Deriving Sub-Pixel Soil Characteristics in Northern Burkina Faso with Spectral Unmixing

Maarten TROMP, Marcel Z. STEENIS

Department of Soil Science and Geology, Wageningen Agricultural University
P.O. Box 37, 6700 AA Wageningen, The Netherlands
E-mail: maarten.tromp@algemeen.beng.wau.nl

Abstract

In this paper the possibilities and advantages of spectral unmixing for information extraction of Landsat TM data are analysed. Raw image data are transferred to field reflectance data with the empirical line method. "Harmattan dust" blown from the Chad basin to the Sahel during the dry periods had, although not visible on hardcopies, a large influence on the reflected radiance represented by the Landsat TM image of May 7, 1988. Resampled field reflectance spectra could be used for the spectral unmixing of Landsat TM. The results of the spectral unmixing resembled the field estimations closely. Units of research showed a tendency of increased coverage with dead grass at the expense of sand in the period between the date of image acquisition and the fieldwork. This corresponded with the presence of deserted agricultural fields due to decreased fertility after intensive agricultural use.

Résumé

Ce rapport fait l'analyse des possibilités et des avantages de la méthode de déconvolution spectrale pour l'interprétation des images Landsat TM. Les données brutes de l'image sont transformées en valeurs de réflectance de terrain par la méthode de la ligne empirique. Sans que cela soit visible sur les *hardcopies*, la poussière d'harmattan emportée du bassin du Tchad au Sahel par le vent pendant les périodes sèches avait une grande influence sur la radiation réfléchie saisie dans l'image Landsat TM du 7 mai 1988. Les spectres de réflectance de terrain ont été utilisés pour l'analyse

des images Landsat TM. Les résultats de la déconvolution spectrale se rapportent bien aux estimations obtenues dans le terrain. Les unités de recherche ont montré une tendance à une couverture d'herbes mortes au détriment du sable nu pendant les dernières années. Cette tendance correspond à la présence des champs agricoles abandonnés à cause d'une fertilité diminuée après une culture intensive.

Introduction

Population growth and man-induced climatic change makes life in the Sahelian countries even tougher. The increased pressure on land is a threat to the ecosystem of the sahel. Overgrazing and decreased soil fertility lead to erosion and poor harvests. To tackle these serious problems, it is necessary to find methods to map these degradation features. This can also help to understand the different processes.

Degradation of land often expresses itself in a change in the amount and distribution of earth-surface properties. For example, increased degradation in semi-arid areas leads to an increased area of crusted surfaces. Traditional image classification techniques do not give quantified measurements of soil surface characteristics.

Desertification is one of the major causes of degradation in the Sahelian countries. It expresses itself as an increase in erosion and a decrease in the vegetated area.

The processes of degradation are often active on a sub-pixel scale. The spatial variation of most landunits is very high, the range of the semivariogram of land qualities and reflectance in these areas is 7 to 10 meters. With pixels larger than 10 meters it is not possible to map terrain differences within the mapping units (EPEMA and BOM, 1994). Special image processing techniques are required to extract this detailed information from satellite images. In this research the possibilities of spectral unmixing with broad available Landsat TM data are examined.

The study area

The study area is located in the department Sanmatenga. The department was mapped as part of a co-operation of the *Antenne Sahélienne* of the University of Wageningen (The Netherlands), and the development organisation PEDI, financed by the Dutch government. Apart from the research presented in this paper, a soil map, vegetation map and land evaluation map are produced at a scale of 1:100,000. The mapping is carried out with the aid of multitemporal satellite images.

The area covers mainly the geologic time zones Birrimien and anti-Birrimien. The main rock of the Birrimien is schist and of the anti-Birrimien granite and granodiorite.

Due to past and present ferralitisation, there are hardened plinthite plateaus and layers throughout the area (ELKENBRACHT *et al.*, 1995). The ferralitisation in combination with ferrolyse, clay transport, erosion and leaching results in large differences in soil characteristics in the horizontal and vertical. Differences in texture and available nutrients lead to differences in fertility and vulnerability to erosion. The soils of the Birrimien are more fertile than those of the anti-Birrimien. The type and amount of clay minerals, organic matter and slope determine the erosivity of soil units.

Spectral characteristics of soil and vegetation

Each object on earth has its own spectral reflectance due to the chemical composition of the object. Objects can be discriminated by their spectral reflectance, caused by the reflectance of solar radiation. Light emitted by the sun interacts with the object and part of that light is reflected in the direction of the sensor. Various algorithms have been developed to differentiate between different objects. In the next part the different characteristics, which influence the reflectance of soils, green and dead vegetation, are briefly summarised.

Soils

The reflectance of soils depends on texture, structure, minerals, organic matter and water content. Some characteristics lead to an overall decrease in reflectance, others absorb radiation in a more specific wavelength. With hyper spectral (airborne and in the future spaceborne) radiometers it is possible to detect individual absorption features. Satellites, like Landsat TM, measure the reflectance in broad absorption bands, hiding these features. Absorption features, present in bare soil spectra, are mostly caused by iron oxides and OH-bearing minerals. Iron oxides cause a broad absorption dip around 900 nm, OH and CO_3-bearing minerals more narrow dips between 2,000 and 2,500 nm.

Healthy vegetation

Healthy vegetation absorbs a lot of radiance in the visible part of the spectrum for photosynthesis. The absorption is lowest in green (500 nm), therefore vegetation is green for the human observer. The reflectance in the near infrared is very high, due to the physiological structure of the leaves.

Dead vegetation

Dead vegetation lacks the high absorption in the visible, and the reflectance is highest around 1,600 nm (Landsat TM band 5). The spectral reflectance gently increases until 1,600 nm after which it decreases.

Data sources

Field data

The spectral reflectance of different surfaces is measured in the field with the mini-IRIS RT (GER2100) field spectrometer of the Geophysical Environmental Research corporation (GER, 1992). This instrument measures in 140 wavelength bands between 400 and 2,500 nm. It is a single field of view instrument, so a calibration plate is used to determine the incoming radiation. Depending on the research, there are two measurement types:

- 76 bands between 400 and 1,100 nm, 64 bands between 1,100 and 2,500 nm;
- 76 bands between 400 and 1,100 nm, 64 bands between 2,000 and 2,500 nm.

