Mapping Sagebrush Distribution Using Fusion of Hyperspectral and Lidar Classifications

Jacob T. Mundt, David R. Streutker, and Nancy F. Glenn

Abstract

The applicability of high spatial resolution hyperspectral data and small-footprint Light Detection and Ranging (lidar) data to map and describe sagebrush in a semi-arid shrub steppe rangeland is demonstrated. Hyperspectral processing utilized a spectral subset (605 nm *to 984* nm*) of the reflectance data to classify sagebrush presence to an overall accuracy of 74 percent. With the inclusion of co-registered lidar data, this accuracy increased to 89 percent. Furthermore, lidar data were utilized to generate stand specific descriptive information in areas of sagebrush presence and sagebrush absence. The methods and results of this study lay the framework for utilizing co-registered hyperspectral and lidar data to describe semi-arid shrubs in greater detail than would be feasible using either dataset independently or by most ground based surveys.*

Introduction

Wyoming big sagebrush *(Artemisia tridentata Spp. wyomingensis)* is an important rangeland shrub associated with the Intermountain West shrub steppe environment. Many vertebrate species utilize habitats within sagebrush to maintain viable populations (e.g., pygmy rabbit *(Brachylagus idahoensis)*, sage grouse *(Centrocercus urophasianus)*, and sharp-tailed grouse *(Tympanuchus phasianellus)*). As sagebrush communities in the Intermountain West become increasingly fragmented due to agricultural and urban growth, range fires, and invasive weeds, critical habitats and historic grazing/browsing regimes become threatened (NPCC, 2004). Thus, land managers and conservation agencies need accurate tools to inventory and assess these semiarid shrubs. This study combines the strengths of passive, high spatial resolution, hyperspectral reflectance imagery and active, small-footprint Light Detection and Ranging (lidar) pulse data to assess sagebrush distribution and stand structure information, exceeding the mapping ability of either dataset alone.

The study area for this project covers nine square kilometers within the Idaho National Engineering and Environmental Laboratory (INEEL) in southeastern Idaho (Figure 1). The vegetative composition of the sagebrush ecosystem in the study area commonly includes Wyoming big sagebrush, green rabbitbrush *(Chrysothamnus viscidiflorus)*, Russian thistle *(Salsola kali)*, and winterfat *(Ceratoides lanata)*. Associated native grasses include thick-spiked wheatgrass *(Elymus lanceolatus)*, bottlebrush squirreltail *(Elymus elymoides)*, Indian ricegrass *(Oryzopsis hymenoides)*, needle-and-thread grass *(Stipa comata)*, and Nevada bluegrass *(Poa secunda)* (ESER, 2004).

Two large fires burned across separate regions of the study area, the first in 1974 and the second in 2000. Wildfires are naturally occurring disturbances in semi-arid rangelands, and it may take decades for sagebrush to re-establish a viable population following a disturbance (Colket, 2003). Thus, in the burned areas, sagebrush is dominantly absent, having been replaced by increased populations of green rabbitbrush, cheatgrass *(Bromus tectorum)*, native grasses, and Russian thistle. Additionally, land managers have seeded some burned areas with crested wheatgrass *(Agropyron desertorum)* in an attempt to increase botanical diversity.

Previous Work

Remote sensing techniques have the unique ability to capture a spatial and temporal snapshot of an ecosystem that can feasibly be used to establish geographic context and perform spatial interrogation of ecological health and inventory (Patil and Myerst, 1999). Remotely sensed data also have the potential to provide large-scale information on vegetation in semi-arid and arid rangelands (Tueller, 1987). Multispectral sensors have been used for more than 30 years to characterize vegetation across the globe, while hyperspectral instruments came into widespread use in the 1990s. Warren and Hutchison (1984) document the utility of remote sensing for holistic change detection in shrub/grass rangeland. Other studies have demonstrated success at mapping ecological resources in semi-arid systems using high spectral resolution data (e.g., Elmore *et al.*, 2000; Lewis *et al.*, 2000; Okin *et al.*, 1999), and the use of multiple or fractional end-members has been shown to increase classification confidence in semi-arid environments (e.g., Roberts *et al.*, 1998). Okin *et al.* (2001), however, encountered complications in discrimination of spectrally indeterminate arid shrubs using high spectral resolution field data. Significant soil exposure and an increased level of non-linear spectral mixing in semi-arid environments likely influenced these complications (Huete and Jackson, 1988; Huete *et al.*, 1985; Musick, 1984; Ray and Murray, 1998).

While lidar is newer technology than passive reflectance remote sensing, recent studies have demonstrated the applicability of small (1 m or less) and large (10 m to 25 m) footprint lidar data to precisely characterize forest canopies. Waveform returns from large-footprint lidar sensors such as the Scanning Lidar Imager of Canopies by Echo Recovery (SLICER) and Laser Vegetation Imaging Sensor (LVIS) have successfully been used to measure variables such as tree height, canopy structure, leaf area index (LAI), and biomass

Department of Geosciences, Boise Center Aerospace Laboratory, Idaho State University, 322 E. Front Street, Suite 240, Boise, ID 83702 (mundjaco@isu.edu; stredavi@isu.edu; glennanc@isu.edu).

Photogrammetric Engineering & Remote Sensing
Vol. 72, No. 1, January 2006, pp. 47–54.

