Geologie en Mijnbouw **77:** 153–160, 1999.
© 1999 *Kluwer Academic Publishers. Printed in the Netherlands.*

Spectral mixture analysis for mapping land degradation in semi-arid areas

Maarten Tromp & Gerrit F. Epema
Centre for Geo-Information, Wageningen Agricultural University, PO Box 339, 6700 AH Wageningen, the Netherlands

Received 1 February 1997; accepted in revised form 17 December 1998

Key words: Burkina Faso, end-members, remote sensing

Abstract

Spectral mixture analysis (SMA) is an image-processing technique used for the analysis of airborne hyperspectral remote-sensing data which consist of a large number of spectral bands, typically over 100. In this paper the possibilities are examined of using SMA for the analysis of Landsat Thematic Mapper satellite data with only six bands in the visible to shortwave-infrared wavelength. We use data from a semi-arid area in the Sanmatenga province of Burkina Faso, an area known to experience land-degradation problems. In SMA, we assume that pixels in an image consist of one or more homogeneous (uniform in character) or pure surfaces, the so called end-members. The end-members were derived from the image data on the basis of specific image characteristics. Field data was collected to describe the characteristics of these end-members in terms of land cover, soil and degradation phenomena. The end-members derived from the image analysis, although statistically pure in terms of image spectral characteristics, prove to be mixtures at a field scale. This lack of purity is explained by the nature of semi-arid areas which is more heterogeneous than that of most other areas. The SMA yielded cover percentages for the end-members from which an unsupervised classification was made. Comparison of the classification with the results of SMA shows that the latter provides information on the amount and spatial distribution of land characteristics such as land degradation.

Introduction

Land degradation is a serious problem in semi-arid areas such as the Sanmatenga province of Burkina Faso (west Africa). Information on land degradation is sparsely available and mapping of degradational features is time-consuming and expensive. Remote sensing, using Landsat Thematic Mapper (TM) or SPOT data, is a promising tool in providing information on the spatial and temporal degradational state of the land because these satellite observations allow a synoptic view over large areas. Traditional image-processing techniques applied to satellite data, including pixel-based classification, are not appropriate because spatial variability of parameters controlling degradation processes occur at scales smaller than the image spatial resolution. Epema & Bom (1994) concluded, on the basis of a geostatistical analysis in a semi-arid area in Niger, that pixel sizes smaller than 10 m would be required to characterize degradational processes in sufficient detail. An alternative approach to analyze Landsat TM data, which have a spatial resolution of 30 by 30 m, is spectral mixture analysis (SMA). SMA strives at finding the sub-pixel abundance of a set of pure materials (i.e. end-members) present within the area contained by the pixel. The end-member selection process is a very important step in the mixture modeling process. The aim of the paper is to evaluate SMA in semi-arid areas and to investigate different selection methods to derive end-members. Use is made of the software package SIPS, the Spectral Image Processing System (CSES 1993, Kruse et al. 1993).

The study area

The Sanmatenga province in Burkina Faso (Figure 1) is located on the old crystalline formations of the Pre-Cambrian. In the northern part of the province, granites and some migmatite of ante-Birimian age (i.e.

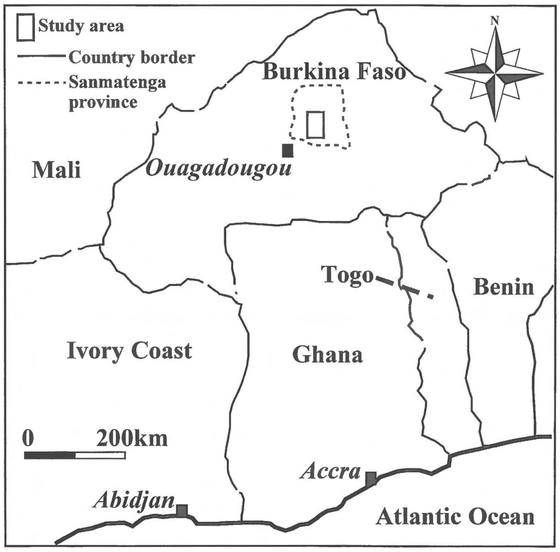

Figure 1. Location of the Sanmatenga study area within Burkina Faso.

