Contract No. and Disclaimer:

Remote Sens. **2011**, *3*, 1-x manuscripts; doi:10.3390/rs30x000x

OPEN ACCESS

Remote Sensing

ISSN 2072-4292

www.mdpi.com/journal/remotesensing

Article

Vegetation Cover Analysis of Hazardous Waste Sites in Utah and Arizona using Hyperspectral Remote Sensing

Jungho Im [1,*], John R. Jensen [2], Ryan R. Jensen [3], John Gladden [4], Jody Waugh [5], and Mike Serrato [4]

[1] Department of Environmental Resources Engineering, State University of New York, College of Environmental Science and Forestry, Syracuse, NY 13210, USA
[2] Department of Geography, University of South Carolina, Columbia, SC 29208, USA
[3] Department of Geography, Brigham Young University, Provo, UT 84605, USA
[4] Savannah River National Laboratory, Department of Energy, Aiken, SC 29808, USA
[5] Environmental Sciences Laboratory, Department of Energy, Grand Junction, CO 81503, USA

* Author to whom correspondence should be addressed; E-Mail: imj@esf.edu; Tel.:+1-315-470-4709; Fax: +1-315-470-6958

Received: / Accepted: / Published:

Abstract: Remote sensing technology can provide a cost-effective tool for monitoring hazardous waste sites. This study investigated the usability of HyMap airborne hyperspectral remote sensing data (126 bands at 2.3 × 2.3 m spatial resolution) to characterize the vegetation at U.S. Department of Energy uranium processing sites near Monticello, Utah and Monument Valley, Arizona. Grass and shrub species were mixed on an engineered disposal cell cover at the Monticello site while shrub species were dominant in the phytoremediation plantings at the Monument Valley site. The specific objectives of this study were to: 1) estimate leaf-area-index (LAI) of the vegetation using three different methods (i.e., vegetation indices, red-edge positioning (REP), and machine learning regression trees), and 2) map the vegetation cover using machine learning decision trees based on either the scaled reflectance data or mixture tuned matched filtering (MTMF)-derived metrics and vegetation indices. Regression trees resulted in the best calibration performance of LAI estimation ($R^2 > 0.80$. The use of REPs failed to accurately predict LAI ($R^2 < 0.2$). The use of the MTMF-derived metrics (matched filter scores and infeasibility) and a range of vegetation indices in decision trees improved the vegetation mapping when compared to the decision tree classification using just the scaled reflectance. Results suggest that hyperspectral imagery are useful for characterizing biophysical characteristics (LAI) and vegetation cover on capped hazardous waste sites. However, it is believed that the

vegetation mapping would benefit from the use of higher spatial resolution hyperspectral data due to the small size of many of the vegetation patches (< 1m) found on the sites.

Keywords: hazardous waste sites; hyperspectral remote sensing; HyMap; vegetation mapping; LAI estimation; decision trees

---

1. Introduction

Humans produce large amounts of hazardous waste. In 2007, the United States generated 47 million tons of hazardous waste with Louisiana and Texas responsible for more than 50% of the total [1]. Hazardous waste from current and historic activities is often isolated in landfills or disposal cells using a capping system consisting of barriers that limit deep percolation of precipitation and mobilization of the hazardous constituents [2, 3]. In addition to the thousands of capping systems in existence, in 2004 the EPA estimated that almost 300,000 additional waste sites were expected to require remediation [4]. Many of these sites, as well as the older sites contain residual contamination or are undergoing some form of *in situ* remediation. In all cases, the management of water infiltration into the waste area is a key issue to prevent the migration of the hazardous constituents into the environment.

Historically, the vegetation component of a hazardous waste capping system has been viewed as a means to stabilize the surface soils and prevent erosion. However, for some capping systems in arid and semi-arid climates, the vegetative cover has taken an increasingly functional role through the construction of evapotranspiration or water balance cover systems [5, 6]. In these systems, the vegetative cover and soil system are constructed to maintain a hydrologic balance with the vegetation withdrawing water from the underlying soils on an annual basis, thereby minimizing deep infiltration. Proper functioning of these types of systems depends on the development and maintenance of a robust plant community that can maintain water withdrawal capacity over the life of the capping system.

Some *in situ* remediation strategies are also being implemented whereby the migration of subsurface contaminants is dependent on the water withdrawal capability of vegetation. Considered to be a type of Phytoremediation [7], these strategies may be applicable where subsurface contaminants are potentially mobile, and management of vegetation can result in reduced infiltration and subsequent hydraulic control of the migration of contaminants in soils and shallow groundwater [8].

In all such cases, the maintenance of a high evapotranspiration capacity through well-adapted and healthy plant communities is key to the proper and long term stabilization of the wastes. Monitoring of these systems is commonly conducted by ground level observations by trained professionals and is becoming a significant cost element in the management of such systems. Consequently, there is a growing demand for an efficient and reliable approach to vegetation monitoring at waste remediation and stabilization sites. Remote sensing technology can provide a cost effective tool for this type of monitoring in harmony with information obtained from in situ investigation.

Multispectral (several bands) and hyperspectral (hundreds of narrow bands) remote sensing has been used for monitoring hazardous sites [3, 9-11] as well as typical environmental resources such as water, land, and vegetation [12-15]. In particular, remote sensing-derived vegetation products can provide valuable information regarding vegetation health and dynamics when monitoring hazardous waste sites [16]. Various classification approaches have been investigated for vegetation mapping, including: maximum likelihood classification [17], subpixel analysis [18], machine learning [19], and object-based methods [20]. The phenological cycle of vegetation has also been studied to better map vegetation dynamics [21].

Several approaches have been investigated for modeling vegetation biophysical parameters such as biomass and leaf-area-index (LAI) using remotely sensed spectral data. These approaches include: empirical methods such as statistical regression, spectral positioning, and artificial intelligence, and physical modeling [22]. Simple linear regression analysis has been widely adopted to correlate vegetation biophysical parameters measured *in situ* with various vegetation indices such as the Normalized Difference Vegetation Index (NDVI) [23]. More advanced regression techniques, including principal component regression and partial least squares regression, have also been examined [24, 25]. Some scientists have focused on identifying the spectral reflectance red-edge position (REP) because of its close association with chlorophyll content and its seasonal variations, which directly influence vegetation health [26]. Artificial intelligence methods such as neural networks and regression trees can incorporate field training samples to estimate the vegetation parameters [27]. These methods are relatively simple, but have some limitations, including the fact that the relationships are based on representative training samples and the methods are sensitive to atmospheric conditions, sensor viewing geometry, and the spatial resolution of the remote sensor data. Therefore, the methods generally need to be calibrated each time a new remote sensing dataset is acquired [28].

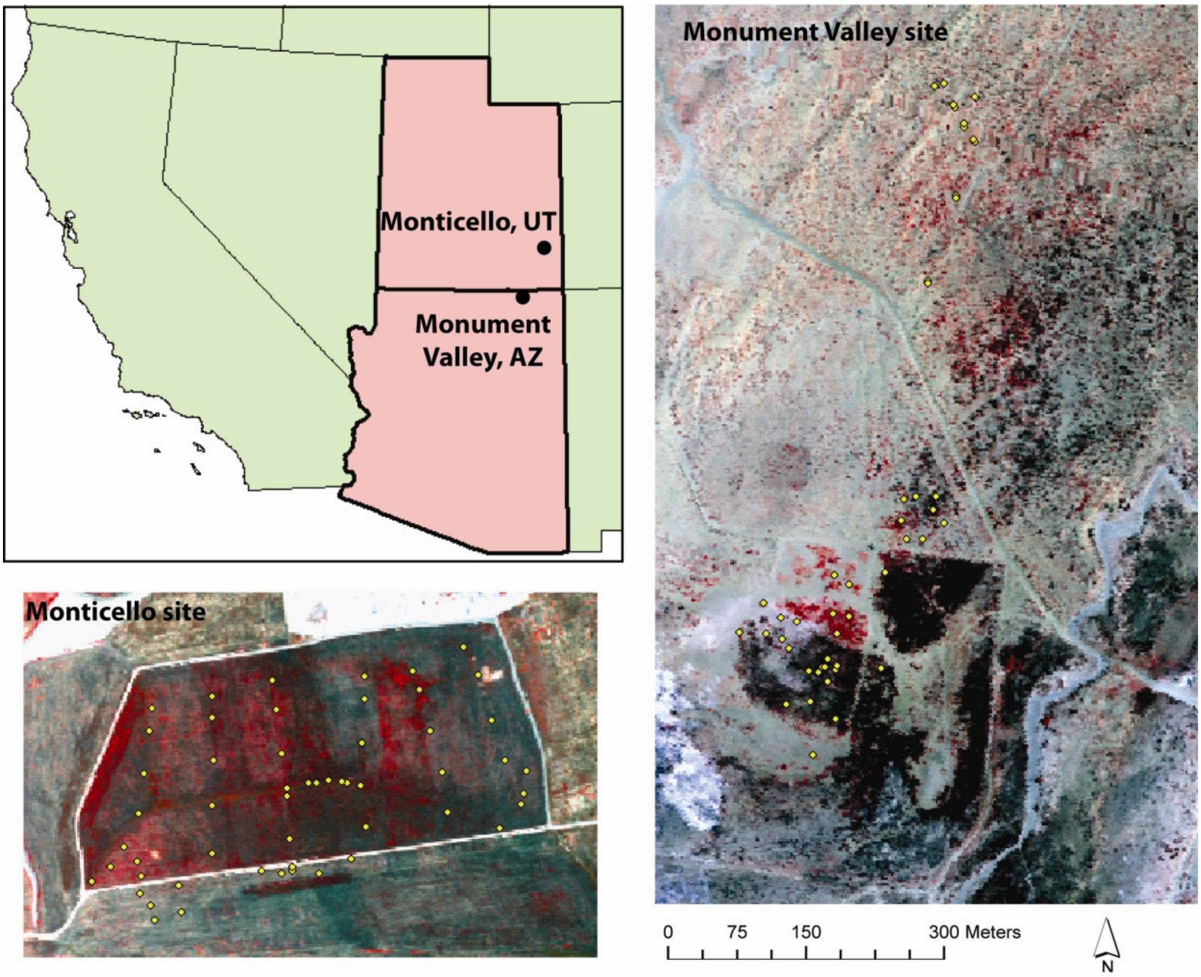
Physical models are theoretically based on leaf scattering and absorption mechanisms associated with biochemistry [29]. A representative type of model is the radiative transfer model, which simulates radiation transfer processes in vegetation by computing the interaction between plants and solar radiation. Vegetation biophysical parameters can be retrieved through inversion of the radiative transfer model (e.g., PROSPECT). Simulated reflectance databases have been frequently used with model inversion techniques [30]. Model inversion approaches may result in a multitude of different possible solutions increasing uncertainties [31].

Vegetation is typically characterized by slow rates of change and affected by other slow processes such as climate change or soil acidification. However, the vegetation cover on a hazardous waste site may rapidly change due to the unanticipated conditions (e.g., soil subsidence, biointrusion). This change must be quickly detected. The objective of this research was to demonstrate the usefulness of hyperspectral remote sensing to provide long term monitoring capability for Department of Energy (DOE) remediation and waste sites. This study investigated characteristics of vegetation cover on hazardous waste sites with regard to species type and LAI distribution using various technical approaches.

2. Study Area and Data

Two U.S. Department of Energy sites were investigated in this study (Figure 1): a) a uranium mill tailings disposal cell capping system near Monticello, UT, and b) a phytoremediation planting of desert shrubs near Monument Valley, AZ.

Figure 1. The HyMap imagery of two study sites (RGB = Hyperspectral bands 24, 17, 11). The yellow symbols are *in situ* sampling locations.



To limit percolation into underlying tailings, the Monticello capping system relies on the water-storage capacity of a 163-cm sandy clay loam soil and rock "sponge" layer overlying a 38-cm coarse sand capillary barrier, and native sagebrush steppe vegetation to seasonally remove stored precipitation [32]. The capillary barrier increases the water-storage capacity of the soil "sponge" [33]. The topsoil has favorable edaphic properties for a sustainable plant community. Percolation flux, measured within a 3-ha embedded lysimeter, was approximately 0.5 mm yr^{-1} from 2000 through 2009 [32]; the capping system has performed well in the short term. Detecting temporal changes and spatial patterns in plant species and LAI on a landscape scale will be important for performance monitoring in the long term.

