bloom—its most distinct phenological state. Three overlapping flightlines totaling 3.5 × 12.0 km were situated lengthwise approximately 0.6 km south of the town of Spencer, north to Stoddard Creek. Two additional flightlines (1.75 × 10 km each) were located in the Medicine Lodge area, of which the first was oriented parallel and the second perpendicular to the Medicine Lodge Creek drainage (Fig. 1).

The HyMap instrument collects calibrated radiance data in 126 near-contiguous spectral bands (0.45–2.48 μm) that range in width from 15 μm in the visible and near-infrared to 20 μm in the shortwave infrared (Kruse et al. 2000). For comparative purposes, a single Landsat 5 image was acquired over the study area on 13 June 2006 (path 39, row 29). The Thematic Mapper (TM) on board the Landsat 5 satellite collects data in seven relatively broad bands: Band 1 (blue, 0.45–0.52 μm), Band 2 (green, 0.52–0.60 μm), Band 3 (red, 0.63–0.69 μm), Band 4 (near-infrared, 0.76–0.90 μm), Band 5 (mid-infrared, 1.55–1.75 μm), Band 6 (thermal infrared, 10.4–12.5 μm), and Band 7 (mid-infrared, 2.08–2.35 μm). The thermal band has a spatial resolution of 120 × 120 m, and the other six bands have a spatial resolution of 28.5 × 28.5 m.
Field Validation

Roaming surveys of leafy spurge infestations focused on capturing a uniformly distributed range of target abundance at sites representative of the ecological variability within the project areas (although forested locations were excluded). The majority of infestation boundaries were roughly mapped, and circular plots were used to collect calibrated, continuous ocular estimates of leafy spurge percentage canopy cover. Beyond North America Weed Management Association mapping standards were used as a guide for field data collection (Stohlgren et al. 2006). The sample design used a 7.32-m radius circle (168.25 m²) with three transects extending from the center of the circle to the perimeter at N30W, N150W, and N270W. Leafy spurge cover was estimated at the plot scale and in three 1-m² plots along each of the three transects (a total of 9 Daubenmire quadrats [Daubenmire 1959] per plot). The size of the sampling plot allowed for the relative comparison of accuracy assessment results at the hyperspectral and multispectral scales because the plots were treated as polygons at the hyperspectral scale and as individual pixels at the multispectral scale. Sampling was initiated at the Spencer site on 16 June 2006, a few days before full bloom, and continued during and shortly after peak phenology, ending on 26 July 2006. A total of 56 plots, 43 with leafy spurge present and 13 with leafy spurge absent, were sampled. Validation samples were collected in Medicine Lodge from 26 July to 13 August 2006, after peak phenology. A total of 55 plots, 43 with leafy spurge present and 12 with leafy spurge absent, were sampled in Medicine Lodge.

To calibrate ocular estimates of leafy spurge percentage canopy cover across a continuous interval, estimates for the first five plots included an initial ocular estimate at the plot scale, followed by estimates at each of the nine quadrats using a point frame (Floyd and Anderson 1982) and a Daubenmire quadrat frame (Daubenmire 1959). Initial estimates at the plot scale were consistently closer to the average quadrat estimations using a point frame (only one of the five calibration plots varied by more than 1%). Estimations using the Daubenmire quadrat were consistently about 20% lower than initial ocular estimates at the plot scale. Although the point frame estimation technique was designed for sagebrush steppe ecosystems and is regarded as a more objective method than visual cover estimation (Bonham 1989), the Daubenmire frame was chosen for its ease of use and speed to estimate cover at the quadrant scale. Regression plots indicated strong agreement between the ocular cover estimation techniques at the plot and quadrant level for leafy spurge ($r^2 = 0.76$; not shown). These plots also suggested that results are less variable when estimating low and high percentage cover than when estimating percentage canopy cover in the midrange (20–60%).

