Analysis and classification of hyperspectral data for mapping land degradation: An application in southern Spain

D.P. Shrestha a,*, D.E. Margate b, F. van der Meer a, H.V. Anh c

a International Institute for Geo-Information Science and Earth Observation (ITC), PO Box 6, 7500 AA Enschede, The Netherlands
b Bureau of Soils and Water Management, Quezon City, Philippines
c Forest Science Institute, Hanoi, Vietnam

Received 19 March 2004; accepted 23 January 2005

Abstract

Desertification is a severe stage of land degradation, manifested by “desert-like” conditions in dryland areas. Climatic conditions together with geomorphologic processes help to mould desert-like soil surface features in arid zones. The identification of these soil features serves as a useful input for understanding the desertification process and land degradation as a whole. In the present study, imaging spectrometer data were used to detect and map desert-like surface features. Absorption feature parameters in the spectral region between 0.4 and 2.5 μm wavelengths were analysed and correlated with soil properties, such as soil colour, soil salinity, gypsum content, etc. Soil groupings were made based on their similarities and their spectral reflectance curves were studied. Distinct differences in the reflectance curves throughout the spectrum were exhibited between groups. Although the samples belonging to the same group shared common properties, the curves still showed differences within the same group.

Characteristic reflectance curves of soil surface features were derived from spectral measurements both in the field and in the laboratory, and mean reflectance values derived from image pixels representing known features. Linear unmixing and spectral angle matching techniques were applied to assess their suitability in mapping surface features for land degradation studies. The study showed that linear unmixing provided more realistic results for mapping “desert-like” surface features than the spectral angle matching technique.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Linear unmixing; Spectral angle mapper; End members; Absorption features; Desert-like surface features

1. Introduction

Although broadband data from sensors, such as SPOT HRV, Landsat MSS and Landsat TM have been used for mapping soil, they do not provide sufficient information to characterise soil differences, because their 100–200 nm bandwidth cannot resolve diagnostic spectral features of terrestrial materials (De Jong, 1994). The development of scanner systems that acquire data in many narrow-wavelength bands allow the use of almost continuous reflectance data in studies
of the Earth’s surface. This not only produces laboratory-like reflectance spectra with absorption bands specific to the object’s properties, but also helps to increase the accuracy of mapping as reported by various researchers (Pieters and Mustard, 1988; Kruse, 1989; Clark et al., 1990). Broadband data are well suited to asking the question “what is there?” but hyperspectral data are well suited to asking the second question: “what proportion is there?” Techniques developed for analysing broadband spectral data are, however, incapable of taking advantage of the full range of information present in hyperspectral imagery (Cloutis, 1996). Hyperspectral image classification requires new approaches, such as spectral matching techniques. With so much data available, the well-known problem of mixed pixels can be also solved using a mixture model, assumed to occur in a linear fashion (Singer and McCord, 1979).

Factors affecting soil reflectance and soil specific absorption features have been described by De Jong and Epema (2001); Van der Meer (2001); Mulders (1985); Baumgardner et al. (1985); Hunt (1980). Spectral reflectance characteristics of soils are the combined result of their physical and chemical properties: texture, structure, mineral composition, iron content, type of clay minerals, organic matter and soil moisture. In addition, topography, sun elevation and varying soil surface conditions influence surface reflectance (De Jong and Epema, 2001; Shrestha and Zinck, 2001). With so many factors playing a role and each individual soil property having its influence on soil reflectance, a particular reflectance curve cannot be representative for a given soil type. Moreover, a soil consists not only of surface materials but also of specific horizons that have been altered by the interaction of climate, relief, and living organisms over time. A taxonomic soil type incorporates its horizons and the corresponding properties. Thus, selecting pure spectra using spectral library of pure minerals is not representative of typical soil types. Although Shepherd and Walsh (2002) have developed spectral libraries for characterization of soil properties, their use in mapping soil distribution cannot be so straightforward. Characterisation of surface features, which indicates certain soil processes, i.e. soil salinity development, erosion, etc., would be more useful for mapping their extent. In arid and semi-arid regions, the climatic conditions together with geomorphologic processes help in moulding the so-called desert-like soil surface features. Desertification is a severe stage of land degradation, manifested by desert-like conditions in dry-land areas outside the desert boundaries (Rapp, 1986). The transportation by either wind or water of surface loose materials leaves behind desert pavement, which is a continuous layer of gravel and small stones on the soil surface. Similarly, due to high evaporation rates and lack of leaching and percolating to deeper horizons many low-lying areas may be saline and/or alkaline. In the same way, calcium carbonate and gypsum may often be present in abundance, forming hard pans and contributing to the formation of surface crusts. The identification of these soil surface features serves as a useful input for assessing desertification and land degradation as a whole. In the present study, imaging spectrometer data has been used to detect and map desert-like surface features. Overall objectives of the study are as follows: (i) analysis of the measured reflectance spectra in order to see if absorption features related to desert-like soil surface features exist in specific wave bands, (ii) grouping of soils based on their chemical and physical properties measured in the laboratory and checking their similarities in spectral measurements and finally, (iii) assessing performance of two classification techniques, namely spectral angle mapper and linear unmixing in mapping land degradation.