3 billion years old) are outcropping. In the southern part of the area the results of different tectonic periods throughout the Birimian cause a complex undulating terrain. In the Eocene until Pliocene the area experienced strong weathering and leaching resulting in plinthitisation on the slopes, now present as elevated laterite caps in the landscape (Hottin & Ouedraogo 1992). The plinthitisation in combination with ferrolysis, clay transport, erosion and leaching during the Pleistocene and Holocene resulted in large differences in soil characteristics. Variations in texture and in available nutrients have resulted in differences in fertility and in vulnerability to erosion. The type and amount of clay minerals, the organic matter and the slope determine the erodibility of the soils.

The northern part of the Sanmatenga province shows flat terrain with footslopes on the laterite plateaus which reach a maximum height of 20 m. The large diversity in geomorphology, land-use and land-cover results in a high spatial variability over short distances of parameters that influence land-degradation processes.

The land-use in the area is a combination of intensively used plots around houses, extensively used arable land at larger distances from the houses and areas used for cattle grazing. Nutrient content of the soils and variability in moisture content are the main constraints for agriculture.

A study site of 11 by 16 km to serve as type locality for the southern part of the Sanmatenga province was selected. Large-scale aerial photographs (scale 1 : 20 000) and ancillary field information were used to map land-use systems. A Landsat TM satellite image acquired over the area on 14 May 1994, was used for the end-member selection and for SMA.

Theory of spectral mixture analysis (SMA)

The ability to identify objects in remotely sensed data depends on the definition of the objects, the complexity of the mapped area and the spatial and spectral resolution of the sensor. Objects often have dimensions smaller than the spatial resolution of the sensor. Therefore the spectral response of ground resolution cells (i.e. pixels) is the result of the interaction of different objects, each with its own spectral characteristics. Reflectance spectra can be modeled as mixtures of pure end-member spectra (Adams et al. 1989). The spectral variation observed in an image often can be modeled with spectra of a small number of surface materials typically related to soil types, vegetation types and relief shading. Spectral mixture analysis is mostly used for the analysis of hyperspectral sensor data (i.e. sensors that acquire imaging data in many contiguous spectral channels; see Van der Meer, this issue) such as AVIRIS and GERIS (e.g. Kruse et al. 1993). SMA calculates the coverage of individual objects within the mixed pixel. The result is a set of images portraying, for each end-member, the aerial coverage. Furthermore, an error image is calculated as the difference of the modeled pixel spectrum and the measured pixel spectrum. The SMA technique allows to more accurately specify the distribution of objects within land units rather than to detect what is present inside a specific pixel (Smith et al. 1990). Three constraints apply to SMA:

1. The number of end-members must be equal to the number of spectral bands minus one resulting from the multiple regression analysis used to deconvolve the data set. Using Landsat TM data, the number of end-members is thus limited to five;
2. Each end-member has a coverage between 0 and 1; and
3. The sum of the coverage of all end-members is 1.

In a mathematical form the model and constraints can be written as:

$$R_i = \sum_{j:=1}^{m} (F_j * Re_{ij}) + e_i \text{ and } \sum_{j:=1}^{m} F_{j=1} \text{ and}$$
$$0 \leq F_j \leq 1$$