Image Preprocessing

All image preprocessing and processing, unless otherwise stated, was performed using the Environment for Visualizing Images version 4.3 software (ITT Visual Information Solutions, Boulder, CO). Hyperspectral radiance values were converted by the vendor to apparent reflectance using the HyCorr (Hyperspectral Correction) absolute atmospheric correction modeling package, which was developed by Commonwealth Scientific and Industrial Research Organization Division of Exploration and Mineral Mining Mapping and is based on the Atmospheric Removal Program (ATREM; Gao and Goetz 1990; ATREM 1992). The multispectral imagery was converted to apparent reflectance with FLAASH (Adler-Golden et al. 1999). The absolute atmospheric corrections produced scaled surface reflectance values that account for scattering and absorption of solar radiation by the earth’s atmosphere. Such corrections are relevant to this study because they enable data recorded at the sensor to be directly compared to data recorded on the ground (i.e., field spectroscopy measurements and oblique ocular cover estimates) and to other remotely sensed images obtained under different atmospheric conditions (i.e., comparisons between HyMap and Thematic Mapper imagery of the study area).

To assess HyMap georegistration error, 10 differentially corrected global positioning system ground control points were collected in the central Spencer flightline and nine in the north-south Medicine Lodge flightline. Ground control points were collected for these two flightlines because they contained the majority of ground reference samples. Directional shifts occurred nonuniformly, with mean errors of 3.39 m and 0.813 m for the Medicine Lodge and Spencer flightlines, respectively. For the purpose of conservative accuracy estimations, buffers were not applied to the validation plot perimeters to accommodate georegistration error.

The influence of topographic error was minimized in the Landsat 5 TM image by automatically registering the image to a corresponding terrain-corrected Multi-Resolution Land Characteristics [MRLC] image (row 39, path 29; 1 July 2001, 30-m pixels). Optimal georegistration corrections decreased positional errors to a range of one to three pixels and were achieved using an area-based rather than a feature-based registration. The apparent reflectance image was warped to the MRLC using band 2 from both images and first degree polynomial resampling with cubic convolution.

Image Processing

The two Medicine Lodge hyperspectral flightlines were processed as a single georeferenced mosaic, and the three Spencer flightlines were processed as a single georeferenced mosaic. After a preliminary evaluation of the data, bands obviously influenced by noise or water absorption were removed from the mosaics. The two study sites were classified independently, because the use of a single endmember from the Medicine Lodge training area, when applied to both sites, produced unrealistic classification results for the Spencer site. An endmember is defined here as a single pure pixel selected within the image. The spectral signature of the endmember is used to unmix and estimate subpixel target abundance in the remaining pixels of the image. In addition, nonparametric Kolmogorov-Smirnov ($P = 0.037$) and Mann-Whitney ($z$-score = 2.832) tests were used to compare the distribution shapes and population statistics for leafy spurge cover data collected at the Medicine Lodge and Spencer sites. Both tests concluded that the ground reference data sampled at the Medicine Lodge and Spencer sites were not statistically similar at a 95% confidence level.

To explore changes in leafy spurge detection performance at coarsened spectral and spatial resolutions, different combina-
tions of spectral and spatial resampling were applied to the Medicine Lodge HyMap mosaic (121 bands, 3.2-m pixels) to produce three additional images: a spatially degraded image, spectrally degraded image, and spectrally and spatially degraded image (Fig. 2). The spatially degraded image was generated using a pixel aggregation method to simulate the Landsat 5 TM spatial resolution (30-m pixels) while retaining hyperspectral resolution (121 bands). The pixel aggregation method averages all of the pixels that contribute to the output pixel. In this study spectral averages of roughly 100 contributing pixels were used to generate 30-m-scale images. This method is considered a square-wave approach because the measurement that is produced assumes that radiance is equally weighted within the sensor’s field of view (Gao and Huete 2000). The spectrally degraded image was generated using a filter function to simulate the Landsat 5 TM spatial resolution (30-m pixels) while retaining hyperspectral resolution (121 bands). The pixel aggregation method averages all of the pixels that contribute to the output pixel. In this study spectral averages of roughly 100 contributing pixels were used to generate 30-m-scale images. This method is considered a square-wave approach because the measurement that is produced assumes that radiance is equally weighted within the sensor’s field of view (Gao and Huete 2000).