0099-1112/06/7201–0047/$3.00/0

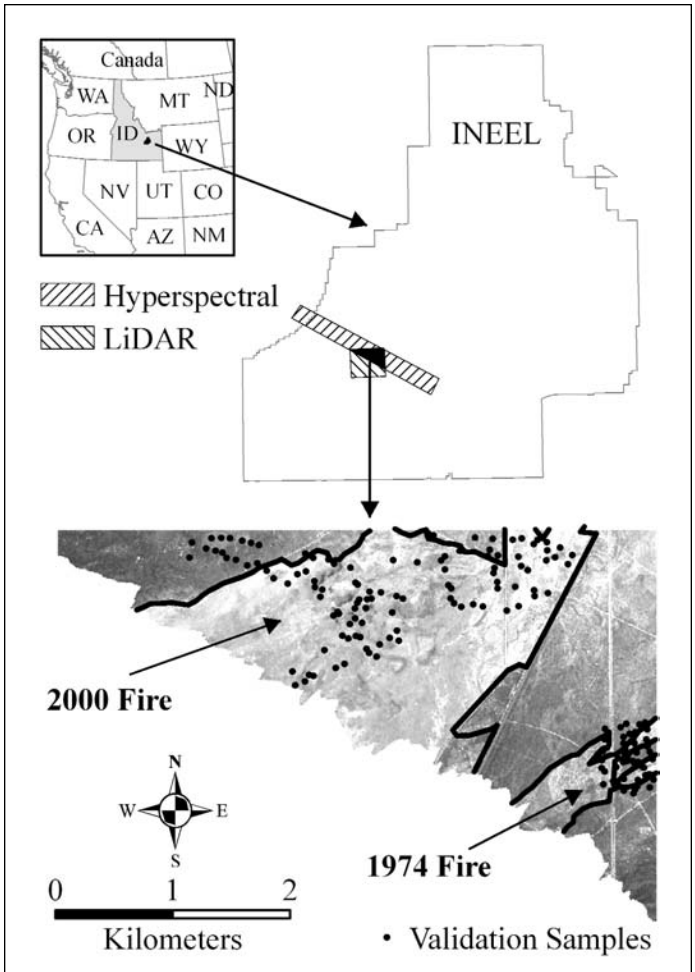

Figure 1. Location of the study area within the Idaho National Engineering and Environmental Laboratory (INEEL). In the hyperspectral image of the study area, dark lines bound the burned areas and points represent field validation sites.

(e.g., Drake *et al.*, 2002; Harding *et al.*, 2001; Hofton *et al.*, 2002). Similarly, small-footprint lidar has been used to determine tree height and LAI in various forest types (e.g., Drake *et al.*, 2002; Popescu *et al.*, 2002; Riaño *et al.*, 2004).

Relative to the number of forestry studies, little work has addressed lidar applications in rangeland ecosystems. Height profiles created by non-scanning lidar sensors have been used to measure plant height and canopy cover in a rangeland environment (Weltz *et al.*, 1994) and to model rangeland surface roughness (Ritchie *et al.*, 1995). Other studies have demonstrated the applicability of scanning lidar combined with multispectral video imagery to map shrub coppice dunes in desert grasslands (Rango *et al.*, 2000).

Hyperspectral and lidar data have fundamentally different characteristics. Lidar data are irregularly spaced monochromatic near-infrared laser pulses (postings) emitted by and reflected back to the sensor. The postings are distributed in a pseudorandom fashion and are thus assumed to statistically represent the terrain characteristics. Conversely, hyperspectral data are intensity values of radiation reflected from a contiguously recorded surface across a broad range of narrow electromagnetic bandwidths. Though these differences imply complications for data fusion, some recent work has quantitatively related lidar and reflectance datasets. Popescu and Wynne (2004) used multispectral imagery to distinguish between deciduous and pine trees, thereby enhancing lidar-dependent tree height models. Additionally, Hudak *et al.* (2002) successfully fused Landsat ETM+ imagery with lidar to improve tree height calculation. Lee and Shan (2003) combined lidar with multispectral Ikonos imagery to improve upon supervised classification accuracies in a coastal environment, and Blackburn (2002) used lidar to mask tree gaps while using hyperspectral Compact Airborne Spectrographic Imager (CASI) data to estimate photosynthetic pigments in forest plantations.

Methods

Data Collection

This study fuses hyperspectral data collected in July 2002 with lidar data collected approximately two months later. The hyperspectral data were acquired by the HyVista Corporation of North Ryde, New South Wales, Australia. The dataset has a spatial resolution of 4.6 m × 4.6 m and contains 126 spectral bands between 450 *n*m and 2,500 *n*m, with spectral bandwidths of approximately 10 *n*m in the visible and near-infrared and 20 *n*m in the short-wave infrared. The lidar data were collected by the Airborne 1 Corporation of El Segundo, California using a 25 KHz scanning pulsed laser at a wavelength of 1,067 *n*m. The dataset contains 13.3 million postings within the study area (approximately 1.2 postings per square meter) where each lidar posting has a diameter of approximately 25 cm. Two returns (first and last) were recorded for each posting, and each return included a time stamp, the *X*, *Y*, and *Z* coordinates, and an intensity value.

A field survey in the summer of 2004 collected 157 differentially corrected Global Positioning System (GPS) validation samples within the study area (Figure 1). Each validation sample was acquired as a point location, attributed by fractional coverage of sagebrush, grass, rabbitbrush, and bare ground (oblique field estimates) to a search radius of approximately 15 m. In addition to the point data samples, GPS polygon samples and line features were also collected around geographically distinctive features in the study area (e.g., rock outcrops, sand dunes, and roads) to assist in the assessment of image registration errors.

Although there was a two-year period between image acquisition and field data collection, we assume negligible changes in the sagebrush ecosystem. Independent test plot measurements of mature sagebrush canopy indicate a variation in cover between 5 percent and 10 percent over two-year intervals (Inouye, personal communication, 2004). Additionally, the study area is neither heavily grazed nor public land, and no major range fires or other disturbances occurred in the area between the time of data collection and the time the validation was performed.

Hyperspectral Processing

The hyperspectral data were atmospherically corrected by the vendor using the MODTRAN4 radiative transfer code (Berk *et al.*, 1999). Exploratory data evaluation identified an inconsistency in cross-track illumination, which was rectified using a multiplicative correction algorithm built into the Environment for Visualizing Images (ENVI) software (RSI, 2004). Further evaluation revealed generally higher reflectance in the visible wavelengths within the fire scars as compared to areas of mature sagebrush. We hypothesized that these differences were the result of variation in vegetative composition and increased soil exposure. Preliminary unsupervised classifications, performed to explore the

spectral variability within the data, determined that areas of sagebrush on the west side of the study area were spectrally distinct from areas of sagebrush on the east side. This artifact was correlated to a change in soil composition, with the east portion of the study area typified by thick floodplain sediments and the west portion by aeolian sands overlying basalt.

Hyperspectral data were classified utilizing the Mixture Tuned Matched Filtering (MTMF) algorithm. MTMF is a partial unmixing algorithm that combines the strengths of linear spectral unmixing and statistical matched filtering to produce a Matched Filter (MF) score (estimate of target abundance) and a measure of infeasibility (Boardman, 1998).