2. Methods and techniques

2.1. Study area

The study area is located in the surroundings of Tabernas in the Province of Almeria in Spain (Fig. 1). The exact site corresponds with the coverage of the HYMAP airborne hyperspectral image, with its flight line starting at 37°02'32"N and 2°30'14"W and ending at 37°04'25"N and 2°16'40"W. The Tabernas basin is a structural depression in the Alpine nappes of the Betic Cordilleras of Southern Spain, which is bounded by major strike-slip fault (Kleverlaan, 1989). The terrain is relatively rugged and sparsely vegetated. The mountain ridges on the northern and southern sides of the basin are acting as main barriers for precipitation, resulting in pronounced dry conditions causing desertification. The climate can be characterised as
semi-arid with long hot summers. Annual precipitation ranges from 115 to 431 mm. In general, soils are shallow (less than 50 cm depth), except in the valleys and, occasionally, in the piedmont area where they are deep. Saline soils with surface crusting occupy the low-lying areas. Main land use on the mountain slopes is grazing, while rainfed cereals and field crops are grown in the piedmont area and valleys. It is reported that many rainfed cultivation areas were abandoned in recent past due to repeated droughts and land degradation in general (Alemayehu, 2001).

2.2. Data collection

The 128-channel HYMAP data acquired on 2 June 1999 with 5 m spatial resolution and covering an area of 4 km × 20 km, was obtained. The data strip covered all the landscape units in the area. Fieldwork was undertaken during the months of September and October in 1999 and 2000 in order to measure and collect data on spectral reflectance for a range of soil surface types. Field observations were made on representative sites, based on geopedologic air photo interpretation approach (Zinck, 1988) and using a GPS receiver (Garmin 12XL). Observations included collecting the following information: geomorphic unit, surface gravel content, topsoil texture and colour (Munsell Colour Charts), calcareousness (using 10% HCl) and land cover/land use type. Soil samples were collected from representative sites for laboratory analysis of soil properties. At each observation point, spectral measurements were recorded using a field spectrometer (GER 3700) with full real-time data acquisition from 0.35 to 2.5 μm, in 647 channels. Reflectance was measured by comparing the radiance of the target sample with the radiance of a reference panel made of barium sulphate (BaSO4). Soil reflectance was also measured in the laboratory using controlled illumination.

2.3. Data analysis

2.3.1. Field and lab spectra

The field and laboratory reflectance spectra were analysed to study the absorption feature parameters as described by Green and Craig (1985). The absorption feature parameters were wavelength position, depth, width, area, and asymmetry. The absorption wavelength is defined as the wavelength of minimum reflectance of an absorption feature. The width, W, of an absorption feature is defined as:

\[ W = \frac{\text{Area}_{\text{left}} + \text{Area}_{\text{right}}}{2D} \]  

where D is the depth of the feature relative to the hull and the Area_{left} and Area_{right} are the percentages of the areas left and right of the line through the centre of the absorption feature (Fig. 2). The upper convex hull is an envelope curve fitted over the original reflectance spectrum, having no absorption features. The asym-
metry of an absorption feature, $S$, is derived as the ratio of the area left (Area\_left) of the absorption centre to the area right (Area\_right) of the absorption centre as:

$$ S = \frac{\text{Area}_{\text{left}}}{\text{Area}_{\text{right}}} $$

(2)

2.3.2. Soil data

The laboratory analysis results of selected soil properties were subjected to statistical groupings of soils. Cluster analysis, a multivariate statistics tool, was used to group the soils according to relevant properties, which included particle size distribution (clay, silt, fine sand and coarse sand contents), gypsum, CaCO$_3$, organic carbon, and organic matter contents, and pH and electrical conductivity values. This resulted in formation of soil groups with similar properties. Similarity was measured using Euclidean distances. Comparison of the reflectance spectra from each resulting soil groups was also carried out.