Image Classification

MTMF is a spectral mixture analysis technique that has been successfully used in previous studies to identify leafy spurge in hyperspectral imagery (Parker Williams and Hunt 2002, 2004; Dudek et al. 2004; Glenn et al. 2005). The MTMF classification generates two data sets: a matched filtering (MF) band and an infeasibility band. Matched filtering scores provide estimates of subpixel target abundance, where a score near zero would be interpreted as background or noise in the image and a score of one would be interpreted as a perfect match to the spectral signature of an endmember (Harsanyi 1993; Harsanyi and Chang 1994; Boardman 1998). Infeasibility values provide estimates of how closely the pixels approximate the endmember pixel or the likelihood that the classified pixel is a false positive (Boardman 1998). Further details on the MTMF method can be found in Mundt et al. (2007).

MTMF classifications were applied to the Spencer and Medicine Lodge HyMap mosaics, the three degraded Medicine Lodge HyMap mosaics, and the Landsat 5 TM image focused on the Medicine Lodge and Spencer areas (Fig. 2). First,
forward minimum noise fraction (MNF) transformations were applied to the reflectance bands of each dataset (Green et al. 1988). These transformations estimated noise statistics using shift differencing over the complete scenes (Mundt et al. 2007). Shift differencing assumes adjacent pixels have the same signal but different noise. Resultant MNF-transformed bands were reordered in terms of unit noise standard deviations such that coherent bands could be separated from noise-dominated bands and retained in later image processing steps (Kruse 2003).

For the Spencer HyMap mosaic, the first 40 MNF bands explained 83% of the data and were retained to identify endmembers for classification. For the Medicine Lodge HyMap mosaic, the first 30 MNF bands explained 82% of the data and were likewise retained for endmember derivation. For images with six multispectral bands, all six bands were retained for subsequent classification steps.

Endmembers were chosen for the Medicine Lodge and Spencer datasets comparing pixels identified from training areas in the images to spectrally pure pixels identified in image processing. Where more than one potential endmember was identified for each dataset, we selected a user-guided endmember pixel with high percentage target cover and an average spectral signature. (Glenn et al. 2005; Mundt et al. 2007). The endmember identified for the MTMF classification of the Medicine Lodge HyMap mosaic was also used to identify endmembers for the MTMF classifications of the two degraded images. For both the spatially degraded HyMap image and the spectrally and spatially degraded HyMap image, a single user-defined endmember pixel was selected for each classification by examining spectral signatures of pixels in the vicinity of the training areas. In each degraded image the spectral signature that most closely resembled that of the previously identified Medicine Lodge endmember was selected as the classification endmember.

To arrive at presence or absence classification thresholds for individual MTMF classifications, spatial subsets within the project area mosaics were used to generate scatterplots of infeasibility values (y) vs. MF scores (x; Fig. 3). Clusters of pixels within each scatterplot were interactively selected to investigate which portions of the pixel cloud corresponded to noise, or spectrally distinct classes such as riparian or bare ground. Once a threshold was determined that represented a balance between errors of omission and commission, linear regression was used to describe the leafy spurge presence or absence threshold line. The matched filtering and infeasibility images were then mathematically combined such that values below the threshold equated to presence and values above the threshold equated to absence.

Figure 3. A scatterplot of infeasibility values (y-axis) vs. matched filter (MF) scores (x-axis) illustrates the use of a threshold to delineate target presence and absence. The MF demonstrates abundance of leafy spurge and the infeasibility value provides a measure of false positives.
A series of classification methods were applied to the entire Landsat 5 TM image, including MTMF, SAM, maximum likelihood, and minimum distance. These classification methods were applied to all six bands, both untransformed and MNF transformed. The MTMF classification method was applied to the Medicine Lodge area of the Landsat 5 image using an endmember pixel from the Medicine Lodge training area, and the MTMF classification method was also applied to the Spencer area of the Landsat 5 image using an endmember pixel from the Spencer training area. Figure 4 depicts spectral signatures for pixels that were selected as final MTMF classification endmembers. Figure 5 depicts spectral signatures for pixels that were selected as final SAM classification endmembers. A total of five image-derived endmembers from the Medicine Lodge and Spencer training areas were used in the SAM classification (Fig. 5).