MTMF classifications utilizing the full spectral range of the hyperspectral dataset overestimated sagebrush presence within the burn scars. Comparison of spectral profiles derived from areas of both sagebrush presence and sagebrush absence illustrated strong similarities across the spectral range of the dataset (Figure 2). In the visible and near-infrared, however, the profiles deviate slightly. Accordingly, the data were reprocessed using only the spectral range from 605 *n*m to 984 *n*m (27 spectral bands), and the resulting classification generated a qualitatively improved representation of sagebrush distribution.

MTMF classification requires the input of spectral training endmembers. While the objective of this study is to map the distribution of sagebrush, it is notable that there is a certain amount of heterogeneity within a sagebrush stand (e.g., dead brush, rabbitbrush, grasses, and soil exposure may vary locally) (Huete and Jackson, 1988). In this study, areas referred to as *sagebrush* are realistically only partially *Artemisia tridentata*, and it is assumed that no sagebrush stand will be an *Artemisia tridentata* monoculture. Thus, from an ecological perspective, it is important to consider any vegetative compositional variance within the sagebrush ecosystem when mapping. To model this variance, training areas were input in block (polygon) form (Chen and Stow, 2002). Iterative data classification demonstrated that a single sagebrush endmember (derived from an area of sand overlying basalt soils) effectively mapped sagebrush locally, but was not effective across the entire study area. We hypothesized that this anomaly was dominantly due to variability in underlying soil texture and composition. The addition of a second block endmember derived from an area of sagebrush underlain by floodplain sediments also performed well locally, and visually complemented the first endmember classification.

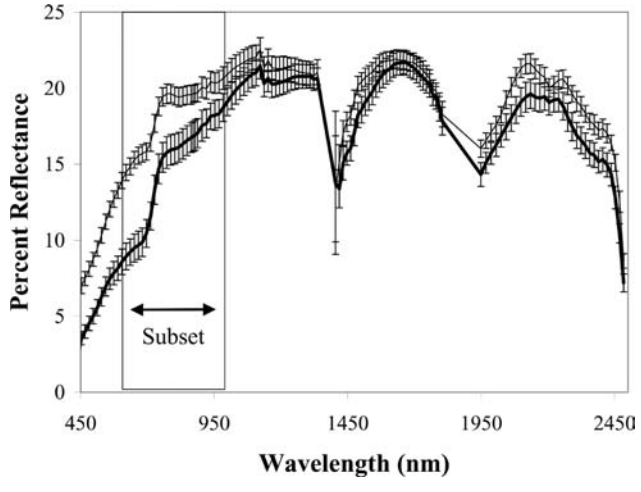

Figure 2. Spectral profiles derived from burned (upper spectra, 22 sample mean) and unburned (lower spectra, 18 sample mean) areas. Typical burned areas lack any sagebrush, while the unburned areas had relatively high percent sagebrush cover. The spectral range used for the final hyperspectral analysis is noted.

As MTMF is a linear operator with MF scores representing sub-pixel target abundance, the two endmember classifications, in the form of MF values, were arithmetically summed to derive a composite sub-pixel abundance of the target. An arithmetic average of the two bands of infeasibility was calculated and combined with the summed MF scores, generating a composite MTMF classification. This composite classification represented both sagebrush endmembers, and more appropriately estimated sagebrush abundance across the entire study area.

Lidar Processing

The lidar data over the study area were collected in 22 separate flight lines, each approximately 300 m wide. The vendor performed a GPS validation survey to constrain the absolute accuracy of the lidar coordinates to approximately 1 m in the horizontal direction and 25 cm in the vertical direction (two-sigma; Flood, 2004). A similar accuracy was determined in the study area using assessment methods modeled after Latypov (2002) which calculated the error between individual flight lines in areas of overlap.

Relative error within individual flight lines was considerably lower, approximately 10 cm in the horizontal direction and 5 cm in the vertical direction. Relative errors of this size can be expected when using a properly calibrated lidar sensor under normal operating conditions (Airborne 1, 2001). After using the time stamp of each posting to separate the data into individual flight lines, the relative vertical error within each flight line was estimated by calculating the standard deviation of the heights of points collected over a flat surface. The roofs of three structures and the surfaces of two cooling ponds, each spanning at least 75 m^2 and assumed to be topographically invariant, were used in this analysis. The horizontal accuracy of the lidar points was determined using a similar methodology applied to five different vertical walls of multi-story buildings, each of which is at least 5 m wide. Coordinates of lidar posts reflected from the walls were extracted and fit to a flat vertical plane, and the relative horizontal accuracy was estimated as the standard deviation of the points about this plane. In order to preserve the highest possible relative accuracy, subsequent processing considered all flight lines individually.

As the study area is relatively devoid of major topographic features and semi-arid rangeland canopies are fairly open, the first return lidar data were used to characterize both the terrain and the vegetative canopy. Using a single lidar dataset instead of attempting to merge the first and last pulse datasets (each of which contain over ten million points within the study area) kept the overall processing requirements to a minimum and avoided the introduction of problems associated with detecting multiple returns at a similar height (Hodgson *et al.*, 2003).

Individual lidar postings were grouped into 5 m $\times$ 5 m cells, where the lowest point in each cell was assumed to represent the ground surface. Using these ground points, an initial ground surface was interpolated across the study area using a thin plate spline (Meinguet, 1979). Heights above this surface were calculated for the remaining data points, and resulting values of zero or less were reclassified as

ground points. The ground surface was re-interpolated iteratively until all lidar points were classified as either ground or non-ground (99 percent of the points converged after two iterations). Of all lidar points, 21 percent were classified as ground, and the remaining 79 percent were classified as non-ground. All resulting ground points were used to interpolate a final ground surface, and above-ground height values were calculated for all remaining (non-ground) points. The final height points were combined from all individual flight lines into a single dataset, and were subsequently grouped into 4.6 m cells (the same spatial resolution as the hyperspectral data). Surface roughness values were derived for each cell by computing the standard deviation of all height values, which generated a surface roughness raster matching the spatial resolution of the hyperspectral data (Figure 3).