At the arid Monument Valley site, two deep-rooted native shrubs, *Sarcobatus vermiculatus* (black greasewood) and *Atriplex canescens* (fourwing saltbush), are part of the remedy for nitrate contamination in soil where a uranium mill tailings pile once stood, and in an alluvial aquifer spreading away from the source area soil [8]. When protected from livestock grazing, populations of these phreatophytic shrubs transpire enough water from the source area soil to limit recharge and nitrate leaching [34], and from the alluvial aquifer to slow the spread of the nitrate plume [35]. Monitoring the long-term performance of phytoremediation at Monument Valley will include tracking responses of phreatophyte health and transpiration rates to changing land management practices over many hectares [36].

HyMap hyperspectral remote sensing data were collected by *HyVista, Inc*. at Monument Valley, AZ on 2 June 2008 and Monticello, UT on 3 June 2008. Ground reference data were collected at these sites at the same time as data acquisition. An additional ground level dataset was collected at the Monticello, UT, site the previous week. Field data included vegetation composition (percent canopy cover) and LAI (n = 54 on the Monticello site and n = 19 on the Monument Valley site; refer to Table 1). The dominant species included *Artemisia tridentata* (big sagebrush), *Ericameria nauseosa* (rubber rabbitbrush), and *Pascopyrum smithii* (western wheatgrass) on the Monticello site and *Sarcobatus vermiculatus* (black greasewood) and *Atriplex canescens* (fourwing saltbush) on the Monument Valley site.

The HyMap hyperspectral data consisted of 126 bands from 440 to 2500 nm at 2.3 × 2.3 m nominal spatial resolution. The HyMap radiance data were radiometrically corrected to scaled reflectance using the HYCORR algorithm with EFFORT spectral polishing [37]. The scaled reflectance data were then geometrically rectified to a Universal Transverse Mercator (UTM) projection using 15 to 20 GCPs collected from the 2006 National Agricultural Imagery Program (NAIP) Digital Orthophoto Quarter Quadrangle (DOQQ) data (1 x 1 m spatial resolution) over the two study sites resulting in root-mean-square-error (*RMSE*) < 1 pixel.

Table 1. Field data characteristics for vegetation mapping and leaf-area-index (LAI) estimation.

	Monticello site	**Monument Valley site**
Strata (classes)*	Big sagebrush (n = 8) Rubber rabbitbrush (n = 12) Western wheatgrass (n = 16) Litter (dead plant materials; mostly grass species) (n = 17)	Black greasewood (n = 12; percent cover available for 8 out of 12) Fourwing saltbush (n = 14; percent cover available for 9 out of 14) Soil (n = 17)
LAI range	0.09 – 5.1 (n = 54)**	0.95 – 6.26 (n = 19)***

* A dominant class at a sampled location was determined based on percent canopy cover.

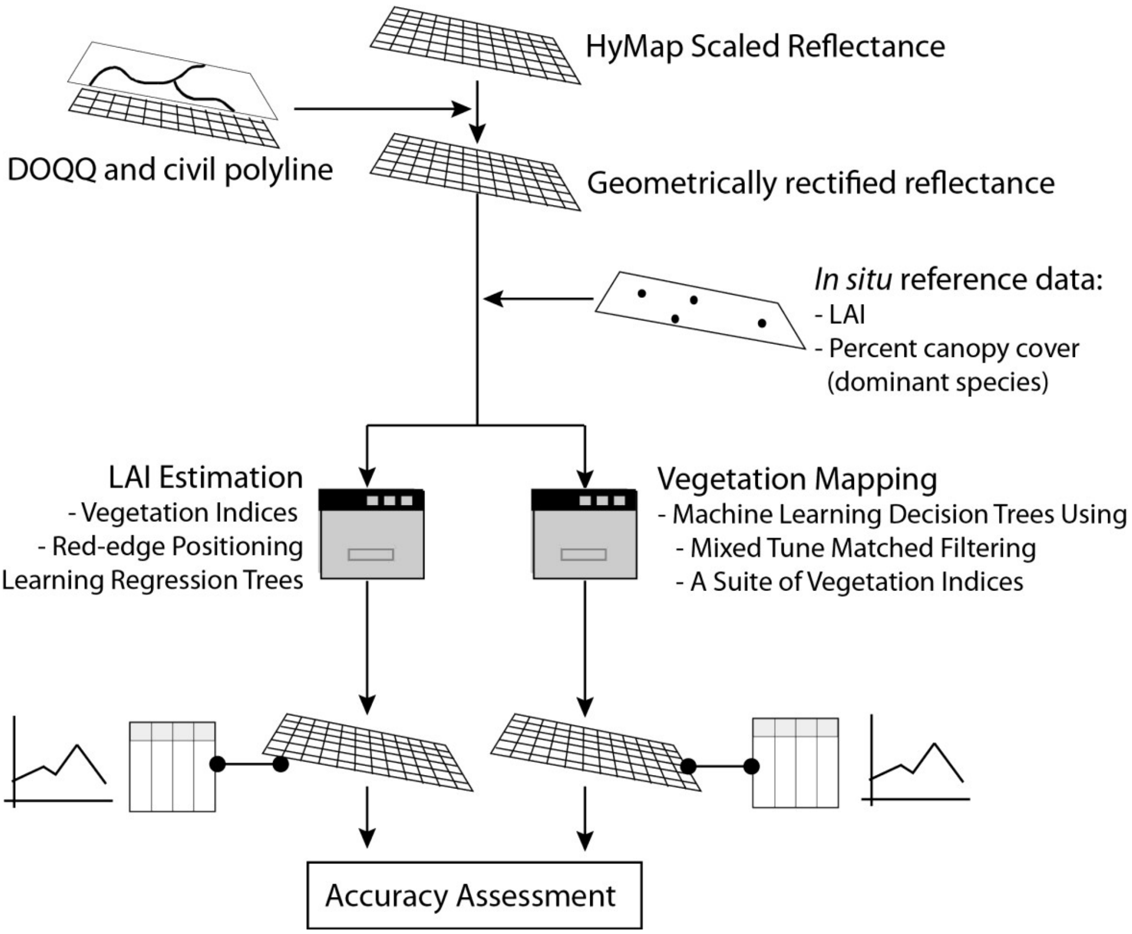
** One of the samples was not used for the classification since the dominant class (soil) was other than the four.

*** LAI was measured at only 19 sample locations of greasewood and saltbush plots.

3. Methods

Figure 2 is a flow diagram of the digital image processing methods used to a) predict the spatial distribution of LAI, and b) map the vegetation species present on the two hazardous waste sites.

Figure 2. Digital image processing flow diagram.


HyMap Scaled Reflectance
DOQQ and civil polyline
Geometrically rectified reflectance
In situ reference data:
- LAI
- Percent canopy cover
(dominant species)
LAI Estimation
- Vegetation Indices
- Red-edge Positioning
- Machine Learning Regression Trees
Vegetation Mapping
- Machine Learning Decision Trees Using
- Mixed Tune Matched Filtering
- A Suite of Vegetation Indices
Accuracy Assessment

3.1. Leaf-area-Index (LAI) Estimation

One of the major goals of this research was to map the spatial distribution of vegetation biomass on the waste sites using remote sensing-derived indices and algorithms in conjunction with *in situ* derived LAI used as a surrogate for vegetation biomass. LAI, the total area of one-sided green leaves in relationship to the ground below them, directly quantifies the vegetation canopy structure and is highly-related to diverse canopy processes including water interception, photosynthesis, evapotranspiration, and respiration.

Three approaches were investigated to estimate vegetation LAI, including: 1) vegetation index (VI)-based methods, 2) the red-edge positioning (REP) methods, and 3) the use of machine learning regression trees. Although there are numerous vegetation indices, we evaluated two basic vegetation indices:

$$VI1 = \frac{B_2}{B_1} \tag{1}$$

$$VI2 = \frac{B_2 - B_1}{B_2 + B_1} \tag{2}$$

These are extended versions of the simple ratio (SR) and normalized difference vegetation index (NDVI), respectively [14, 38]. Unlike SR and NDVI which only use one red band and one near-infrared band, we evaluated VI1 and VI2 with all 126 × 125 (= 15,750) possible band combinations in the entire spectral range from 400 to 2,500 nm as long as the value of B_1 was smaller than B_2. While the correlation between the VI2 and LAI results in a symmetric pattern when B_1 and B_2 are switched, the correlation between the VI1 and LAI is a bit asymmetric with the switch of B_1 and B_2. However, the asymmetric characteristic of the VI1-LAI relationship were not considered in the study because the asymmetry was slight and we were only interested in the band combinations that resulted in the highest correlation.

The REP is the spectral position between the red (~680 nm) and near-infrared (~800 nm) wavelengths where the maximum slope is found [29]. The REP is sensitive to the biophysical and biogeochemical properties of vegetation such as LAI and leaf nitrogen content. Phenological change and/or vegetation stress can affect the REP. There are several methods for computing REP, including: derivative function-based methods [39] and Gaussian model-based methods [40]. Some studies have reported that the REP is not a single location, but multiple wavelengths [39, 41]. In this study, three REP techniques were tested to estimate LAI, including: 1) a linear four-point interpolation (LI_REP) [42], 2) a three-point LaGrange interpolation (LG_REP) [43], and 3) a linear extrapolation (LE_REP) [26].

LI_REP assumes that the reflectance curve at the red-edge can be simplified to a straight line centered near the midpoint between the NIR reflectance and the minimum reflectance of the chlorophyll absorption. It was computed using:

$$LI\_REP = 700 + 40\left[\frac{\rho_{red\,edge} - \rho_{700nm}}{\rho_{740nm} - \rho_{700nm}}\right], \text{ where } \rho_{red\,edge} = \frac{\rho_{670nm} + \rho_{780nm}}{2} \tag{3}$$

LG_REP uses the point with the maximum first derivative reflectance FDR_{λ_i}, ($\lambda_i, FDR_{\lambda_i}$), and two points on both sides, ($\lambda_{i-1}, FDR_{\lambda_{i-1}}$) and ($\lambda_{i+1}, FDR_{\lambda_{i+1}}$). A second-order polynomial was fit using the three points and the wavelength where the second derivative equals zero was determined to be the REP:

$$LG\_REP = \frac{A(\lambda + \lambda) + B(\lambda + \lambda) + C(\lambda + \lambda}{2(A + B + C)} \quad (4)$$

where $A = \frac{FDR_{\lambda}}{(\lambda - \lambda)(\lambda - \lambda)}$, $B = \frac{FDR_{\lambda}}{(\lambda - \lambda)(\lambda - \lambda)}$, and $C = \frac{FDR_{\lambda +}}{(\lambda - \lambda)(\lambda - \lambda)}$.

LE_REP was based on a double-peak feature in the first derivative reflectance resulting from the discontinuity in REP and foliar-nitrogen relationship. Four points, two at the far-red peak (680 to 700 nm) and two at the NIR peak (725 to 760 nm), were used to create two straight lines. The wavelength corresponding to the intersection of the two lines was the REP:

$$LE\_REP = \frac{-(c_{far-red} - c_{NIR})}{(m_{far-red} - m_{NIR})} \quad (5)$$

where *c* and *m* were the intercept and slope of the lines, respectively.

Machine learning regression trees typically use a binary recursive partitioning process to generate rule-based models based on user-supplied training samples for estimating a target variable such as LAI [27, 44]. Cubist by *RuleQuest Inc.* was used in this study. The usefulness of Cubist for creating robust regression trees has been documented in the remote sensing literature [27, 45, 46].

The coefficient of determination (R^2), representing the goodness-of-fit of a model, and *RMSE* were used to measure calibration performance. Due to the limited number of field data points, it was not possible to perform cross-validation associated with the three approaches.