Accuracy Assessment

This study assessed the accuracy of each presence/absence classification by generating error matrices and calculating overall accuracy, user’s accuracy (percentage of pixels that are correctly classified on the ground), and producer’s accuracy (percentage of a given class that is correctly identified on a map; Congalton and Green 1999; Congalton 2004; Foody 2004). No training data were used in the error matrices. The error matrices were also used to compute a kappa coefficient of agreement for each classification. The kappa statistic is a measure of how well the classified map agrees with the validation samples compared to chance agreement. An incremental cover technique was applied to each classified Medicine Lodge image to quantify the minimum percentage cover of leafy spurge necessary to detect leafy spurge under various resolution scenarios using the producer’s accuracy (Mundt et al. 2006). This method assumes that pixels containing high percentage target cover are more likely to classify correctly than pixels containing low percentage target cover. The producer’s accuracy is calculated by dividing the number of samples classified as leafy spurge by the total number of presence (field) samples. Consequently, the producer’s accuracy for the presence category should increase as infestations with lower percentage cover are successively removed. Changes in producer’s accuracy are evaluated in cumulative 10% cover increments.

RESULTS

The true HyMap Medicine Lodge classification had an overall accuracy of 85% and a kappa value of 0.65, which indicated that agreement in the error matrix was significantly greater than chance agreement (Table 1). The true HyMap Spencer classification had an overall accuracy of 67% and a kappa value of 0.31 (Table 1). Comparing the Medicine Lodge and Spencer classifications (both HyMap and Landsat 5 TM tuned to each area), the Medicine Lodge classifications outperformed the Spencer classifications (Table 1). The MTMF classification of the Landsat 5 image tuned to the Medicine Lodge area produced a high overall accuracy (62%) but was accompanied by a low kappa value (0.38). For classifications of the entire Landsat 5 scene (i.e., minimum distance, maximum likelihood, and SAM using both untransformed and MNF transformed bands as input), the only classification method realistic enough...
to quantify with an accuracy assessment was the SAM classification using MNF transformed bands. This classification method produced an overall accuracy of 46%, and the kappa value of 0.15 indicated poor agreement. The high user’s accuracy for the present category (96%) and the high producer’s accuracy for the absent category (96%) are artifacts of the relatively low number of absence validation samples (Table 1).

The Medicine Lodge simulated Landsat 5 TM classification outperformed the true Landsat 5 TM classification of the Medicine Lodge area (Table 2). An acceptable producer’s accuracy for this project was determined by land managers to be 70% given the risk associated with committing an error of omission. The simulated classification produced acceptable results for detecting leafy spurge infestations for incremental cover classes with canopy cover greater than 20%. The Landsat 5 TM classification failed to produce acceptable results for detecting leafy spurge infestations with the exception of infestations with 71–100% cover (corresponding to a producer’s accuracy of 71%). Incremental cover evaluations of the Medicine Lodge spectrally degraded HyMap mosaic resulted in producer’s and overall accuracies that were higher than the HyMap classification (Table 2). The simulated Landsat 5 tended to perform slightly better than the spatially degraded mosaic, even though the spatially degraded image retained high spectral resolution (Table 2).

**DISCUSSION**

Baseline classifications of leafy spurge using HyMap imagery produced high overall accuracies in Spencer (67%) and Medicine Lodge (85%). These accuracies were influenced by the inclusion of noise, which was considered an acceptable tradeoff for minimizing the risk of committing an error of omission. The Spencer classification accuracies are lower than