The advantage of grating 2 is the very high spectral resolution (8 nm) in the part of the spectrum where the discrimination of different minerals is possible.

The viewing angle is 6.5 degrees. A 120 cm instrument height results in a pixel size of 27 x 27 cm.

Image data

For this research it was possible to use different Landsat TM images. Available were images of the following dates: January 10, 1991; May 7 1988 and September 20, 1988. Since most interest went to bare soil surfaces, the January and May images, both of the dry season, were found most suitable. Bare soil differences are well expressed in those images.

Data processing and discussion

Resampling of field spectra

To use field spectra as endmember for unmixing of Landsat TM data, it is necessary to resample the spectra to TM wavelength bands. The overall form of the spectral curve is preserved but the specific absorption features, detectable with the field spectrometer, are not visible after resampling. In figure 1, typical spectral curves of objects present in the study area with the positions of the TM wavelength bands are shown.

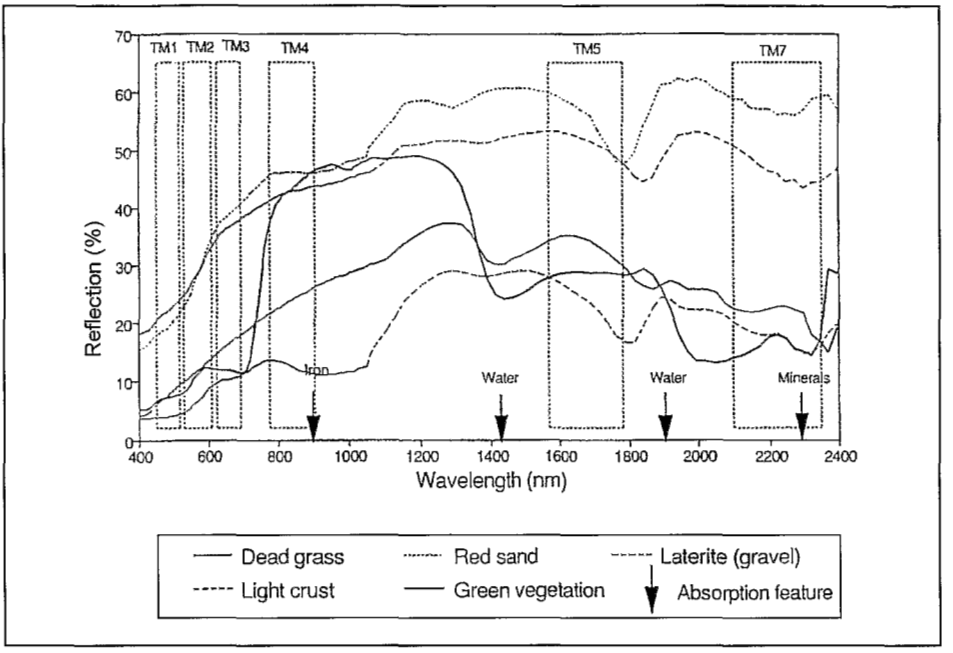

Figure 1. Different spectral curves of surfaces in the study area.

Geometric and spectral correction of image

Spectra collected in the field can only be used when correlation between the satellite data and the field data is high. To determine this correlation, the "empirical line method" is used. This, mathematically, relative simple spectral correction method proved to be very reliable (e.g. CONEL *et al.*, 1987 and FARRAND *et al.*, 1994). The regression parameters can be used to calculate field reflectance with the dimensionless Digital Number (DN) values of the satellite data.

Resampled field reflectance spectra of different surfaces are collected in the field and compared with the Landsat TM data. The location of the areas of interest is determined visually on hardcopies of the image and with the aid of a Global Position System. Areas with high and low reflectance are used to find the gain and offset for each of the Landsat TM bands. In figures 2 and 3, two examples of the correlation between mini-IRIS data and Landsat TM data of May 1988 are shown.

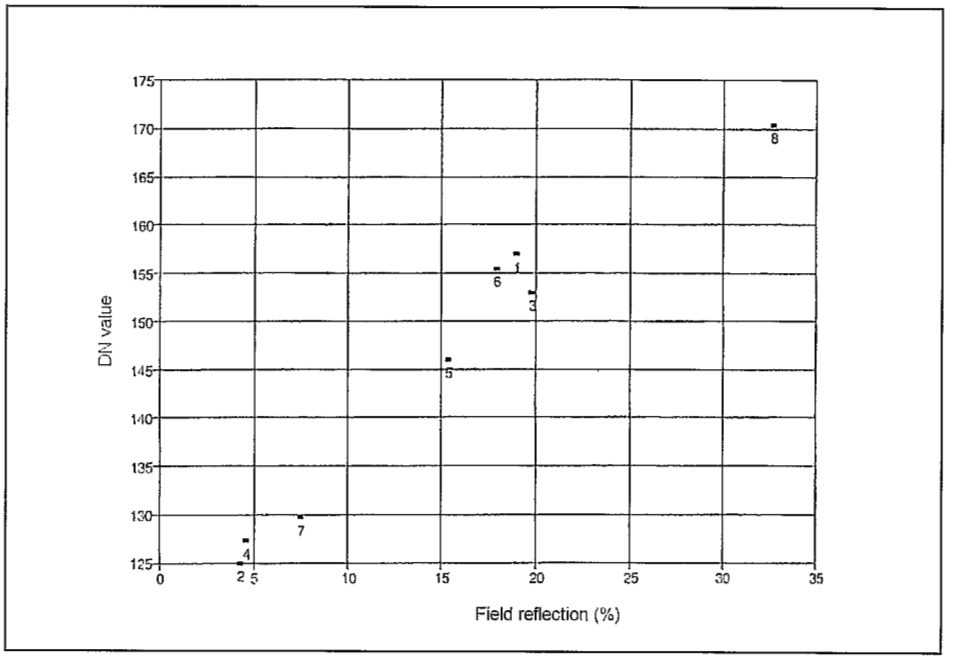

Figure 2. Correlation between Landsat band 1 of May 7, 1988 and field reflectance.

