Remote Sensing of Environment 127 (2012) 139–152

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

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

Comparing endmember selection techniques for accurate mapping of plant species and land cover using imaging spectrometer data

Keely L. Roth [a,*], Philip E. Dennison [b], Dar A. Roberts [a]

[a] Department of Geography, University of California Santa Barbara, 1832 Ellison Hall, Santa Barbara, CA 93106, United States
[b] Department of Geography and Center for Natural and Technological Hazards, University of Utah, 260 S Central Campus Dr, Room 270, Salt Lake City, UT 84112, United States

ARTICLE INFO

Article history:
Received 25 January 2012
Received in revised form 23 August 2012
Accepted 23 August 2012
Available online xxxx

Keywords:
Multiple endmember spectral mixture analysis
AVIRIS
Imaging spectroscopy
Species classification
Endmember selection

ABSTRACT

Building representative spectral libraries and quantitatively selecting a subset of spectra for mapping plant species and land cover/land use within remotely sensed imagery remain challenging for accurate classification. Multiple Endmember Spectral Mixture Analysis (MESMA) can be used for both classification and modeling fractional composition, and has been applied to map multiple biogeophysical variables. Our major objectives in this research were to 1) test a sampling design for building independent and representative training and validation spectral libraries; 2) compare endmember selection by a combination of two established techniques (count-based selection (CoB) and endmember average root mean square error (EAR)) with a recently introduced one (iterative endmember selection (IES)); and 3) develop and test a hybrid method, which combines the strengths of the previous two methods. We applied CoB/EAR, IES, and the new hybrid technique to mapping plant species and cover types in the Santa Ynez Mountains and Santa Barbara urban area, California, USA, using Airborne Visible Infrared Imaging Spectrometer (AVIRIS) data. For all endmember selection techniques, the number of selected endmembers varied across 25 random training samples. IES was consistently more accurate than CoB/EAR, but resulted in spectral libraries more than twice as large and failed to model rare species. The hybrid endmember selection technique resulted in the highest overall accuracy and kappa values and proved to be least sensitive to the random sampling protocols, but also produced the largest spectral libraries. A modified hybrid method, in which the number of endmembers selected was limited, produced the second highest accuracies, combining the strengths of the more parsimonious endmember selection by CoB/EAR with improved endmember selection by IES. Both IES and the hybrid methods selected endmembers that successfully classified a wide range of plant species and cover types, indicating their usefulness for these applications.

1. Introduction

Spectral mixture analysis (SMA) and its variants are powerful algorithms which determine the relative proportions of pure or 'endmember' spectra within a mixed spectrum (Adams et al., 1993; Settle & Drake, 1993). These methods have been used with imagery from a range of sources with varying spatial, spectral and radiometric resolutions over an extensive number of natural and urban ecosystems (e.g. Elmore et al., 2000; Lu & Weng, 2004; Roberts et al., 1993, 2006; Small, 2001). A critical component for successfully applying SMA is selecting appropriate endmembers (Dennison & Roberts, 2003a; Somers et al., 2011; Tompkins et al., 1997).

Endmember spectra most often represent a target class or cover type believed to be present in the data to be unmixed. Potential endmember spectra can be collected from reference materials in a laboratory setting (e.g. Roberts et al., 1993), extracted from imagery (Bateson et al., 2000; Dennison & Roberts, 2003a), measured in the field (Herold et al., 2004; Roberts et al., 2004) or simulated with radiative transfer modeling (e.g. Dennison et al., 2006; Eckmann et al., 2008; Painter et al., 1998; Sonnentag et al., 2007). Incorporating endmember variability into SMA is particularly challenging, since spectra belonging to a single class can vary through space and time. Spectral variability is present both within endmember classes and between classes. Selecting endmember sets which adequately represent intra- and inter-class variability is important for accurately mapping sub-pixel cover fractions and for identifying components of a scene. Several variations of SMA have been developed that incorporate or suppress endmember variability in the modeling process (e.g., MESMA, AutoMCU, Stable Zone Unmixing, Sparse Unmixing, Bayesian Spectral Unmixing, and others (see Somers et al., 2011). Each of these techniques assumes an endmember library that captures endmember variability—the spectral variability of endmember classes.

While capturing endmember variability is important, it is also desirable to select the smallest optimal subset of available endmember spectra for many applications of SMA. Using a smaller, representative

* Corresponding author. Tel.: +1 805 893 4434; fax: +1 805 893 3146.
E-mail address: klroth@geog.ucsb.edu (K.L. Roth).

0034-4257/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.rse.2012.08.030

set of endmembers can eliminate redundancy and increase computational efficiency in modeling. Fewer endmembers used in modeling leads to more straightforward interpretation of model results and reduces the extent of combinatorial iterations when increasing model complexity (i.e., more than 2 endmembers in a pixel). Moreover, endmember selection can help identify relevant endmembers for areas that lack extensive reference data.

Multiple techniques have been developed to select a reduced set of endmembers that also capture endmember variability. Selection of endmember spectra collected from one source or multiple sources can use extremes in the data (Boardman et al., 1995), constraints based on field or photographic estimates of sub-pixel fractions (Roberts et al., 1997), minimization of modeling error through application to a spectral library (Dennison and Roberts, 2003a), minimization of modeling error through creation of virtual endmembers (Tompkins et al., 1997), or optimization of modeling accuracy (Schaaf et al., 2011).

While endmember extraction techniques (locating spectra which represent pure endmember classes) have been extensively compared in the literature (Martinez et al., 2006; Plaza et al., 2004; Veganzones & Grana, 2008), few studies have compared different endmember selection techniques (selecting a subset of endmember spectra for use in analysis) to determine their relative strengths and weaknesses. Quantitative comparisons between endmember selection techniques can provide a clearer picture of when certain techniques are more appropriate and can stimulate the design of improved selection methods. Our objective was to quantitatively compare endmember selection techniques originally developed for Multiple Endmember Spectral Mixture Analysis (MESMA) and to test a newly developed hybrid method. However, it is important to realize these techniques could be applied to many other SMA algorithms that seek to incorporate endmember variability. Additionally, the creation of independent, robust training and validation libraries has been largely ignored in previous studies. Here, we sought to address these issues by answering the following research questions:

1. What are the differences in the number and identity of endmembers selected using different selection techniques?
2. How do the selected endmember libraries impact classification performance as evaluated using an independent data set?
3. Can hybridizing endmember selection techniques improve classification performance?
4. How sensitive is each selection technique to a stratified-random sampling design?

2. Background

2.1. Spectral mixture analysis

SMA decomposes measured reflectance or radiance from several materials within the instantaneous field of view of an instrument (pixel) as a mixture of a fixed set of endmembers (Adams et al., 1993). Mixing is often assumed to be linear, allowing a pixel spectrum to be modeled as the sum of endmembers multiplied by their fractional contribution to a best-fit mixed spectrum. Linear SMA uses the equation:

$$\rho'_{\lambda} = \sum_{i=1}^{N} f_i * \rho_{i\lambda} + \varepsilon_{\lambda} \quad (1)$$

where ρ'_{λ} is reflectance at wavelength λ, $\rho_{i\lambda}$ is the spectral reflectance of endmember i, and f_i is the fractional cover of endmember i. N is the total number of endmembers in the model, and ε_{λ} is the model residual error. Model fit is assessed using root mean square error (RMSE):

$$RMSE = \left(\frac{\sum_{b=1}^{M} (\varepsilon_{\lambda})^2}{M} \right)^{1/2} \quad (2)$$

where b is band number and M is the total number of bands.

2.2. Incorporating endmember variability

A major advantage of using SMA with spectroscopic data is that it utilizes the entire spectral response rather than a small number of bands and allows for variations in both composition and illumination in an image (Dennison & Roberts, 2003a). SMA also provides better estimates of fractional cover than many commonly used vegetation indices (Elmore et al., 2000; Riano et al., 2002). However, SMA has some shortcomings, including the assumption that spectral mixing within a pixel is linear. Several studies have determined that multiple scattering can cause non-linear mixing, which may be significant (Borel & Gerstl, 1994; Huete, 1986; Ray & Murray, 1996; Roberts et al., 1993). Additionally, simple SMA, which uses the same set of endmembers for an entire image, cannot account for spectral variability introduced by multiple types of vegetation or soils, and cannot account for within-class spectral variability (Roberts et al., 1998; Somers et al., 2011).

Numerous variants of SMA have been developed to address the issue of endmember variability. Multiple Endmember Spectral Mixture Analysis (MESMA) was developed by Roberts et al. (1998) to allow the number and type of endmembers to vary on a per-pixel basis. Other variants of SMA designed to address endmember variability include fuzzy un-mixing techniques such as Automated Monte Carlo Unmixing (Asner & Lobell, 2000), endmember bundles (Bateson & Curtiss, 1996), and Bayesian Spectral Mixture Analysis (BSMA) (Song, 2005). Somers et al. (2011) provide an extensive review of these techniques, along with others which incorporate spectral feature selection (e.g., PCA-based SMA, Stable Zone Unmixing), spectral weighting, spectral transformations (e.g., normalized SMA, Derivative Spectral Unmixing) and spectral modeling (e.g., MODTRAN-generated endmembers, Soil Modeling Mixture Analysis). While some of these approaches have been successfully applied over a range of imagery and ecosystems, MESMA remains the most widely used SMA technique for dealing with endmember variability (Somers et al., 2011).

2.3. Multiple Endmember Spectral Mixture Analysis

In practice, MESMA iteratively computes linear models using different sets of endmembers, and the model with the lowest root mean square error (RMSE) is selected for each individual pixel. Two endmember MESMA (one shade endmember and one non-shade endmember) accounts for variability in reflectance, and can be used as a classification algorithm where the non-shade endmembers are assigned to specific classes (e.g. Dennison & Roberts, 2003a, 2003b). MESMA using three or more endmembers can be used to estimate fractional cover of the non-shade endmembers (e.g. Franke et al., 2009).

The flexibility of MESMA has permitted its application to mapping a wide variety of vegetated and non-vegetated land cover. In natural ecosystems, MESMA has been used to discriminate natural land cover classes, separate soils from vegetation, assess fire fuel distribution, and map snow grain size (Jia et al., 2006; Linn et al., 2010; Okin et al., 2001; Painter et al., 1998, 2003). Urban land cover has also been mapped with MESMA. Examples include applications in Brazil

(Powell & Roberts, 2010; Powell et al., 2007), Los Angeles County (Rashed et al., 2003), Phoenix, Arizona (Myint & Okin, 2009) and Bonn, Germany (Franke et al., 2009).

The two-endmember case of MESMA has also been successfully used to map plant species and functional types in various ecosystems. In the chaparral shrublands of California, species-level maps have been created with accuracies up to 89% (Dennison & Roberts, 2003a, 2003b). Roberts et al. (1999) used MESMA to map major species in the Canadian boreal forest, and Plourde et al. (2007) quantified tree species abundance in a temperate deciduous forest. In the Wasatch Mountain Range of Utah, Schaaf et al. (2011) mapped four classes of plant functional type at several spatial resolutions with 78–88% overall accuracy. In Australian eucalyptus forest, Youngentob et al. (2011) were able to discriminate two Eucalyptus subgenera using HyMap data with 83% accuracy. In salt marshes and wetlands, MESMA has been used to map native and invasive species (Li et al., 2005; Rosso et al., 2005; Underwood et al., 2006). The results of these studies demonstrate that MESMA is capable of mapping both the composition of complex urban environments, from simple pervious vs. impervious classifications to detailed maps of building materials and tree species, and vegetation patterns within natural ecosystems to the species level.