---

**Table 1.** Comparison of HyMap and Landsat 5 Thematic Mapper (TM) classifications using mixture-tuned matched filtering (MTMF) and spectral angle mapper (SAM) for the Medicine Lodge and Spencer study sites.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Medicine Lodge HyMap using MTMF</th>
<th>Medicine Lodge area Landsat 5 TM using MTMF</th>
<th>Spencer HyMap using MTMF</th>
<th>Spencer area Landsat 5 TM using MTMF</th>
<th>Entire Landsat 5 TM image using SAM and transformed input bands</th>
</tr>
</thead>
<tbody>
<tr>
<td>User’s accuracy (present)</td>
<td>100%</td>
<td>89%</td>
<td>64%</td>
<td>31%</td>
<td>96%</td>
</tr>
<tr>
<td>User’s accuracy (absent)</td>
<td>60%</td>
<td>35%</td>
<td>77%</td>
<td>62%</td>
<td>30%</td>
</tr>
<tr>
<td>Producer’s accuracy (present)</td>
<td>80%</td>
<td>59%</td>
<td>90%</td>
<td>72%</td>
<td>30%</td>
</tr>
<tr>
<td>Producer’s accuracy (absent)</td>
<td>100%</td>
<td>75%</td>
<td>40%</td>
<td>22%</td>
<td>96%</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>85%</td>
<td>62%</td>
<td>67%</td>
<td>38%</td>
<td>46%</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.65</td>
<td>0.38</td>
<td>0.31</td>
<td>−0.05</td>
<td>0.15</td>
</tr>
</tbody>
</table>

**Figure 5.** Spectral signatures of endmember pixels used in spectral angle mapper (SAM) and mixture-tuned matched filtering (MTMF) classifications of the entire Landsat 5 Thematic Mapper image.
those of Medicine Lodge because of lower spectral contrast between background and target vegetation in the Spencer area. The 2003 Deep Fire in Medicine Lodge reduced vegetation cover and exposed burned soil throughout the study site. Leafy spurge reestablished quickly and was a prominent landscape feature. Field reference samples in the Spencer area had lower average leafy spurge cover, denser shrub coverage, and less bare ground. Accordingly, pixels containing leafy spurge in the Spencer area were difficult to identify in the MF vs. infeasibility plots because they tended to occur in the dense center of pixel clouds, along with a large number of other mixed and background pixels (Fig. 3). These factors may explain why leafy spurge was frequently confused with riparian areas and mixed shrub communities in the Spencer classifications.

Poor classification performance results for the entire Landsat 5 TM image were attributed to greater spectral variability and mixing associated with a regional study area. In addition, the sample size for the accuracy assessment was not ideal in relation to the size of the Landsat 5 image. The simulated Landsat 5 and actual Landsat 5 TM classification results are inconsistent with previous work that compared simulated Landsat Enhanced Thematic Mapper Plus (ETM+) classifications derived from AVIRIS imagery (174 bands, 4-m pixels) to a Landsat ETM+ classification of six vegetation community types containing target invasive species (Underwood et al. 2007). Rather, the authors reported comparable results for the simulated and true classifications and attributed small differences in mapping accuracies to instrument differences and calibration and geocorrection issues associated with the AVIRIS flightlines. It is inferred from the results presented herein that the Landsat 5 TM simulated classification derived from the HyMap sensor performed better than the Landsat 5 TM because of HyMap’s higher geometric resolution and better SNR. SNRs for the HyMap sensor approach 1000:1 (Cocks et al. 1998), whereas SNRs for the Landsat 5 TM are less than 100:1. This suggests that a sensor with comparable spectral and spatial resolutions but improved instrumentation could be a viable weed management resource. This inference is supported by a recent study by Stitt et al. (2006), who used the new Landsat prototype sensor ALI to detect leafy spurge (which has a higher SNR than Landsat 5 TM). Accuracy assessment results indicated low omission errors (high producer’s accuracy) and demonstrated the potential for regional leafy spurge distribution mapping using multispectral sensors with improved instrumentation.

Higher results for the simulated Landsat 5 TM classification compared to the spatially degraded HyMap classification and higher results for the spectrally degraded HyMap classification compared to the HyMap classification suggest that fewer broad bands may have an advantage for leafy spurge discrimination. Within Medicine Lodge, the HyMap and the spectrally degraded HyMap (six bands, 3.2-m pixels) provided the highest producer’s and overall accuracies at all cover levels. Interestingly, even when low cover values were included \( n = 53, 0–100\% \), both the producer’s and overall accuracies were above 80\% for both of these images. Additional research is needed to identify which, if not all, bands were essential for the discrimination of leafy spurge in the spectrally resampled images.