2.3.3. Pre-processing HYMAP data

Signal-to-noise ratio (SNR) gives an indication on data quality for which various methods such as homogenous area (Duggin et al., 1985), and variogram parameter (Curran and Dungan, 1989; Chappell et al., 2001) are available. Although the use of spatial dependency seems to be appropriate for estimating noise, one should not forget that the nugget value depends not only on measurement error but also on the choice of mathematical model fitted to the sample variogram (Atkinson, 1997), which would mean tedious computing time in obtaining SNR values for all the bands. Homogenous area method was thus selected, which uses scenes in dark and bright locations to calculate signal mean and its standard deviation, and their ratio to obtain the SNR value (Fig. 3). For some selected bands, SNR values were calculated using variogram method and compared with the ones obtained using homogenous area method. The values from the two methods were similar. The calculated SNR values were low at 0.403 and 1.401 $\mu$m wavelengths (Fig. 3) and thus the spectral bands at these wavelengths were considered not useful and discarded from further analysis.

Since the HYMAP data was obtained under moderate weather conditions, the data was thus corrected to reflectance using ATCOR-A atmospheric correction method (Richter, 1996). It consisted of four main correction procedures: (1) calculating atmospheric correction functions for a specified atmosphere and geometry, (2) calculating a ground reflectance image for the reflective spectral bands, and emissive images for the thermal bands, (3) deriving the calibration coefficients for each reflective band based on measured ground reflectance data and the specified atmospheric data, and (4) interactively displaying spectra of user-specified profiles.

2.3.4. Image classification

In order to classify hyperspectral data characteristic spectral curves, the so-called end-members representing surface features are required. For this both field and laboratory spectrometer data, and HYMAP data
were used. Use was also made of the maximum noise fraction (MNF) transformation to get the purest pixels in the image (Green et al., 1988) by plotting of the lower MNF images to delineate the extreme pixels (Boardman et al., 1995). End-members were also derived from known areas in the image using the function, region-of-interest in ENVI. Finally characteristic spectral curves were established for each of the surface features by comparing the image-derived spectra, and the field and laboratory derived spectra of target objects (Fig. 4).

The HYMAP data was subjected to two classification techniques: spectral angle mapper (SAM) and linear unmixing. SAM determines the similarity between two spectra by calculating the “spectral angle” between them, treating them as vectors in a space with dimensionality equal to the \(n\) number of bands (Kruse et al., 1993) (Fig. 5). Since it uses only the “direction” of the spectra, and not their “length,” the method is insensitive to the unknown gain factor, thus avoiding requirement for any pre-processing technique such as normalization of data for uniform intensity (Shrestha and Zinck, 2001). Spectral angle between an unknown spectrum \(t\) to a reference spectrum \(r\), is calculated by applying the following equation:

\[
\cos^{-1}\left(\frac{\hat{t} \cdot \hat{r}}{||\hat{t}|| \cdot ||\hat{r}||}\right)
\]  

which can be written as:

\[
\cos^{-1}\left(\frac{\sum_{j=1}^{n} t_j r_j}{\left(\sum_{i=1}^{n} t_i^2\right)^{0.5} \cdot \left(\sum_{i=1}^{n} r_i^2\right)^{0.5}}\right)
\]  

The resulting value, in radians, is assigned to the corresponding pixel in the output SAM image, one output image for each reference spectrum. The derived spectral angle maps form a new data cube with the number of bands equal to the number of reference spectra used in the mapping.

Ground surfaces constituting individual pixels of remotely sensed imagery often contain more than one cover type, each type contributing to the overall spectral response to that pixel. Spectral mixing occurs in a linear fashion if mixing is large (Singer and McCord, 1979) and non-linear for microscopic mixing (Nash and Conel, 1979). Assuming that spectral mixing takes place in linear fashion, unmixing algorithm helps in mapping the abundances of each end-members in a given pixel. In a linear model, the reflectance \(R_i\) of a pixel in \(i\)th band is given by Shimabukuro and Smith (1991) as follows:

\[
R_i = \sum_{j=1}^{n} (F_j R_{ij}) + \epsilon_i
\]

with the condition that \(\sum_{j=1}^{n} F_j = 1\), where \(i = 1, \ldots, m\) and \(j = 1, \ldots, n\), \(R_i\) the reflectance of the mixed spectrum in image band \(i\) for each pixel, \(F_j\) the fraction
of each end-member \( j \) calculated by band, \( \text{RE}_j \)
the reflectance of end-member spectrum \( j \) in band \( i \),
i the band number, \( j \) each of the end-members and \( \epsilon \)
the residual error, \( m \) the number of spectral bands
while \( n \) stands for the number of components in the pixel.