2.4. Endmember selection methods

Multiple approaches for selecting representative endmembers from an available endmember library have been developed for MESMA. These approaches balance the selection of a reduced set of endmember spectra while still capturing endmember variability. Roberts et al. (2003) proposed a count-based (CoB) approach to endmember selection. For a given class, CoB selects the endmembers that modeled the largest number of spectra within each class. This method can be altered to account for both number of spectra modeled within a class (inCoB) and number of spectra modeled that belong to other classes (outCoB). Whether or not a spectrum is considered to be 'successfully modeled' by an endmember is determined by an RMSE fit threshold. Dennison and Roberts (2003a) developed Endmember Average RMSE (EAR) to select endmembers for MESMA. EAR uses each spectrum within a class to model all other spectra in the class using linear SMA. EAR is calculated as:

$$EAR_{A_i} = \frac{\sum_{j=1}^{n} RMSE_{A_i A_j}}{n-1} \quad (3)$$

where A is a single class, A_i is the spectrum being tested as an endmember, and A_j is a modeled spectrum. n is the total number of spectra in class A, and the term "$n-1$" accounts for the spectrum modeling itself. EAR values for each spectrum in the class are compared, and the spectrum with the minimum EAR value is then selected as an endmember for that class. Dennison et al. (2004) introduced Minimum Average Spectral Angle (MASA), which functions similarly to EAR but uses spectral angle (Kruse et al., 1993) in place of RMSE. CoB, EAR, and MASA are relatively simple approaches to selecting endmembers for MESMA, but they only account for intra-class spectral variability and do not account for inter-class spectral variability or the resulting interactions between selected endmembers. These techniques also require some degree of subjectivity in selecting the single "best" set of endmembers for MESMA (Schaaf et al., 2011).

Schaaf et al. (2011) proposed a new method for selecting endmembers for MESMA that does take into account interactions between endmembers across multiple classes. Iterative Endmember Selection (IES) uses an iterative process to select representative endmembers based on their ability to accurately classify an entire spectral library, rather than individual classes. The training spectral library is composed of all potential endmembers from each class. This is particularly well-suited for cases where MESMA is applied as a classifier, since high classification accuracy is a primary goal. IES uses two-endmember linear SMA to classify the spectral library, similar to EAR, MASA, and CoB. An RMSE threshold is used to determine which endmember models which spectra within the training library. Accuracy for classifying the library is based on Cohen's kappa across all classes, which accounts for agreement by chance (Cohen, 1960) and is superior to overall accuracy when class representation is not uniform within a dataset (Congalton, 1991).

IES begins with a comparison of all possible pairs of endmembers to find the two endmembers that result in the highest kappa value for classifying the training library. All spectra remaining in the library are tested to determine which third endmember, in combination with the previously selected two, increases kappa the most. This endmember is then added to the selected set. Once an endmember has been added to the selected set, all endmembers within the set are subtracted from the set to determine whether kappa can be increased by removing an endmember. Addition to and subtraction from the selected set continues iteratively as long as kappa increases. The final endmember library is determined when kappa no longer increases, although the endmember addition process can be arbitrarily ended once a specified number of endmembers has been reached. The selected set of endmembers is optimized for the training library, and an independent accuracy assessment library should be used to confirm that the selected set of endmembers can produce an accurate classification. Unlike EAR, MASA, and CoB, IES is entirely automated. However, in the process of maximizing kappa, IES can produce very large endmember libraries. These methods of endmember selection have not been compared, so potential tradeoffs, including endmember library size and classification accuracy, are unknown. Furthermore, it may be beneficial to hybridize methods in order to retain their individual advantages while reducing disadvantages.

3. Data and methods

3.1. Study site

Our study region is located in the Santa Ynez Mountain Range and coastal area of Santa Barbara County, California, USA (approximately 34.5° N 119.5° W) (Fig. 1). The region includes the urban areas of Santa Barbara and Goleta, agricultural lands (mainly orchards) and natural vegetation communities including chaparral, coastal sage scrub, oak woodland, and grassland. Elevation ranges from sea level up to approximately 1400 m on several peaks. The climate is Mediterranean-type, characterized by warm, dry summers and cool, moist winters, with the first rains arriving in October or November. Average annual rainfall measured at downtown Santa Barbara is 472 mm, and mean monthly temperature varies between 11 and 21 C throughout the year. As is typical of vegetation in Mediterranean-type climates, plant biodiversity is high (Cowling et al., 1996), with several community types present. Grasslands are of mixed species composition, dominated by annual non-native grasses and forbs. Oak woodlands favor broadleaf species, *Quercus agrifolia* (California coast live oak) being dominant. *Umbellularia californica* (California bay laurel), *Platanus racemosa* (sycamore), and other species (e.g., *Salix* spp.) dominate riparian zones and canyon drainages. Coastal sage scrub is composed mainly of *Artemisia californica* (California sagebrush) and *Salvia* spp. (sage). Major chaparral species include *Adenostoma fasciculatum* (chamise), *Arctostaphylos glauca* and *glandulosa* (mazanita), several *Ceanothus* species, and *Quercus dumosa* (scrub oak).

3.2. Image and reference data

3.2.1. Image data

Airborne Visible Infrared Imaging Spectrometer (AVIRIS) data were acquired over the study area on August 6, 2004 at approximately 19:08

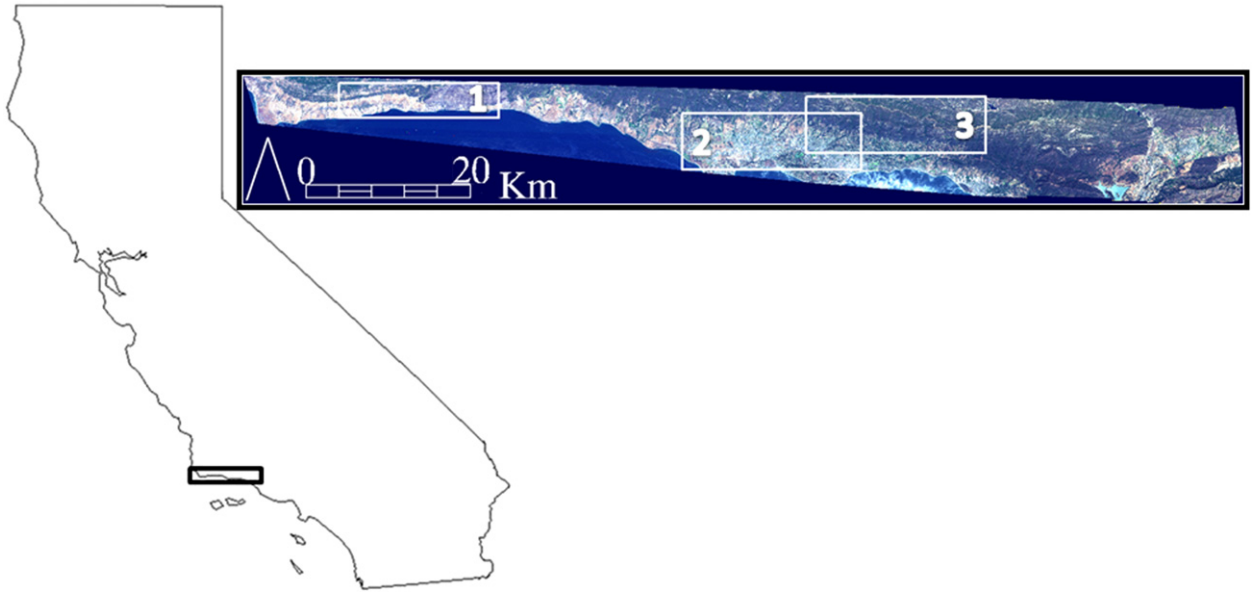

Fig. 1. The study region (outlined in black) and true color composite image of the AVIRIS data acquired. The boxes on the image denote the three subregions discussed in the results section.

UTC (solar zenith 22.1°; solar azimuth 141.2°) (Fig. 1). The AVIRIS sensor collects spectral radiance data over 224 bands from 370 to 2500 nm with 10 nm sampling (Green et al., 1998). The sensor was flown aboard the ER-2 platform at approximately 16 km altitude, yielding an image swath of ~11 km with a ground instantaneous field of view of 16 m. Preliminary data processing and orthorectification was done by the NASA Jet Propulsion Laboratory (Boardman, 1999). The image was corrected to reflectance using a modified version of the MODTRAN radiative transfer model (Berk et al., 2003; Green et al., 1993; Roberts et al., 1997), and a reference spectrum from a ground target (sand) was used to further correct image reflectance following Clark et al. (2002). The image was orthorectified using a 1 m Digital Orthophoto Quarter-Quadrangle and triangulation with nearest neighbor resampling. Bands with strong water vapor absorption and poor signal-to-noise were removed, leaving a subset of 186 bands.

3.2.2. Reference data

Reference data on the spatial distribution of dominant species and land cover types were collected both in the field and using aerial imagery. During two field campaigns (2003 and 2009), we used a composition estimation method adapted from Meentemeyer and Moody (2000), where patches of dominant plant species and their relative composition were collected using a high-power spotting scope from remote vantage points. Patches having greater than 75% single species composition were recorded and digitized on the AVIRIS flight line. Subsequently, a 1 m Digital Orthophoto Quarter-Quadrangle was used to digitize urban, agricultural, golf course and orchard cover polygons. Dominant classes and their abbreviations are listed in Table 1.

3.3. Spectral library development

All reference polygons were merged into a single geospatial layer, and hierarchical metadata for each polygon, including perviousness, cover type, functional type, and species, were recorded. Broad classes were used to assign functional type based on lifeform (e.g., tree, shrub, etc.), leaf duration (i.e., evergreen vs. deciduous) and leaf type (i.e., broadleaf vs. needleleaf). Spectral library creation and processing was all done using code developed with the ENVI/IDL (ITT Visual Information Solutions) image processing software package. The spectra within all reference polygons were extracted to a master spectral library linked to metadata for each spectrum (Roberts et al., 2007).

The spectral library was partitioned into training and validation pairs using random sampling within each class. With the goal of including spectra from each polygon in the training library, spectra were randomly selected from the polygons until one of two pre-defined sampling limits was met for all polygons within the class. A proportion limit ensured that an adequate number of spectra were reserved for validation from each polygon (a concern for small polygons). A sampling absolute limit set the maximum number of spectra selected from a polygon to offset the impact of polygon size on the training sample (a concern for large polygons). Multiple proportional and absolute limits were tested, and a proportional limit of 50% and an absolute limit of 10 spectra were empirically found to produce the most balanced training libraries while still preserving an acceptable number of spectra for validation. Fig. 2 compares training library class proportion with and without the absolute limit. While the sampling procedure did not create completely balanced class representation, the dominance of the largest classes with in the library was reduced by adding the absolute limit (Fig. 2).

The training/validation partitioning process was iterated 25 times, creating 25 unique training/validation library pairs to determine the sensitivity of the endmember selection results to the random selection of training data. We chose 25 iterations to test the sampling design as a conservative upper end estimate and this was supported by our findings (see Results). In each iteration, different spectra may have been selected for the training sample, but the total number

Table 1
Dominant classes mapped and corresponding abbreviations used.