3.2. Vegetation mapping

Two different classification approaches were investigated to map vegetation on the Monticello, UT, and Monument Valley, AZ, hazardous waste sites. Both approaches employed machine learning decision trees, but one used scaled reflectance data as input variables while the other used mixture-tuned-matched-filtering (MTMF)-derived metrics and a suite of vegetation indices as input variables.

Decision trees have wide application for classification problems because they divide a complex decision into a hierarchy of simple and interpretable decisions [47-52]. See5 by *RuleQuest Research Inc.*, a widely used machine learning decision tree software, was used to generate decision trees for image classification.

MTMF is a hybrid classification method based on a combination of linear mixture theory and matched filtering, which is based on a partial unmixing approach with user-defined targets [53]. One of the advantages of MTMF is that the endmembers (i.e., spectral reflectance characteristics for spectrally pure materials) within a scene do not need to be identified because MTMF uses each endmember independently and models the pixel at each endmember as a mixture of the endmembers and an undefined background material [54, 55]. MTMF typically uses the minimum noise fraction (MNF) results extracted from the reflectance data. In this study, the cumulative 80% variation threshold was used to determine the subset of MNF results to be used in the MTMF analysis for each site. Consequently, the first 18 and 25 MNF transformed images were used in the MTMF analysis for the Monticello and Monument Valley sites, respectively. The MTMF output includes a matched filter (MF) score and an infeasibility value for each endmember. Ideally, pixels with a high MF score value and a low infeasibility value have a high percent cover of each endmember (e.g., sagebrush). Pixels with high MF score values and high infeasibility values may be false alarms. We used two image-derived endmembers for each class based on the percent cover data, which resulted in 8 endmembers for the Monticello site and 6 endmembers for the Monument Valley site.

The MF scores and infeasibility values were used as input variables along with a suite of vegetation indices. A total of 11 vegetation indices were used and two (i.e., VI1 and VI2) of them were developed from the LAI estimation in this study (Table 2). The original scaled reflectance data (126 bands) were also used as input variables in the decision tree classifications for comparison.

Table 2. Vegetation indices used in the decision tree classification.

Index	Equation	Reference
Simple ratio (SR)	$\frac{R_{845}}{R_{665}}$	Tucker [55]
Normalized difference vegetation index (NDVI)	$\frac{R_{845} - R_{665}}{R_{845} + R_{665}}$	Tucker [55]
Modified NDVI (MNDVI)	$\frac{R_{750} - R_{705}}{R_{750} + R_{705}}$	Fuentes et al. [56]
Photochemical reflectance index (PRI)	$\frac{R_{531} - R_{570}}{R_{531} + R_{570}}$	Gamon et al. [57]
Normalized difference water index (NDWI)	$\frac{R_{860} - R_{1240}}{R_{860} + R_{1240}}$	Gao [58]
Water band index (WBI)	$\frac{R_{900}}{R_{970}}$	Penuelas et al. [59]
Normalized difference nitrogen index (NDNI)	$\frac{\log(R_{1680}/R_{1510})}{\log(1/R_{1680}R_{1510})}$	Serrano et al. [60]
Normalized difference lignin index (NDLI)	$\frac{\log(R_{1680}/R_{1754})}{\log(1/R_{1680}R_{1754})}$	Serrano et al. [60]
Cellulose absorption index (CAI)	$\frac{R_{2020} + R_{2220}}{2} - R_{2100}$	Nagler et al. [61]
(VI1) for the Monticello site	$\frac{R_{1583.8}}{R_{1746.6}}$	From this study
(VI2) for the Monticello site	$\frac{R_{1746.6} - R_{1583.8}}{R_{1746.6} + R_{1583.8}}$	From this study
(VI1) for the Monument Valley site	$\frac{R_{1329.5}}{R_{1187.8}}$	From this study
(VI2) for the Monument Valley site	$\frac{R_{1329.5} - R_{1187.8}}{R_{1329.5} + R_{1187.8}}$	From this study

Decision trees are known to be sensitive to the characteristics of the training samples, especially when there are a limited number of training samples [47]. To improve the stability of the decision tree classifier, an aggregation approach was introduced [55, 63-66] where multiple decision trees were generated with different sets of training samples and a majority rule was used to determine the class for each pixel. Due to the relatively small number of training data samples, we used 50% of the reference data through random selection to train the decision tree and the remaining 50% to test the decision tree. A total of 40 decision trees were generated and 20 trees were selected based on the testing results for each site (i.e., the decision trees with higher testing accuracy were selected). These 20 trees were used for the voting process to generate the final vegetation map for each site.

Accuracy assessment of the vegetation maps included the determination of commission and omission errors for each class, overall accuracy (%), and the Kappa Coefficient of Agreement (κ). A Kappa *Z*-test was used to determine if there was a significant difference between two Kappa Coefficients associated with each site.

4. Results and Discussion

4.1. LAI Estimation

The results of the correlation matrices produced between the vegetation indices and LAI for all possible band combinations are shown in Figure 3. The greater the correlation of the ground reference LAI value with the two bands used to compute the vegetation index, the more blue the pixel in the diagram. Both VI1 and VI2 exhibited very similar patterns for each site. While relatively high correlations between the vegetation indices and LAI were found in the region between 1500 and 1800 nm for the Monticello site, they were found in the region between 900 and 1400 nm for the Monument Valley site (Figure 3). Interestingly, the bands in the red and near-infrared regions, which are assumed to have more information regarding vegetation health, did not produce higher correlations than the bands in the middle-infrared region. This might be because the study sites are located in semi-arid/arid areas and thus the spectral response of each vegetation type may be distinguishable in the middle-infrared region, where the spectral response of vegetation to water is more sensitive.

Table 3 summarizes the best band combination and associated statistics (i.e., R^2, *RMSEs* from calibration and cross-validation) of the two vegetation indices for each site. It was not surprising that one of the two best bands was found in the water absorption region (~1,200 nm) [67] for the Monument Valley site because some of the sample locations were irrigated while the others were not at this site. The vegetation index approach resulted in better performance for the Monument Valley site than the Monticello site (e.g., R^2 = 0.501 accounting for 50% of the variance with r = 0.7). This might be because the vegetation distribution and structure were more dynamic in the Monticello site than in the Monument Valley site. In particular, the grass and shrub species were highly mixed in the Monticello site, and this made it difficult to calibrate the *in situ* LAI data with the HyMap data at the 2.3 × 2.3 m resolution. Slight difference between the field and the pixel location might have also introduced errors in LAI estimation.

Figure 3. The correlation matrices using the vegetation index approach to estimate LAI: using (a) VI1, and (b) VI2 for the Monticello, UT site; and using (c) VI1 and (d) VI2 for the Monument Valley, AZ site.