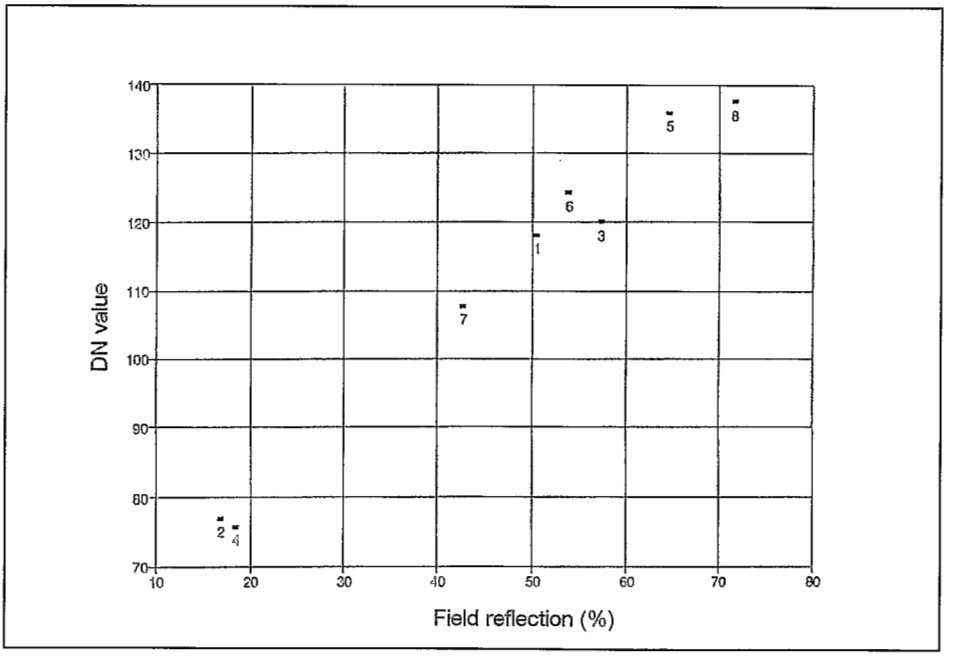

Figure 3. Correlation between Landsat band 7 of May 7, 1988 and field reflectance.

The regression parameters found with the empirical line method seemed very good. However, large differences were found between field reflectance spectra and image spectra outside the area where the training pixels were sampled. In spite of the fact that no albedo differences were visible on hardcopies of the May 1988 image, there were considerable differences in DN values of the same objects in different locations in the image. After careful examination of the raw TM data, differences of 5 - 25 DN were found. Surfaces with a low albedo, located in the North of the area, showed the largest differences. The DN differences were highest in band 5. In figures 4 and 5 respectively the DN values of laterite (gravel)- and crusts-pixels are shown for various locations. The locations of the villages range from Kaya in the South to Dablo in the North. The differences of light coloured, crusted surfaces are less than the differences of the laterite (gravel) surfaces. The differences are probably caused by the harmattan dust. This dust is blown from the Chad basin to the Sahel during the dry periods (MCTRAINSH and WALKER, 1982). Because the dust is very light coloured, the influence on the reflection is higher above dark surfaces.

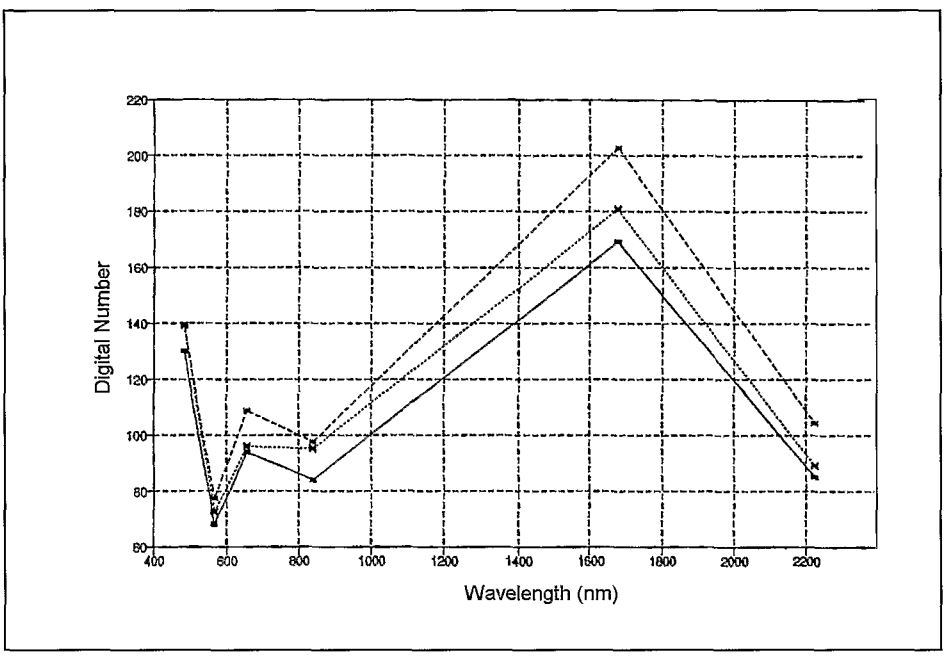

Figure 4. Differences in reflectance of (laterite) gravel due to the influence of harmattan dust in the air.

For the unmixing process Landsat TM bands 1, 2, 3, 4, 5 and 7 of January 10 1991 are used. In this image no influences of the harmattan dust were found. The pixel DN values were processed to field reflection using regression parameters found with the empirical line method.

Selection and comparison of endmembers

During the fieldwork, surfaces present in the different landunits are detected and the reflectance of these surfaces is measured with a field reflectance meter.