Dominant class	Abbreviation	Dominant class	Abbreviation
Adenostoma fasiculatum	ADFA	Young orchard	ORCHARD
Mixed agriculture	AG	Mature orchard	ORCHARD_LS
Artemisia californica	ARCA	*Platanus racemosa*	PLRA
Arctostaphylos glauca/glandulosa	ARGL	*Quercus agrifolia*	QUAG
Burn scar	BURN	*Quercus dumosa*	QUDU
Ceanothus cuneatus	CECU	Riparian	RIPARIAN
Ceanothus megacarpus	CEME	Rock	ROCK
Ceanothus spinosus	CESP	Soil	SOIL
Senesced *Brassica nigra*	dBRNI	*Umbellularia californica*	UMCA
Senesced annual grasses	dGRASS	Urban	URBAN
Golf course	GC	*Yucca whipplei*	YUWH

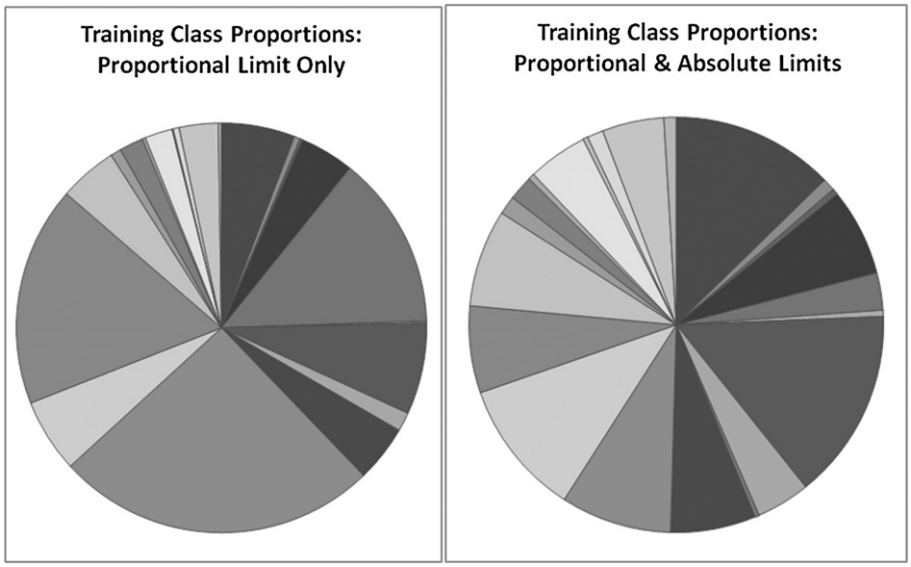

Fig. 2. Training data class proportions with 50% limit (left) and 50% or 10 spectra limit (right). Each "wedge" represents a different class (see Table 1). A greater balance in the number of spectra within each class is illustrated by more equality in the size of "wedges" in the pie diagram on the right.

and class distribution of these spectra remained the same based on the sampling limits.

3.4. Endmember selection

3.4.1. CoB/EAR selection

Endmember selection based on EAR, MASA, and CoB are aimed at selecting endmembers which best represent individual classes. In our work, we selected endmembers using CoB, with the minimum EAR value used as a tiebreaker as in Roberts et al. (2012). The within-class CoB value of an endmember is the number of spectra uniquely modeled by that endmember, as defined by an RMSE constraint. Within a class, the endmember that modeled the largest number of class spectra was assigned the highest CoB value. The endmember with the next highest CoB value must then model the largest unique set of spectra not modeled by the highest CoB endmember. For each individual class, all endmembers with non-zero CoB values (indicating that they modeled unique spectra from the training library that were not modeled by other spectra) were selected. When ties between more than one endmember occurred, indicating that endmembers modeled the same number of spectra (but not necessarily the same spectra), they were broken by selecting the endmember with the minimum EAR value.

3.4.2. Iterative Endmember Selection (IES)

IES aims to select endmembers which maximize the accuracy of a two-endmember MESMA training library classification (Schaaf et al., 2011). In the IES procedure, spectra are iteratively added and subtracted from the endmember set until the highest kappa value is achieved. Maximizing overall kappa may reduce accuracy in individual classes. For example, no endmembers would be selected for a class if the addition of a class endmember results in lower overall accuracy. IES tends to select large endmember libraries as kappa is incrementally improved. Because greater accuracies may be achieved due solely to the larger number of endmembers selected by IES, we also built a set of endmember libraries using IES in which the total number of endmembers was limited to match the number of spectra in the CoB/EAR library for the same iteration. These libraries are denoted as 'IES_nlim'.

3.4.3. Hybrid IES–CoB/EAR selection

To see if the strengths of CoB/EAR and IES could be combined to create improved endmember libraries, we also designed a hybrid selection process. For iterations where endmembers were not selected for some classes using IES, IES was forced to include endmembers selected by CoB/EAR for those classes. For each iteration, these endmembers were injected into the IES process after the first 10 endmembers had been selected. This was done to allow the IES algorithm to begin the kappa-optimization process and then adapt the best initial set of endmembers selected to include the forced endmembers. A set of size-limited (based on the number of endmembers in each CoB/EAR library) libraries were also created with this hybrid approach, denoted 'IES–CoB/EAR_nlim'.

3.5. Multiple Endmember Spectral Mixture Analysis and statistics

For each of the 25 iterations, the endmember library from each selection technique was used in two-endmember MESMA to model the validation library. Kappa, overall accuracy, user's and producer's accuracy, and an error matrix were calculated for each MESMA run. Distributions (across iterations) of kappa, overall accuracy, and user's and producer's accuracy (by class) were tested for normality using a Lilliefors test (Lilliefors, 1969). The null hypothesis of the Lilliefors test states that the data are normally distributed. This hypothesis was not rejected for the kappa and overall accuracy values, and thus a one-way repeated measures ANOVA (Girden, 1992) was used to test for significant differences across endmember selection techniques in kappa and overall accuracy. Because the producer's and user's accuracies for several classes failed the Lilliefors test for normality, we chose to use Friedman's test (Friedman, 1937). The Friendman's test does not assume a normal distribution and uses ranks to analyze differences across endmember selection techniques for the producer's and user's accuracies. Both the ANOVA and Friedman's tests only indicate if significant variance in the variable of interest (i.e., accuracy metric) occurs among endmember selection techniques, but do not indicate which pairs of methods are significantly different from each other. Thus, we applied the Tukey-Kramer method (Hochberg & Tamhane, 1987) to all pairs of selection techniques to determine which were significantly different for each metric.

To visualize the differences between endmember selection techniques, the set of endmembers from the iteration with the highest overall accuracy was applied to the AVIRIS data to produce classification maps. Endmember fractions were restricted to fall between −0.05 and 1.05, and the maximum allowable RMSE was set at 2.5% (Dennison & Roberts, 2003a). Three sub-regions of the flight line

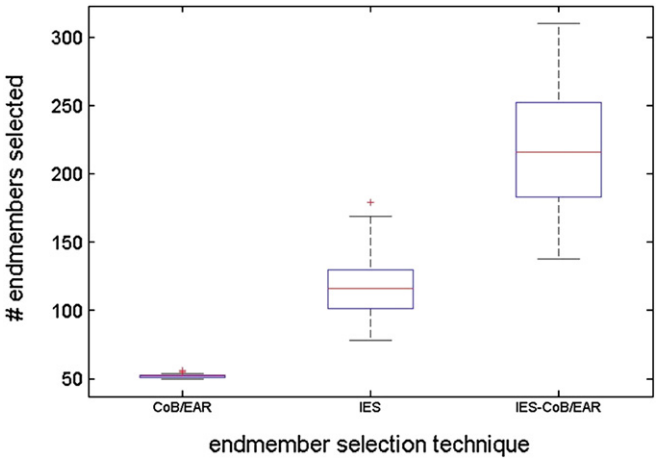

Fig. 3. The distribution of the # of endmembers selected for each method. The horizontal central line denotes the median number of endmembers, the box edges mark the 25th and 75th percentiles, and the whiskers extend to the most extreme values not considered outliers.

were selected for comparing the mapping performance and agreement among endmember selection techniques (Fig. 1).

4. Results

4.1. Training and validation libraries with stratified random sampling

Two factors controlled the number of spectra from each class within the training spectral libraries: the number of polygons per class and the size of the polygons in a class. Training sample sizes ranged from 7 to 309 spectra, representing from 2.4 to 50% of the total number of spectra within each class. Classes with only a few, small polygons contributed roughly 50% of their spectra to the training library. Examples include dBRNI, PLRA and YUWH. By contrast, classes with only a few, large polygons (i.e., many spectra available for selection) had only a small percentage of their potential spectra chosen for the training library due to the absolute limit of 10 spectra per polygon. These classes included BURN, GC and ORCHARD. Intermediate class sampling percentages (~20 to 30%) resulted for classes with intermediate sized polygons (average between 30 and 40 pixels). Classes with many, intermediate sized polygons made up the highest percentage of the final training library. Examples include ADFA, CEME, Orchard_LS and QUAG.

The validation libraries consisted of all spectra not selected as training spectra for each iteration, and thus, no attempt was made to balance the number of spectra across classes. This is representative of most real world mapping scenarios in which species and cover types are not found in equal abundance on the landscape. The number of spectra in the validation libraries ranged from 8 (senesced BRNI and PLRA) to 4369 (GC) (50 to 97.6% of total class spectra). The classes with the largest and/or most polygons made up the largest proportion of the validation library.

4.2. Size and composition of endmember libraries

The size of endmember libraries varied across selection techniques and iterations (Fig. 3). CoB/EAR yielded the smallest library sizes, around 52 endmembers for each iteration, and library size did not vary much across sampling iterations (standard deviation = 2). The endmember libraries selected via IES were, on average, over twice as large (mean = 117) and library size varied by 20% across iterations. Libraries selected using the hybrid technique (IES–CoB/EAR) were even larger (mean = 218) and more variable in size. The standard deviation of the hybrid selection technique across sampling iterations was nearly equal to the mean size of the CoB/EAR-selected libraries (51).

The composition and number of endmembers per class varied across techniques and sampling iterations (Table 2). Using the CoB/EAR technique, endmembers were selected for all 22 classes from each training sample. The mean number of endmembers per class ranged between 1 (e.g., AG, CECU, ROCK) and 5 (e.g., GC, ORCHARD_LS) with very low standard deviation values among sampling iterations (<1). Because CoB/EAR selection is designed to select the most representative endmembers within a class, the number of endmembers chosen for partially indicates the level of spectral variability in that class. Golf course, young orchard, soil and urban classes had the largest number of endmembers. While training sample size may partially explain the higher number of selected endmembers, these classes could be

Table 2
Summary statistics for the number of endmembers selected by each method. Numbers in parentheses indicate the number of iterations in which at least one endmember for the class was selected (if not 25).