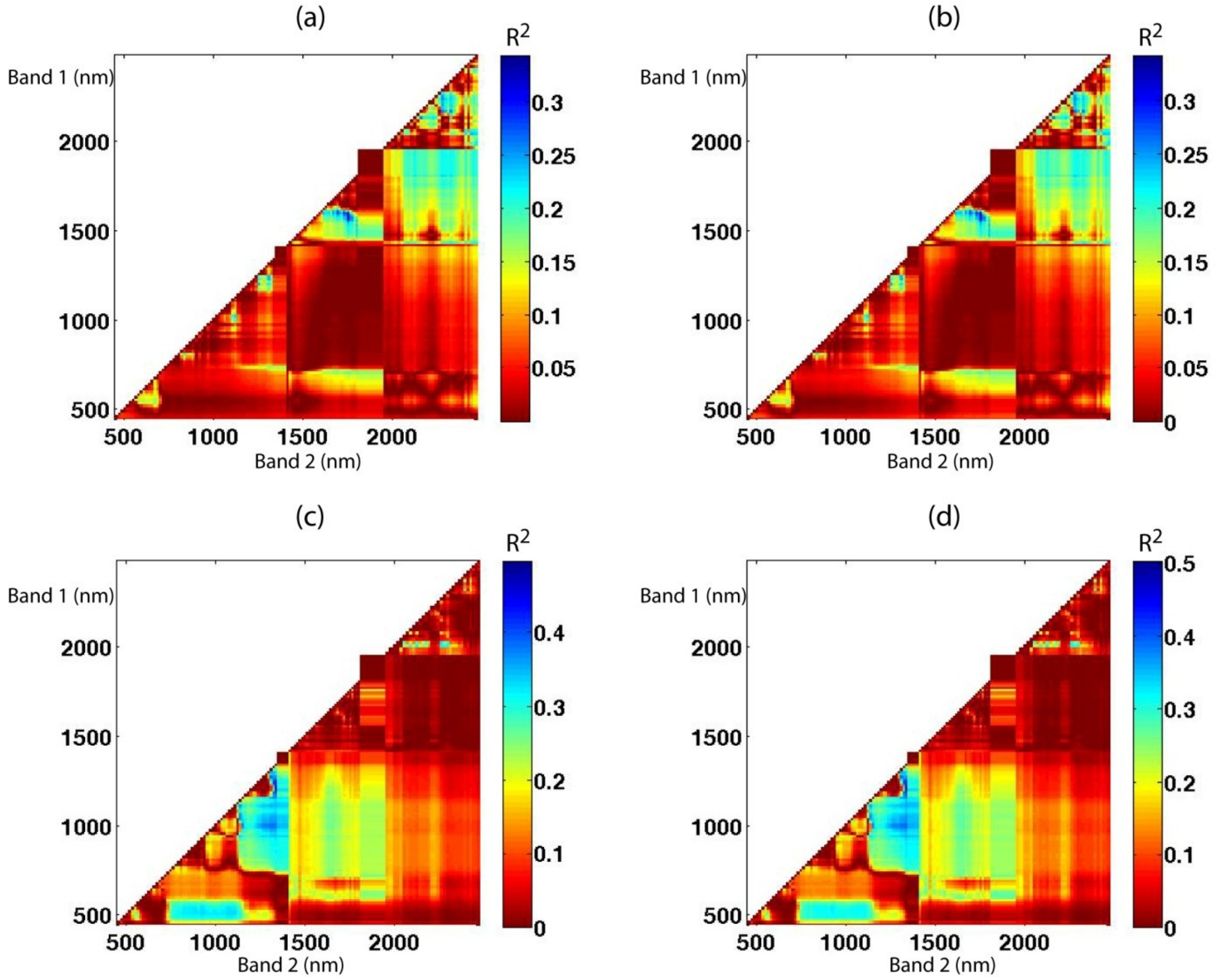


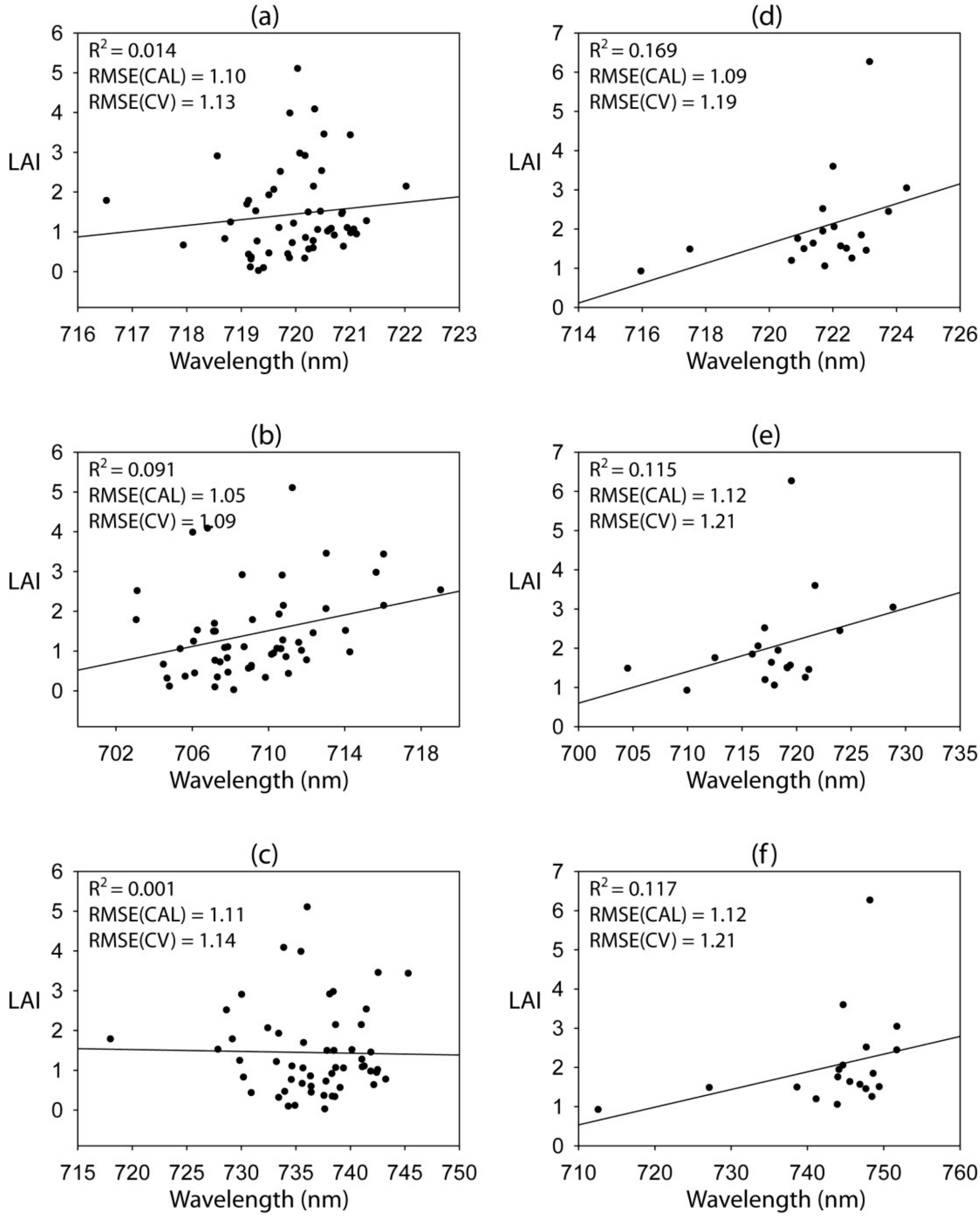
Table 3. The best band combination and associated statistics of both VI1 and VI2 for the Monticello and Monument Valley sites.

Site	Vegetation index	Band 1 (nm)	Band 2 (nm)	R^2	RMSE
Monticello	**VI1**	1583.8	1746.6	0.343	0.90/0.93*
	VI2	1583.8	1746.6	0.341	0.90/0.93
Monument Valley	**VI1**	1187.8	1329.5	0.495	0.85/1.03
	VI2	1187.8	1329.5	0.501	0.84/1.03

* RMSE from calibration/RMSE from cross validation

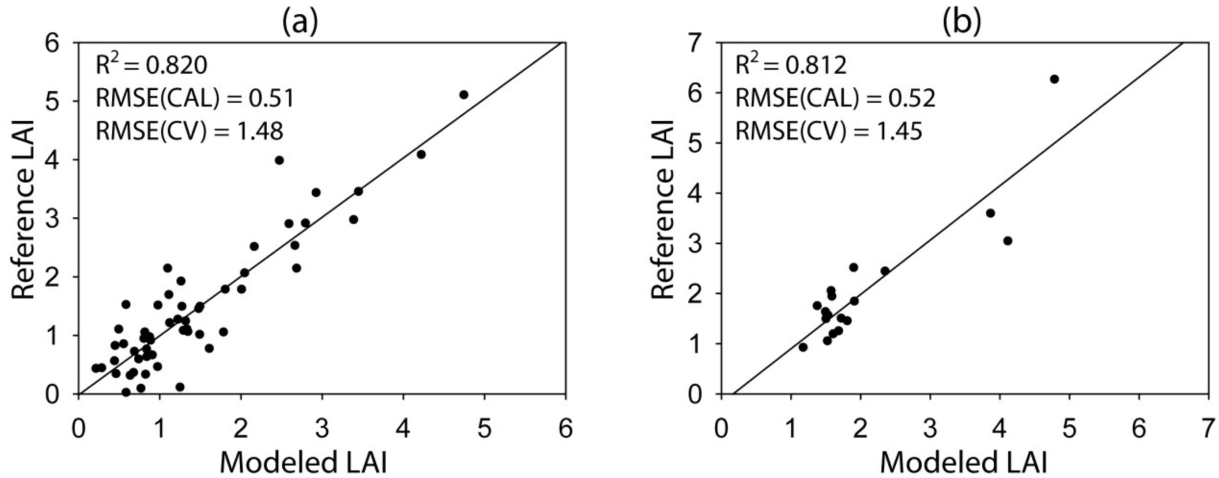
The scatterplots between each of the three REP approaches and LAI are shown in Figure 4. The REP approach did not predict LAI well, resulting in low correlations (< 0.2). Similar to the vegetation index approach, REP methods resulted in better performance for the Monument Valley site than the Monticello site. While the LG_REP resulted in the best performance for the Monticello site, the LI_REP produced the highest accuracy in LAI estimation for the Monument Valley site. However, there was no significant relationship between any of the REPs and LAI ($p > 0.05$). One of the reasons for the poor performance of the REP approach might be background soil spectral influence within pixels. The REP approach typically works well for vegetation with full canopies [68, 69]. Many of the sample pixels had vegetation cover between 70% and 90%, and thus the background soil could have influenced the spectral response associated with the red-edge. That most of the sample pixels contained multiple vegetation species may have also caused the REP approach fail to estimate LAI because each species typically has a unique REP characteristic.

Figure 4. The scatterplots between each of the Red-edge position (REP) and LAI: using (a) LI_REP, (b) LG_REP, and (c) LE_REP for the Monticello, UT site; and using (d) LI_REP, (e) LG_REP, and (f) LE_REP for the Monument Valley, AZ site. The R^2 and RMSEs from calibration (CAL) and cross-validation (CV) are also provided.



Unlike the results of the VI- and REP-based LAI estimation, the regression trees resulted in very good LAI estimation performance ($R^2 > 0.8$) (Figure 5). The regression trees generated three rules for the Monticello, UT, site and two rules for the Monument Valley, AZ, site. Interestingly, one of the three rules for the Monticello site were associated mainly with the grass species samples (i.e., wheatgrass and litter) while the other two rules were applied to most of shrub species samples (i.e., sagebrush and rabbitbrush). Eight bands were used to generate the multivariate equations for the Monticello site, which was not efficient and resulted in inflation of the fitness of the models. Conversely, only two bands (709 and 754 nm) in the red-edge region were used to generate the multivariate equations in the regression trees for the Monument Valley site. One rule was applied to relatively high LAI (> 2) samples for the Monument Valley site, while the other was applied to the lower LAI samples. The red-edge bands effectively divided the samples into the two groups.

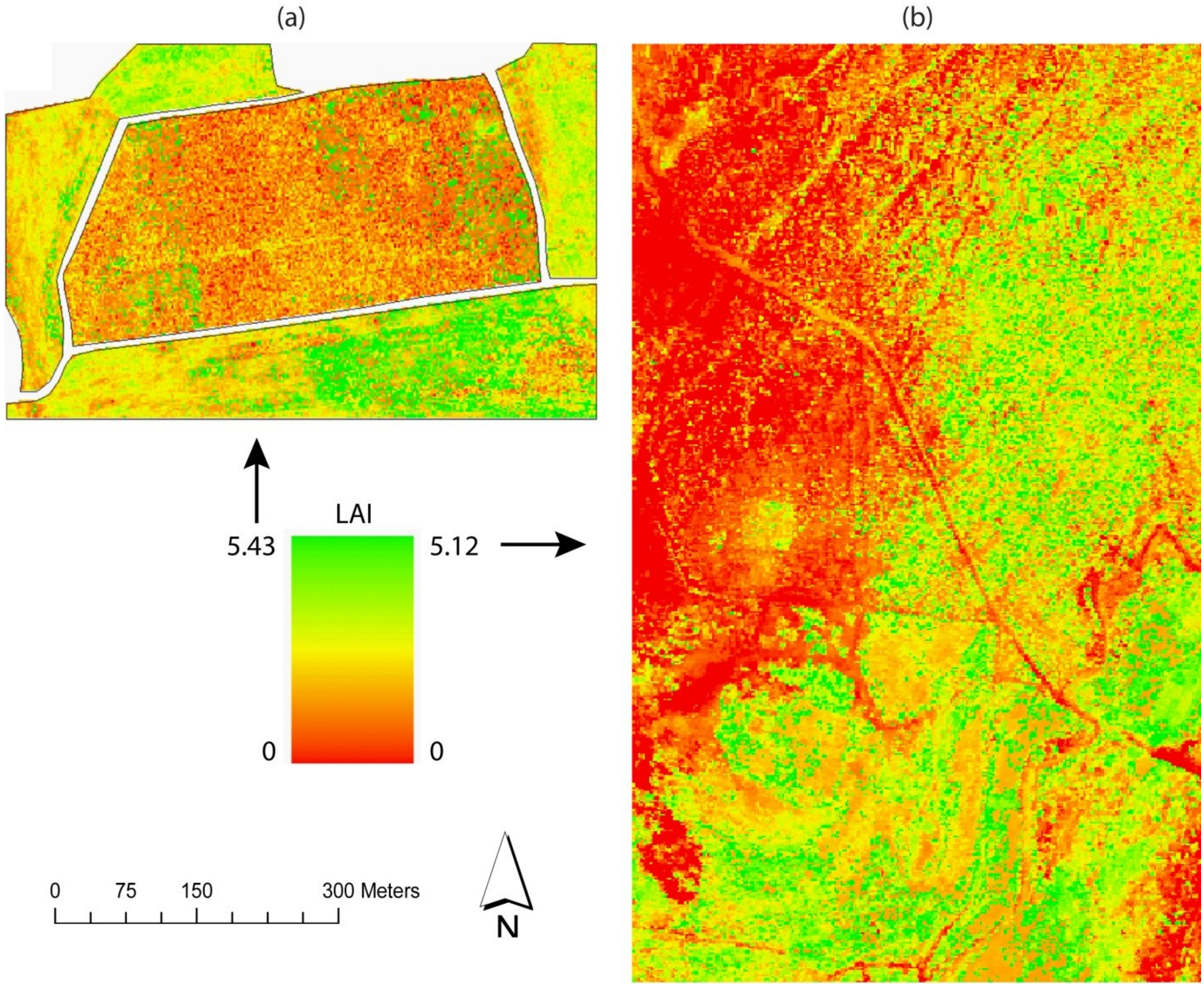
Figure 5. LAI estimation using the regression tree approach for (a) the Monticello, UT site, and (b) the Monument Valley, AZ site. The R^2 and RMSEs from calibration (CAL) and cross-validation (CV) are summarized in the plots.



Although the calibration using the regression trees outperformed those using the vegetation index and REP data, the cross validation did not correspond to the calibration results. The RMSEs through cross-validation using the regression trees were higher than those using the other approaches. While most of the folds for cross-validation generally resulted in low errors, some of the folds (~ 25%) resulted in high LAI estimation errors (e.g., > 2), which consequently increased the total RMSE. The overfitting problem of the regression trees, especially when a small number of samples are used, often occurs [27]. This is also related to the well known problem of decision/regression trees, which are sensitive to the training data configuration [70]. Glenn *et al*. [71] investigated black greasewood and fourwing saltbush over the Monument Valley site using 2007 MODIS data. They measured LAI using the traditional direct method and found a good agreement with the scaled Enhanced Vegetation Index (EVI) from MODIS data based on a simple linear regression ($R^2 = 0.77$; $n = 55$; $P < 0.001$). Figure 6 shows the LAI distribution maps estimated using the regression trees for the Monticello and Monument Valley sites.

Although all possible two-band combinations were tested for the VI1 and VI2 in this study, the most valuable VI might consist of more than two narrow hyperspectral bands. We tested a few narrow band-derived VIs such as the Vogelmann Red Edge Index 2 [72], which uses spectral data at more than two wavelengths. However, the additionally tested indices did not outperform the VIs used in this study. Optimization of multiple wavelengths associated with such narrow band-derived VIs might further improve the performance of LAI estimation.