Application of SAM algorithm results in the so-
called rule images, their values indicating spectral
angles. On the other hand, linear unmixing results in
abundance images. For classification accuracy
assessment, an error matrix or contingency table
was constructed and the estimate of a measure of
overall agreement between classification result
and ground truth data was carried out by \( \kappa \) statistics
(Cohen, 1960), which is given by:

\[
k = \left( \frac{p_0 - p_c}{1 - p_c} \right)
\]

where \( p_0 \) is the proportion of units in which there is
agreement between ground truth and the classification
result, and \( p_c \) the proportion for which agreement is
expected by chance. \( p_0 \) and \( p_c \) can be calculated using
the observation numbers in the row and columns from
the error matrix as follows:

\[
p_0 = \frac{\sum_{i=1}^{r} X_{ii}}{N}
\quad \text{and} \quad
p_c = \frac{\sum_{i=1}^{r} X_{+i} X_{i+}}{N^2}
\]

where \( X_{ii} \) is the count of observations at row \( i \)
and column \( i \), \( X_{+i} \) the sum of the \( i \)th row and \( X_{i+} \)
the sum of the \( i \)th column, \( r \) the number of rows and
columns in the error matrix, and \( N \) the total number
of observations.

3. Results and discussions

3.1. Absorption features

Prominent absorptions at around 1450 and
1950 nm wavelengths in most soil spectra were
attributed to water and hydroxyl ions. Occasionally,
weaker water absorptions also occurred at 970, 1200,
and 1770 nm. Absorptions at 1800 and 2300 nm were
attributed to gypsum, while strong absorption feature
near 2350 nm were assigned to calcite (CaCO_3). The
correlation matrix in Table 1 explains the relationship
between absorption features and the gypsum content.

Depth of the absorption feature varied significantly
with gypsum content in almost all selected wavelength
intervals. At wavelength position near 2200 nm,
depth, width, and area of absorption feature changed
significantly. Calcium carbonate content of the soil
and absorption feature parameters, however, did not
show any correlation.

Soils were also grouped according to their degree
of salinity. Salinity classes were derived from USDA
soil survey manual (Soil Survey Division Staff, 1993),
as follows: (1) non-saline soils (EC_e values between 0
and 2 dS m\(^{-1}\)); (2) saline soils (EC_e values between 2
and 16 dS m\(^{-1}\); and (3) strongly saline soils (EC_e
values more than 16 dS m\(^{-1}\)). Wavelength positions of
absorption features related to salinity were noted.
Absorption features at 390–400, 615–625, 685–695,
800–810, 950–960, 1410–1420, 1935–1945, and
2350–2360 nm wavelengths were selected and corre-
lated with electrical conductivity values (Table 2).
Significant correlation coefficients were observed for
EC_e values with depth, width, and asymmetry at
wavelength of 800–810 nm. This implied that the
depth of the absorption feature at this specific
wavelength interval increased with the degree of

<table>
<thead>
<tr>
<th>Wavelength</th>
<th>Wavelength</th>
<th>Depth</th>
<th>Width</th>
<th>Area</th>
<th>Asymmetry</th>
</tr>
</thead>
<tbody>
<tr>
<td>982–1002</td>
<td>0.519/0.057</td>
<td>0.650/0.012</td>
<td>−0.391/0.167</td>
<td>0.083/0.777</td>
<td>0.830/0.000</td>
</tr>
<tr>
<td>1174–1210</td>
<td>0.509/0.243</td>
<td>0.826/0.022</td>
<td>−0.590/0.163</td>
<td>−0.481/0.274</td>
<td>0.370/0.414</td>
</tr>
<tr>
<td>1434–1454</td>
<td>0.169/0.689</td>
<td>0.975/0.000</td>
<td>−0.231/0.583</td>
<td>0.248/0.554</td>
<td>−0.313/0.450</td>
</tr>
<tr>
<td>1734–1754</td>
<td>0.419/0.120</td>
<td>0.831/0.000</td>
<td>−0.516/0.049</td>
<td>−0.058/0.837</td>
<td>0.220/0.430</td>
</tr>
<tr>
<td>1926–1946</td>
<td>0.053/0.802</td>
<td>0.280/0.175</td>
<td>−0.121/0.563</td>
<td>0.045/0.830</td>
<td>−0.212/0.310</td>
</tr>
<tr>
<td>2162–2214</td>
<td>0.105/0.538</td>
<td>0.723/0.000</td>
<td>0.660/0.000</td>
<td>0.897/0.000</td>
<td>−0.064/0.707</td>
</tr>
</tbody>
</table>