### Table 2

Incremental cover evaluations (Mundt et al. 2006) of leafy spurge using Medicine Lodge HyMap (actual and degraded) and Landsat 5 Thematic Mapper (TM) images with mixture-tuned matched filtering.

<table>
<thead>
<tr>
<th>Sample size</th>
<th>53</th>
<th>51</th>
<th>49</th>
<th>44</th>
<th>41</th>
<th>37</th>
<th>34</th>
<th>29</th>
<th>24</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leafy spurge % canopy cover classes</td>
<td>&gt;0%</td>
<td>&gt;10%</td>
<td>&gt;20%</td>
<td>&gt;30%</td>
<td>&gt;40%</td>
<td>&gt;50%</td>
<td>&gt;60%</td>
<td>&gt;70%</td>
<td>&gt;80%</td>
<td>90%</td>
</tr>
<tr>
<td>HyMap (121 bands, 3.2-m pixels)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Producer’s accuracy</td>
<td>80%</td>
<td>85%</td>
<td>89%</td>
<td>88%</td>
<td>93%</td>
<td>96%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>85%</td>
<td>88%</td>
<td>92%</td>
<td>91%</td>
<td>95%</td>
<td>97%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.65</td>
<td>0.72</td>
<td>0.80</td>
<td>0.79</td>
<td>0.89</td>
<td>0.94</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>HyMap: spectrally degraded (6 bands, 3.2-m pixels)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Producer’s accuracy</td>
<td>88%</td>
<td>92%</td>
<td>95%</td>
<td>97%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>91%</td>
<td>94%</td>
<td>96%</td>
<td>98%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.77</td>
<td>0.85</td>
<td>0.90</td>
<td>0.94</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>HyMap: spatially degraded (121 bands, 30-m pixels)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Producer’s accuracy</td>
<td>59%</td>
<td>62%</td>
<td>65%</td>
<td>69%</td>
<td>72%</td>
<td>76%</td>
<td>77%</td>
<td>82%</td>
<td>83%</td>
<td>100%</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>68%</td>
<td>71%</td>
<td>73%</td>
<td>77%</td>
<td>80%</td>
<td>84%</td>
<td>85%</td>
<td>90%</td>
<td>92%</td>
<td>100%</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.39</td>
<td>0.43</td>
<td>0.47</td>
<td>0.55</td>
<td>0.61</td>
<td>0.67</td>
<td>0.71</td>
<td>0.79</td>
<td>0.83</td>
<td>1.00</td>
</tr>
<tr>
<td>HyMap TM simulation (6 bands, 30-m pixels)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Producer’s accuracy</td>
<td>63%</td>
<td>67%</td>
<td>70%</td>
<td>72%</td>
<td>72%</td>
<td>72%</td>
<td>82%</td>
<td>82%</td>
<td>83%</td>
<td>67%</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>72%</td>
<td>75%</td>
<td>78%</td>
<td>80%</td>
<td>80%</td>
<td>81%</td>
<td>88%</td>
<td>90%</td>
<td>92%</td>
<td>93%</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.44</td>
<td>0.48</td>
<td>0.54</td>
<td>0.58</td>
<td>0.61</td>
<td>0.63</td>
<td>0.76</td>
<td>0.79</td>
<td>0.83</td>
<td>0.76</td>
</tr>
<tr>
<td>Landsat 5 TM: Medicine Lodge Area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Producer’s accuracy</td>
<td>59%</td>
<td>56%</td>
<td>57%</td>
<td>53%</td>
<td>55%</td>
<td>64%</td>
<td>68%</td>
<td>71%</td>
<td>67%</td>
<td>33%</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>62%</td>
<td>61%</td>
<td>61%</td>
<td>59%</td>
<td>61%</td>
<td>68%</td>
<td>71%</td>
<td>72%</td>
<td>71%</td>
<td>67%</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.38</td>
<td>0.37</td>
<td>0.38</td>
<td>0.39</td>
<td>0.42</td>
<td>0.50</td>
<td>0.55</td>
<td>0.60</td>
<td>0.62</td>
<td>0.65</td>
</tr>
</tbody>
</table>
classification. For example, multispectral bands in the mid-infrared (i.e., 1.55–1.75 μm) may actually decrease unsupervised leafy spurge classifications (Stitt et al. 2006).