Figure 6. The estimated LAI distribution maps for (a) the Monticello site, and (b) the Monument Valley site. The dirt road and other land cover classes were masked out for the Monticello site.



4.2. Vegetation Mapping

Figures 7 and 8 show the relationship between the percent cover and the matched filter scores for the Monticello, UT, and Monument Valley, AZ, sites, respectively. The percent covers of wheatgrass and litter were correlated with the matched filter scores at the 90% and 95% confident levels, respectively. However, the percent covers of sagebrush and rabbitbrush failed to be significantly correlated with the corresponding matched filter scores. There might be two reasons for this: the spectral separability of sagebrush and rabbitbrush from other species was not strong. In addition, most of the sagebrush and rabbitbrush samples were dominated by the corresponding species (i.e., percent cover > 70%), while a few of them were mixed with other species (i.e., percent cover < 60%). MTMF is known to be sensitive to the amount of green vegetation present within a pixel [38, 55]. Litter (dead plant materials) for the Monticello site and soil for the Monument Valley site could affect determining the matched filter scores of the healthy vegetation species. Locational errors could also influence the relatively lower correlation between the percent covers and the matched filter scores. Interestingly, the percent covers of saltbush were well correlated with the matched filter scores (Figure 8b). Excluding one outlier, the percent covers of greasewood were also well correlated with the scores. The MTMF approach has been successfully applied to map single vegetation species resulting in good relationships between percent cover and matched filter scores [73], but it typically results in more variation and confusion between species when multiple species are considered [74]. In this study, while several vegetation species were mixed at each sampling location in the Monticello site, one species was dominant at each sampling location in the Monument Valley site. Consequently, there was less influence from other species in the relationship between the percent cover and the matched filter scores for the Monument Valley site than for the Monticello site. However, because the matched filter scores were not sufficient for separating each species from the others, the use of a suite of vegetation indices was expected to improve classification accuracy using decision tree logic.

Figure 7. The relationships between the matched filter scores and the percent cover of the vegetation species for the Monticello site: (a) sagebrush, (b) rabbitbrush, (c) wheatgrass, and (d) litter.

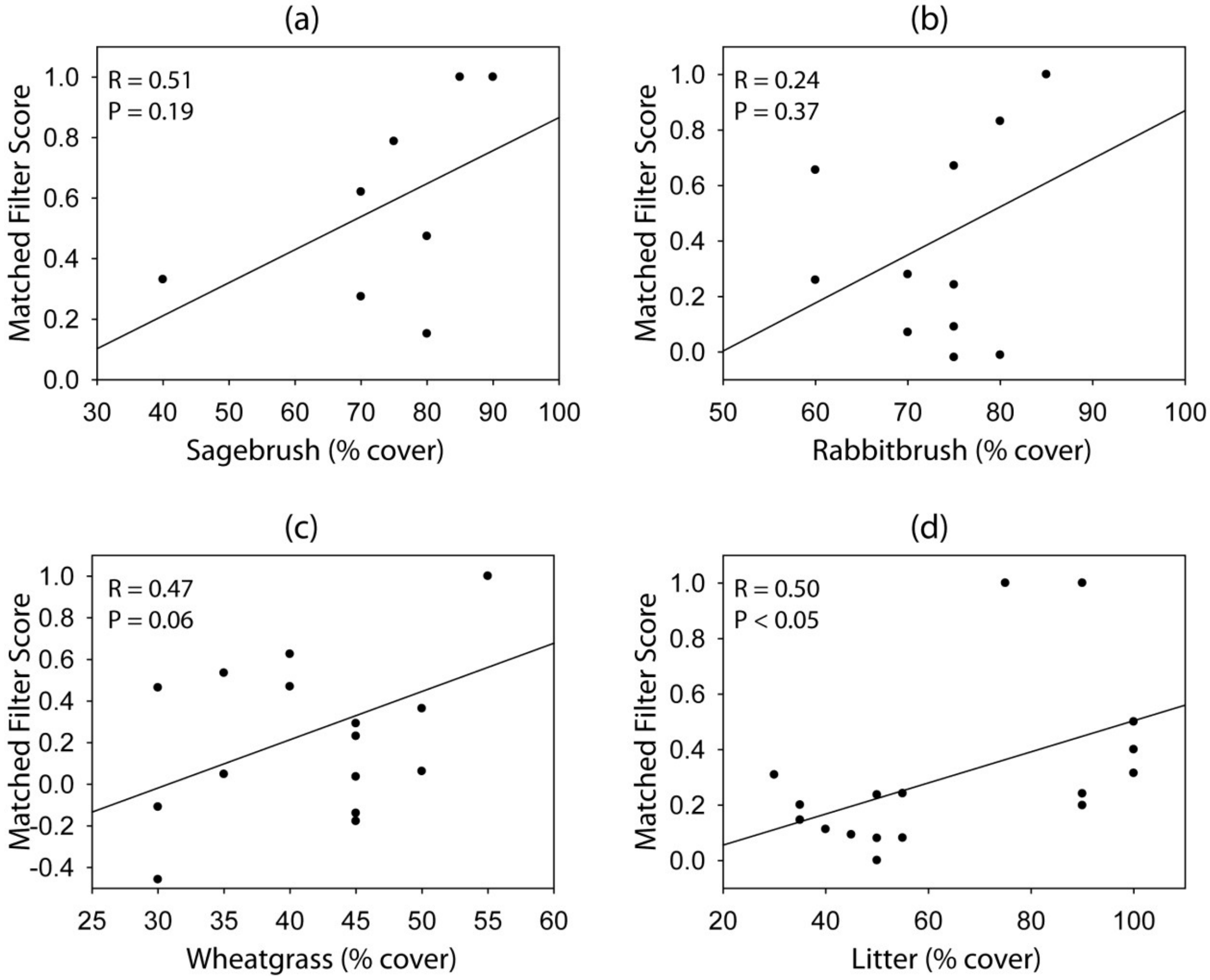


Figure 8. The relationships between the matched filter scores and the percent cover of the vegetation species for the Monument Valley site: (a) greasewood and (b) saltbush.

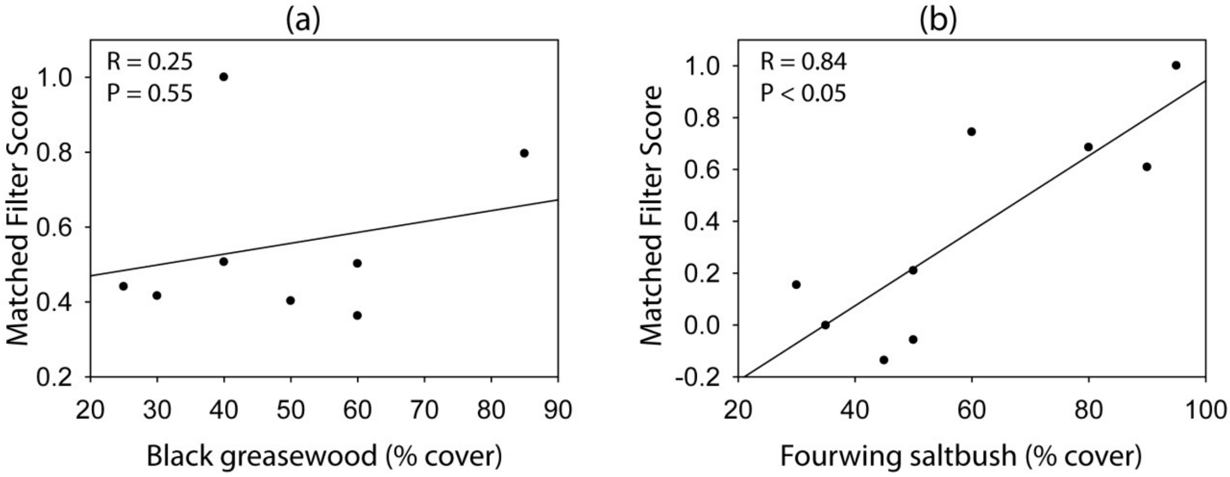


Figure 9 shows the performance variations of the multiple (i.e., 20) decision trees using different sets of training samples for each site. Since the size of the reference data was small (i.e., 53 samples of four classes for the Monticello site and 43 samples of three classes for the Monument Valley site), the performance variation of the multiple decision trees for training and testing was a bit large. For both sites, the MTMF-derived metrics and the vegetation indices (labeled MV in Figure 9) resulted in better performance (for both training and testing) than the original scaled reflectance (labeled REF in Figure 9) in decision tree classification.

Table 4 lists the key input variables to the decision trees classifications for each site. The PRI and NDNI among the vegetation indices were very useful in the decision tree classification for the Monticello site. The MF scores of sagebrush and rabbitbrush also contributed to the decision tree generation for the Monticello site. When the original scaled reflectance data were used in the decision tree classification for the Monticello site, the red (663.4 nm) and red-edge bands (709 nm) contributed most to the classification, followed by the blue (443.3 nm) and middle-infrared band (2477.5 nm). A different pattern of contributing variables was found for the Monument Valley site: The WBI, NDLI, and NDWI contributed most to the MTMF and vegetation index-based decision tree classification. The infeasibility of greasewood was also very useful. When the reflectance data were used, the bands near the water absorption features (i.e., around 1400 and 1940 nm) contributed most to the classification. This may be because the Monument Valley site is located in an arid area and some of the sampling locations were irrigated while others were not. That is why the water-related vegetation indices and reflectance were very useful in vegetation mapping for the Monument Valley site. The chlorophyll absorption features and related vegetation indices contributed moderately to the decision trees for both sites.

Figure 9. Box plots showing the performance variation of the multiple decision trees using different sets of training and testing samples: (a) for the Monticello site, and (b) for the Monument Valley site. MV represents the decision trees using the MTMF-derived metrics and vegetation indices and REF represents the decision trees using the original scaled reflectance data.

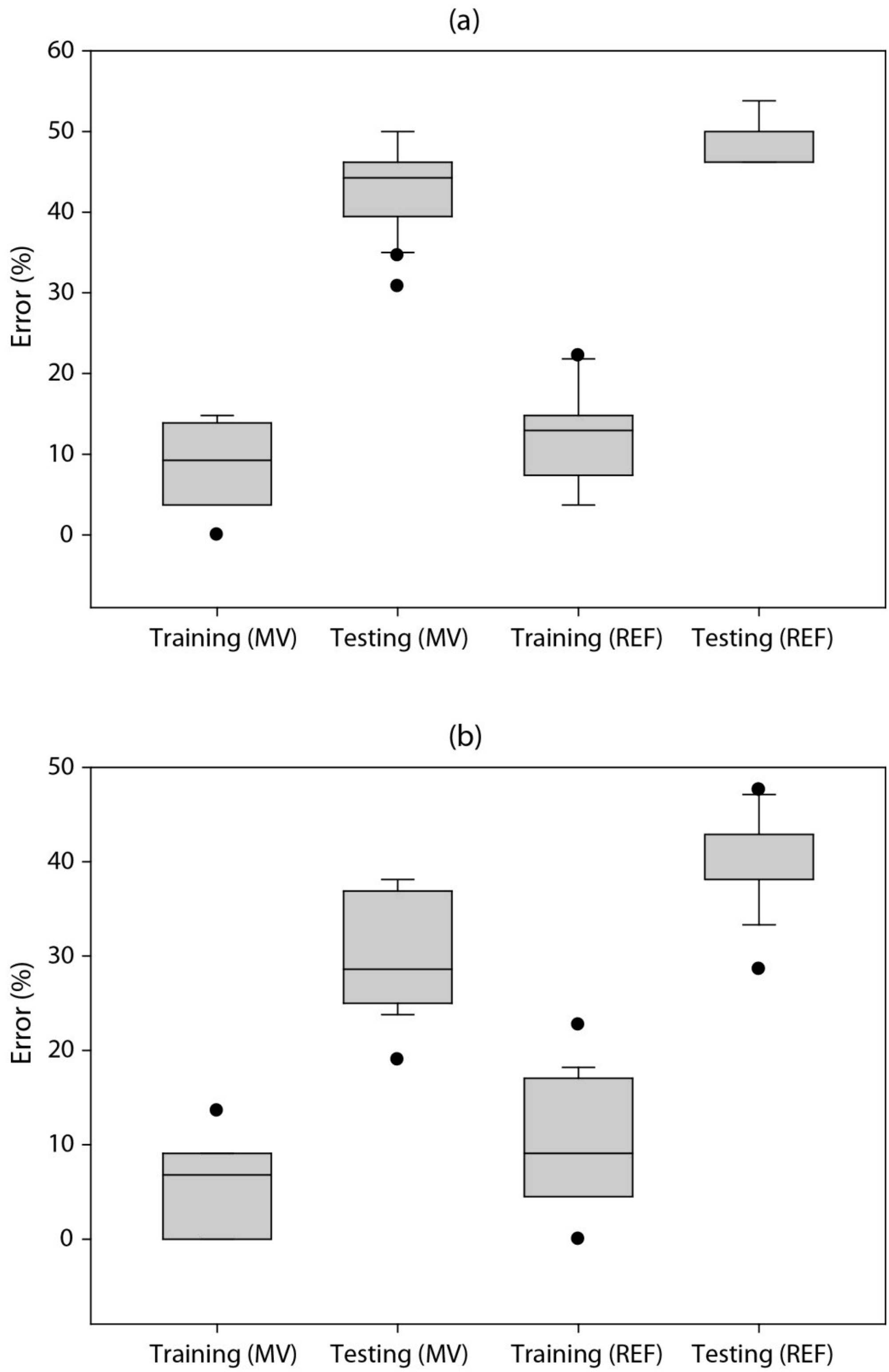


Table 4. Key input variables to the decision trees classifications: (a) for the Monticello site, and (b) for the Monument Valley site.

