
Mona A.A. Dawelbait and Francesco Morari

Abstract: The savanna region in Sudan is defined as a rangeland and rain-fed croplands region. Degradation in vegetation cover by overgrazing and cutting of woody plants are the common desertification processes. Remote sensing techniques have been applied in the study of desertification to monitor land cover degradation and characterize the dynamism of sand dunes. Three Landsat images (Landsat5 TM and two Landsat7 ETM+), acquired in 1987, 1999 and 2008, were used to evaluate the development of degradation processes in Central North Kurdufan State (Sudan), part of the savanna region in the Sahel belt. Traditional methods to extract vegetation and soil information from remote sensing data in semi-arid areas, such as classification techniques and vegetation indices, were found to be inaccurate. In this work, Spectral Mixture Analysis (SMA) and multitemporal comparison techniques were therefore applied to emphasize vegetation loss, soil change and the growth of village areas in the study area. In order to have powerful strategies to combat desertification, accurate information in estimation the driving factors of land cover degradation at local scale are necessary. To identify the soundest strategies, more study in the interpretation techniques and high-resolution tools must be applied. In this study, the application of spectral mixture analysis along with change vector analysis to Landsat data appeared to be a consistent and low-cost technique to obtain information on vegetation cover, soil surface type and identify risk areas.

Key words: Remote sensing • Spectral mixture analysis • Change vector analysis • Landsat • Desertification • Savannah

INTRODUCTION

Regional and national planning to monitor land cover degradation in semiarid areas can be performed more efficiently if accurate spatial information is provided. Maps which show such information can be produced through the use of satellite images and digital data and field indicators [1]. Remote sensing of vegetation cover and soil is thus critical for regional scale monitoring. In order to use remote sensing in mapping land cover degradation in semi-arid areas, study of optical properties of the vegetation and soil elements in these areas is needed [2, 3]. The meaning and value of remote sensing data are enhanced through skilled interpretation, in conjunction with conventionally mapped information and ground-collected data [4].

Most remote sensing in arid regions has concentrated on optical remote sensing techniques which use data from sensors that collect radiation in the reflected solar spectrum. Earth observation data, particularly Multispectral Scanner (MSS) imagery and Landsat Thematic Mapper (TM) have been widely used in semiarid environments to show up temporal and spatial variations in land cover using their reflectance characteristics [5].

Since the launch of the first Earth Resource Technology Satellite on July 23, 1972, the analysis of data has advanced from simple visual observation to sophisticated interpretations based on first principles of spectroscopy and electromagnetic radiation [2]. The Normalized Difference Vegetation Index (NDVI) has been most commonly used to map spatial and temporal variation in vegetation [6]. NDVI has only limited success...
in providing accurate estimates of shrubland cover in arid areas and limited utility in an arid ecosystem. This is due to spectral variability of background materials such as soil and surface litter and the strength and variation of soil spectral albedo, which causes nonlinearity in the relationship between NDVI and vegetation characteristics[7, 8].

Spectral mixture analysis (SMA) is a sub-pixel classification technique which could be used to unmix the soil-plant canopy measurements into the respective soil, vegetation and non-photosynthetic vegetation [5]. SMA depends on the spectral response of land cover components. The spectral response in remote sensing from open canopies is a function of the number and type of reflecting components, their optical properties and relative proportions [9]. SMA appears to be the most efficient technique to obtain information on vegetation cover, soil surface type and vegetation canopy characteristics in semiarid areas because the scale of variability of the principle landscape elements in semiarid areas is larger than the pixel size in most of the remote sensing satellite imageries [9, 3, 4].

Change detection methods are commonly used in monitoring land degradation. Change can be identified explicitly either as change in the number of environmental components or as a change in percentages of the components [9]. Visual interpretation and direct measurement using map-algebra are widely used in change detection. Change Vector Analysis (CVA) by Malila [10], which allows the direction and magnitude of change between two time periods to be evaluated comes to be important in land cover degradation.

The percentages of vegetation and soil vulnerable to erosion fraction images are used to monitor the vegetation re-growing and degradation.

The savanna region in Sudan is defined as a rangeland and rain-fed croplands region. Degradation in vegetation cover by overgrazing and cutting of woody plants are widespread in savanna region. Since limited funds are provided to Sudanese research institutions, remote sensing can be a reliable tool to study land cover degradation without incurring high costs [11, 12, 13]. More studies are therefore necessary in order to define accurate techniques in the remote sensing field to study driving variables affecting the degradation and adopt efficient site-specific strategies to combat it.

This paper aimed to test the application of SMA and CVA to Landsat images as tools to study land cover variation during 2 decades in a savannah region in the central part of Sudan.

MATERIAL AND METHODS

Study Site: The study site is located in the north of Umrowaba in North Kordodan State, central Sudan, in the Sahelian eco-climatic zone (between latitude 12°56'35" and 13°3'49" N and longitude 31°0'51" and 31°58'51" E) (Figure 1). The climate is semi-arid with annual rainfall ranging from 200 to 750 mm, concentrated during a few summer months (June to September), with a peak in August. Mean annual temperature is about 20°C, but the daytime temperature can rise as high as 45°C during summer.

Fig. 1: Study site position and main landscape elements
The soil is a Cambric Arenosols [14], coarse sandy, of Aeolian origin with high infiltration rates and inherent low fertility. Sand sheets and sand dunes stabilized by vegetation are the main natural formations. Natural vegetation consists of trees (Acacias spp.), bushes and grass, Aristida pallida Steud. on crests of dunes, Eragrostis termula Tnismert. in the troughs and Cenchrus biflorus Roxb., which grows after crop cultivation. Rangeland and rain-fed croplands are the most important land use systems. The main crops are sorghum (Sorghum vulgare Pers.), millet (Panicum miliaceum L.), sesame (Sesamum indicum L.) and watermelon (Citrullus lanatus (Thunb.) Matsum & Nakai). The rainy season usually leads to a short growing period followed by a long dry season with a reduction in green vegetation.