Cell contents: Pearson product moment correlation coefficient/P-value.
salinity while the width and asymmetry of this feature decreased with increasing EC_e values.

Absorption feature was also correlated with soil colour. A strong absorption feature at around 380–410 nm for all soil samples was correlated with soil colour parameters namely hue, value and chroma (Table 3). Significant correlation existed between value and depth of the absorption feature, value and area of absorption feature, value and asymmetry of the absorption feature, and hue and width of the absorption feature. The change in colour value significantly affected the depth of the absorption feature. The same held true for the correlations of colour value and area of absorption feature, of colour value and asymmetry of absorption feature, as well as of colour hue and width of absorption feature.

3.2. Soil clustering

Clustering of soils based on their properties yielded nine groups with 80% similarity. Each cluster resulted in having varied number of soil members, which also included four soil clusters each having only one soil member (Table 4). Clustering was applied using a complete linkage method, squared Euclidean distance, and variables were standardised. Reflectance spectra from 350 to 2500 nm wavelengths for each soil groups are shown in Fig. 6. Distinct differences in the reflectance curves throughout the spectrum could be seen between groups. Samples belonging to the same group shared common properties to some extent, resulting in a moderate degree of similarity among their spectral characteristics. Some differences within a group could be attributed to the constituents of the soil material that were not included in the computation but could be expected to strongly affect the spectral characteristics of a particular soil. For instance, iron content and clay minerals both known to affect the absorption features, were not included in the present study.

3.3. Classification

The results showed that areas classified as calcareous and gypsiferous soils were similar in both classifications (Fig. 7 and Table 5). The linear unmixing results showed more area (1113 ha) under desert pavement as compared to that of SAM classification. SAM result showed 16% of the total area under saline conditions whereas it was negligible (<1%) in case of linear unmixing. The unknown (unclassified) area in SAM classification was 22% whereas it was 36% in linear unmixing. The unknown pixels were the ones, which fell beyond the threshold

Table 2
Correlation matrix of electrical conductivity (EC_e) and absorption feature parameters at selected absorption bands

<table>
<thead>
<tr>
<th>Wavelength (10 nm intervals)</th>
<th>Wavelength</th>
<th>Depth</th>
<th>Width</th>
<th>Area</th>
<th>Asymmetry</th>
</tr>
</thead>
<tbody>
<tr>
<td>390–400</td>
<td>−0.431/0.015</td>
<td>0.346/0.057</td>
<td>−0.325/0.075</td>
<td>0.269/0.144</td>
<td>−0.176/0.342</td>
</tr>
<tr>
<td>615–625</td>
<td>0.437/0.462</td>
<td>0.997/0.000</td>
<td>0.432/0.467</td>
<td>0.603/0.282</td>
<td>0.344/0.570</td>
</tr>
<tr>
<td>685–695</td>
<td>−0.222/0.566</td>
<td>0.279/0.467</td>
<td>−0.168/0.666</td>
<td>0.173/0.656</td>
<td>−0.322/0.397</td>
</tr>
<tr>
<td>800–810</td>
<td>0.343/0.251</td>
<td>0.891/0.000</td>
<td>−0.901/0.000</td>
<td>−0.526/0.065</td>
<td>−0.845/0.000</td>
</tr>
<tr>
<td>950–960</td>
<td>0.531/0.114</td>
<td>0.593/0.071</td>
<td>−0.009/0.981</td>
<td>0.637/0.048</td>
<td>−0.443/0.200</td>
</tr>
<tr>
<td>1410–1420</td>
<td>0.127/0.495</td>
<td>0.054/0.773</td>
<td>−0.176/0.343</td>
<td>−0.115/0.537</td>
<td>−0.159/0.393</td>
</tr>
<tr>
<td>1935–1945</td>
<td>0.233/0.296</td>
<td>0.420/0.052</td>
<td>−0.066/0.771</td>
<td>0.101/0.655</td>
<td>−0.186/0.408</td>
</tr>
<tr>
<td>2350–2360</td>
<td>0.365/0.137</td>
<td>−0.089/0.726</td>
<td>−0.067/0.790</td>
<td>0.112/0.658</td>
<td>−0.288/0.246</td>
</tr>
</tbody>
</table>