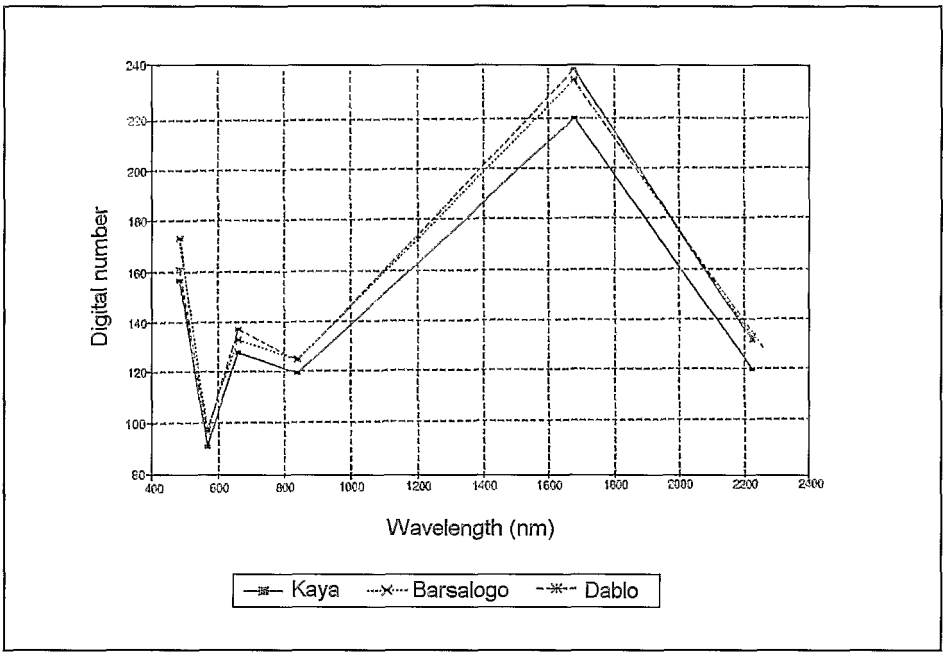

Figure 5. Differences in reflectance of light coloured crusts due to the influence of harmattan dust in the air.

The selection of endmembers is based on the following considerations:

- Due to the co-operation with PEDI, the most important areas of research are those where agriculture can be initialised or is present. For this reason, unsuitable units, e.g. hardened plinthite plateaus, have been excluded as endmembers.

- Shade endmembers are often used to reduce the effects of differences in illumination through differences in slope steepness and direction (GILLESPIE *et al.*, 1990). As the relief in the study area is limited, the exclusion of this endmember is permissible.

- Endmembers often absent in pixels are not included, while this leads to an overestimation of this endmember.

- As mentioned before, the number of endmembers is limited by the number of spectral bands.

To determine the similarity of the endmembers, two methods are used. The euclidean distance is sensitive for similarity in overall reflectance of endmembers. The spectral angle determines the similarity of the form of spectra (KRUSE *et al.*, 1993). Figures 6 and 7 show

respectively the euclidean distance and the spectral angle between the different endmembers. The values can forecast which endmembers are likely to be confused with each other.

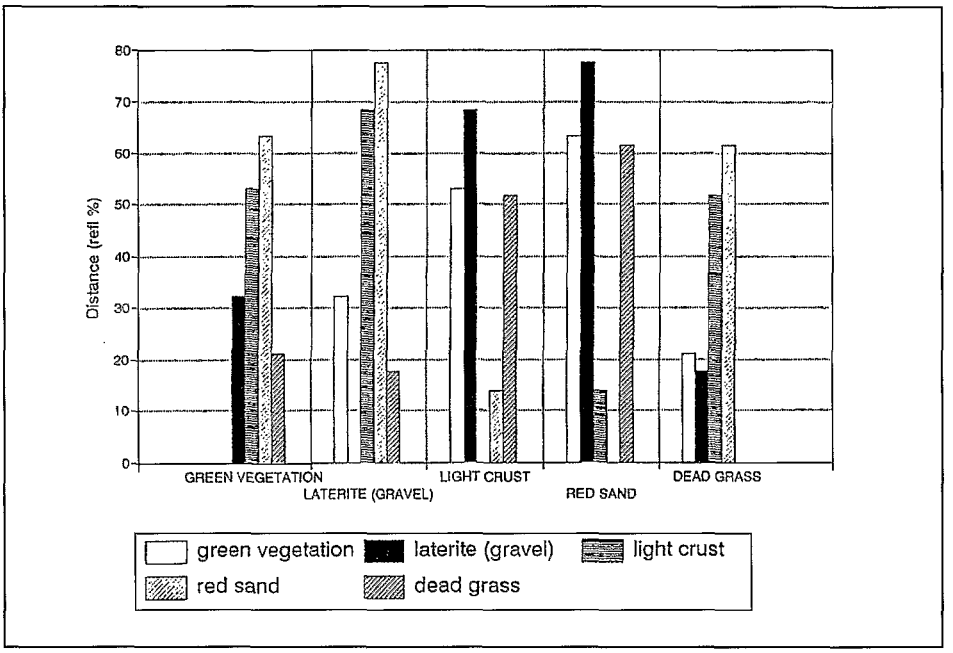

Figure 6. Euclidean distance between endmembers.

Albedo differences cause high differences in euclidean distance between light and dark endmembers. In general, green vegetation and laterite resemble dead grass and light crust resembles red sand. A low spectral angle between two endmembers is likely to cause problems with the calculation of the fractions of the two endmembers.

Spectral unmixing

Most objects on earth have a dimension smaller than the spatial resolution of present satellites. Pixels are composed of different objects each with its own spectral characteristics. Reflectance spectra can be modelled as a mixture of a few, so called endmember spectra (ADAMS *et al.*, 1989). Endmember spectra are the individual reflectance spectra of the different objects present in the pixel. The spectral variation in an image is caused by a limited number of surface materials like soil types, vegetation types and shade. Spectral mixture analysis is, up to now, mostly used with airborne hyper-spectral airborne data like AVIRIS and GERIS. Less emphasis is put on the use of this relative new technique with satellite data. SPOT data has the disadvantage of having only three spectral bands, which will (due to constraints in the calculations of

unmixing) limit the amount of used endmembers to two. With this limited amount of endmembers it is not possible to describe the variability of the image. On the other hand, Landsat TM has, for most occasions, enough spectral bands to execute spectral unmixing.