(a)

Ranking	**MTMF + Vegetation indices**		**Reflectance**	
	Variable	**Average contribution (%)**	**Variable**	**Average contribution (%)**
1	PRI	65.9	Band 16 (663.4 nm)	30.0
2	NDNI	50.0	Band 19 (709.0 nm)	26.9
3	MF-Sagebrush 1	49.4	Band 1 (443.3 nm)	25.9
4	CAI	35.0	Band 126 (2477.5 nm)	25.2
5	MF-Sagebrush 2	31.1	Band 21 (739.2 nm)	22.8
6	MF-Rabbitbrush 1	21.3	Band 2 (451.1 nm)	22.4
7	MF-Rabbitbrush 2	15.9	Band 3 (464.5 nm)	16.7
8	INF-Sagebrush 2	14.8	Band 28 (844.4 nm)	14.1
9	VI1	11.9	Band 64 (1419.9 nm)	11.3
10	MF-Litter 1	10.2	Band 94 (1805.8 nm)	10.2

(b)

Ranking	**MTMF + Vegetation indices**		**Reflectance**	
	Variable	**Average contribution (%)**	**Variable**	**Average contribution (%)**
1	WBI	50.0	Band 63 (1405.4 nm)	35.0
2	NDLI	38.0	Band 97 (1987.0 nm)	30.0
3	NDWI	30.0	Band 31 (892.3 nm)	20.5
4	INF-Greasewood 2	20.5	Band 1 (443.3 nm)	17.5
5	MF-Saltbush 1	18.7	Band 17 (678.4 nm)	15.0
6	NDNI	15.0	Band 61 (1329.5 nm)	15.0
7	VI1	6.9	Band 4 (480.6 nm)	12.1
8	MF-Greasewood 1	6.0	Band 22 (754.1 nm)	10.0
9	INF-Greasewood 1	5.5	Band 94 (1805.8 nm)	10.0
10	MF-Soil 1	5.0	Band 23 (769.3 nm)	8.4

A majority rule was applied to produce the final species distribution map based on the multiple decision trees. Because some pixels had multiple maximum votes, additional processing was necessary. A three-step approach was used: a majority rule was first applied using all of the 20 decision trees, and then another majority rule was applied to the undecided pixels using the top 10 decision trees based on their testing performance. Finally, the best decision tree was used to determine the classes for the still undecided pixels and the final vegetation maps were created for each site.

Figure 10 shows the vegetation maps over the Monticello site using the two sets of decision trees (i.e., one set using the MTMF variables and vegetation indices and the other using the scaled reflectance). Sagebrush appeared to be somewhat overestimated through visual inspection when the MTMF variables and vegetation indices were used (Figure 10a). On the other hand, rabbitbrush and litter were generally overestimated when the scaled reflectance data were used for decision tree classification (Figure 10b). The vegetation maps of the Monument Valley site using the two sets of the decision trees are shown in Figure 11. The classes were more clumped in the map using the scaled reflectance than using the MTMF variables and vegetation indices. Soil appeared to be overestimated when the reflectance data were used. These classification maps exhibited some discrepancies with the actual field conditions. For example, while the disposal cell cover (central region) in Figure 10 appears close to the field conditions, a monoculture of rabbitbrush on the side slopes does not agree with the field conditions. At the Monument Valley site, greasewood appears to be over classified. The limited quantity and quality of the field reference data might account for the discrepancies, including 1) the small sample size resulted in variations in performance for multiple decision tree classifications; and 2) each ground reference sample was measured using a circular plot with a diameter of 1 m, which is smaller than the hyperspectral image pixel size (2.3 x 2.3 m). Given the small shrub and grass patches in the sites, this could result in significant confusion in classification.

Figure 10. The vegetation species distribution maps for the Monticello site based on the decision trees (a) using the MTMF-derived metrics and vegetation indices, and (b) using the original scaled reflectance data. The road and other land cover classes were masked out.

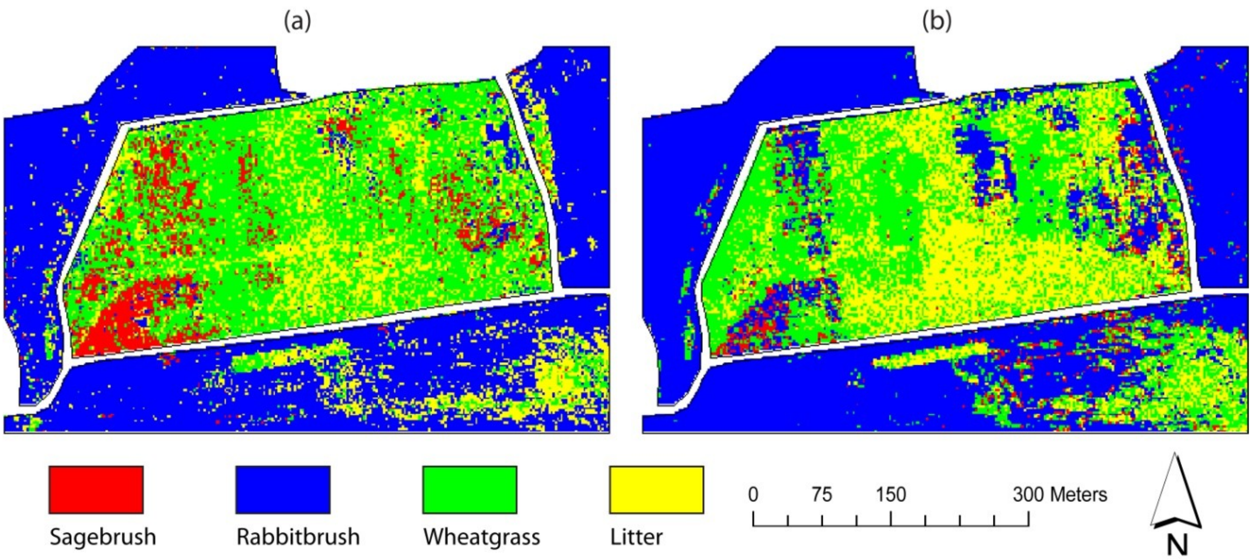
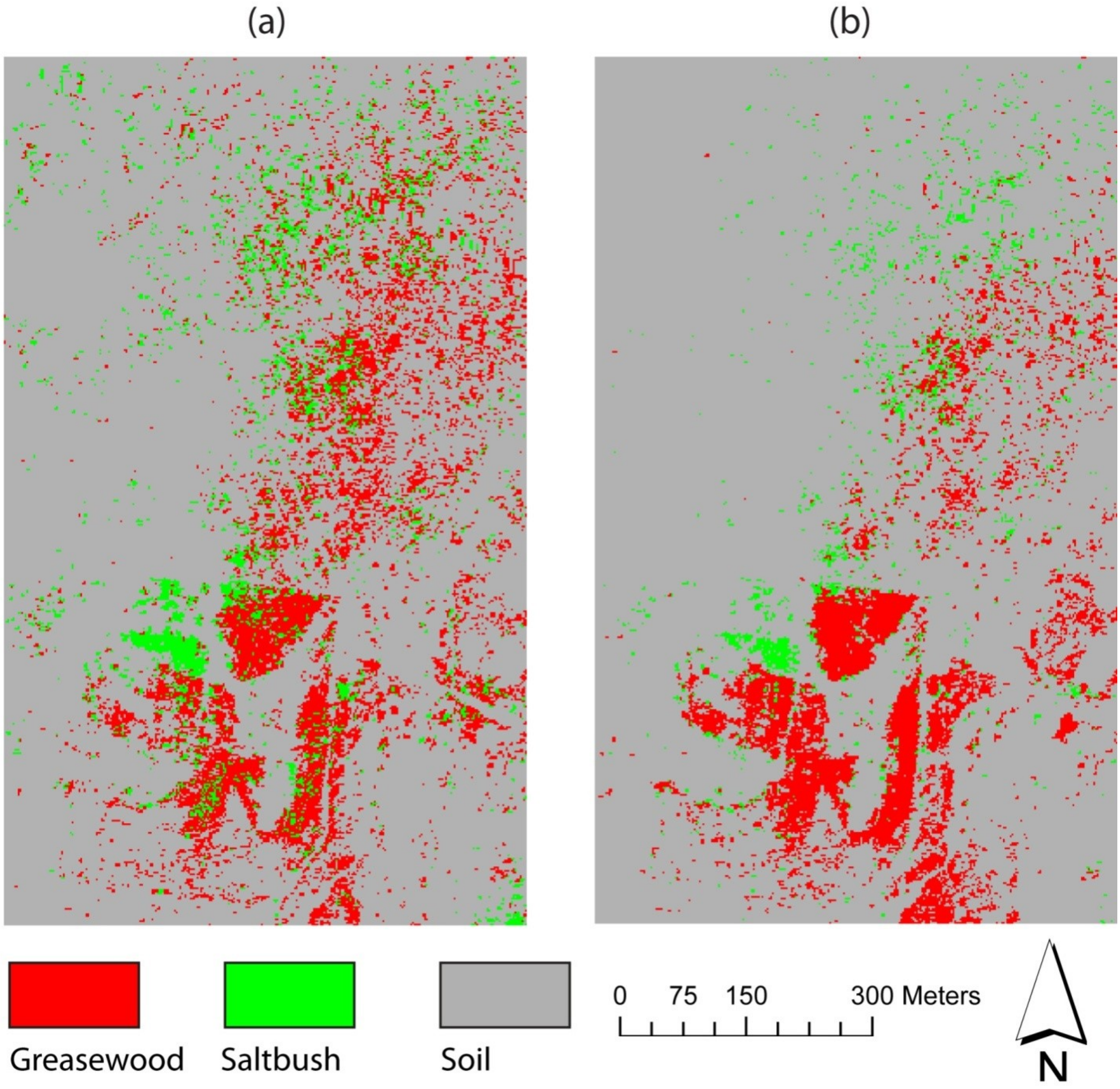


Figure 11. The vegetation species distribution maps for the Monument Valley site based on the decision trees (a) using the MTMF-derived metrics and vegetation indices, and (b) using the original scaled reflectance data.