Table 3
Correlation matrix of colour parameters and absorption feature parameters at 380–410 nm wavelengths

<table>
<thead>
<tr>
<th>Colour parameters</th>
<th>Wavelength</th>
<th>Depth</th>
<th>Width</th>
<th>Area</th>
<th>Asymmetry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hue</td>
<td>−0.169/0.310</td>
<td>−0.098/0.557</td>
<td>−0.441/0.006</td>
<td>−0.141/0.398</td>
<td>0.106/0.527</td>
</tr>
<tr>
<td>Value</td>
<td>−0.068/0.685</td>
<td>0.608/0.000</td>
<td>−0.012/0.942</td>
<td>0.569/0.000</td>
<td>−0.451/0.005</td>
</tr>
<tr>
<td>Chroma</td>
<td>0.239/0.148</td>
<td>0.363/0.025</td>
<td>0.181/0.277</td>
<td>0.376/0.020</td>
<td>−0.337/0.039</td>
</tr>
</tbody>
</table>

Cell contents: Pearson product moment correlation coefficient/P-value.
Table 4
Mean values and class range of soil properties for different soil groupings

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Clay (%)</th>
<th>Silt (%)</th>
<th>Fine sand (%)</th>
<th>Coarse sand (%)</th>
<th>pH (H₂O)</th>
<th>pH (KCl)</th>
<th>ECₑ (dS m⁻¹)</th>
<th>Gypsum (%)</th>
<th>CaCO₃ (%)</th>
<th>OC (%)</th>
<th>OM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>Mean 28.94</td>
<td>46.38</td>
<td>20.10</td>
<td>4.59</td>
<td>7.85</td>
<td>7.54</td>
<td>2.305</td>
<td>10.27</td>
<td>29.0</td>
<td>0.45</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>Range 19.44–38.45</td>
<td>40.13–52.63</td>
<td>7.20–33.00</td>
<td>1.72–7.46</td>
<td>7.75–7.94</td>
<td>7.49–7.58</td>
<td>2.280–2.330</td>
<td>9.75–10.79</td>
<td>26.7–31.3</td>
<td>0.28–0.61</td>
<td>0.5–1.0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>Mean 12.76</td>
<td>50.73</td>
<td>32.77</td>
<td>3.74</td>
<td>8.57</td>
<td>8.06</td>
<td>1.038</td>
<td>1.662</td>
<td>12.6</td>
<td>0.26</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Range 10.58–17.07</td>
<td>34.35–67.64</td>
<td>14.71–51.66</td>
<td>0.59–12.52</td>
<td>8.26–8.79</td>
<td>7.90–8.22</td>
<td>0.271–2.800</td>
<td>1.00–2.48</td>
<td>6.9–18.6</td>
<td>0.10–0.30</td>
<td>0.3–0.7</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>Mean 7.21</td>
<td>22.74</td>
<td>52.43</td>
<td>17.62</td>
<td>8.13</td>
<td>7.71</td>
<td>1.254</td>
<td>3.02</td>
<td>7.7</td>
<td>0.40</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>Range 5.35–8.38</td>
<td>16.47–26.52</td>
<td>45.85–57.31</td>
<td>11.19–25.88</td>
<td>7.95–8.42</td>
<td>7.51–7.89</td>
<td>0.194–2.320</td>
<td>1.00–6.63</td>
<td>2.1–</td>
<td>0.28–0.56</td>
<td>0.5–1.0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>Mean 9.60</td>
<td>27.39</td>
<td>44.00</td>
<td>19.00</td>
<td>8.54</td>
<td>8.09</td>
<td>0.265</td>
<td>1.15</td>
<td>17.10</td>
<td>0.71</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>Range 6.12–11.54</td>
<td>26.38–28.13</td>
<td>41.70–45.69</td>
<td>16.11–20.52</td>
<td>8.44–8.66</td>
<td>8.05–8.13</td>
<td>0.219–0.312</td>
<td>1.00–1.44</td>
<td>10.6–22.8</td>
<td>0.62–0.81</td>
<td>1.1–1.4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>Mean 5.10</td>
<td>19.07</td>
<td>47.27</td>
<td>28.56</td>
<td>8.64</td>
<td>8.24</td>
<td>0.498</td>
<td>1.05</td>
<td>7.4</td>
<td>0.36</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>Range 2.18–8.92</td>
<td>10.74–27.75</td>
<td>35.84–54.21</td>
<td>14.87–39.53</td>
<td>8.24–8.91</td>
<td>8.00–8.46</td>
<td>0.175–3.260</td>
<td>1.00–1.65</td>
<td>2.1–16.9</td>
<td>0.20–0.53</td>
<td>0.3–0.9</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>Mean 10.31</td>
<td>22.44</td>
<td>58.44</td>
<td>8.80</td>
<td>9.51</td>
<td>8.44</td>
<td>0.281</td>
<td>1.00</td>
<td>11.1</td>
<td>0.22</td>
<td>0.4</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>Mean 12.60</td>
<td>42.52</td>
<td>41.79</td>
<td>3.09</td>
<td>8.88</td>
<td>8.05</td>
<td>24.100</td>
<td>4.27</td>
<td>15.7</td>
<td>0.52</td>
<td>0.9</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Mean 6.27</td>
<td>29.35</td>
<td>45.83</td>
<td>18.54</td>
<td>7.63</td>
<td>7.49</td>
<td>2.350</td>
<td>95.58</td>
<td>2.5</td>
<td>0.40</td>
<td>0.7</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>Mean 12.94</td>
<td>34.03</td>
<td>42.50</td>
<td>10.53</td>
<td>8.05</td>
<td>7.81</td>
<td>2.310</td>
<td>15.75</td>
<td>18.1</td>
<td>1.63</td>
<td>2.8</td>
</tr>
</tbody>
</table>
limits. Both techniques showed classification problems: SAM classification result showed that saline soils occurred in all the geomorphic units (Table 6). In the Tabernas area, the development of salinity was due to evaporation of ground water, which rises to the surface by capillary action. It is thus very unlikely that saline soils occur in the highlands, e.g. hills or piedmonts. It appeared that SAM classification had more overlapping cases between the classes while linear unmixing underestimated the salinity problem since only pixels having dominant abundance (more than 0.50) were classified.