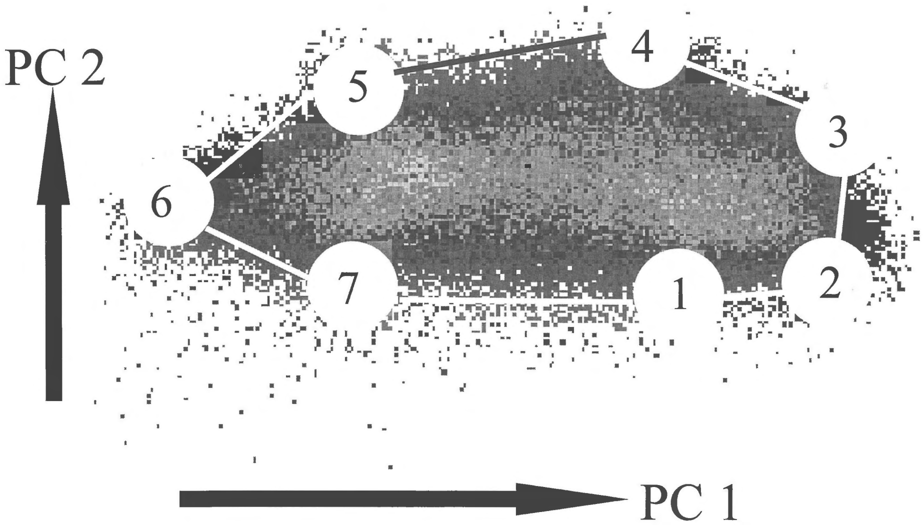

Figure 2. Example of the procedure used to select end-members in the analysis of the Landsat TM data from Sanmatenga. End-members in a principal component feature space appear at the corners on the perimeter of a volume bounding the scatter of points. The scattergram of principal components 1 and 2, and the spanning polygon showing the location in the PCA feature space of the end-members, numbered 1 to 7, are shown.

where R_i is the reflectance of the mixed spectrum in band i, Re_{ij} is the reflectance in band i of end-member j, F_j is the fraction.

Two methods are generally applied in order to collect end-members: 1) end-members can be collected from spectral measurements in the field or laboratory, or 2) end-members can be derived from the image data using principal component analysis (PCA). A disadvantage of the first method (Tromp & Steenis 1996) is that much effort is needed to gather and atmospherically correct field or laboratory spectra. Examples of the second method can be found in Smith et al. (1985) and Bryant (1996). Bryant (1996) demonstrates good results with the method in a playa area in Tunisia. The first step of his approach involves the calculation of principal components for the Landsat TM data. Principal components are orthogonal and linear transformations of the original image values, in which the N original, highly correlated bands (i.e. six bands in the case of Landsat TM) are transformed into N new, un-correlated bands (principal components) by successive rotations of the axes of the original bands. The first two principal components, typically explaining more than 90% of the original variance, are plotted in a scattergram. The corners of the polygon bounding the scatter of data points, represent the location of the purest, and thus end-member pixels.

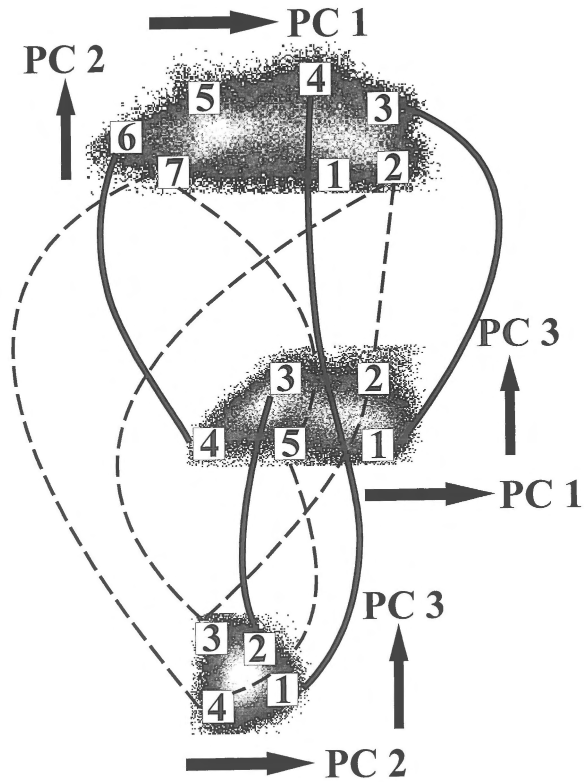

Figure 3. Relations between end-members of the Sanmatenga site, found in different scattergrams of the principal component axes. Thick lines represent strong similarity between the end-members, while thin lines express weak relationships.

Table 1. End-members from scattergram with principal components 1 and 2.

End-member	Soil surface	Soil material	Land use	Vegetation	Erosion	End-member purity*
1	Crusted layer with a bright (reddish yellow) colour	Varying from clayey to sandy	No use, bare soil	Mostly bare, few shrubs or trees	Strongly eroded, deep gullies	Low
2	Bright yellowish brown to greyish brown, slaking crust with some fine sand on top	Silty with low amounts of clay and fine sand and some more clayey materials in the southern part	Arable land	Here and there some trees and shrubs	Some sheet erosion	Intermediate
3	Red colored, thin slaking crust largely covered with a thin layer of white, fine sand	Silt with fine sand and very low amounts of clay	Arable land (mainly millet)	Scarce trees: 3% 3–7%, shrubs	Some sheet erosion	Intermediate
4	Red or reddish brown crust, sometimes covered with some brighter fine sand	Silt and clay with a little fine sand (silty clay)	Arable land	Little natural vegetation, except for fallow fields	Some sheet erosion	High
5	Recently burned surface: burned vegetation, ashes, dark ironstone and ironstone gravel	N/a	N/a	Burned vegetation	N/a	High
6	Shadow or internal shaded surfaces	N/a	N/a	N/a	N/a	Intermediate
7	Varying	Varying	No use	Densely vegetated	N/a	High

* End-member purity gives an indication of the homogeneity within the end-member pixels. If a selected end-member contains several different components within itself, these components should be treated as separate end-members.