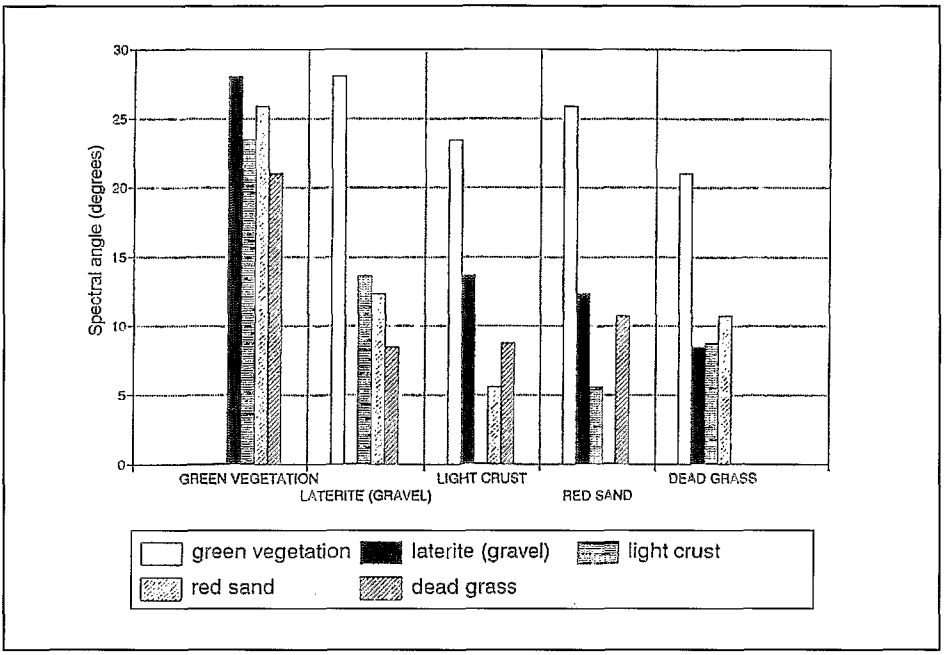

Figure 7. Spectral angle between endmembers.

As an approximation, the spectral mixing can be modelled as a linear combination of the pure endmember spectra, as follows:

$$R_i = \sum_{j=1}^{m} (F_j \cdot RE_{ij}) + e_i \wedge \sum_{j=1}^{m} F_j = 1 \wedge 0 \leq F_j \leq 1 \qquad (1)$$

where:

R_i : reflectance of the mixed spectrum in band *i;*

RE_{ij} : reflectance in band *i* of endmember *j;*

F_j : fraction of endmember *j;*

m : number of endmembers;

e_i : residual error in band *i;*

The above expression has a solution when the number of endmembers does not exceed the number of spectral bands minus one. Linear mixing occurs when the components are large or opaque enough to allow photons to interact with only one component (SINGER and MCCORD, 1979). The mathematical solution of the above

expression can be validated by examining the difference between the calculated and measured reflectance by the root-mean-squared (RMS) error.

$$RMS = \sqrt{\sum_{i=1}^{n} (R_i - R'_i)^2} \qquad (2)$$

where:

R_i : modelled or predicted reflectance of the pixel in band *i;*

R'_i : measured reflectance of the pixel in band *i;*

n : number of spectral bands.

The RMS error gives information about the amount of spectral variability explained by the selected endmembers. When the reflectance of a pixel can not be modelled by linear mixing with the selected endmembers, this will lead to a high RMS error.

As an example, the unmixing results of an area near Dablo in the North of Sanmatenga are examined. Resampled spectral curves, shown in figure 1, are used as endmembers in the unmixing. The processing of the image is done with the program Spectral Image Processing System (CSES, 1993) on a DEC Alpha 166 MHz. The processing time was 420 pixels per second with a 100% CPU time.

In plate 8, five images with each the distribution of a endmembers are shown. The minimum and maximum values used for stretching are given. The lighter the grey tone, the higher the coverage with the corresponding endmember. Some surfaces like green vegetation, light crusts and dead grass show high differences in coverage throughout the area. Laterite (gravel) and red sand are found in the whole area. This has to do with the processes that cause the forming of these surfaces. In the dead grass image, a striping is seen. Dead grass has a typical high reflection in band 5. The original striping is present in band 5. A high band 5 DN, caused by the low signal to noise ratio of this band, gives the spectral curve of a pixel the form of dead grass. The contribution of dead grass in that pixel is then overestimated. A too low band 5 DN, results in a spectrum that differs from dead grass. The dead grass coverage is then underestimated.

The crusted surface is highest at the south-west side of the laterite plateau. This side of the plateau is characterised by a crest. The geologic layer underneath the hardened plinthite layer is rich in kaolinite. This kaolinite, in combination with the slope, results in erosion and crusting of the soil.

During the fieldwork the coverage of endmembers is estimated for different sites. The estimates are compared with the unmixing results. In figures 8 to 11 these comparisons are visualised for four sites.

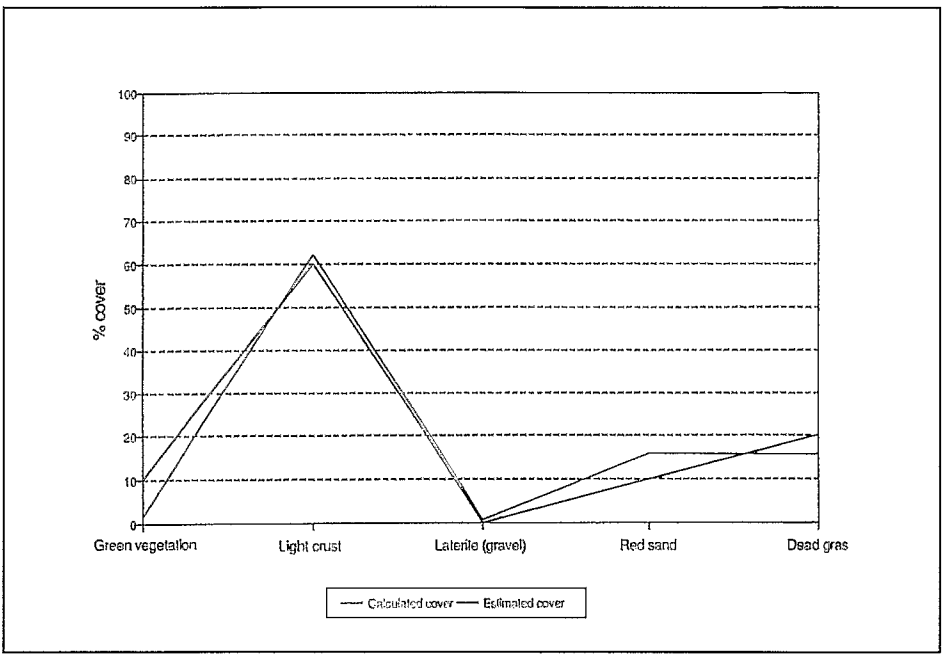

Figure 8. Ground cover near Barsalogo.