In addition to surface roughness, lidar data were used to calculate several variables describing the physical properties of all vegetation in the study area. The mean height of each pixel was calculated as the average of all heights within the pixel, including the ground postings, while the mean vegetation height was calculated as the average height of only the non-ground postings. The highest single posting in each pixel was recorded (tallest vegetation) as was the fraction of lidar postings within each pixel that were classified as bare ground.

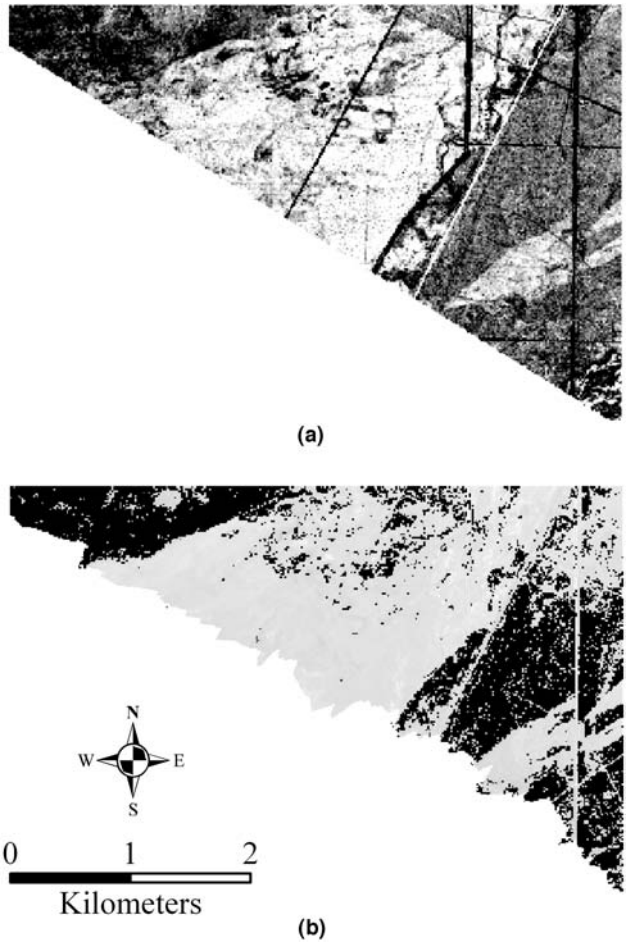

Figure 3. Map of surface roughness within the study area (a), and a final MTMF sagebrush presence distribution map (b). Surface roughness ranges from 0 cm (white regions) to over 300 cm (dark regions).

Co-registration

Co-registration is the process of assigning exactly the same spatial context to two raster geographic datasets such that when overlain, pixels correspond to exactly the same geographic location. Previous work demonstrated that small differences in co-registration can lead to large errors, and that for quantitative analysis in semi-arid environments, co-registration must be accurate to between 0.5 and 1 pixel (Townshend, 1992).

Lidar intensity data were interpolated into a raster with the same spatial resolution as the hyperspectral reflectance data (4.6 m). The hyperspectral band 43 (1,062 *n*m) corresponded most closely with the wavelength of the lidar laser (1,067 *n*m). This band was overlaid with both the interpolated lidar intensity data and five differentially corrected GPS ground control points. The lidar data were internally consistent and representative of GPS features (e.g., roads discernable in the image were straight and correctly positioned as compared to the GPS centerlines), and demonstrated less than one pixel error with respect to the GPS ground control points. The hyperspectral data expressed an image shift from the lidar dataset that averaged approximately 20 m between GPS ground control points and image projected locations. A first-order polynomial transform was applied to the hyperspectral dataset using sixteen image-derived ground control points discernable in both datasets (lidar intensity data used as the base image). Following the transform, agreement between the lidar and hyperspectral images was within one pixel at the GPS ground control points.

Away from the GPS ground control points, the hyperspectral image displayed geometric anomalies with a local shift between one and three pixels (compared to the lidar image). The lower geometric accuracy of the hyperspectral data is likely due to turbulent atmospheric conditions during data collection. Further, lidar geocorrection routines utilize differentially corrected GPS transects and real-time communication with mobile base stations for enhanced accuracy. Nine image control points were used to calculate a mean residual geometric error of 11.5 m (standard deviation of 3 m) between lidar and hyperspectral equivalent locations. Because these anomalies were small scale and randomly distributed across the study area, both datasets were re-sampled to three hyperspectral pixels (13.8 m). The hyperspectral data were re-sampled using a pixel aggregate function, while lidar point height data were interpolated to an equivalent raster resolution. The resulting datasets did not have quantifiable geometric errors with respect to each other or to any of the GPS ground control data.

Due to the possibility of uncertainties being generated and propagated by repeatedly projecting and re-sampling data as necessary for co-registration, both the hyperspectral and lidar data were processed in raw (reflectance and point, respectively) format. The analysis products were then geometrically transformed with the intent of preserving the integrity of the raw data while minimizing processing requirements.

Classification and Fusion

Hyperspectral MTMF classification products were evaluated using scatterplots of MF versus infeasibility values. Pixels with MF values greater than zero and infeasibility values less than seven were considered to contain sagebrush. Classification accuracies for sagebrush presence and absence were calculated using error matrices and methods according to Congalton and Green (1999). Because field validation

surveys (15 m diameter plots) were collected at approximately the same spatial scale as the resolution of the resampled and co-registered imagery (13.8 m pixels), validation plots were buffered by 0.5 pixels (6.9 m) to ensure that representative accuracies were determined. MF scores were extracted for the true positive field validation plots (field observed sagebrush coinciding with image classified sagebrush) as estimates of sub-pixel sagebrush components. Where multiple image pixels occurred within a single plot, the average value of all corresponding MF scores was used. The extracted MF values were correlated to field estimated sagebrush cover for each validation plot to evaluate the ability of MTMF to predict sub-pixel abundances of sagebrush. Positive MF values were rescaled to range between zero and one (representing 0 to 100 percent cover).

Lidar roughness values were extracted for all field validation plots and were considered on a pixel-by-pixel basis. Separate distributions were drawn for sagebrush presence and sagebrush absence validation plots, and were compared using significance testing. The distribution of roughness values for sagebrush presence and absence plots overlap, yet are found to be statistically distinguishable according to an F-test for sample variances (Table 1).