This work demonstrates that leafy spurge is spectrally distinct in portions of the electromagnetic spectrum that are captured in the bandwidths of Landsat 5 TM. For example, Lewis et al. (2000), using discriminate analysis to identify spectral regions most critical to the discrimination of arid Australian vegetation types, found that relative differences across broad spectral regions were more relevant than a select number of narrow bands. The high spectral dimensionality of the HyMap classification could have resulted in greater spectral noise, or less background separability, whereas spectral noise was averaged out for the wider Landsat 5 bandwidths. Another possible noise factor to consider is that the shift differencing technique used to estimate noise during the MNF transformation may have reduced different net quantities of noise in the multispectral data than in the hyperspectral data (Kruse 2003; Mundt et al. 2007). Additional factors that could have influenced results include resampling algorithms, the order in which the simulated image was created (i.e., spectral resampling was applied first and spatial resampling was applied second), and subjectivity associated with MTMF presence or absence thresholding. Although this study spatially degraded imagery by applying a pixel aggregation technique to the imagery, applying a modular transfer function to the instrument could produce different results (Gao and Huete 2000).

The results of this study were primarily limited by georegistration error and variables associated with field data collection, such as sample size, sampling design, timing of data collection, and the ocular method of estimating percent cover. As indicated earlier, results are considered conservative in that buffering was not applied to reference samples to accommodate locational errors in the HyMap imagery and topographic displacement in the Landsat scene. Maximum positional errors were estimated at 6.11 m for the HyMap imagery and 90 m for the Landsat imagery. The rule of thumb for the minimum number of suggested samples within an error matrix land cover class is 50 (Jensen 2005). This study used 41–42 samples for the presence category and a disproportionately low number of samples for the absence category (12–13). Sample size was limited by the collection of additional field data at sample locations (e.g., species diversity and microhabitat measurements) as well as the need to validate each study area independently. Ground reference samples were not completely random due to challenges associated with the need for a priori knowledge of infestation locations to ensure a representative number of presence locations were sampled within a reasonable amount of time. Although we were able to sample the Spencer study area at the same time the imagery was acquired and leafy spurge was in peak bloom, cover may have been underestimated at the Medicine Lodge site because sampling occurred after image acquisition and peak phenology. In addition, ocular estimates of percent canopy cover are subjective and prone to observer bias (Elzinga et al. 1998; Booth et al. 2003). Uncertainties associated with ocular estimation in this study were constrained by the use of a single observer and the collection of subplot cover estimates using a Daubenniere frame.

Although study results were limited by the aforementioned factors, they provide insight into the suitability of broadband multispectral imagery and the relative influence of spectral and spatial resolutions on leafy spurge discrimination. The Landsat MTMF classification that was tuned to the Medicine Lodge area produced the best Landsat 5 TM classification, but accuracies and agreement (kappa value) were not high enough to warrant exploration in applications such as historical mapping, and surveying and monitoring of core infestations. Significantly, the study sites in Clark County, Idaho, represent ideal demonstration areas for leafy spurge mapping in terms of both extent and percent cover. Therefore, if data from a true Landsat image did not produce acceptable results here, it is unlikely that the sensor would perform better elsewhere.

**IMPLICATIONS**

Although one may expect the high spectral resolution of a hyperspectral sensor to increase the ability to distinguish leafy spurge from its surrounding background, the results of this study suggest that fewer broad bands may have an advantage for leafy spurge discrimination. From a rangeland management perspective, the prospect of multispectral sensors with improved radiometric and SNR capabilities could mean cost-effective tools for monitoring leafy spurge infestations at the regional scale. Although sensors with these capabilities are not readily available today, there is potential in the near term with the Landsat Data Continuity Mission and other proposed sensors.

**LITERATURE CITED**