Since all of the reference data might have been used to train decision trees (50% on average), the classification accuracy based on the assessment using the reference data might be a bit inflated. The classification accuracy assessment for the Monticello site is presented in Table 5. Interestingly, although the MTMF variables and the vegetation indices outperformed the original scaled reflectance based on the individual decision trees for both training and testing (refer to Figure 9), the accuracy assessment results were similar between the two maps, resulting in the overall accuracy around 87% and Kappa around 0.82. Similar to the visual inspection of the classification maps, the commission error of sagebrush when using the MTMF variables and vegetation indices was large (~36.4%), while the omission error of sagebrush when using the reflectance data was large (~37.5%). Due to the small size of the reference data, the overestimation of rabbitbrush and litter when using the scaled reflectance was not clearly indicated by the error matrix. Wheatgrass was confused with sagebrush and litter in the classification using the reflectance data. There was no significant difference between the two Kappa values (Table 5c).

Table 5. Accuracy assessment results of the decision tree classifications for the Monticello site: (a) using the MTMF variables and vegetation indices, (b) using the scaled reflectance data, and (c) Kappa Z-test between the two classifications.

(a)

Ref. / Class.	Sagebrush	Rabbitbrush	Wheatgrass	Litter	Sum	Commission Errors (%)
Sagebrush	**7**	2	1	1	11	36.4
Rabbitbrush	0	**10**	1	0	11	9.1
Wheatgrass	1	0	**14**	1	16	12.5
Litter	0	0	0	**15**	15	0
Sum	8	12	16	17	**53**	
Omission Errors (%)	12.5	16.7	12.5	11.8		

Overall accuracy: 86.79%
Kappa Coefficient of Agreement: 0.821

(b)

Ref. / Class.	Sagebrush	Rabbitbrush	Wheatgrass	Litter	Sum	Commission Errors (%)
Sagebrush	**5**	1	1	0	7	28.6
Rabbitbrush	1	**11**	0	0	12	8.3
Wheatgrass	2	0	**15**	2	19	11.1
Litter	0	0	0	**15**	15	0
Sum	8	12	16	17	**53**	
Omission Errors (%)	37.5	8.3	6.3	11.8		

Overall accuracy: 86.79%
Kappa Coefficient of Agreement: 0.819

(c)

	Kappa	ASE	MTMF + VIs	Reflectance
MTMF + VIs	0.821	0.0622	NA	
Reflectance	0.819	0.0629	0.0276*	NA

* no significant difference between the two Kappa values at the 95% confidence level

Table 6 presents the accuracy assessment results of the classification maps for the Monument Valley site. Saltbush was confused with soil for both maps. This might be because some of the saltbush sample locations were grazed and not irrigated, causing their spectral response to be similar to soil. Saltbush was also confused with greasewood when the reflectance data were used in the decision tree classification. However, the confusion was much improved when the MTMF variables and vegetation indices were used in the decision tree classification. There was no significant Kappa difference between the two classifications due to the relatively large asymptotic standard errors (ASE), a measurement of uncertainty, even though there was a 10% difference between the two Kappa values (Table 6c).

Table 6. Accuracy assessment results of the decision tree classifications for the Monument Valley site: (a) using the MTMF variables and vegetation indices, (b) using the scaled reflectance data, and (c) Kappa Z-test between the two classifications.

(a)

Class. \ Ref.	Greasewood	Saltbush	Soil	Sum	Commission Errors
Greasewood	**10**	1	0	11	9.1
Saltbush	1	**10**	1	12	16.7
Soil	1	3	**16**	20	20
Sum	12	14	17	**43**	
Omission Errors	16.7	28.6	5.9		

Overall accuracy: 83.72%

Kappa Coefficient: 0.751

(b)

Class. \ Ref.	Greasewood	Saltbush	Soil	Sum	Commission Errors
Greasewood	**11**	3	1	15	26.7
Saltbush	1	**7**	0	8	12.5
Soil	0	4	**16**	20	20
Sum	12	14	17	**43**	
Omission Errors	8.3	50	5.9		

Overall accuracy: 79.07%

Kappa Coefficient: 0.682

(c)

	Kappa	ASE	MTMF + VIs	Reflectance
MTMF + VIs	0.751	0.0856	NA	
Reflectance	0.682	0.0905	0.5541*	NA

* no significant difference between the two Kappa values at the 95% confidence level

For vegetation mapping over hazardous waste sites, the omission errors of shrub could be more serious than the commission errors. The commission errors could be false alarms for biointrusion on the capped materials, but the omission errors might indicate undetected biointrusion, which requires quick response and treatment. From this point of view, the use of the MTMF-derived metrics and vegetation indices was better than the use of the scaled reflectance for vegetation mapping. The omission errors of rabbitbrush and saltbush (both shrub species) were relatively large when the scaled reflectance data were used.

At the Monticello site, rabbitbrush is an early successional shrub adapted to disturbed, unstructured soils. Sagebrush is a later successional shrub that appears to be increasing in abundance as soil structure develops in the engineered soil cover, creating preferential flow pathways for water to move deeper in the profile and, hence, gradually creating a more favorable habitat for sagebrush. The phytoremediation study at the Monument Valley site was designed, in part, to compare the two dominant native desert phreatophytes: the obligate black greasewood and facultative four-wing saltbush. Consequently, for long-term monitoring of these sites, differentiating rabbitbrush and sagebrush at the Monticello site, and greasewood and saltbush at the Monument Valley site are critical.

When considering the small shrub and grass patches (~ 1 m and sometimes < 1 m in size) found in the study sites, the spatial resolution of the HyMap imagery (2.3 × 2.3 m) appears to be a bit coarse. Although there is a concern that higher spatial resolution data may actually reduce classification accuracy by increasing within-class spectral variability [75], the Monticello and Monument Valley sites should benefit from higher spatial resolution data (e.g., ~ 1 × 1m) for vegetation mapping because small grass and shrub patches (not tall vegetation) are dominant and their cover is relatively dense (percent cover > 70%).

5. Summary and Conclusions

This study evaluated the usefulness of HyMap hyperspectral data for characterizing the vegetation cover (i.e., LAI estimation and vegetation species mapping) on two hazardous waste sites. The findings of this study are:

- The vegetation index approach to estimating LAI revealed that reflectance data in the middle-infrared region were more useful than reflectance data in the red or near-infrared region. The REP approach failed to estimate LAI mainly because the 2.3 x 2.3 m pixels often contained multiple species and background soil spectral influence. The regression trees resulted in the best calibration accuracy for estimating LAI ($R^2 > 0.80$), but the instability of the models due to the small sample size was a concern. More sophisticated narrow band-derived vegetation indices need to be investigated further.
- Aggregated decision trees were successfully used to map the vegetation species with a limited amount of reference data. Overall accuracies exceeded 85% in both study areas with Kappa Coefficients of Agreement > 0.81. The MTMF approach and the associated metrics improved the classification accuracy. However, if the endmember pixels contained only one species (i.e., they were pure pixels), the MTMF approach resulted in improved classification. In other words, the MTMF approach would benefit from the use of higher spatial resolution hyperspectral remote sensor data. A suite of vegetation indices were also useful for the vegetation mapping. In particular, the water-related indices were especially useful for classification of the Monument Valley site.
- The site characteristics influenced the performance of the LAI estimation and vegetation mapping. While multiple species were generally found in the individual pixels in the Monticello site, only two species were dominant within the scene in the Monument Valley site, which was more arid than the Monticello site and included irrigation/non-irrigation treatments. These site characteristics affected the performance of the LAI estimation and vegetation

mapping, and provided an explanation of the performance to some extent.

Automated monitoring of vegetation cover on hazardous waste sites using hyperspectral remote sensing data and modeling techniques appears feasible, but requires further investigation using different remote sensing data sources, higher spatial resolution hyperspectral data, and more advanced modeling techniques. Site characteristics must be carefully considered when determining the remote sensing data to be collected and the approaches to be used. Future research includes applications of multi-sensor data fusion (e.g., high density LiDAR data + hyperspectral imagery) and/or different modeling techniques (e.g., artificial immune networks, support vector machines, and artificial neural networks) for monitoring hazardous waste sites. In addition, while this study provided the preliminary results and single-date baseline data associated with monitoring of the phytoremediation systems at the Monticello and Monument Valley sites, linking remote sensing methods with actual monitoring tasks in a hazardous waste context should be further examined. Such links include: (1) detecting changes in the spatial distribution of plant species and LAI through time at the landscape scale using < 1 x 1 m multiple-date hyperspectral remote sensor data, and (2) tracking the response of phreatophyte health and evapotranspiration rates to changing land management practices.

Acknowledgements

This research was funded by the Department of Energy.

References and Notes

1. USEPA. *National Biennial RCRA Hazardous Waste Report: Based on 2007 Data*; USEPA, 2008; pp.1-68.
2. USEPA. *Evaluation of Subsurface Engineered Barriers at Waste Sites*; EPA 542-R-98-005, USEPA, 1998; pp. 1-148.
3. Jensen, J.; Hodgson, M.; Garcia-Quijano, M.; Im, J.; Tullis, J. A remote sensing and GIS-assisted spatial decision support system for hazardous waste site monitoring. *Photogramm. Eng. Rem. S.* **2009**, *75*, 169-178.
4. USEPA. Cleaning Up the Nations Waste Sites: Markets and Technology Trends. EPA 042-R-04-015. Cincinnati, Ohio, 2004.
5. Albright, W.; Benson, C.; Waugh, W. *Water Balance Covers for Waste Containment: Principles and Practices*. ASCE Press, Reston, VA, 2010.
6. USEPA. Fact Sheet on Evapotranspiration Cover Systems for Waste Containment. EPA 542-F-11-001. Office of Solid Waste and Emergency Response, Cincinnati, Ohio, 2011.
7. USEPA. Introduction to Phytoremediation. EPA/600/R-99/107, Cincinnati, Ohio, 2000.
8. Waugh, W.; Miller, D.; Glenn, E.; Moore, D.; Carroll, K.; Bush, R. Natural and enhanced attenuation of soil and groundwater at the Monument Valley, Arizona, DOE Legacy Waste Site. In *Proceedings of Waste Management 2010 Symposium, Phoenix, AZ*, 2010.
9. Herman J.; Waites, J.; Ponitz, R.; Etzler, P. A temporal and spatial-resolution remote-sensing study of a Michigan Superfund Site. *Photogramm. Eng. Rem. S.* **1994**, *60*, 1007-1017.
10. Vincent, R. Remote sensing for solid-waste landfills and hazardous-waste sites. *Photogramm. Eng.*

Rem. S. **1994**, 60, 979-982.