Roughness values representing sagebrush were determined iteratively. The lowest roughness values were associated with smooth surfaces (e.g., paved roads or short grass), while the highest roughness values represented tall anthropogenic features (e.g., power poles). An upper roughness threshold of 75 cm (approximately 18 standard deviations above the mean: Table 1) was determined appropriate for removing most tall anthropogenic features. When plotted against MF values, roughness values generate two distinct groupings, one with slightly lower mean and lower MF values (correlated to sagebrush absence) and the other with a slightly higher mean and higher MF values (correlated to sagebrush presence: Figure 4). The lower roughness threshold was incrementally adjusted and accuracies calculated at each increment, and an optimal lower roughness value of 6.5 cm was determined.

Because both MF scores and lidar roughness were correlated to sagebrush presence (Table 1; Figure 4), sagebrush was re-classified using a minimum MF value of zero, a maximum infeasibility value of seven, and roughness values between 6.5 cm and 75 cm. Image re-classification using fused hyperspectral and lidar variables resulted in a more accurate and detailed sagebrush distribution map for the study area. The sagebrush classification was then augmented by the vegetative structure information derived using the lidar data. Population statistics for each of the previously derived lidar variables (roughness, mean height, mean vegetation height, tallest vegetation, and percent bare ground) were tabulated for all pixels classified as either sagebrush present or sagebrush absent (Table 2; Figure 5). As in image classification, areas with very high surface roughness (greater than 75 cm) were eliminated from this analysis.

TABLE 1. THE STATISTICAL PROPERTIES OF THE LIDAR-DERIVED SURFACE ROUGHNESS VALUES CORRESPONDING TO THE SAGEBRUSH PRESENT AND SAGEBRUSH ABSENT FIELD VALIDATION SITES

Lidar Roughness	Sagebrush Absent	Sagebrush Present
Mean (cm)	7.0	9.1
Median (cm)	5.8	8.1
Integer Mode (cm)	5	8
Standard Deviation (cm)	4.5	3.6
Number of Samples	88	64

F Statistic = 1.57.
F_{Crit} (α = 95%) = 1.48.

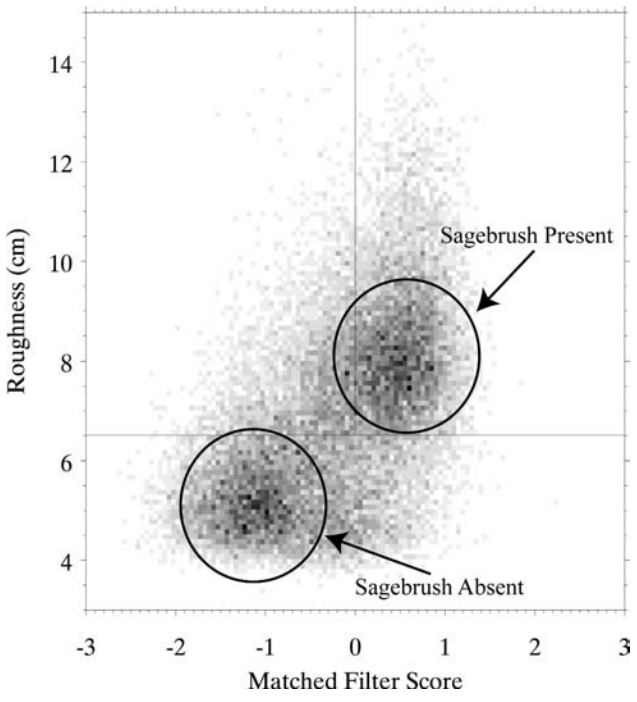

Figure 4. Scatterplot of Matched Filter (MF) score versus surface roughness for the entire study area, illustrating the relationship between areas of sagebrush presence (higher MF and higher roughness) and sagebrush absence (lower MF and lower roughness). Lines represent the roughness threshold of 6.5 cm and the MF threshold of zero that were used to generate sagebrush classifications.

TABLE 2. THE MEAN, MEDIAN, VARIANCE, AND F-VALUE (95% CONFIDENCE LEVEL) OF LIDAR HEIGHT PRODUCTS FOR SAGEBRUSH PRESENT AND SAGEBRUSH ABSENT IMAGE CLASSIFIED AREAS OF THE ENTIRE STUDY AREA. THESE VALUES WERE CALCULATED USING 15,273 PIXELS FROM AREAS CLASSIFIED AS SAGEBRUSH PRESENT AND 28,601 PIXELS FROM AREAS CLASSIFIED AS SAGEBRUSH ABSENT. VALUES ARE PRESENTED AS SAGEBRUSH "NOT PRESENT/PRESENT"

Lidar Variable	Mean	Median	Standard Deviation	F-Statistic (F_{crit} = 1.02)
Mean Pixel Height (cm)	7.8/9.6	7.1/9.2	3.5/2.8	1.52
Mean Vegetation Height (cm)	10.0/11.9	9.0/11.3	4.2/3.2	1.72
Percent Bare Soil	7.0/19.4	6.0/18.2	4.5/4.3	2.50
Roughness (cm)	6.8/9.2	5.9/8.5	3.8/3.5	1.17
Tallest Vegetation (cm)	34.1/53.2	30.0/51.0	15.2/16.4	1.17

Results

All classification accuracies are summarized in Table 3. Sagebrush was successfully mapped from hyperspectral imagery using MTMF, resulting in producer's, user's, and overall accuracies of 91 percent, 64 percent, and 75 percent, respectively, with no infeasibility constraint. The inclusion of a maximum infeasibility threshold of seven resulted in producer's, user's, and overall accuracies of 72 percent, 75 percent, and 74 percent, respectively. Subsequent attempts at lowering the infeasibility threshold generated unacceptable

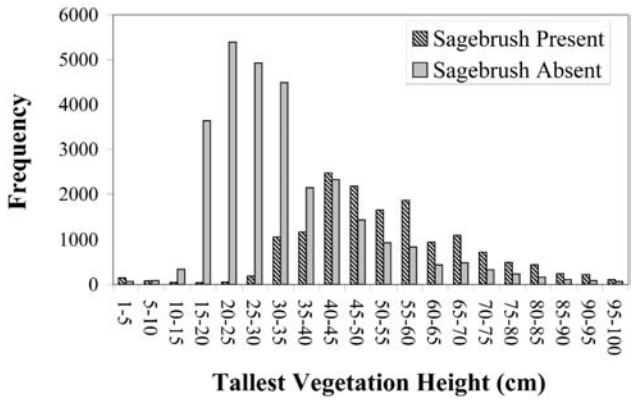

Figure 5. Histograms of lidar-derived tallest vegetation heights for both the sagebrush present (higher mean) and sagebrush absent (lower mean) classified areas.