The error matrices are shown in Tables 7 and 8. Overall accuracy of linear unmixing classification seemed to be better (0.75) as compared to that of SAM (0.60). The $\kappa$ value, which took into account not only complete agreement between the ground truths but also the agreements by chance, showed that a large portion of the class agreement for SAM could be due to chance agreement since its $\kappa$ value (0.45) was lower than its overall accuracy. For linear unmixing, the $\kappa$ value was higher (0.63).

### Table 5
Classification results

<table>
<thead>
<tr>
<th>Classes</th>
<th>SAM Area (ha)</th>
<th>Percentage</th>
<th>Linear unmixing Area (ha)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desert pavement</td>
<td>805</td>
<td>10</td>
<td>1113</td>
<td>14</td>
</tr>
<tr>
<td>Saline soil</td>
<td>1247</td>
<td>16</td>
<td>2</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Calcareous soil</td>
<td>2827</td>
<td>36</td>
<td>2610</td>
<td>34</td>
</tr>
<tr>
<td>Gypsiferous soil</td>
<td>1204</td>
<td>16</td>
<td>1252</td>
<td>16</td>
</tr>
<tr>
<td>Unknown</td>
<td>1732</td>
<td>22</td>
<td>2838</td>
<td>36</td>
</tr>
</tbody>
</table>

Fig. 6. Spectral reflectance curves of soils belonging to different clusters: A for group 1; B for group 2; C for group 3; D for group 4; E for group 5; and F for groups with one soil member.
To test whether the two classification results were significantly different, the method described by Cohen (1960), and elaborated by various authors (Congalton et al., 1983; Skidmore, 1999; and Rossiter, 2004) was used. The method uses the normal curve deviate statistics ($z$) and the $\kappa$ values ($\kappa_1$, $\kappa_2$) and their associated variance ($\sigma^2_{\kappa_1}$, $\sigma^2_{\kappa_2}$) as follows:

$$z = \frac{\kappa_1 - \kappa_2}{\sqrt{\sigma^2_{\kappa_1} + \sigma^2_{\kappa_2}}}$$  \hspace{1cm} (8)

with $\kappa_1 = 0.45$, $\kappa_2 = 0.63$, $\sigma^2_{\kappa_1} = 0.000479$ and $\sigma^2_{\kappa_2} = 0.000262$, we found that the result was 6.6, which substantially exceeded $Z_t = 1.96$ (at $\gamma = 0.05$). Thus we may conclude here that there was a significant difference between the two classification results.