The calculated coverage with red sand is in all test sites higher than estimated in the field. On the other hand is the dead grass coverage underestimated. The image used in this research is from February 1991, the fieldwork was in October 1994. The time difference is therefore about four years. During the fieldwork, several sites are found with deserted agricultural fields. The sand of these fields seems to be replaced by dead grass. The high grass coverage is also caused by the relative wet rainy season of 1994. Other possible causes of the differences found between field estimations and unmixing calculations are mentioned below:

- Errors in the determination of the regression parameters used for spectral correction;
- The spectra of the endmembers are based on a good average spectra measured in the field. These average spectra might not be used for description of surfaces which have a slightly other reflection. The deviant reflection can cause wrong unmixing results.
- The combination of endmembers is not such that it is possible to describe the different units.

After the unmixing process, the five endmember images were unsupervised classified. The clusters represent relative homogeneous units. Figure 12 gives the average composition of 6 from in total 10 clusters of the Dablo area. The statistics of these clusters give information about the average distribution of endmembers. With this method, the classified image contains direct information about the constitution of the units. This information can be very useful for applications where specific, quantitative information about units is necessary. Possible applications are: decisions about the position of mapping boundaries, more specific remote sensing data input for erosion and other models.

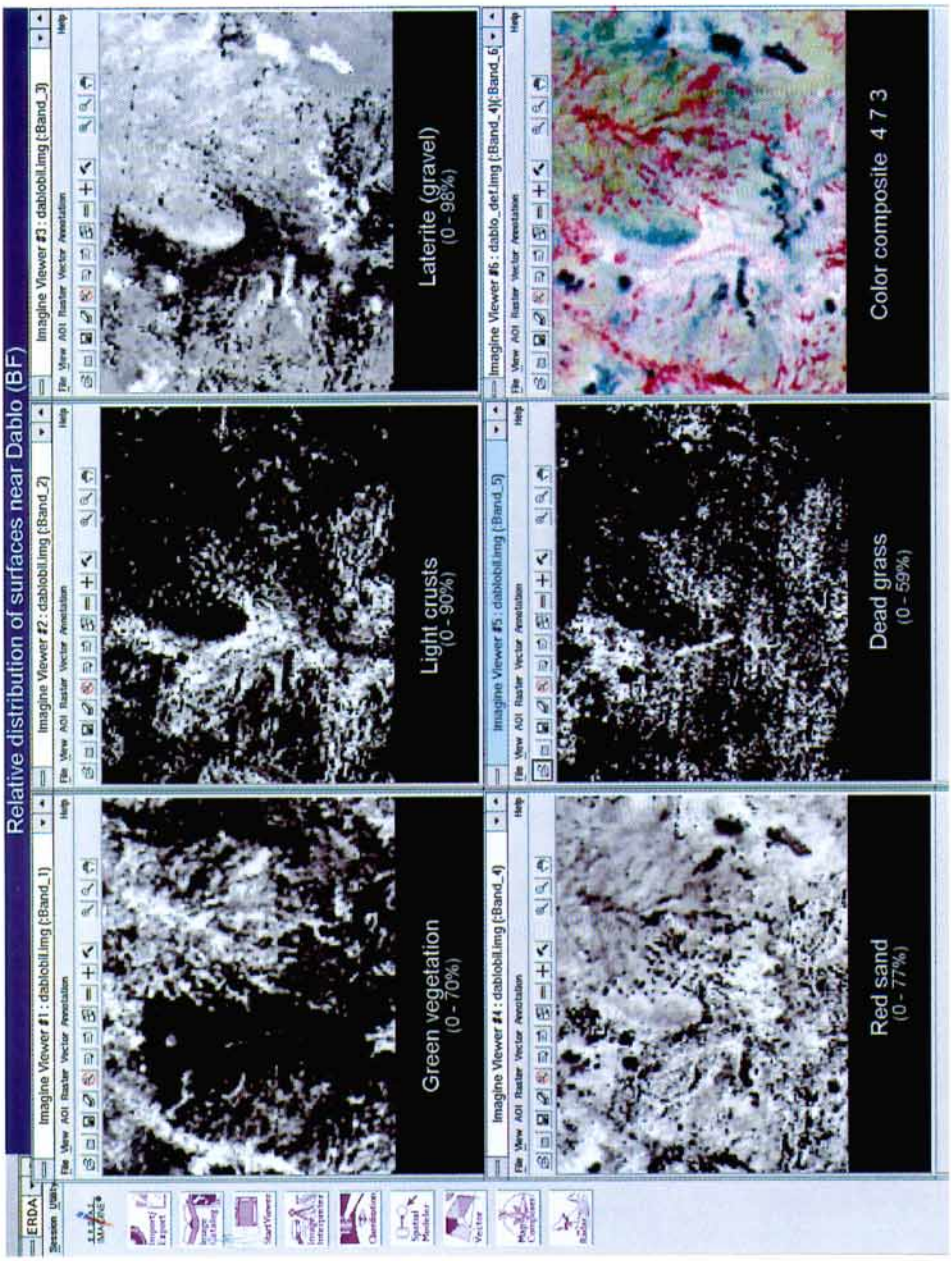

Plate 8. Spatial distribution of endmembers in the test site near Dablo in Burkina Faso (p. 279).