TABLE 3. ACCURACY ASSESSMENT RESULTS FOR MULTIPLE SAGEBRUSH CLASSIFICATION STRATEGIES ARE INVESTIGATED. ACCURACIES WERE CALCULATED USING 88 VALIDATION POLYGONS WITH SAGEBRUSH PRESENCE AND 152 POLYGONS WITH SAGEBRUSH ABSENCE

Classification Strategy	Producer's Accuracy	User's Accuracy	Overall Accuracy
MF greater than 0	91%	64%	75%
MF greater than 0 AND Infeasibility less than 7	72%	75%	74%
MF greater than 0 AND Roughness between 6.5 and 75 cm	88%	86%	87%
MF greater than 0 AND Infeasibility less than 7 AND Roughness between 6.5 and 75 cm	84%	92%	89%

producer's accuracies. Fusion of lidar roughness values as a third variable in sagebrush classification increased producer's, user's, and overall accuracies to 84 percent (12 percent increase), 92 percent (17 percent increase), and 89 percent (15 percent increase), respectively. However, no significant correlation ($R^2 = 0.02$) was found between field estimates of sagebrush canopy cover and MF scores.

With relatively good constraint on sagebrush distribution, vegetative structure information derived from lidar data was cross-tabulated against the final classification of sagebrush presence and absence. All tested populations were found to be significantly distinguishable according to an F-test for sample variances (95 percent confidence: Table 2). Overall, pixels with sagebrush present had significantly greater mean heights, mean vegetation heights, roughness, and percent bare ground than did pixels with no sagebrush present. The maximum vegetation height for areas of sagebrush presence was also significantly greater in general than in areas of sagebrush absence. Further, histograms of tallest vegetation height clearly indicate that areas with sagebrush have higher population means and medians than do areas with no sagebrush (Figure 5).

Discussion

Hyperspectral classification utilizing the full electromagnetic range of the data resulted in highly confused classifications of the distribution of sagebrush in a semi-arid rangeland. This corresponds with results published by Okin *et al.* (2001) in that semi-arid partial-cover canopies with bright soil backgrounds are difficult to discriminate. The separation of sagebrush from background vegetation was optimized based on empirical evaluations of the spectral relationships shown in Figure 2, which substantially increased the quality of the sagebrush classification. We hypothesize that spectrally subsetting the data removed a significant portion of soil-dominated reflectance ranges confusing the classifications. Further, the use of multiple target endmembers increased classification accuracy in this study. It is notable that in large areas where holistic variability becomes difficult to model, the inclusion of all necessary endmembers may be difficult. These results indicate that future studies in areas of low variability or highly constrained parameters can produce useful results (e.g., indicator plots or key ecologic zones), though this methodology would be difficult over large and diverse areas.

MTMF classification did not discern representative sub-pixel abundances of sagebrush. While it is likely that some degree of error in sagebrush percent cover estimation confused the relationship between the MF score and the actual sub-pixel component, we infer that there are additional complications. For example, MTMF calculates the image mean as spectral background, which is expected to be distinct from the target (Mundt *et al.*, in review). In this study, the target is a large component of the dataset, and it is possible that some of the target spectral properties were interpreted as background, confusing the classification. This limitation is demonstrated in Figure 4, where the cluster of sagebrush present pixels crosses into negative MF space. While MTMF may not be the best approach to sub-pixel component estimation in this scenario, it still visually outperformed Linear Spectral Unmixing and Spectral Angle Mapping in iterative sagebrush presence and absence classifications.

This study used a simple sum of individual classification MF values to represent the composite sagebrush abundance (MF) distribution. Combination of linear unmixing endmembers in this context is not well documented, and we hypothesize that alternative approaches may also be effective (e.g., non-linear or weighted combinations). The summation of MF variables in this study is justified as we were able to generate a distribution map in which we have high confidence (Figure 3). Further, other empirical combinations of MF and infeasibility classifications (e.g., using geometric instead of arithmetic means) did not produce significantly different products.

While no correlation was found between MF scores and field estimates of sagebrush canopy cover, pixels with sagebrush presence generally had higher MF scores than pixels with sagebrush absence (Figure 4). Further, in most areas of sagebrush presence, field validation estimated 30 percent cover, which is near the mean (as well as the median and mode) of the rescaled MF distribution (Table 2). In light of this, we hypothesize that the MF scores may potentially be useful, though they are likely confused by high degrees of mixing in an indeterminate spectral endmember (Okin *et al.*, 2001).

Lidar data significantly improved hyperspectral classifications and added useful stand structure information in a semi-arid rangeland. Though typical rangeland vegetation heights are of the same order as the absolute vertical accuracy of small-footprint lidar scanners (Table 2), it is possible to measure such heights if attention is paid to preserving the high relative accuracy throughout the processing. In this study, high vertical accuracy was maintained by processing the lidar data in individual flight lines and as raw, irregu-

larly spaced posting data. This allowed for the measurement of heights as small as 5 cm and for the production of height and roughness maps which effectively describe a large number of geographic features in addition to sagebrush presence, including roads, railroads, power lines, drainage canals, and fences. It is further notable that most sagebrush in the study area was significantly taller than the vertical precision of the lidar data (Figure 5).

Co-registration of high spatial resolution data is processing and time intensive, yet provides a significantly more useful product than either of the parent models for discrimination and description of sagebrush. Fusion of classification products increased overall accuracies by 14 percent over standard hyperspectral processing. We assert that the fundamental differences between the two sensors adds significantly to the potential discrimination of semi-arid shrubs as long as the vegetative ecosystems are relatively static or the datasets are acquired in rapid succession. The increase in accuracy resulting from data fusion may justify the added processing time and cost in areas where precision monitoring is necessary or key environmental indicators are present.