Dominant class	CoB/EAR		IES		IES–CoB/EAR		IES_nlim		IES–CoB/EAR_nlim	
	mean	std	mean	std	mean	std	mean	std	mean	std
Overall	52	2	117	24	218	51	52	2	52	2
ADFA	3	0.2	11	5.4	20	6.5	4	1.4	4	1.0
AG	1	0.0	1 (9)	0.3	2 (24)	1.0	1 (7)	0.0	1 (19)	0.3
ARCA	1	0.0	1	0.4	2 (24)	0.9	1 (11)	0.3	1 (5)	0.0
ARGL	3	0.4	3 (23)	1.8	15	6.2	2 (19)	0.8	3	0.9
BURN	2	0.5	3	0.6	3	0.6	2	0.6	2	0.4
CECU	1	0.0	0	0.0	2	0.7	0	0.0	1	0.0
CEME	3	0.4	6	2.9	19	9.3	3	1.0	4	1.3
CESP	3	0.5	5 (24)	2.1	13	6.0	2 (21)	0.6	2	0.9
dBRNI	1	0.0	1 (8)	0.5	2 (23)	0.8	1 (1)	0.0	1 (17)	0.0
dGRASS	3	0.5	8	2.4	12	3.0	4	1.0	3	1.1
GC	5	0.8	14	4.4	26	5.2	7	2.2	7	0.9
ORCHARD	3	0.6	12	3.8	18	5.9	5	0.9	4	0.9
ORCHARD_LS	5	0.7	10	3.2	15	4.0	5	0.9	5	0.9
QUAG	3	0.4	9	4.1	26	9.6	3	1.1	5	1.3
QUDU	2	0.6	4 (23)	1.6	6	2.5	2 (22)	0.7	1 (22)	0.4
RIPARIAN	1	0.3	4 (24)	1.6	5	2.4	2 (23)	0.6	1 (24)	0.3
ROCK	1	0.0	1	0.2	1	0.0	1	0.0	1	0.0
SOIL	4	0.6	10	2.1	11	2.5	5	1.4	4	0.8
PLRA	1	0.5	1 (12)	0.4	2 (23)	1.1	1 (6)	0.0	1 (14)	0.5
UMCA	1	0.2	3 (23)	1.3	5	2.2	1 (13)	0.5	1 (12)	0.3
URBAN	4	0.8	8	1.9	8	2.0	5	1.5	3	0.8
YUWH	2	0.5	4	2.5	4 (24)	2.3	2 (21)	0.8	1 (9)	0.3

expected to have higher spectral variability as well. For example, golf courses (GC) represent mixed turfgrass in a range of conditions from fully senesced to highly irrigated and fertilized, and the mature orchard (ORCHARD_LS) class includes several different tree species and large variations in percent canopy cover. Classes with fewer endmember spectra selected were those that are fairly homogenous (e.g., ROCK) or those represented by a low number of small reference polygons. Examples include UMCA, ARCA, PLRA and CECU. This may indicate a larger reference sample is needed to adequately characterize the variability of these classes. Overall, the standard deviations of total library size and the number of endmembers per class across sampling iterations for CoB/EARselection were low, suggesting random training samples for each class were representative.

Endmember libraries selected with IES contained a higher number of endmembers per class (Table 2). The mean number of endmembers per class (averaged over the number of libraries in which the class was included) ranged from 1 (e.g., AG, dBRNI, ROCK) to 14 (GC). The standard deviation of endmembers selected per class across sampling iterations ranged from 1 to more than 5 (ADFA). Classes with a high average of endmembers (10 +) included GC, ORCHARD_LS, ADFA, ORCHARD, and SOIL. Perhaps most importantly, the IES technique did not select endmembers for all classes in every sampling iteration. No endmembers were ever selected for CECU, and endmembers for AG, dBRNI and PLRA were selected in less than half the iterations. Other classes which were occasionally left out included ARGL, CESP, QUDU, RIPARIAN and UMCA. The reason IES does not select endmembers for all classes is due to spectral confusion. In trying to maximize kappa (cross-class discrimination), the algorithm will not select an endmember that reduces kappa (by increasing confusion among classes).

In order to address the missing classes in the IES libraries, we created and applied the IES–CoB/EAR hybrid method, which allows the IES algorithm to select the best set of 10 endmembers, and then injects the endmembers selected by CoB/EAR for the missing classes. The IES algorithm then continues to select endmembers to maximize kappa with all classes present in the endmember library. Because IES alone does not select endmembers for frequently confused classes, the hybrid technique results in more endmembers selected overall to correct for this confusion. A marked increase (>100%) in the average number of class endmembers selected when using the hybrid method was observed for QUAG, CEME, CESP and ARGL, indicating that these classes may be frequently confused with the missing classes (Table 2). The average number of endmembers selected for spectrally unique classes or classes with low spectral variability, such as BURN and ROCK, were similar to the numbers from CoB/EAR and IES. Notably, several classes were missing from the hybrid-selected libraries for a small number of sampling iterations (Table 2). These classes were likely left out of the selection process because of increased confusion with the forced endmember classes, which decreases kappa values. This demonstrates that although the hybrid method does better at including all classes than IES alone, cases can still arise in which no endmembers are selected for a class.

When IES and IES–CoB/EAR were limited to the same total number of endmembers as selected via CoB/EAR, similar patterns in the number of endmembers per class were observed (Table 2). Similar to IES, no CECU endmembers were selected by IES_nlim. The number of sampling iterations in which endmembers were selected decreased (compared to IES) for 10 classes, including some classes for which IES had always selected endmembers (i.e., ARCA and YUWH). Limiting the total number of endmembers also decreased the range of the class means and standard deviations for the hybrid method (IES–CoB/EAR_nlim). These results demonstrate that limiting the total number of endmembers to match that of libraries selected via CoB/EAR strongly impacts the composition of the library selected by IES or IES–CoB/EAR. In the case of IES, endmembers for the missing classes may have been selected if the process was allowed to continue. For IES–CoB/EAR, the classes with no endmembers selected were likely those most frequently confused with other classes, and because of the limit, the algorithm could not select enough endmembers to correct for this confusion.

4.3. Two-endmember MESMA classifications

4.3.1. Overall class separation and accuracy

The distributions of kappa values and overall accuracy across sampling iterations for each method are presented in the Figs. 4 and 5. CoB/EAR-selected libraries performed the worst of all the selection techniques with an average kappa value of 0.57 and overall accuracy of 62.4%. IES-selected libraries performed better, on average, with an average kappa of 0.69 and overall accuracy of 72.5%. The hybrid selection method yielded libraries that performed the best (average kappa = 0.77 and average overall accuracy = 79.9%). Average kappa and overall accuracy values dropped for libraries selected using sizelimited IES (0.66; 69.9%) or IES–CoB/EAR (0.71; 75.3%) . Still, both 'n_lim' techniques outperformed CoB/EAR. According to the one-way, repeated measures ANOVA, both overall accuracy and kappa values differed significantly across all five techniques ($\alpha = 0.05$), and the Tukey–Kramer comparison confirmed these differences were significant between each pair of techniques ($\alpha = 0.05$).

The variation in kappa and overall accuracy across sampling iterations illustrated a technique's sensitivity to the random sampling design. The most variable were IES and IES_nlim, followed by CoB/EAR. Both the full and size-limited hybrid selection methods showed the lowest variation across sampling iterations. These results indicate the hybrid methods are not very sensitive to random sampling, whereas the IES methods are fairly sensitive to sampling.

To test whether higher overall accuracy is a product of spectral library size, we plotted overall accuracy against library size (Fig. 6). This shows overall accuracy and library size are only weakly correlated for each selection technique. For example, for the CoB/EAR and size-limited libraries, significant differences in overall accuracy exist despite the same number of endmembers in each sampling iteration. Accuracy does increase with increasing endmembers for IES and IES–CoB/EAR, but the relationship is not linear. Accuracies for IES are widely distributed across differences in library size. For the hybrid method, accuracies do not vary greatly despite the wide range in library size.

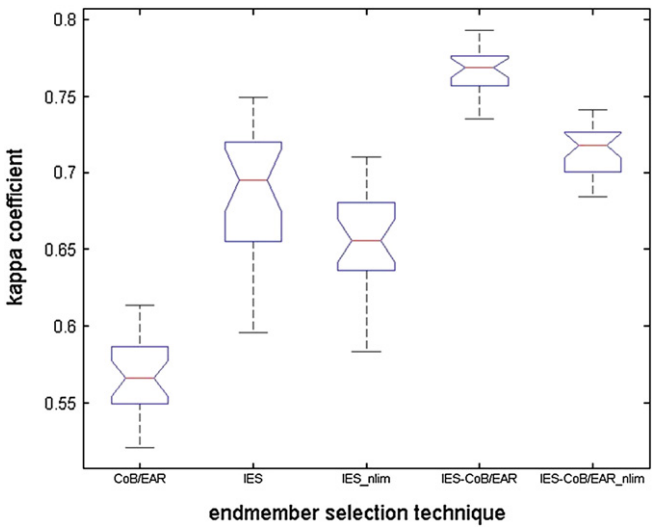

Fig. 4. The distributions of the validation kappa coefficients for each selection technique. The central horizontal line denotes the median value, the box edges mark the 25th and 75th percentiles, and the whiskers extend to the most extreme values not considered outliers. The notches signify 95% confidence intervals on the median.

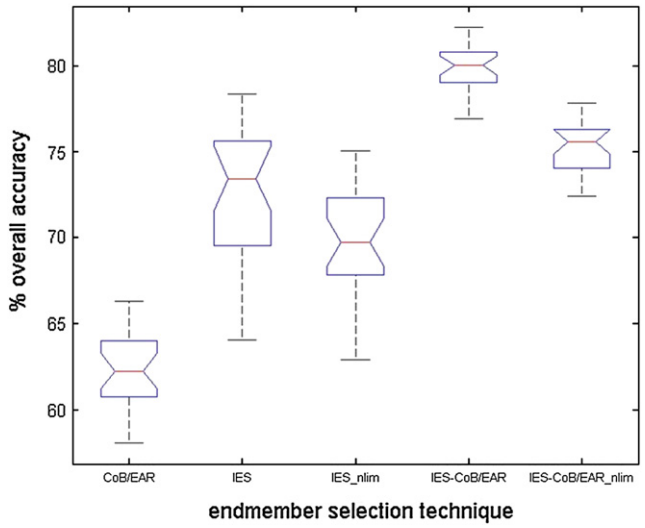

Fig. 5. The distributions of the overall accuracy for each selection technique. The central horizontal line denotes the median value, the box edges mark the 25th and 75th percentiles, and the whiskers extend to the most extreme values not considered outliers. The notches signify 95% confidence intervals on the median.

4.3.2. Class accuracies

We compared class producer's and user's accuracies across sampling iterations for each endmember selection technique. Figs. 7 and 8 summarize the mean producer's and user's accuracies over 25 sampling iterations for each endmember selection technique. Friedman's Test verified that significant differences in producer's accuracy existed among the five endmember selection methods for all dominant classes ($\alpha = 0.05$) except ROCK. Significant differences in the mean user's accuracies also existed among the methods ($\alpha = 0.05$) for all classes except BURN, which had a mean user's accuracy of nearly 100% across all methods.

All five techniques performed well in selecting endmembers to classify BURN, dGRASS, ORCHARD_LS, ROCK, SOIL and URBAN (mean producer's and user's accuracies $\geq 70\%$). Classes for which all selection techniques performed poorly included CESP, CECU, PLRA, UMCA and YUWH (mean producer's and user's accuracies below 50%). A combination of between-class confusion and small sample size was the likely cause for the low producer's accuracies for these classes. For AG and dBRNI, mean producer's accuracies were high when endmembers were selected with CoB/EAR, hybrid or size-limited hybrid techniques, but mean user's accuracies were very low for all techniques.

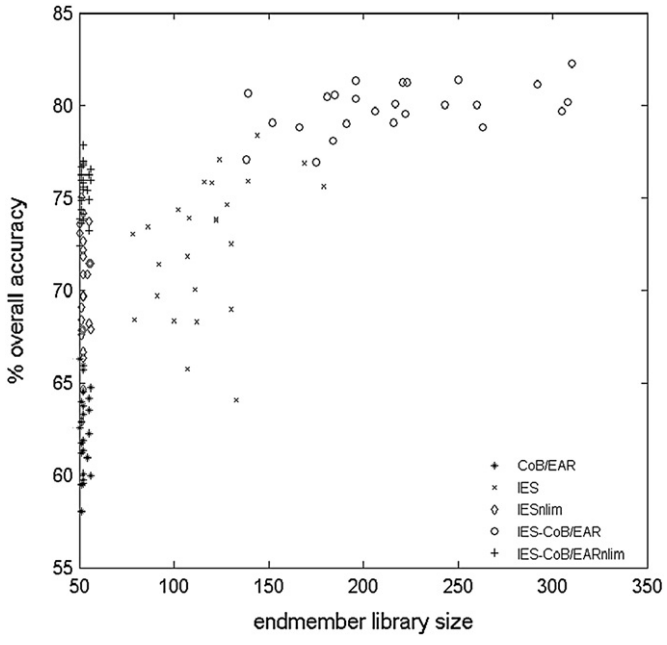

Fig. 6. Endmember library size vs. overall accuracy for all sampling iterations and all five selection methods.