11. Schuerger, A.; Capelle, G.; Di Benedetto, J.; Mao, C.; Thai, C.; Evans, M.; Richards, J.; Blank, T.; Stryjewski, E. Comparison of two hyperspectral imaging and two laser-induced fluorescence instruments for the detection of zinc stress and chlorophyll concentration in bahia grass (*Paspalum notatum Flugge.*). *Remote Sens. Environ.* **2003**, *84*, 572-588.
12. Harris, A.; Bryant, R. A multi-scale remote sensing approach for monitoring northern peatland hydrology: Present possibilities and future challenges. *J. Environ. Manage*. **2009**, *90*, 2178-2188.
13. Nonomura, A.; Kitahara, M.; Masuda, T. Impact of land use and land cover changes on the ambient temperature in a middle scale city, Takamatsu, in Southwest Japan. *J. Environ. Manage*. **2009**, *90*, 3297-3304.
14. Gong, B.; Im, J.; Jensen, J.; Coleman, M.; Nelson, E. Characterization of forest crops with a range of nutrient and water treatments using AISA hyperspectral imagery. *GISci. Remote Sens.* **2011**, in press.
15. Thenkabail, P.; Lyon, J.; Huete, A. Hyperspectral Remote Sensing of Vegetation. 2011, CRC Press, 781 p.
16. Slonecker, T.; Fisher, G.; Aiello, D.; Haack, B. Visible and infrared remote sensing of hazardous waste: a review. *Remote Sens.* **2010**, *2*, 2474-2508.
17. Rogan, J.; Franklin, J.; Roberts, D. A comparison of methods for monitoring multitemporal vegetation change using Thematic Mapper imagery. *Remote Sens. Environ.* **2002**, *80*, 143-156.
18. Tooke, T.; Coops, N.; Goodwin, N.; Voogt, J. Extracting urban vegetation characteristics using spectral mixture analysis and decision tree classifications. *Remote Sens. Environ.* **2009**, *113*, 398-407.
19. Ke, Y.; Quackenbush, L.; Im, J. Synergistic use of QuickBird multispectral imagery and LiDAR data for object-based forest species classification. *Remote Sens. Environ.* **2010**, *114*, 1141-1154.
20. Laliberte, A.; Rango, A.; Havstad, K.; Paris, J.; Beck, R.; McNeely, R.; Gonzalez, A. Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. *Remote Sens. Environ.* **2004**, *93*, 198-210.
21. Tuanmu, M.; Vina, A.; Bearer, S.; Xu, W.; Quyang, Z.; Zhang, H.; Liu, J. Mapping understory vegetation using phenological characteristics derived from remotely sensed data. *Remote Sens. Environ.* **2010**, *114*, 1833-1844.
22. Im, J.; Jensen, J. Hyperspectral remote sensing of vegetation, *Geography Compass*, **2008**, *2*, 1943-1961.
23. Tilling, A.; O'Leary, G.; Ferwerda, J.; Jones, S.; Fitzgerald, G.; Rodriguez, D.; Belford, R. Remote sensing of nitrogen and water stress in wheat. *Field Crops Res.* **2007**, *104*, 77-85.
24. Goodenough, D.; Chen, H.; Dyk, A.; Han, T.; Li, J. Multisensor data fusion for aboveground carbon estimation. *Proceedings*, XXVIII General Assembly of the International Union of Radio Science (URSI), New Delhi, India, vol. CD, 2005; pp. 1-4.
25. Hansen, P.; Schjoerring, J. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens. Environ.* **2003**, *86*, 542-553.
26. Cho, M.; Skidmore, A. A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method. *Remote Sens. Environ.* **2006**, *101*, 181-193.

27. Im, J.; Jensen, J.; Coleman, M.; Nelson, E. Hyperspectral remote sensing analysis of short rotation woody crops grown with controlled nutrient and irrigation treatments. *Geocarto Int.* **2009***, 24*, 293-312.
28. Maire, G.; Francois, C.; Soudani, K.; Berveiller, D.; Pontailler, J.; Breda, N.; Genet, H.; Davi, H.; Dufrene, E. Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index, and leaf canopy biomass. *Remote Sens. Environ.* **2008**, *112*, 3846-3864.
29. Gong, P.; Pu, R.; Heald, R. Analysis of in-situ hyperspectral data for nutrient estimation of giant sequoia. *Int. J. Remote Sens.* **2002**, *23*, 1827-1850.
30. Bacour, C.; Baret, F.; Beal, D.; Weiss, M.; Pavageau, K. Neural network estimation of LAI, fAPAR, fCover and LAIxC(ab), from top of canopy MERIS reflectance data: principles and validation. *Remote Sens. Environ.* **2006**, *105*, 313−325.
31. Combal, B.; Baret, F.; Weiss, M.; Trubuil, A.; Mace, D.; Pragnere, A.; Myneni, R.; Knyazikhin, Y.; Wang, L. Retrieval of canopy biophysical variables from bidirectional reflectance: using prior information to solve the ill-posed inverse problem. *Remote Sens. Environ.* **2003**, *84*, 1−15.
32. Waugh, W.; Benson, C.; Albright, W. Sustainable covers for uranium mill tailings, USA: Alternative design, performance, and renovation. In *Proceedings of 12th International Conference on Environmental Remediation and Radioactive Waste Management, Liverpool, UK,* 2009; pp. 639-648.
33. Stormont, J.; Morris, C. Method to estimate water storage capacity of capillary barriers. *J. of Geotech. And Geoenviron Eng ASCE*, **1998**, *124*, 297−302.
34. McKeon, C.; Glenn, E.; Waugh, J.; Eastoe, C.; Jordan, F.; Nelson, S. Growth and water and nitrate uptake patterns of grazed and ungrazed desert shrubs growing over a nitrate contamination plume. *J. of Arid Environ.* **2006**, *61*, 119-136.
35. Jordan, F.; Waugh, J.; Glenn, E.; Sam, L.; Tompson, T.; Tompson, T.L. Natural bioremediation of a nitrate-contaminated soil-and-aquifer system in a desert environment. *J. of Arid Environ*. **2008**, *72*, 748-763.
36. Glenn, E.; Morino, K.; Didan, K.; Jordan, F.; Carroll, K.; Nagler, P.; Hultine, K.; Sheader, L.; Waugh, J. Scaling sap flux measurements of grazed and ungrazed shrub communities with fine and coarse-resolution remote sensing. Ecohydrology **2009**, *1*, 316-329.
37. HyVista. HyMap data product. http://www.hyvista.com/wp_11/wp-content/uploads/2011/02/hvc_data_products.pdf, last accessed on September 10th, 2011.
38. Jensen, J. *Introductory Digital Image Processing.* Upper Saddle River, Prentice-Hall, 2005; 526 p.
39. Horler, D.; Dockray, M.; Barber, J. The red edge of plant leaf reflectance. *Int. J. Remote Sens.* **1983**, *4*, 273−288.
40. Pinar, A.; Curran, P. Grass chlorophyll and the reflectance red edge. *Int. J. Remote Sens.* **1996**, *17*, 351-357.
41. Smith, K.; Steven, M.; Colls, J. Use of hyperspectral derivative ratios in the red edge region to identify plant stress responses to gas leak. *Remote Sens. Environ.* **2004**, *92*, 207−217.
42. Guyot, G.; Baret, F. Utilisation de la haute résolution spectrale pour suivre l'état des couverts végétaux, *Proceedings of the 4th International colloquium on spectral signatures of objects in remote sensing. ESA SP-287, Assois, France*, 1988; pp. 279−286.

43. Dawson, T.; Curran, P. A new technique for interpolating the reflectance red edge position. *Int. J. Remote Sens.* **1998**, *19,* 2133 – 2139.
44. Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. *Classification and regression trees.* Chapman and Hall, New York, 1984.
45. Huang, C.; Townshend, J. A stepwise regression tree for nonlinear approximation: Applications to estimating subpixel land cover. *Int. J. Remote Sens.* **2003**, *24*, 75-90.
46. Moisen, G.; Freeman, E.; Blackard, J.; Frescino, T.; Zimmermann, N.; Edwards, T. Predicting tree species presence and basal area in Utah: a comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. *Ecol. Model.***2006,** *199*, 176-187.
47. Li, R.; Belford, G. Instability of decision tree classification algorithms. In *Proceedings of the 8th ACM SIGKDD international conference on knowledge discovery and data mining,* Edmonton, Alberta, Canada, 2002; pp. 570 – 575.
48. Im, J.; Jensen, J. A change detection model based on neighborhood correlation image analysis and decision tree classification. *Remote Sens. Environ.* **2005**, *99*, 326-340.
49. Im, J.; Jensen, J.; Hodgson, M. Object-based classification using high posting density lidar data, *GISci. Remote Sens.* **2008**, *45*, 209-228,
50. Im, J.; Jensen, J.; Tullis, J. Object-based change detection using correlation image analysis and image segmentation techniques. *Int. J. Remote Sens.* **2008**, *29*, 399-423.
51. Im, J.; Lu, Z.; Rhee, J.; Quackenbush, L. Impervious surface quantification using a synthesis of artificial immune networks and decision/regression trees from multi-sensor data. *Remote Sens. Environ.* **2011**, in press.
52. Lu, Z.; Im, J.; Quackenbush, L.; Halligan, K. Population estimation based on multi-sensor data fusion. *Int. J. Remote Sens.* **2010**, *31*, 5587-5604
53. Williams, A.; Hunt, E. Estimation of leafy Spurge cover from hyperspectral imagery using mixture tuned matched filtering. *Remote Sens. Environ.* **2002**, *82*, 446-456.
54. Boardman, J.; Kruse, F.; Green, R. Mapping target signatures via partial unmixing of AVIRIS data, Summaries of the fifth JPL Airborne Geoscience Workshop JPL Publication, Vol. 95–1. (pp. 23−26) Pasadena, CA: NASA Jet Propulsion Laboratory, 1995.
55. Andrew, M.; Ustin, S. The role of environmental context in mapping invasive plants with hyperspectral image data. *Remote Sens. Environ.* **2008**, *112*, 4301-4317.
56. Tucker, C. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127−150.
57. Fuentes, D.; Gamon, J. Qiu, H.; Sims, D.; Roberts, D. Mapping Canadian boreal forest vegetation using pigment and water absorption features derived from the AVIRIS sensor. *J. Geophys. Res.* **2001***, 106*, 33565−33577.
58. Gamon, J.; Peñuelas, J.; Field, C. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* **1992**, *41*, 35−44.
59. Gao, B. NDWI — A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* **1996**, *58*, 257−266.
60. Peñuelas, J.; Piñol, J.; Ogaya, R.; Filella, I. Estimation of plant water concentration by the reflectance Water Index WI (R900/R970). *Int. J. Remote Sens.* **1997**, *18*, 2869−2875.
61. Serrano, L.; Peñuelas, J.; Ustin, S. Remote sensing of nitrogen and lignin in Mediterranean

vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote Sens. Environ.* **2002**, *81*, 355−364.

62. Nagler, P.; Daughtry, C.; Goward, S. Plant litter and soil reflectance. *Remote Sens. Environ.* **2000**, *71*, 207−215.
63. DeFries, R.; Chan, J. Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. *Remote Sens. Environ.* **2000**, *74*, 503−515.
64. Briem, G.; Benediktsson, J. Sveinsson, J. Multiple classifiers applied to multisource remote sensing data. *IEEE T. Geosci. Remote* **2002**, *40*, 2291−2299.
65. Leverington, D.; Moon, W. On the use of consensus algorithms to address variability in the results of neural network classifications: Preliminary tests involving two northern study areas. *Can. J. Remote Sens.* **2005**, *31*, 269−273.
66. Carreiras, J.; Pereira, M.; Shimabukuro, Y. Land-cover mapping in the Brazilian Amazon using SPOT-4 vegetation data and machine learning classification methods. *Photogramm. Eng. Rem. S.* **2006**, *72*, 897−910.
67. Curran, P. Remote sensing of foliar chemistry. *Remote Sens. Environ.* **1989**, *30*, 271-278.
68. Elvidge, C.; Lyon, R. Influence of rock-soil spectral variation on the assessment of green biomass. *Remote Sens. Environ.* **1985**, *17*, 265–269.
69. Huete, A.; Jackson, R.; Post, D. Spectral response of a plant canopy with different soil backgrounds. *Remote Sens. Environ.* **1985**, *17*, 37–53.
70. Gey S.; Poggi, J. Boosting and instability for regression trees, *Comput. Stat. Data An.* **2006**, *50*, 533-550.
71. Glenn, E.; Morino, K.; Didan, K.; Jordan, F.; Carroll, K.; Nagler, P.; Hultine, K.; Sheader, L.; Waugh, J. Scaling sap flux measurements of grazed and ungrazed shrub communities with fine and coarse-resolution remote sensing. *Ecohydrology* **2008**, *1*, 316-329.
72. Vogelmann, J.; Rock, B.; Moss, D. Red edge spectral measurements from sugar maple leaves. *Int. J. Remote Sens*. **1993**, *14*, 1563-1595.
73. Glenn, N.; Mundt, J.; Weber, K.; Prather, T.; Lass, L.; Pettingill, J. Hyperspectral data processing for repeat detection of small infestations of leafy spurge. *Remote Sens. Environ.* **2005**, *95*, 399−412.
74. Jensen, J.; Garcia-Quijano, M.; Hadley, B.; Im, J.; Wang, Z.; Nel, A.; Teixeira, E.; Davis, B. Remote sensing agricultural crop type for sustainable development in South Africa. *Geocarto Int.* **2006**, *21*, 5-18.
75. Underwood, E.; Ustin, S.; Ramirez, C. A comparison of spatial and spectral image resolution for mapping invasive plants in coastal California. *Environ. Manage.* **2007**, *39*, 63−83.