Table 6

<table>
<thead>
<tr>
<th>Landscape</th>
<th>Desert-like features</th>
<th>SAM (ha)</th>
<th>Linear unmixing (ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hills</td>
<td>Desert pavement</td>
<td>148</td>
<td>252</td>
</tr>
<tr>
<td></td>
<td>Saline soils</td>
<td>483</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Calcareous soils</td>
<td>1412</td>
<td>1447</td>
</tr>
<tr>
<td></td>
<td>Gypsiferous soils</td>
<td>743</td>
<td>673</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>929</td>
<td>1341</td>
</tr>
<tr>
<td>Piedmont</td>
<td>Desert pavement</td>
<td>126</td>
<td>171</td>
</tr>
<tr>
<td></td>
<td>Saline soils</td>
<td>267</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Calcareous soils</td>
<td>673</td>
<td>599</td>
</tr>
<tr>
<td></td>
<td>Gypsumiferous soils</td>
<td>126</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>210</td>
<td>542</td>
</tr>
<tr>
<td>Valley</td>
<td>Desert pavement</td>
<td>509</td>
<td>689</td>
</tr>
<tr>
<td></td>
<td>Saline soils</td>
<td>509</td>
<td>&lt;1</td>
</tr>
<tr>
<td></td>
<td>Calcareous soils</td>
<td>749</td>
<td>554</td>
</tr>
<tr>
<td></td>
<td>Gypsumiferous soils</td>
<td>321</td>
<td>485</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>589</td>
<td>948</td>
</tr>
</tbody>
</table>

Table 7

<table>
<thead>
<tr>
<th>Classification</th>
<th>Desert pavement</th>
<th>Saline soils</th>
<th>Calcereous soils</th>
<th>Gypsumiferous soils</th>
<th>Test pixels</th>
<th>User accuracy</th>
<th>Commission error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desert pavement</td>
<td>149</td>
<td>65</td>
<td>33</td>
<td>2</td>
<td>249</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>Saline soils</td>
<td>28</td>
<td>29</td>
<td>4</td>
<td>0</td>
<td>61</td>
<td>0.48</td>
<td>0.52</td>
</tr>
<tr>
<td>Calcereous soils</td>
<td>17</td>
<td>26</td>
<td>59</td>
<td>4</td>
<td>106</td>
<td>0.56</td>
<td>0.44</td>
</tr>
<tr>
<td>Gypsumiferous soils</td>
<td>0</td>
<td>7</td>
<td>38</td>
<td>95</td>
<td>140</td>
<td>0.68</td>
<td>0.32</td>
</tr>
<tr>
<td>Total</td>
<td>194</td>
<td>127</td>
<td>134</td>
<td>101</td>
<td>556</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Producer accuracy</td>
<td>0.77</td>
<td>0.23</td>
<td>0.44</td>
<td>0.94</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Omission error</td>
<td>0.23</td>
<td>0.77</td>
<td>0.56</td>
<td>0.06</td>
<td></td>
<td></td>
<td>$\kappa$ value = 0.45</td>
</tr>
</tbody>
</table>

Fig. 7. Classification results.
4. Conclusion

It can be concluded that the selection of endmembers is of vital importance for adequate hyperspectral classification when applied to soils. Special care should be taken when selecting soil-characteristic spectra as soil reflectance is governed by many factors and their combined effects play an important role, especially in the way they manifest themselves under field conditions. Soils representing similar properties may show varying reflectance patterns. In considering these, selecting end-members based on homogenous areas representing certain soil processes of interest, therefore seems to be most appropriate for application in soils.

On the other hand, soil-landscape relationships are often well known and actively used by the soil surveyor for efficient mapping. A soil map unit may consist of associations of soils, which may be similar in terms of surface properties yet fall in different taxonomic units. A given soil taxonomic unit again allows some variation of its properties from its main concept. Since linear unmixing helps map abundances, it seems thus logical to use this technique for the purpose of mapping soils. Although spectral angle mapper algorithm is simple and insensitive to the effect caused by topographic variations there is a problem of overlap between classes, similar to the problem associated with box classification of multispectral broadband data.

Acknowledgements

The authors would like to acknowledge two anonymous reviewers for their comments. Rob Hennemann is thanked for improving the language.

References