CoB/EAR-selected endmembers yielded the highest producer's accuracies for 4 classes (AG, CECU, dBRNI, ROCK), as did IES-selected endmembers (ADFA, dGRASS, SOIL, YUWH). Endmembers selected via the IES–CoB/EAR hybrid technique had the highest mean producer's accuracy for the remaining 14 classes. CoB/EAR stood out among selection methods for high mean producer's accuracy of dBRNI (82%), but in general had lower mean user's accuracies than other selection methods across classes. IES had noticeably higher mean user's accuracy for ARCA. The hybrid method yielded markedly higher mean producer's accuracies for GC, QUDU, CESP, and UMCA, though these values were low overall with the exception of GC (81%). For mean user's accuracies, the hybrid method clearly outperformed all other methods for AG, CECU, dBRNI and QUDU, though again, these values were low overall with the exception of QUDU (64%).

When both IES-selected and IES–CoB/EAR-selected library size was limited, the mean producer's accuracies dropped for some classes and remained stable for others. For IES, the limited libraries showed a significant decrease in mean producer's accuracy for ARCA (−30.7%), BURN (−2%) and SOIL (−5.1%). The impact was more severe between the hybrid and limited hybrid classifications, with mean producer's accuracies decreasing by 3.1% (BURN) to 49% (ARCA) for 10 classes. The mean user's accuracies between the IES libraries and the size limited versions decreased significantly for ARCA and UMCA, and as with the producer's accuracies, limiting the endmember library size for the hybrid method resulted in significant decreases for 11 classes. The largest decrease was for ARCA (65.9%) and the smallest for ORCHARD_LS (2.7%).

4.3.3. Image classification

For each endmember selection technique, the set of endmembers with the highest overall accuracy was used to classify the 2004 AVIRIS scene. All methods successfully modeled greater than 98% of the image (percentage of pixels modeled with an RMSE<2.5%). Overall map agreement between pairs of endmember selection techniques ranged from 50 to 60%. All five methods agreed on class for 27.9% of the total image and for over 50% of the reference pixels for BURN, dGRASS, GC, ORCHARD, SOIL and URBAN. The most common case (31% of image) was class agreement among three techniques. Areas where substantial disagreement occurred were still common with 18.7% of the image showing agreement between only two methods. Classes with low agreement among methods included CESP, QUDU, PLRA and UMCA. None of these classes were mapped very accurately with endmembers from any of the techniques, which may be a result of poor training data (rare classes) or suggests that MESMA is not an effective technique for mapping these species. Still, only 0.84% of the image had no class agreement between the five selection methods.

Within the classification images, we examined three subregions (outlined in Fig. 1). Subregion 1 is composed mainly of senesced grasslands grading up slope into sage scrub and chaparral and includes a large portion of the 2004 Gaviota Fire burn scar (Fig. 9). Agreement among selection techniques was high for both the burn scar and senesced grasses, but was weaker in the bands of mixed sage scrub and chaparral that stretch east–west and in riparian areas. Despite the lower agreement in these areas, the species selected on a broader scale are correct (i.e., mostly *Ceanothus* and *Artemisia* in the sage scrub/chaparral zones and *Platanus*, *Umbellularia* and *Quercus* in the riparian zones). Man-made features, such as the fire station and roads were correctly mapped as urban. CoB/EAR incorrectly mapped golf courses in several riparian areas and orchards were mapped along many riparian zones in all maps.

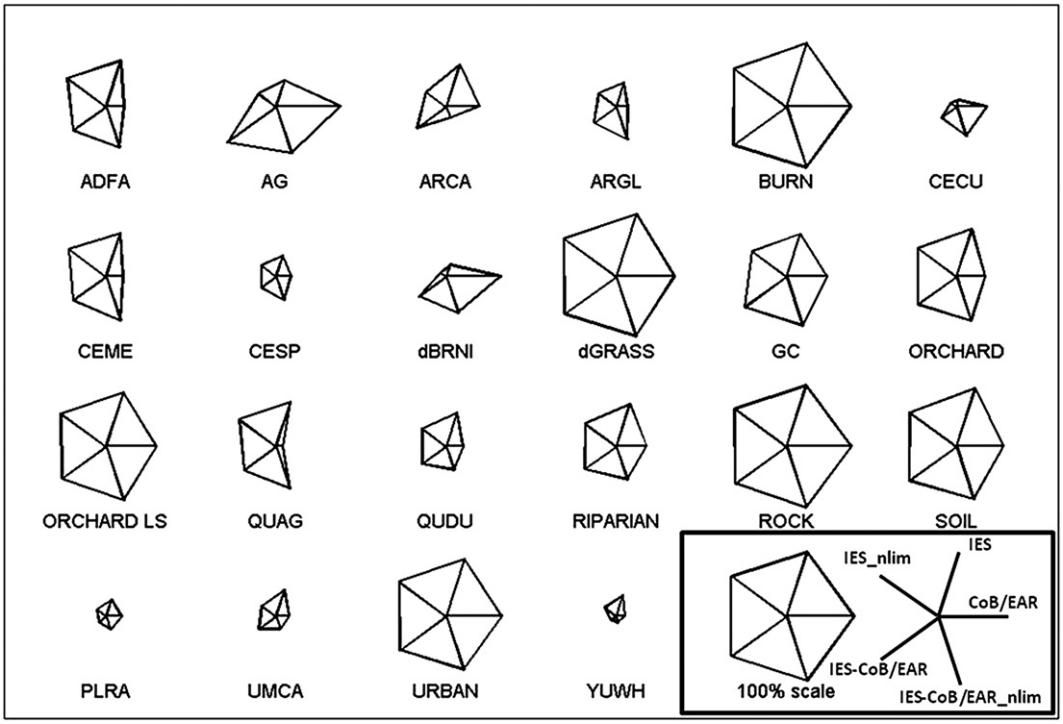

Fig. 7. Mean producer's accuracies for each class using endmembers selected with each technique. The length of the line emanating from the center of each polygon denotes accuracy while the direction of the line indicates the technique.

The second subregion (Fig. 10) covers major portions of the urban areas of Santa Barbara and Goleta and extends slightly into the foothills (oak and chaparral). Again, man-made features (i.e., roadways and golf courses) were mapped well using all methods; however, golf courses appear to be overmapped in the hybrid and n-limited methods. Orchards were highly overmapped by all methods. All methods mapped the foothills as a mixture of CEME, ADFA, and QUAG, though both IES and IES_nlim incorporated more ARGL into this region. Overall, the methods showed high agreement in urban and senesced grass areas, and low agreement in riparian zones.

The third subregion covers a section of chaparral, oak woodland and riparian areas on the south-facing side of the Santa Ynez range up to the ridgeline and across to north-facing slopes (Fig. 11). Again, orchards were highly overmapped by all methods and across

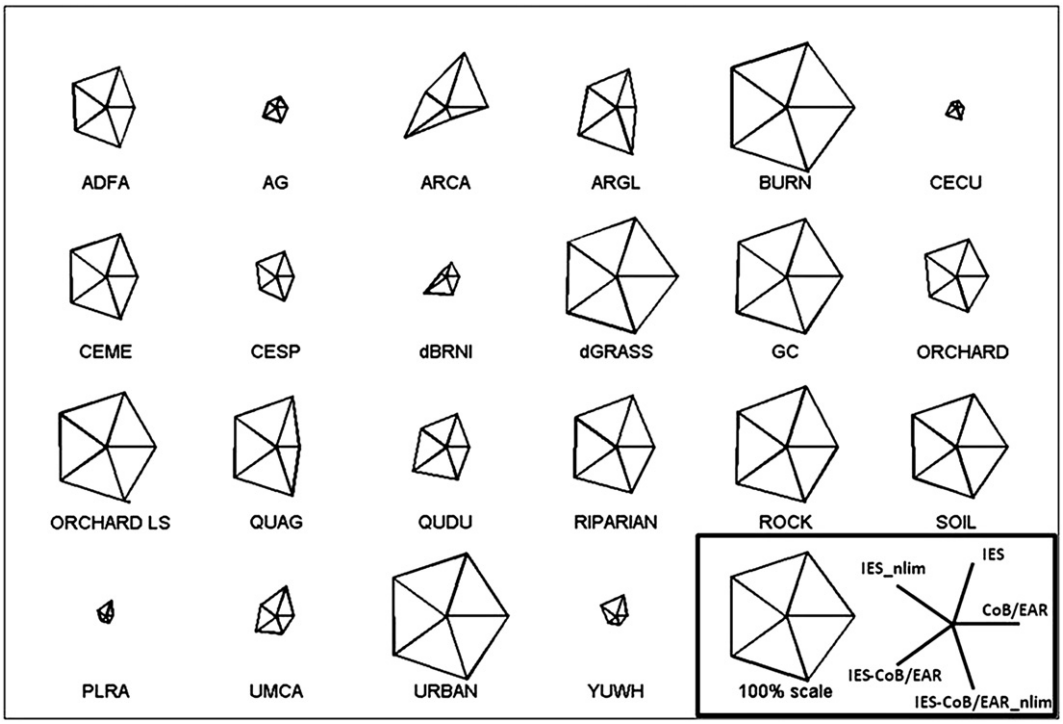

Fig. 8. Mean user's accuracy for all methods by class.