During fieldwork, field-reflection data were gathered using a fieldspectrometer in order to spectrally correct the Landsat data. Furthermore, we followed the approach of Bryant (1996) to analyze our data using the following steps:

1. The first three principal components, explaining most of the original variance, were plotted in scattergrams. Based on the method of Smith et al. (1985) and Bryant (1996) seven end-members were identified;
2. In the field the locations of the end-members were visited and their land surfaces described. For verification, some field reflectance measurements were performed (see Tromp & Steenis 1996);
3. Based on the field observations a selection was made of the five most suitable end-members explaining the maximum amount of variability for the Landsat TM data;
4. SMA was applied using the five end-members and in addition an unsupervised clustering of the end-member images was carried out; and
5. The different images based on SMA were compared with a traditional maximum likelihood classification of the original bands. As input for this classification, units were selected derived from a traditional mapping of soil and vegetation, concurrent with this research.

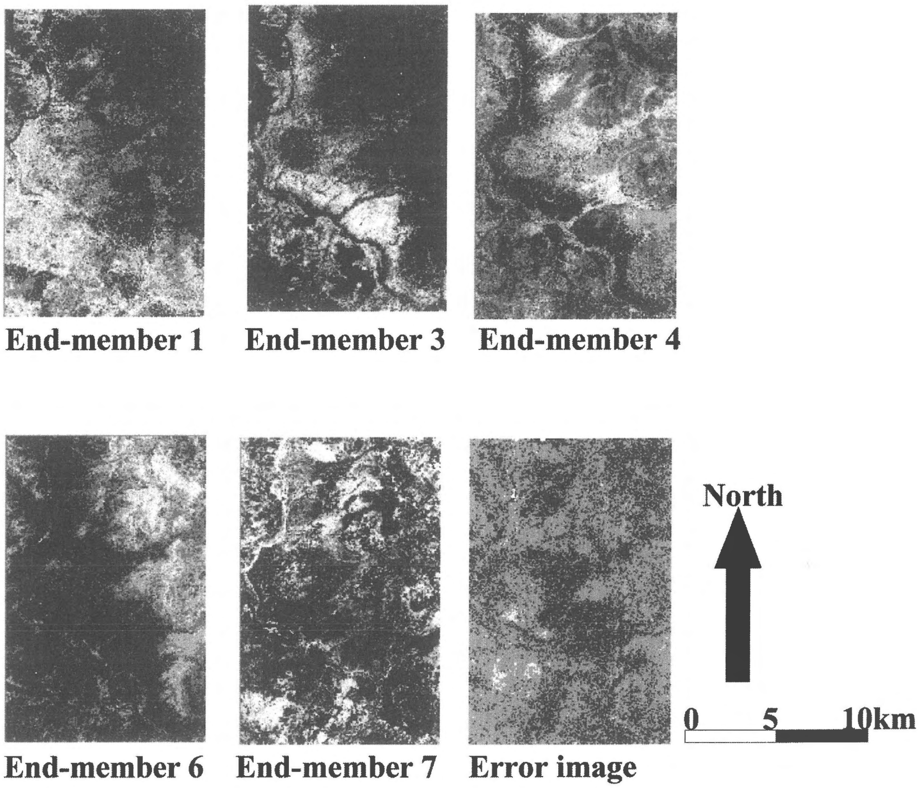

Figure 4. Output images of the spectral mixture analysis, including an error image using a Landsat TM data set acquired in May 1994 over the Sanmatenga site. A description of the end-members is given in Table 1.

Results and discussion

End-member selection

A similar approach as that described by Bryant (1996) was used in order to select end-members. In a scattergram of principal components 1 and 2 (Figure 2) the pixels located at the seven corners in the bounding polygon were selected as pure end-members, thus yielding seven end-members.