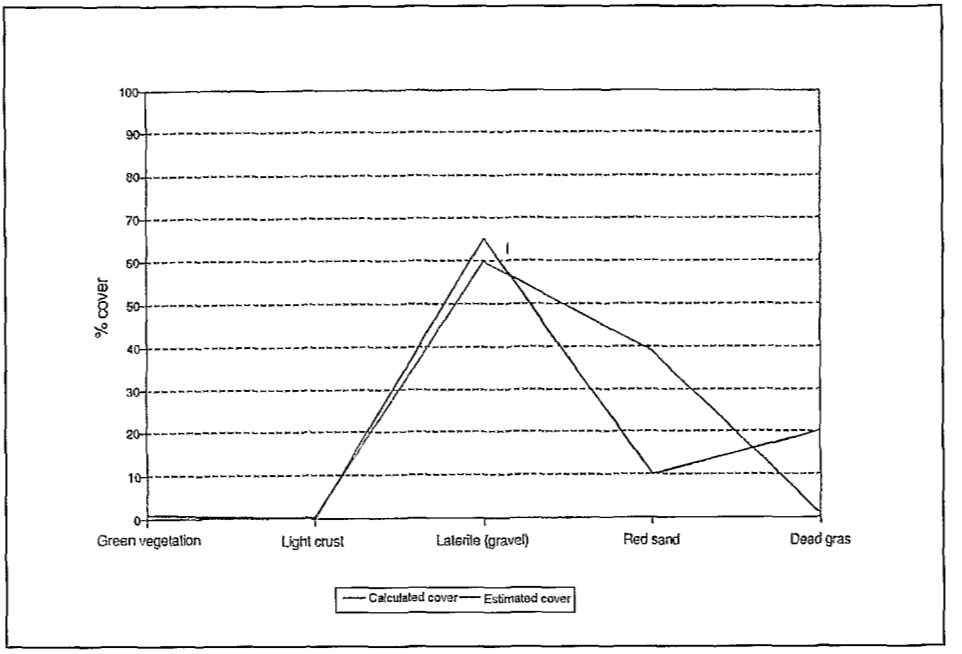

Figure 9. Ground cover Dablo site 1.

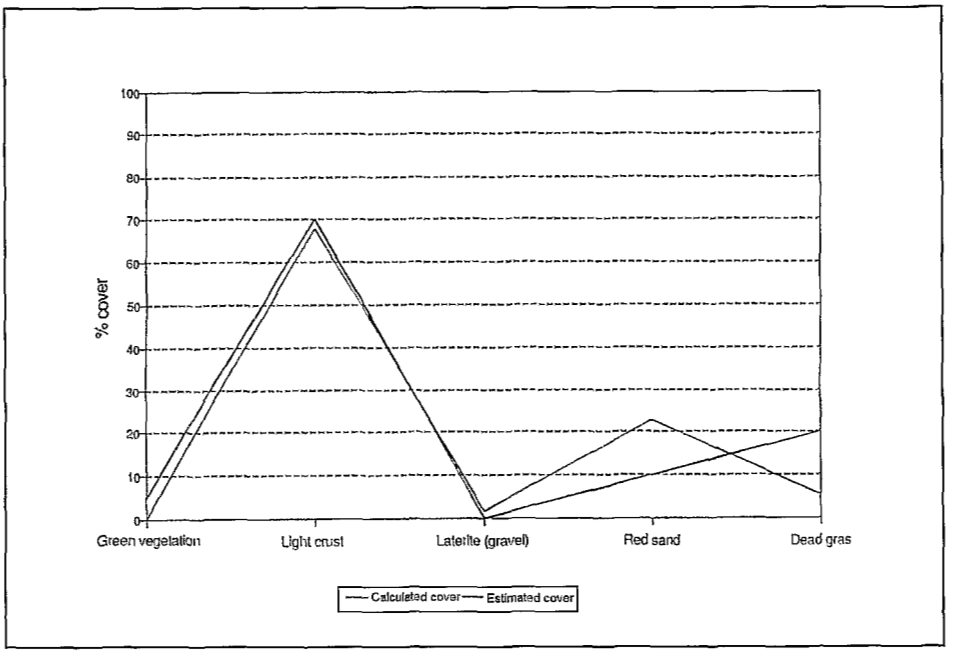

Figure 10. Ground cover Dablo site 2.

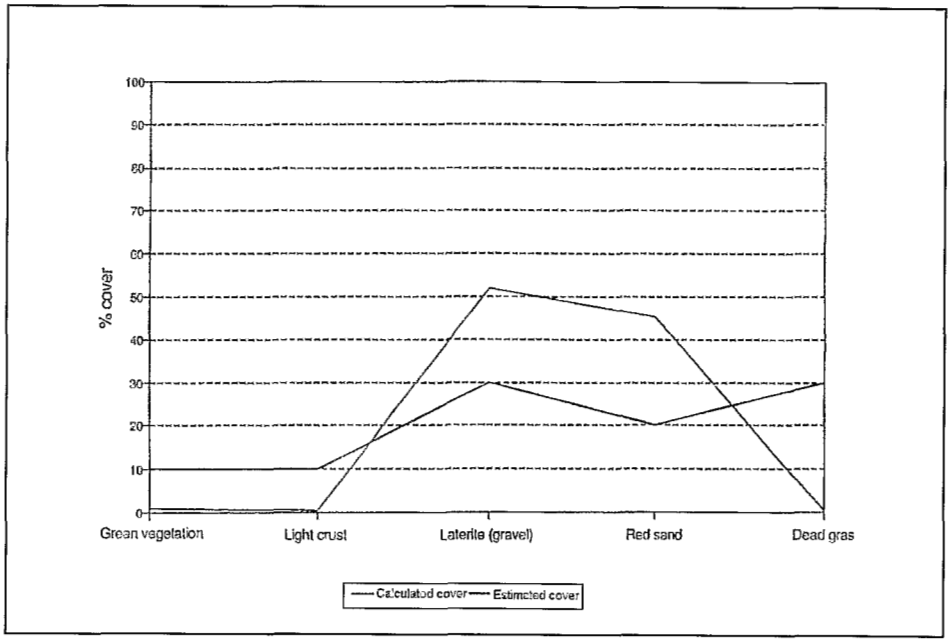

Figure 11. Ground cover Dablo site 5.