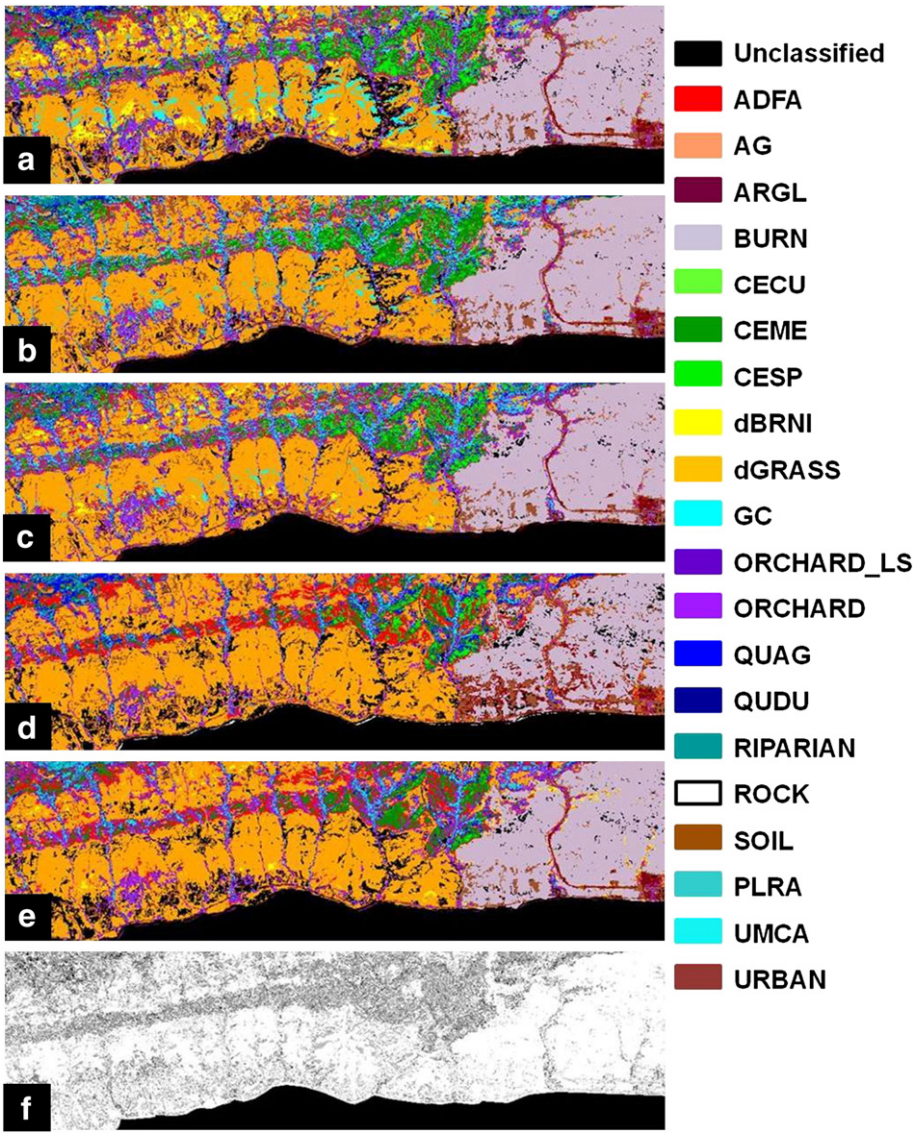

Fig. 9. Classification maps of subregion 1 (a–e) and class agreement map among selection methods (f). (a) CoB/EAR; (b) IES; (c) IES–CoB/EAR; (d) IES_nlim; (e) IES–CoB/EAR_nlim. In the class agreement image (f), lighter tone indicates greater agreement.

the entire subregion. Agreement among maps was highest on the south-facing slopes at lower elevations. This area is mainly composed of *Ceanothus* spp. and *Quercus* and *Umbellularia* (along drainages). Moving up in elevation, *Adenostoma* and *Artcostaphylos* spp. become more prevalent (seen in all maps). The areas of lowest agreement among maps occurred on the north-facing slopes just over the ridgeline. These areas are composed of both mixed chaparral and oak woodland (well-represented in the hybrid and n-limited maps). Of note, IES successfully mapped the shift in dominance from *Ceanothus megacarpus* to *C. cuneatus* that takes place from the south to north side of the range.

5. Discussion

5.1. Developing a protocol for building robust spectral libraries

The main motivations for selecting endmembers from a sample are to reduce the labor and time necessary for implementing a selection technique and to retain some data for validation.

Our results support the use of a class and polygon-stratified random sampling technique to create independent training and validation libraries. CoB/EAR selection resulted in consistent library size across sampling iterations (Fig. 3). On the other hand, both IES and IES–CoB/EAR library sizes varied greatly. Un-even class sample sizes appeared to impact endmember selection. The number of training spectra and the mean number of endmembers selected for a class were significantly positively correlated ($\alpha = 0.01$) for all selection methods. Ideally, the number of endmember spectra selected for a class will vary based solely on the spectral variability within that class. While it is true that a larger sample size is more likely to capture more of the spectral variability of a class, our findings suggest the number of endmembers selected per class is partly an artifact of the original training sample size. An in-depth examination of the spectral variability of both the class-level endmembers and training spectra could further inform this finding, but was deemed beyond the scope of this study.

A class and polygon-stratified random sampling also appeared to impact library performance for some selection techniques. Lower standard deviations of kappa and accuracy indicated some selection techniques were less sensitive to changes in the training libraries resulting from the random sampling process (Figs. 4 and 5). Library performance by IES and its size-limited version were fairly sensitive to the random sampling. The opposite was true of the hybrid and size-limited hybrid libraries, which had much tighter kappa and

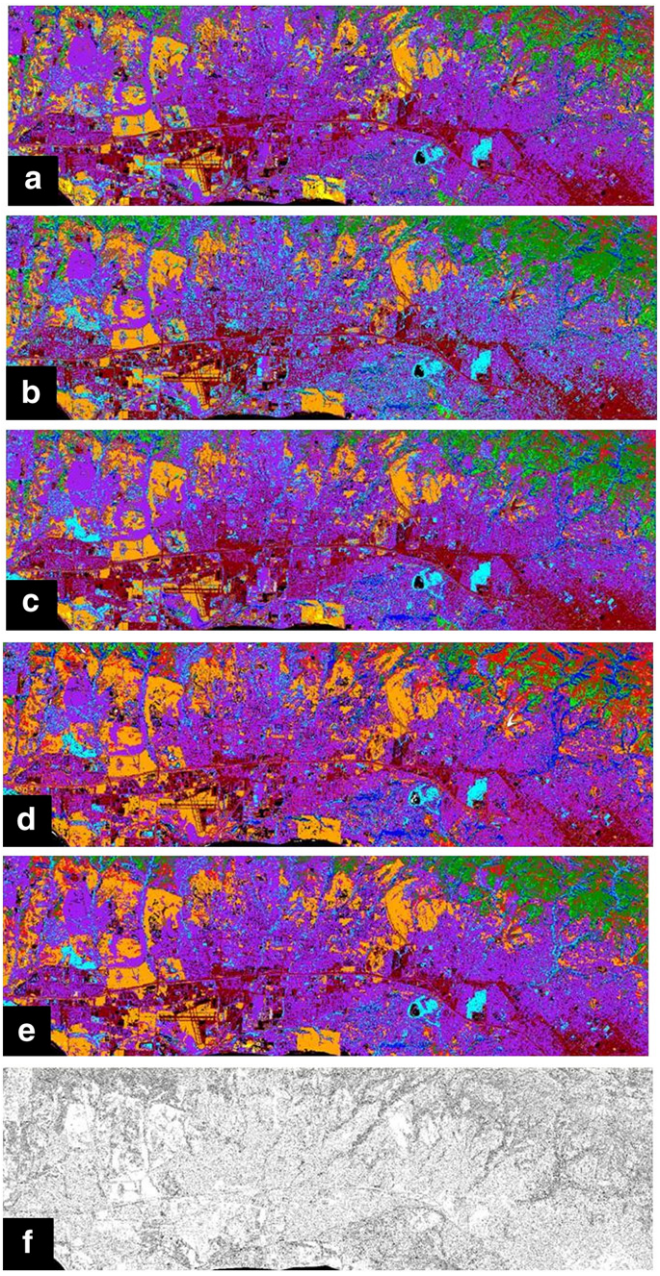

Fig. 10. Classification maps of subregion 2 (a–e) and class agreement map among selection methods (f). (a) CoB/EAR; (b) IES; (c) IES–CoB/EAR; (d) IES_nlim; (e) IES–CoB/EAR_nlim. See Fig. 9 for classification legend. In the class agreement image (f), lighter tone indicates greater agreement.

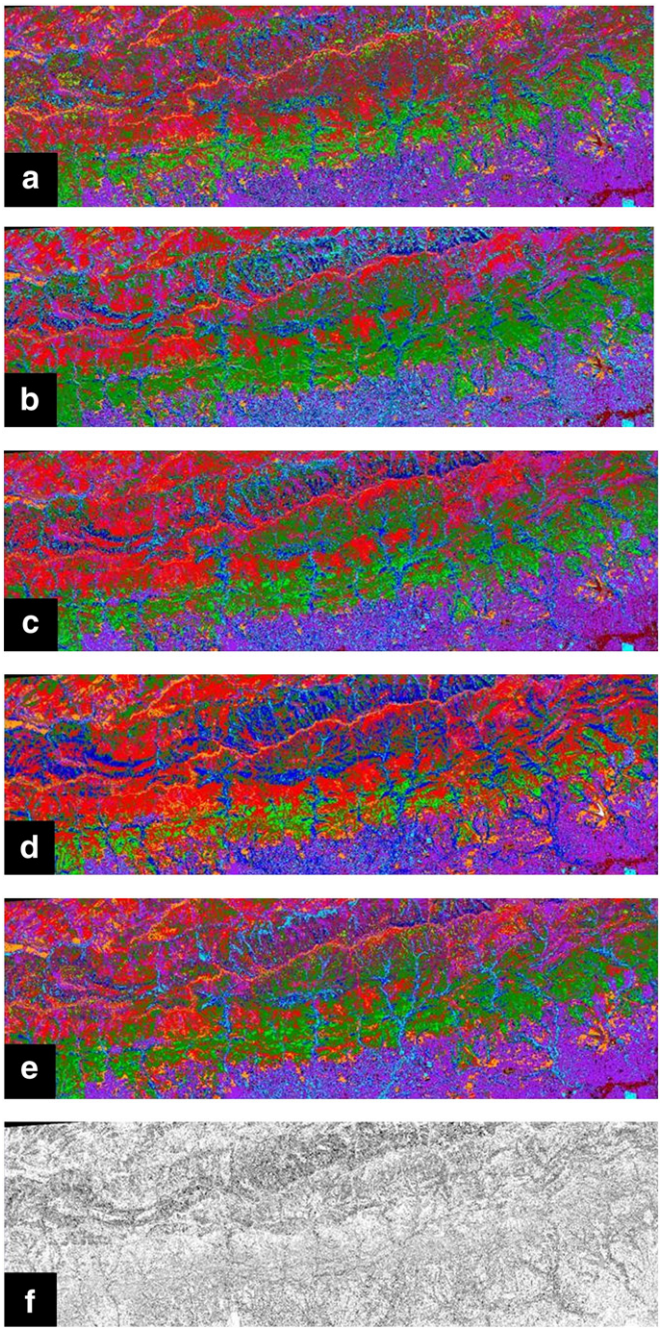

Fig. 11. Classification maps of subregion 3 (a–e) and class agreement map among selection methods (f). (a) CoB/EAR; (b) IES; (c) IES–CoB/EAR; (d) IES_nlim; (e) IES–CoB/EAR_nlim. See Fig. 9 for classification legend. In the class agreement image (f), lighter tone indicates greater agreement.

overall accuracy distributions. One potential explanation for the higher variability in IES results could be over-fitting to the training library, resulting in lower kappa values when the endmember library was applied to the validation data. However, a plot of training kappa values against validation kappa values does not support this hypothesis (Fig. 12). IES values fall very close to the 1:1 line, and in most cases, validation kappa values are higher than those achieved within the training library. On the other hand, over-fitting does seem to have occurred for the hybrid method. Despite achieving high training kappa values for most sampling iterations, validation kappa values never exceed the highest training values.

Another way of examining each technique's sensitivity to the sampling method is illustrated in Fig. 13. Here we see when the maximum overall accuracy for endmembers selected by each technique was achieved. Because the samples are randomly drawn, the exact position of the maximum is not informative, and there is no guarantee that a global maximum was actually found in the 25 iterations we analyzed. However, it is useful to visualize the number of random samples that were necessary to achieve a stable accuracy. High overall accuracy values were reached in only a few sampling iterations for the hybrid and size-limited hybrid methods. Accuracy stabilized after 5 to 10 iterations for the other three methods.