Since other principal component axes can also highlight a specific characteristic object or phenomenon of the area, other scattergrams were constructed using combinations of the first three principal components (Figure 3). Most end-members found in these scattergrams represent the same objects found with the scattergram of principal components 1 and 2.

Manual selection of end-members in multi-dimensional space is presented by Bateson & Curtiss (1996). Recently available software packages allow visual selection of pixels in multi-dimensional spaces thus significantly improving the speed and accuracy of end-member selection.

The end-members found with the PCA technique were identified in the image and visited in the field. With the aid of a Global Positioning System (GPS) the locations could be found within a few meters. Soil type, land cover and land use at the sites were described. The end-members we identified in the image as pure objects proved to consist in the field of mixtures of several objects indicating the complexity of the terrain. The aerial coverage of the different surface components was estimated.

Spectral reflection measurements were carried out on the different surface components. These measurements indicated that, especially after convolving the hyperspectral curves to the Landsat TM wavelength bands, the soil spectra strongly resemble each other, differing only in albedo. Using these convolved spectra as end-members in an SMA will introduce large errors (Bateson & Curtiss 1996). A description of the areas represented by the selected end-members is given in Table 1. Although not all end-members are pure, they provide an indication of the surfaces useful in the classification and description of the character-

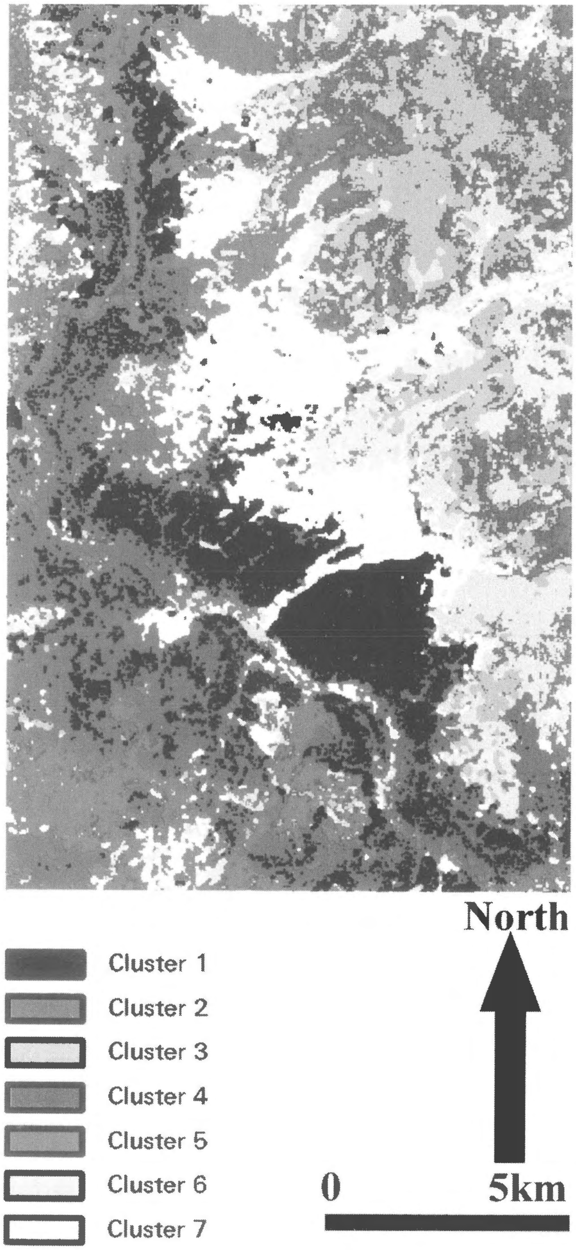

Figure 5. Unsupervised classification with the output images of the SMA using a Landsat TM data set acquired in May 1994 over the Sanmatenga site. (For full colour reproduction of this figure, see Appendix at the end of this issue.)

istics of the area. Described surfaces represent land qualities which reflect the conditions and processes acting upon the area: erosion, variations in organic matter content and differences in parent material.