One of the major limitations of the co-registration process is the requirement to re-sample both datasets to a lower spatial resolution to accommodate geometric error. Although the hyperspectral imagery in this study was re-sampled to one-third of its acquired spatial resolution, it still maintained higher spatial resolution than high altitude Airborne Visible Infrared Spectrometer (AVIRIS) instrument, which has had documented success in the context of vegetation discrimination (Merenyi *et al.*, 2000). The lidar data were spatially re-sampled to a small fraction of their original resolution, yet were significantly useful in this study. Furthermore, because lidar data are geometrically precise, invaluable ecological descriptions for high-resolution inventory and monitoring can be derived from the lidar data at a finer spatial scale than classifications from hyperspectral data.

General field observations of sagebrush in the study area found a mean height of approximately 50 cm and a maximum height of approximately 1 m, while vegetation in areas of sagebrush absence was most commonly shorter than 50 cm (with rabbitbrush often reaching heights of approximately 30 cm). These field relationships are supported by the derived lidar products (Figure 5). Analyses using lidar-derived variables such as this have implications for high-resolution habitat monitoring, biomass/fuel load calculations, and numerous other objectives that may not be efficiently interpreted by field survey crews.

Conclusions and Future Work

This study introduces the utility of hyperspectral data for mapping sagebrush communities, demonstrates the applicability of lidar data to describe semi-arid shrubs, and successfully fuses the two datasets to generate improved classification accuracies and stand structure descriptions. These products are desirable in the context of long term monitoring, and may potentially be useful in habitat analysis or biomass calculations. Figure 4 clearly indicates the correlation between surface roughness and sagebrush presence; and fires scars (including sagebrush islands) are plainly visible in the final sagebrush distribution map (Figure 3). Spectral subsetting and multiple spectral endmember classification helped to minimize the spectral confusion introduced by partial cover canopies and soil variability. Sub-pixel components could not be accurately modeled in this study, likely due to complex mixing in semi-arid partial canopies. While accuracies from MTMF classification of sagebrush distribution were intermediate (75 percent overall accuracy), the inclusion of lidar-derived roughness notably increased the ability to map the distribution of sagebrush (89 percent overall accuracy). Following the generation of accurate and co-registered maps of sagebrush distribution and vegetative structure, quantitative statistics were derived that described sagebrush stands in terms of vegetation heights and percent bare ground. The methods and results of this study lay the framework for utilizing co-registered hyperspectral and lidar data to describe semi-arid shrubs in greater detail than would be feasible for most ground based surveys. Although it is currently not reasonable to use these methods to inventory or monitor large scale vegetative ecosystems (e.g., semi-arid shrubs in the Southwest U.S.), we assert that these methods are useful where small-scale precision has implications for large-scale inventory or land management objectives.

Acknowledgments

This work was supported by the Idaho National Environmental and Engineering Laboratory (INEEL) through the Idaho State University – INEEL Partnership for Integrated Environmental Analysis Education Outreach Program. Additional support was provided by a grant from the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center and the NOAA Environmental Technology Laboratory (ETL). ISU would also like to acknowledge the Idaho Delegation for their assistance in obtaining this grant. We thank Ron Rope, Shane Cherry, and Tanya Johnson for field and office assistance. Finally, we thank our two anonymous reviewers for their helpful comments.

References

Airborne 1, 2001. *Briefing Note BN#01; LiDAR Accuracy*, Airborne 1 Corporation, El Segundo, California, 24 p.

Berk, A., G.P. Anderson, L.S. Bernstein, P.K. Acharya, H. Dothe, M.W. Matthew, S.M. Adler-Golden, J.H. Chetwynd, S.C. Richtsmeier, B. Pukall, C.L. Allred, L.S. Jeong, and M.L. Hoke, 1999. MODTRAN4 radiative transfer modeling for atmospheric correction, *Proceedings of the AVIRIS Airborne Geosciences Workshop*, Pasadena, California (JPL, Publication 99–17, Pasadena, California).

Blackburn, G.A., 2002. Remote sensing of forest pigments using airborne imaging spectrometer and LIDAR imagery, *Remote Sensing of Environment*, 82:311–321.

Boardman, J.W., 1998. Leveraging the high dimensionality of AVIRIS data for improved sub-pixel target unmixing and rejection of false positives: Mixture Tuned Matched Filtering, *Proceedings of the 7th Annual JPL Airborne Geoscience Workshop* (JPL, Publication 97–1, Pasadena, California), p. 55.

Chen, D., and D. Stow, 2002. The effect of training strategies on supervised classification at different spatial resolutions, *Photogrammetric Engineering & Remote Sensing*, 68(11):1155–1161.

Colket, E.C., 2003. *Long-Term Vegetation Dynamics and Post-Fire Establishment Patterns of Sagebrush Steppe*, M.S. thesis, University of Idaho, Moscow, Idaho, 144 p.

Congalton, R.G., and K. Green, 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Lewis Publishers, Boca Raton, Florida, 137 p.

Drake, J.B., R.O. Dubayah, D.B. Clark, R.G. Knox, J.B. Blair, M.A. Hofton, R.L. Chazdon, J.F. Weishampel, and S.D. Prince, 2002. Estimation of tropical forest structural characteristics using large-footprint LiDAR, *Remote Sensing of Environment*, 79:305–319.

Elmore, A.J., J.F. Mustard, S.J. Manning, and D.B. Lobell, 2000. Quantifying vegetation change in semi-arid environments: Precision and accuracy of spectral mixture analysis and the Normalized Difference Vegetation Index, *Remote Sensing of Environment*, 73:87–102.

ESER, 2004. INEEL Environmental Surveillance, Education and Research program; INEEL vegetation, an overview, URL:

http://www.stoller-eser.com/Flora/vegetation.htm, S.M. Stoller Corporation, Idaho Falls, Idaho (last date accessed: 12 October 2005).

Flood, M. (editor), 2004. *American Society for Photogrammetry and Remote Sensing Guidelines – Vertical Accuracy Reporting for Lidar Data*, ASPRS, Bethesda, Maryland, 15 p.

Harding, D.J., M.A. Lefsky, G.G. Parker, and J.B. Blair, 2001. Laser altimeter canopy height profiles: Methods and validation for closed-canopy, broadleaf forests, *Remote Sensing of Environment*, 76:283–297.