5.2. Comparing endmember selection techniques

Clear tradeoffs exist among the endmember selection methods tested here. While it has been demonstrated that CoB/EAR is an effective method for selecting endmembers to map fractional cover of green

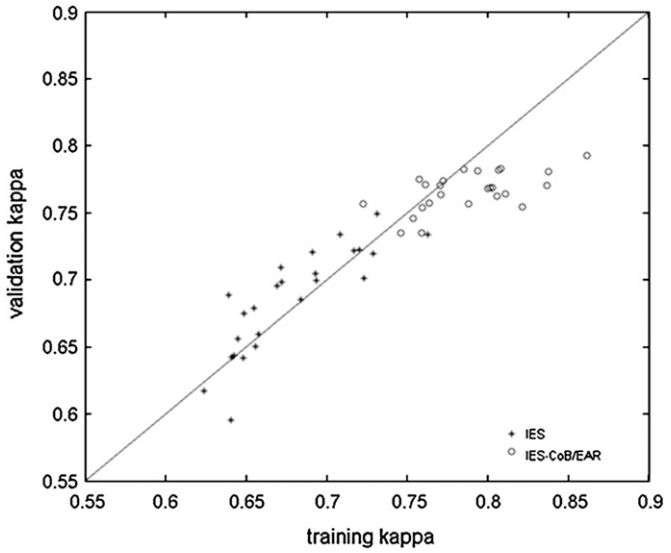

Fig. 12. Training library kappa values vs. validation kappa values for IES and hybrid endmember selection techniques.

and non-photosynthetic vegetation and soils (Roberts et al., 2012), in this study it produced the least accurate libraries for classification of plant species and land cover. It is also moderately sensitive to the random sampling process. Still, it produces consistent, small library sizes, which are an advantage computationally, especially when moving to more complex models (i.e., 3 and 4 endmember models). The selection process for CoB/EAR is currently manual, but could be automated. IES and the hybrid selection techniques are both automated and yield more accurate, but significantly larger libraries. Run time for library selection in these techniques is a concern. Selecting the first pair of endmembers is the most time-consuming step in the IES algorithm due to comparison of all possible pairs of endmembers. More computationally efficient methods could be substituted for this first step. Subsequent steps required approximately 30 s for addition of an endmember and 1–2 s for subtraction of an endmember. IES is also very sensitive to the random sampling and does not select endmembers for all classes. To address the missing classes, we implemented a hybrid technique. This technique yielded the most accurate libraries and was the least sensitive to random sampling. Because smaller libraries and shorter selection run times are more desirable, we limited the size of the IES and hybrid-selected libraries to match those from CoB/EAR. The size-limited hybrid

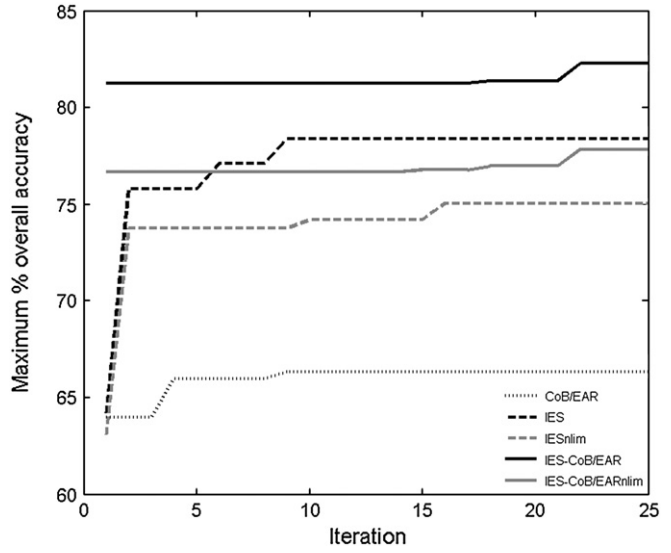

Fig. 13. Maximum overall accuracy over all iterations for each selection method.

method offers the best compromise of library size, accuracy and sensitivity to sampling.

Though the hybrid method yielded the best results, it is only useable in cases where an IES library has missing classes. Recall in the hybrid method that IES is first run, and then CoB/EAR is used to select endmembers for missing classes. In libraries with no small classes (i.e., many reference pixels for each class), IES is likely to select at least one endmember per class, and thus, the hybrid method as analyzed here would not be applied. It may be possible to gain the benefits of the hybrid method in cases where IES selects endmembers for all classes by using CoB/EAR to first select one endmember for each class, and then running IES as in the previously used hybrid method. We ran this modified hybrid technique for one sampling iteration, and the resulting library produced the same overall accuracy as the full hybrid technique, with a similar final library size. When we limited the library size as with the other methods, the overall accuracy dropped about 10%. This was still higher than the accuracies of the CoB/EAR, IES and IES_nlim-selected libraries for this sampling iteration, but slightly lower than the size-limited hybrid library. These results show that this size-limited CoB/EAR-seeded hybrid method may yield libraries which are smaller, more complete, less sensitive to random sampling, reasonably more accurate and more quickly selected than libraries created with the other selection methods. Still, a full sensitivity analysis would be necessary to confirm this.

Within individual classes, endmember performance varied. For the most spectrally distinct classes, areas of burn scar, senesced annual grasses, rock, soil and urban, endmembers chosen by each selection method performed well. The only class of vegetation with high producer's and user's accuracy across all techniques was mature orchard. Classes with the lowest producer's accuracies were those for which very little reference data existed. In general, CoB/EAR achieved higher producer's accuracies than the other methods for only three small classes: senesced BRNI, CECU and mixed agriculture. The other methods frequently did not select endmembers for these classes, mainly because they are easily confused with other spectrally similar classes. Despite the higher producer's accuracies from CoB/EAR-selected endmembers, these classes still had low user's accuracies (reflecting the confusion). Within the image, the endmembers from the tested selection techniques had high levels of agreement. Patterns of disagreement and class confusion illustrated the difficulty of separating similar cover types with differing uses. For example, golf courses are composed of irrigated green grass and senesced grass. However, senesced grass is a class on its own. Another class, riparian, could be composed of a mixture of several other species-specific classes, including QUAG, PLRA and UMCA. Spectral similarity between classes is a challenge faced by all classification methods, including 2-endmember MESMA. Methods which can be used to address this include spectral feature analysis (e.g., band subsetting or feature selection by PCA). The resulting spectra can be input to the endmember selection techniques used here. In this study, we only analyzed endmember selection for classification with a two endmember mixing model. The performance of each selection technique for more complex models (i.e., three and four-endmembers) needs to be explored in future research. Because spectral unmixing with MESMA can be quite computationally intensive, the large libraries yielded by the full IES or the hybrid methods would have to be further subset. The size limited versions of these techniques are one option for doing this, but other approaches may yield subsets better suited for mapping fractional cover. Furthermore, these selection techniques could be used with libraries developed from other endmember extraction techniques, such as VCA (Nascimento & Dias, 2005) or N-Finder (Winter, 1999).

6. Conclusions

Endmember selection is critical for successfully applying MESMA and other SMA techniques for mapping species, cover types and

fractions. Quantitative selection methods, such as EAR, CoB, MASA and now IES, enable scientists to select a set of endmembers based on specific quantitative criteria (e.g., best fit, maximizing kappa). However, the endmembers selected using each technique will vary in both number and identity, directly impacting modeling performance, including both endmember library size and accuracy. We quantitatively compared two existing endmember selection techniques, CoB/EAR & IES, and designed and tested a hybrid IES–CoB/EAR selection technique. Our study also implemented a class and polygon-stratified random sampling technique for partitioning potential endmember spectra into training and validation libraries.

CoB/EAR-selected libraries achieved the lowest classification accuracies of all the techniques. Accuracies for IES-selected libraries were higher, but highly sensitive to random sampling, and library size was large. Furthermore, IES frequently did not select endmembers for rare classes, underrepresented classes, and/or classes that may increase confusion and reduce kappa. The hybrid selection method achieved the highest accuracies and appeared insensitive to random sampling, but produced the largest libraries. When IES and hybrid library size was limited, the hybrid technique still out-performed the other selection methods, though with slightly lower accuracies than the full hybrid. Additionally, there is a need for sampling protocols which yield robust training and validation spectral libraries to allow for independent performance assessment.

The choice of endmember selection technique will depend on end-user priorities regarding the tradeoffs between classification accuracy, library size, and computation time. If high accuracies are desired, and run time and model complexity are not limiting factors, the hybrid selection technique yields the best endmember libraries. The size-limited hybrid technique represents a good compromise if smaller libraries are desired. In future research, we will aim to increase the computational efficiency of the IES algorithm and further explore the potential of novel hybrid endmember selection techniques.

Acknowledgments

Support for this work was provided by the NASA Terrestrial Ecology program, grant ##NNX08AM89G. We would like to acknowledge Seth Peterson and Michael Toomey for their assistance in collecting the reference data for this study and to thank Seth for his assistance in the reflectance retrieval. Special thanks to the NASA JPL AVIRIS team for data collection and pre-processing.

References

Adams, J. B., Smith, M. O., & Gillespie, A. R. (1993). Imaging spectroscopy: Interpretation based on spectral mixture analysis. In C. M. Pieters, & P. A. J. Englert (Eds.), *Remote geochemical analysis: Elemental and mineralogical composition*. New York: Press Syndicate of the University of Cambridge.

Asner, G. P., & Lobell, D. B. (2000). A biogeophysical approach for automated SWIR unmixing of soils and vegetation. *Remote Sensing of Environment*, *74*, 99–112.

Bateson, C. A., Asner, G. P., & Wessman, C. A. (2000). Endmember bundles: A new approach to incorporating endmember variability into spectral mixture analysis. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(2), 1083–1094. http://dx.doi.org/10.1109/36.841987.

Bateson, A., & Curtiss, B. (1996). A method for manual endmember selection and spectral unmixing. *Remote Sensing of Environment*, *55*, 229–243.

Berk, A., Anderson, G. P., Acharya, P. K., Hoke, M. L., Chetwynd, J. H., Bernstein, L. S., et al. (2003). *MODTRAN4 Version 3 Revision 1 User's manual*. Hanscom AFB: Airforce Research Laboratory, Airforce Material Command.

Boardman, J. W. (1999). Precision geocoding of low altitude AVIRIS data: Lessons learned in 1998. *Summaries of the eighth JPL Airborne Earth Science Workshop, 99-17* (pp. 63–68). Pasadena, CA: Jet Propulsion Laboratory.

Boardman, J. W., Kruse, F. A., & Green, R. O. (1995). Mapping target signatures via partial unmixing of AVIRIS data. *Summaries of the Fifth Annual JPL Airborne Geoscience Workshop* (pp. 23–26). Pasadena, CA: Jet Propulsion Laboratory.

Borel, C. C., & Gerstl, S. A. W. (1994). Nonlinear spectral mixing models for vegetative and soil surfaces. *Remote Sensing of Environment*, *47*(3), 403–416. http://dx.doi.org/10.1016/0034-4257(94)90107-4.

Clark, R. N., Swayze, G. A., Livo, K. E., Kokaly, R. F., King, T. V., Dalton, J. B., et al. (2002). Surface reflectance calibration of terrestrial imaging spectroscopy data: A tutorial using AVIRIS. *Proceedings of the 10th Airborne Earth Science Workshop, 02–1*. Pasadena, CA: Jet Propulsion Laboratory.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. http://dx.doi.org/10.1177/001316446002000104.

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*, 35–46.

Cowling, R. M., Rundel, P. W., Lamont, B. B., Arroyo, M. K., & Arianoutsou, M. (1996). Plant diversity in Mediterranean-climate regions. *Trends in Ecology & Evolution*, *11*, 362–366.

Dennison, P. E., Charoensiri, K., Roberts, D. A., Peterson, S. H., & Green, R. O. (2006). Wildfire temperature and land cover modeling using hyperspectral data. *Remote Sensing of Environment*, *100*, 212–222.

Dennison, P. E., Halligan, K. Q., & Roberts, D. A. (2004). A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper. *Remote Sensing of Environment*, *93*, 359–367.