In the area, the heterogeneity does not allow for 'pure' end-members on basis of the PCA method; each pixel in itself is a mixed pixel. The image e

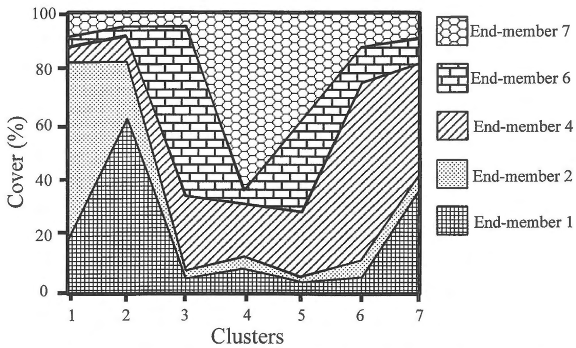

Figure 6. The coverage of each end-member for the clusters used in Figure 5.

members identified using the PCA technique did not coincide with the field end-members described during the field campaign. In case the area of field end-members covers multiple pixels, the PCA technique can detect an image end-member which resembles this field end-member.

End-member results

For the SMA with Landsat TM only five end-members can be selected. Therefore, two of the seven selected end-members were not used. End-members 2, 3 and 5 appeared spectrally very similar, thus we decided arbitrarily to exclude end-members 2 and 5 from the analysis.

The output of the SMA is five images with the coverage (i.e. abundance) for each end-member for all pixels (Figure 4). A white or light color in this image indicates a high coverage of the end-member within the area contained by a pixel. In case of a black or dark color, the coverage of that end-member is low. The end-member images provide insight in the location and distribution of degraded areas and the quantitative amount of degradation.

The error image indicates the quality of the modeled reflection with respect to the measured reflection. The few white areas in this image are locations where end-members 2 or 5 dominate, indicating that these end-members are spatially relatively unimportant. In addition to these end-member images, an unsupervised classification of the output images of the SMA was performed (Figure 5). Clustering of the output images allows to examine the distribution of the end-members within the clusters. Figure 6 shows the percentage of each end-member within the generated clusters. Some clusters such as 1, 2 and 4 contain a majority of a single end-member. As previously men-

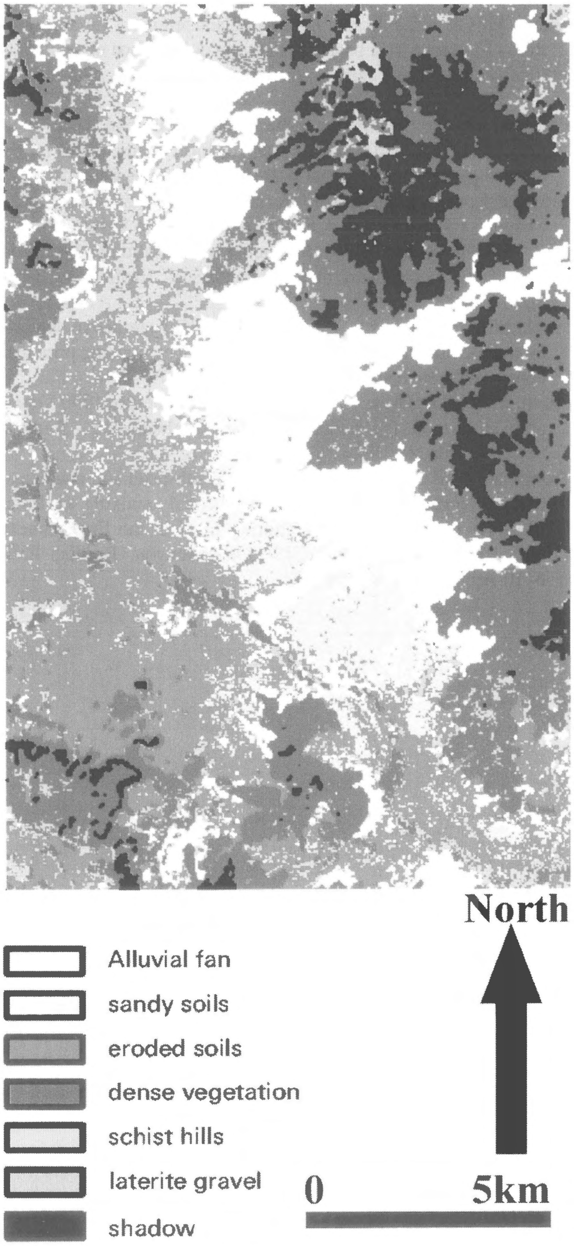

Figure 7. Maximum likelihood classification with the original spectral bands using a Landsat TM data set acquired in May 1994 over the Sanmatenga site. (For full colour reproduction of this figure, see Appendix at the end of this issue.)

tioned, certain end-members were not 'pure' and the heterogeneity of the clusters is therefore higher than presented in Figure 6.