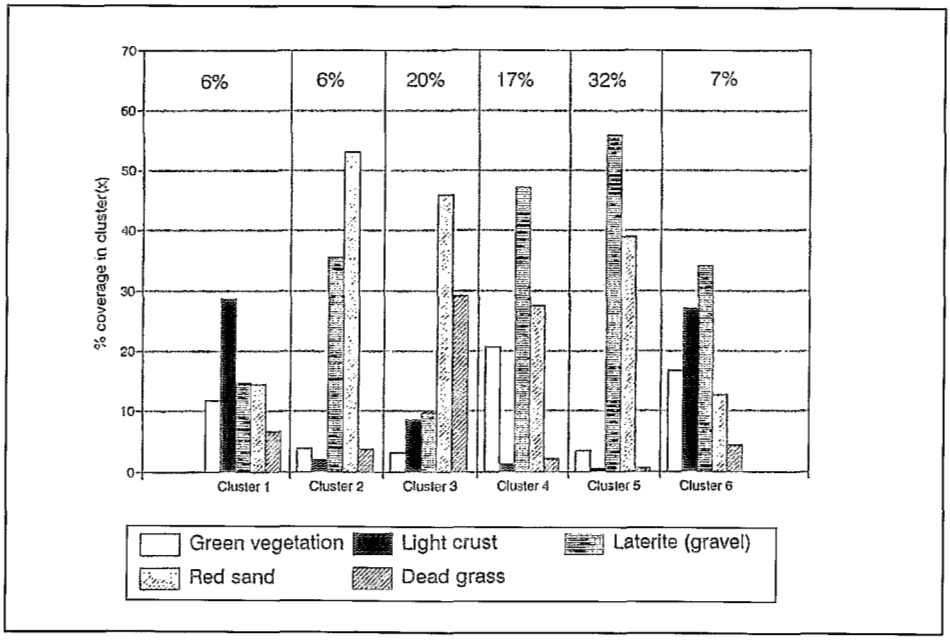

Figure 12. Average composition of 6 clusters. Numbers above bars give the total area in percentage. covered by the clusters.

Cluster 5 has, with 32% of the area, the highest coverage. This cluster represents the laterite plateaus and areas with a high coverage of laterite gravel. Clusters 3 and 4 also have a relative high coverage with laterite gravel. Cluster 4 is located in the so called *bas-fond*, the valley with relative dense (tree) vegetation. The 21% vegetation might be underestimated because of the shade of the trees. This shade will, due to the lowering in overall reflectance of the pixel, contribute to the percentage laterite.

Conclusions and recommendations

In the Sahel, the presence of harmattan dust in the air can influence the reflection of the surface, even if there is no influence visible at first sight. The data should be carefully examined before any data processing is started.

The use of field reflectance spectra is possible, when good regression parameters are found, e.g. with the empirical line method, to correct the digital numbers of the satellite image.

The first unmixing results show a quite good correspondence between field estimates and calculations. Differences can be caused by:

- Changes in landcover during the time of recording of the image and the time of fieldwork;
- The regression parameters found with e.g. the empirical line method can be wrong;
- The endmembers used are not representative for surfaces that have a slightly different reflection.

Clustering of the images with the distribution of endmembers, results in useful additive information about classification units.

To unmix Landsat TM images of complex landscapes, a first subdivision in mayor units is necessary. The units can then be unmixed each with their own selection of endmembers.

In further research, a sensitivity analysis is necessary to determine the influence on the final results of each step in the unmixing procedure.

References

ADAMS J.B., SMITH M.O., GILLESPIE A.R. (1989). "Simple models for complex natural surfaces: a strategy for the hyperspectral era of remote sensing", in: *Proc. IEEE Int. Geosci. and Remote Sensing Symp. '89 I*, IEEE, New York, pp. 16-21.

CONEL J.E., GREEN R.O., VANE G., BRUEGGE C.J., ALLEY R.E. (1987). "AIS-2 radiometry and a comparison of methods for the recovery of ground reflectance", in: *Proc. 3rd Airborne Imaging Spectrometer Data Analysis Workshop* (G. Vane, Ed.), JPL Publication 90-54, Jet Propulsion Laboratory, Pasadena, CA, pp. 18-47.

CSES (1993). *SIPS User's guide*, Spectral Image Processing System Version 1.2, Center for the study of Earth from space (CSES) and Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado, Boulder.

ELKENBRACHT E. HOLTETEN A., OTTER L., SLAA T. (1995). "Remote sensing and soil science in the Kaya area (Burkina Faso)". *Antenne Sahélienne report 34*, Wageningen Agricultural University, The Netherlands.

EPEMA G.F., BOM B.C.J. (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, pp. 23-28.

FARRAND W.H., SINGER R.B., MERENYI E. (1994). "Retrieval of apparent surface reflectance from AVIRIS data: A comparison of empirical line, radiative transfer, and spectral mixture methods", *Remote Sens. Environ.* 47:311-321

Geophysical & Environmental Research corp. (1993). Operation manual GER'S mini-IRIS RT spectroradiometer, 23 p.

GILLESPIE A.R., SMITH M.O., ADAMS J.B., WILLIS S.C., FISCHER A.F., SABOL D.E. (1990). "Interpretation of residual images: Spectral mixture analysis of AVIRIS images", Owens Valley, California, in: *Proc. 2nd Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Workshop* (R.O. Green. Ed.), JPL Publication 90-54, Jet Propulsion Laboratory, Pasadena, CA, pp. 243-270.

KRUSE F.A., LEFKOFF A.B., BOARDMAN J.W., HEIDEBRECHT K.B., SHAPIRO A.T., BARLOON P.J., GOETZ A.F.H. (1993). "The Spectral Image Processing system (SIPS) - Interactive visualisation and analysis of imaging spectrometer data", *Remote Sens. Environ.* 44:145-163

MCTRAINSH G.H., WALKER P.H. (1982). Nature and distribution of harmattan dust, *Z. Geomorph.* N.F. 26 (4):417-435, Berlin.

SINGER R.B., MCCORD T.B. (1979). "Mars: large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance", in: *Proc. 10th Lunar Planetary Sci. Conf., J. Geophys. Res. Suppl.*: 1835-1848.