Hodgson, M.E., J.R. Jensen, L. Schmidt, S. Schill, and B. Davis, 2003. An evaluation of LIDAR- and IFSAR-derived digital elevation models in leaf-on conditions with USGS Level 1 and Level 2 DEMs, *Remote Sensing of Environment*, 84:295–308.

Hofton, M.A., L.E. Rocchio, J.B. Blair, and R. Dubayah, 2002. Validation of vegetation canopy LiDAR sub-canopy topography measurements for a dense tropical forest, *Journal of Geodynamics*, 34:491–502.

Hudak, A.T., M.A. Lefsky, W.B. Cohen, and M. Berterretche, 2002. Integration of LiDAR and Landsat ETM+ data for estimating and mapping forest canopy height, *Remote Sensing of Environment*, 82:397–416.

Huete, A.R., and R.D. Jackson, 1988. Soil and atmosphere influences on the spectra of partial canopies, *Remote Sensing of Environment*, 29:89–105.

Huete, A.R., R.D. Jackson, and D.F. Post, 1985. Spectral response of a plant canopy with different soil backgrounds, *Remote Sensing of Environment*, 17:37–53.

Latypov, D., 2002. Estimating relative LiDAR accuracy information from overlapping flight lines, *ISPRS Journal of Photogrammetry and Remote Sensing*, 56:236–245.

Lee, D.S., and J. Shan, 2003. Combining LiDAR elevation data and IKONOS multispectral imagery for coastal classification mapping, *Marine Geodesy*, 26:117–127.

Lewis, M., V. Jooste, and A. DeGasparis, 2000. Hyperspectral discrimination of arid vegetation, *Proceedings of the 28th International Symposium on Remote Sensing of Environment*, 27–31 March, Cape Town, South Africa, pp. 148–151.

Meinguet, J., 1979. Multivariate interpolation at arbitrary points made simple, *Journal of Applied Mathematics and Physics*, 30:292–304.

Merenyi, E., W.H. Farrand, L.E. Stevens, T.S. Melis, and K. Chhibber, 2000. Mapping Colorado River ecosystem resources in Glen Canyon: Analysis of hyperspectral low-altitude AVIRIS imagery, *ERIM Proceedings of the 14th* International *Conference on Applied Geologic Remote Sensing*, 06–08 November, Las Vegas, Nevada, pp. 44–51.

Mundt, J.T., D.R. Streutker, and N.F. Glenn, In Review. Partial unmixing of hyperspectral imagery: Theory and methods, *Remote Sensing of Environment*.

Musick, H.B., 1984. Assessment of Landsat multispectral scanner spectral indexes for monitoring arid rangeland, *IEEE Transactions on Geoscience and Remote Sensing*, GE-22(6):512–519.

NPCC, 2004. *Upper Snake Subbasin Assessment*, Northwest Power and Conservation Council, Portland, Oregon, 75 p.

Okin, G.S., D.A. Roberts, B. Murray, and W.J. Okin, 2001. Practical limits on hyperspectral vegetation discrimination in arid and semi-arid environments, *Remote Sensing of Environment*, 77:212–225.

Okin, W.J., G.S. Okin, D.A. Roberts, and B. Murray, 1999. Multiple endmember spectral mixture analysis: Endmember choice in an arid shrubland, *Proceedings of the 8th Annual AVIRIS Airborne Geoscience Workshop* (JPL, Publication 99–17), Pasadena, California, pp. 323–332.

Patil, G.P., and W.L. Myerst, 1999. Environmental and ecological health assessment of landscapes and watersheds with remote sensing data, *Environmental Health*, 5(4):221–224.

Popescu, S.C., and R.H. Wynne, 2004. Seeing the trees in the forest: Using LiDAR and multispectral data fusion with local filtering and variable window size for estimating tree height, *Photogrammetric Engineering & Remote Sensing*, 70(5):589–604.

Popescu, S.C., R.H. Wynne, and R.F. Nelson, 2002. Estimating plot-level tree heights with LiDAR: Local filtering with a canopy-height based variable window size, *Computers and Electronics in Agriculture*, 37:71–95.

Rango, A., M.J. Chopping, J.C. Ritchie, K. Havstad, W. Kustas, and T. Schmugge, 2000. Morphological characteristics of shrub coppice dunes in desert grasslands of southern New Mexico derived from scanning LiDAR, *Remote Sensing of Environment*, 76:26–44.

Ray, T.W., and B.C. Murray, 1998. Nonlinear spectral mixing in desert vegetation, *Remote Sensing of Environment*, 55:59–64.

Riaño, D., F. Valladares, S. Condés, and E. Chuvieco, 2004. Estimation of Leaf Area Index and covered ground from airborne laser scanner (LiDAR) in two contrasting forests, *Agricultural and Forest Meteorology*, 124:269–275.

Ritchie, J.C., K.S. Humes, and M.A. Weltz, 1995. Laser altimeter measurements at Walnut Gulch watershed, Arizona, *Journal of Soil and Water Conservation*, 50(5):440–442.

Roberts, D.A., M. Gardner, R. Church, S.L. Ustin, G.J. Scheer, and R.O. Green, 1998. Mapping chaparral in the Santa Monica mountains using multiple endmember spectral mixture models, *Remote Sensing of Environment*, 65:267–279.

RSI, 2004. *ENVI User's Guide (Version 4.1)*, Research Systems Incorporated, Boulder, Colorado, 1,084 p.

Townshend, J.R.G., 1992. The impact of misregistration on change detection, *IEEE Transactions on Geoscience and Remote Sensing*, 30(5):1054–1060.

Tueller, P.T., 1987. Remote sensing science applications in arid environments, *Remote Sensing of Environment*, 23:143–154.

Warren, P.L., and C.F. Hutchison, 1984. Indicators of rangeland change and their potential for remote sensing, *Journal of Arid Environments*, 7:107–126.

Weltz, M.A., J.C. Ritchie, and H.D. Fox, 1994. Comparison of laser and field measurements of vegetation height and canopy cover, *Water Resources Research*, 30(5):1311–1319.

(Received 20 May 2004; accepted 05 October 2004; revised 10 December 2004)