For comparison, a maximum likelihood classification has been carried out using the original band and with the same number of classes as we used to create the clustered image (Figure 7). The maximum likelihood classification generates intrinsic classes with the average spectral response per class. The clustered image based on bands of the SMA, in combination with the end-member output (Figure 4) and the coverage percentage of the end-members (Figure 6), provides more information on the thematic content of the classes. Products of SMA can be used to extract quantitative information on existing mapping units. This kind of information can be used to identify areas with erosion features or potentially more fertile soils.

Recommendations for future research and application

In the previous sections, the results of end-member selection and of SMA were discussed. For the selection of end-members we recommend further studies to investigate:

1. the correlation between field-derived end-members and end-members from SMA;
2. how field-derived end-members can be extracted from images in areas with high spatial variability following the approach of Bateson & Curtiss (1996);
3. different methods for stratification of images prior to end-member selection for example based on the spectral characteristics or on external information such as geological maps;
4. the role of the results of SMA in defining units for mapping at a reconnaissance scale.

Conclusions

End-members derived through the principal component scattergram method do not yield pure end-members in the studied area. Semi-arid areas are very heterogeneous which may explain this observation. For a proper description of end-members, field observations are essential. Field reflectance measurements help in this description although their incorporation implies spectral correction of the data. The different output images of the SMA provide useful information in better understanding the morphology of the area. Furthermore, these images provide information about the amount and the spatial distribution of land characteristics, such as land-degradation features. This is an advantage over traditional image-processing techniques such as supervised classification.

We expect that in the future SMA will become an important image-processing technique in many areas. New satellite missions as for example ESA's ENVISAT (cf. Van der Meer, this issue) will be launched in 2000 carrying a 15 bands imaging spectrometer with a high spectral resolution and low spatial resolution, thus opening many opportunities for SMA.

References

Adams, J.B., M.O. Smith & A.R. Gillespie 1989 Simple models for complex natural surfaces: a strategy for the hyperspectral era of remote sensing. In: Proc. IEEE Internat. Geosci. Remote Sens. Symp. '89, IEEE, New York: 16–21

Bateson, A. & B. Curtiss 1996 A method for manual end-member selection and spectral unmixing – Remote Sens. Environ. 60: 229–243

Bryant, R.G. 1996 Validated linear mixture modeling of Landsat TM data for mapping evaporate minerals on a playa surface: methods and applications – Int. J. Remote Sens. 2: 331–350

CSES 1993 SIPS User's guide, Spectral Image Processing System Version 1.2 – Center for the Study of Earth from Space (CSES), Boulder, 122 pp

Epema, G.F. & B.C.J. Bom 1994 Spatial and temporal variability of field reflectance as a basis for deriving soil surface characteristics from multiscale remote sensing data in Niger – ITC Journal 1994(1): 23–28

Hottin, G. & O.F. Ouedraogo 1992 Notice explicative de la Carte Géologique au 1 : 100 000 du Burkina Faso – Bureau des mines et de la géologie du Burkina BUMIGEB, Burkina Faso, 50 pp

Kruse, F.A., A.B. Lefkoff, J.W. Boardman, K.B. Heidebrecht, A.T. Shapiro, P.J. Barloon & A.F.H. Goetz 1993 The Spectral Image Processing System (SIPS); Interactive visualisation and analysis of imaging spectrometer data – Remote Sens. Environ. 44: 145–163

Smith, M.O., P.E. Johnson & J.B. Adams 1985 Quantitative determination of mineral types from reflectance spectra using principal components analysis – J. Geophys. Res. 90: 797–804

Smith, M.O., S.L. Ustin, J.B. Adams & A.R. Gillespie 1990 Vegetation in deserts. I: a regional measure of abundance – Remote Sens. Environ. 31: 1–26

Tromp, M. & M.Z. Steenis 1996 Deriving sub-pixel soil characteristics in Northern Burkina Faso with spectral unmixing. In: Proc. ISSS Internat. Sympos. 'Monitoring Soils in the Environment with Remote Sensing and GIS'. Ouagadougou, Burkina Faso 6–10 Feb. 1995: 269–284