Dennison, P. E., & Roberts, D. A. (2003a). Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE. *Remote Sensing of Environment*, *87*, 123–135.

Dennison, P. E., & Roberts, D. A. (2003b). The effects of vegetation phenology on endmember selection and species mapping in southern California chaparral. *Remote Sensing of Environment*, *87*, 295–309.

Eckmann, T., Roberts, D., & Still, C. (2008). Using multiple endmember spectral mixture analysis to retrieve subpixel fire properties from MODIS. *Remote Sensing of Environment*, *112*(10), 3773–3783. http://dx.doi.org/10.1016/j.rse.2008.05.008.

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

Franke, J., Roberts, D. A., Halligan, K., & Menz, G. (2009). Hierarchical Multiple Endmember Spectral Mixture Analysis (MESMA) of hyperspectral imagery for urban environments. *Remote Sensing of Environment*, *113*, 1712–1723.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, *32*(200), 675–701.

Girden, E. R. (1992). *ANOVA: Repeated measures.* : Sage Publications.

Green, R., Conel, J., & Roberts, D. (1993). Estimation of aerosol optical depth and additional atmospheric parameters for the calculation of apparent surface reflectance from radiance as measured by the Airborne Visible-Infrared Imaging Spectrometer (AVIRIS). *Summaries of the Fourth Annual JPL Airborne Geosciences Workshop* (pp. 73–76). Pasadena, CA: Jet Propulsion Laboratory.

Green, R. O., Eastwood, M. L., Sarture, C. M., Chrien, T. G., Aronsson, M., Chippendale, B. J., et al. (1998). Imaging spectroscopy and the Airborne Visible Infrared Imaging Spectrometer (AVIRIS). *Remote Sensing of Environment*, *65*, 227–248.

Herold, M., Roberts, D. A., Gardner, M. E., & Dennison, P. E. (2004). Spectrometry for urban area remote sensing—Development and analysis of a spectral library from 350 to 2400 nm. *Remote Sensing of Environment*, *91*, 304–319.

Hochberg, Y., & Tamhane, A. C. (1987). *Multiple comparison procedures.* John Wily & Sons.

Huete, A. R. (1986). Separation of soil–plant spectral mixtures by factor analysis. *Remote Sensing of Environment*, *19*, 237–251.

Jia, G. J., Burke, I. C., Goetz, A. F. H., Kaufmann, M. R., & Kindel, B. C. (2006). Assessing spatial patterns of forest fuel using AVIRIS data. *Remote Sensing of Environment*, *102*(3–4), 318–327.

Kruse, F. A., Lefkoff, A. B., Boardman, J. W., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J., et al. (1993). The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, *44*(2–3), 145–163.

Li, L., Ustin, S. L., & Lay, M. (2005). Application of multiple endmember spectral mixture analysis (MESMA) to AVIRIS imagery for coastal salt marsh mapping: A case study in China Camp, CA, USA. *International Journal of Remote Sensing*, *26*, 5193–5207.

Lilliefors, H. W. (1969). On the Kolmogorov–Smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association*, *64*, 387–389.

Linn, R. M., Rolim, S. B. A., & Galvao, L. S. (2010). Assessment of the Multiple Endmember Spectral Mixture Analysis (MESMA) model applied to the Hyperion/EO-1 hyperspectral data of the coastal plain of Rio Grande do Sul, Brazil. In W. Wagner, & B. Szekely (Eds.), *ISPRS TC VII Symposium*, *Vol. XXXVIII*. (pp. 134–138)Vienna, Austrua: IAPRS.

Lu, D., & Weng, Q. (2004). Spectral mixture analysis of the urban landscape in Indianapolis with Landsat ETM+ Imagery. *Photogrammetric Engineering and Remote Sensing*, *70*(9), 1053–1062.

Martinez, P. J., Perez, R. M., Plaza, A., Aguilar, P. L., Cantero, M. C., & Plaza, J. (2006). Endmember extraction algorithms from hyperspectral images. *Annals of Geophysics*, *49*, 93–101.

Meentemeyer, R. K., & Moody, A. (2000). Rapid sampling of plant species composition for assessing vegetation patterns in rugged terrain. *Landscape Ecology*, *15*(8), 697–711.

Myint, S. W., & Okin, G. S. (2009). Modelling land-cover types using multiple endmember spectral mixture analysis in a desert city. *International Journal of Remote Sensing*, *30*, 2237–2257.

Nascimento, J. M. P., & Dias, J. M. (2005). Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, *43*(4), 898–910.

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

Painter, T. H., Roberts, D. A., Green, R. O., & Dozier, J. (1998). The effect of grain size on spectral mixture analysis of snow-covered area from AVIRIS data. *Remote Sensing of Environment*, *65*, 320–332.

Painter, T. H., Dozier, J., Roberts, D. A., Davis, R. E., & Green, R. O. (2003). Retrieval of subpixel snow-covered area and grain size from imaging spectrometer data. *Remote Sensing of Environment*, *85*, 64–77.

Plaza, A., Martinez, P., Perez, R., & Plaza, J. (2004). A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, *42*(3), 650–663.

Plourde, L. C., Ollinger, S. V., Smith, M. L., & Martin, M. E. (2007). Estimating species abundance in a northern temperate forest using spectral mixture analysis. *Photogrammetric Engineering and Remote Sensing*, *73*, 829–840.

Powell, R. L., & Roberts, D. A. (2010). Characterizing urban land-cover change in Rondônia, Brazil: 1985 to 2000. *Journal of Latin American Geography*, *9*(3), 29.

Powell, R. L., Roberts, D. A., Dennison, P. E., & Hess, L. L. (2007). Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. *Remote Sensing of Environment*, *106*(2), 253–267. http://dx.doi.org/10.1016/j.rse.2006.09.005.

Rashed, T., Weeks, J. R., Roberts, D. A., Rogan, J., & Powell, R. L. (2003). Measuring the physical composition of urban morphology using multiple endmember spectral mixture models. *Photogrammetric Engineering and Remote Sensing*, *69*(9), 1011–1021.

Ray, T. W., & Murray, B. C. (1996). Nonlinear spectral mixing in desert vegetation. *Remote Sensing of Environment*, *55*(1), 59–64. http://dx.doi.org/10.1016/0034-4257(95)00171-9.

Riano, D., Chuvieco, E., Ustin, S., Zomer, R., Dennison, P., Roberts, D., et al. (2002). Assessment of vegetation regeneration after fire through multitemporal analysis of AVIRIS images in the Santa Monica Mountains. *Remote Sensing of Environment*, *79*, 60–71.

Roberts, D. A., Dennison, P. E., Gardner, M., Hetzel, Y. L., Ustin, S. L., & Lee, C. (2003). Evaluation of the potential of Hyperion for fire danger assessment by comparison to the Airborne Visible Infrared Imaging Spectrometer. *IEEE Transactions on Geoscience and Remote Sensing*, *41*, 1297–1310.

Roberts, D. A., Dennison, P. E., Peterson, S., Sweeney, S., & Rechel, J. (2006). Evaluation of Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and Moderate Resolution Imaging Spectrometer (MODIS) measures of live fuel moisture and fuel condition in a shrubland ecosystem in southern California. *Journal of Geophysical Research*, *111*(G4), 1–16. http://dx.doi.org/10.1029/2005JG000113.

Roberts, D. A., Gamon, J., Keightley, K., Prentiss, D., Reith, E., & Green, R. (1999). AVIRIS land-surface mapping in support of the BOReal Ecosystem–Atmosphere Study (BOREAS). *Proceedings of the 8th AVIRIS Earth Science Workshop* (pp. 355–364). Pasadena, CA: Jet Propulsion Laboratory.

Roberts, D. A., Green, R. O., & Adams, J. B. (1997). Temporal and spatial patterns in vegetation and atmospheric properties from AVIRIS. *Remote Sensing of Environment*, *62*, 223–240.

Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G., & Green, R. O. (1998). Mapping Chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sensing of Environment*, *65*, 267–279.

Roberts, D. A., Halligan, K. Q., & Dennison, P. E. (2007). ViperTools. http://www.vipertools.org/

Roberts, D. A., Quattrochi, D. A., Hulley, G. C., Hook, S. J., & Green, R. O. (2012). Synergies between VSWIR and TIR data for the urban environment: An evaluation of the potential for the Hyperspectral Infrared Imager (HyspIRI) Decadal Survey mission. *Remote Sensing of Environment*, *117*, 83–101.

Roberts, D. A., Smith, M. O., & Adams, J. B. (1993). Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data. *Remote Sensing of Environment*, *44*, 255–269.

Roberts, D. A., Ustin, S. L., Ogunjemiyo, S., Greenberg, J., Dobrowski, S. Z., Chen, J. Q., et al. (2004). Spectral and structural measures of northwest forest vegetation at leaf to landscape scales. *Ecosystems*, *7*, 545–562.

Rosso, P. H., Ustin, S. L., & Hastings, A. (2005). Mapping marshland vegetation of San Francisco Bay, California, using hyperspectral data. *International Journal of Remote Sensing*, *26*, 5169–5191.

Schaaf, A. N., Dennison, P. E., Fryer, G. K., Roth, K. L., & Roberts, D. A. (2011). Mapping plant functional types at three spatial resolutions using multiple endmember spectral mixture analysis. *GIScience and Remote Sensing*, *48*, 324–344.

Settle, J. J., & Drake, N. A. (1993). Linear mixing and the estimation of ground cover proportions. *International Journal of Remote Sensing*, *14*, 1159–1177.

Small, C. (2001). Estimation of urban vegetation abundance by spectral mixture analysis. *International Journal of Remote Sensing*, *22*(7), 1305–1334.

Somers, B., Asner, G. P., Tits, L., & Coppin, P. (2011). Endmember variability in spectral mixture analysis: A review. *Remote Sensing of Environment*, *115*(7), 1603–1616. http://dx.doi.org/10.1016/j.rse.2011.03.003.

Song, C. H. (2005). Spectral mixture analysis for subpixel vegetation fractions in the urban environment: How to incorporate endmember variability? *Remote Sensing of Environment*, *95*, 248–263.

Sonnentag, O., Chen, J. M., Roberts, D. A., Talbot, J., Halligan, K. Q., & Govind, A. (2007). Mapping tree and shrub leaf area indices in an ombrotrophic peatland through multiple endmember spectral unmixing. *Remote Sensing of Environment*, *109*, 342–360.

Tompkins, S., Mustard, J. F., Pieters, C. M., & Forsyth, D. W. (1997). Optimization of endmembers for spectral mixture analysis. *Remote Sensing of Environment*, *59*(3), 472–489.

Underwood, E. C., Mulitsch, M. J., Greenberg, J. A., Whiting, M. L., Ustin, S. L., & Kefauver, S. C. (2006). Mapping invasive aquatic vegetation in the Sacramento–San Joaquin Delta using hyperspectral imagery. *Environmental Monitoring and Assessment*, *121*, 47–64.

Veganzones, M. A., & Grana, M. (2008). Endmember extraction methods: A short review. *Knowledge-Based Intelligent Information and Engineering Systems, 12th International Conference, KES 2008* (pp. 400–407).

Winter, M. E. (1999). N-finder: An algorithm for fast autonomous spectral endmember determination in hyperspectral data. *Proc. SPIE-Image Spectrometry V*, *Vol. 3753*. (pp. 266–277).

Youngentob, K. N., Roberts, D. A., Held, A. A., Dennison, P. E., Jia, X., & Lindenmayer, D. B. (2011). Mapping two Eucalyptus subgenera using multiple endmember spectral mixture analysis and continuum-removed imaging spectrometry data. *Remote Sensing of Environment*, *115*(5), 1115–1128.