Data Acquisition and Preprocessing: Three Landsat images were selected and analyzed to application of SMA in Landsat images in monitoring land cover and use in the study site: one Landsat Thematic Mapper scenes (TM5 1987) acquired on September 15th 1987, Landsat Enhanced Thematic Mapper plus (ETM+7 SLC-on) scene acquired on November 11th 1999 (ETM+7 1999) and Landsat Enhanced Thematic Mapper plus (ETM+7 SLC-off) scene acquired on October 18th 2008 (ETM+7 2008). The selected ETM+7 1999 was in November since there was no available image free of cloud from August to October.

While TM5 1987 and ETM+7 2008 were acquired in periods of comparable rainfall amount (4.6 mm in September 1987 and 8.4 mm in October 2008), ETM+7 1999 images were selected to study the effect of rain. Indeed the mean annual rainfall in 1999 was 581 mm in the study site relativity higher than the mean annual rainfall (470 mm). Landsat images were selected because they are free of charge, with high monitoring frequency and cover areas appropriate for monitoring the environment in a large geographic zone. Landsat TM5 and ETM+7 have a temporal revisit time of 16 days and a spatial resolution of 30 m with six visible/near infrared bands and one thermal band. The gaps in ETM+7 SLC-off were filled using the localized linear histogram mach (LLHM) method [15]. Landsat ETM+7 SLC-off, November 3rd 2008, were used to fill the gaps in selected ETM+7 SLC-off 2008 image since the gaps were not overlapping and the time lag between the two images was only 15 days.

ETM+7 1999 and ETM+7 2008 were co-registered to TM5 1987 to undertake comparative analysis. Images were not referenced to a standard map base, since the only available map had a coarser resolution (scale 1:250,000). They were geometrically rectified using 13 ground control points to accurately mach them to ground reference data. The nearest neighbour assignment [16] was applied yielding a root mean square (RMS) error of 0.34 pixels. Subsets covering only the study area were then extracted from each image. To apply SMA the digital number (DN) of the images band1-5 and 7 recorded in 8 bits were converted to exo-atmospheric reflectance units according to Markham and Barker, [17]. The conversion also improved the image quality [18]. No atmospheric correction techniques, such as empirical line calibration [19] or dark object subtraction [20] were applied since they have no significant effect on the modelling [21].

Spectral Mixture Analysis and Endmembers: In remote sensing images of arid and semi-arid environments, the pixel contains mixed spectral information due to the high variability in the distribution of land cover components. SMA is based on the concept that the variance across a given scene is dominated by the relative proportion of a few spectrally distinct components [22]. SMA transforms radiation or reflectance data into fractions of a few dominant endmembers, which are fundamental physical components of the scene and not themselves a mixture of other components [22]. Fraction images represent the mixing proportions of these endmember spectra [23, 24]. SMA generally involves three steps [25]: a) assessment of dimensionality or number of unique reflecting materials in a landscape to obtain the endmembers; b) identification of the physical nature of each endmember within a pixel; c) determination of the amounts of each endmember in each pixel.

The basic linear spectral mixture analysis (LSMA) equation is [4].

\[
R_s(\lambda) = \sum_{i=1}^{n} f_i R_i(\lambda) + \epsilon(\lambda)
\]  

(1)

Where \(R_s(\lambda)\) is the apparent surface reflectance of a pixel in an image, \(f_i\) is the weighting coefficient (\(\sum_{i=1}^{n} f_i = 1\)) interpreted as fraction of the pixel made up of the endmember \(i = 1, 2, \ldots, n\), \(R(\lambda)\) is the reflectance spectrum of spectral endmember in an \(n\)-endmember model and \(\epsilon(\lambda)\) is the difference between the actual and modelled reflectance. \(f_i\) represents the best fit coefficient that minimizes RMS error given by the following equation:
\[ \text{RMS} = \sqrt{\frac{\sum_{j=1}^{m} (e_j)^2}{m}}^{0.5} \]  

(2)

where \( e_j \) is the error term for each of the \( m \) spectral bands considered.

One problem related to the application of SMA is nonlinear mixing, which can hinder the SMA applications [26, 27]. Nonlinear mixing occurs when photons interact with more than one type of object on the earth before returning to the sensor [8]. However, the importance of the effect is not widely recognised since other studies [28, 29] showed that nonlinear mixing is a secondary feature.

Some SMA approaches use endmember spectra derived from the image (image endmember) [30, 22], whereas others employ libraries of endmember spectra (library endmember), which are produced from reflectance measurement in a laboratory [5]. Tompkins et al. [31] pointed out that endmembers selection is the most critical step in SMA to provide a physically meaningful fraction. While library endmembers would undoubtedly represent a purer endmember spectrum and would possibly have given a more accurate absolute abundance, image endmembers simply produce a different scaling and can thus be used for change detection [22]. Bateson and Curtiss [32] and Bateson et al. [33] generated SMA models using PCA to explore image data in multiple dimensions, although in drylands it is exceedingly difficult to locate image pixels containing 100% cover of each appropriate endmember. One advantage of this technique is that the selection of the endmember spectra is based on inherent spectral variability of the image data without requiring homogeneous pixels of each endmember [8]. The approach of Johnson et al. [34] and Smith et al. [23] was used to select the endmembers in this paper. The method is based on PCA application to identify the endmembers of multiple surface components. The authors observed that for a mixture of three individual substances (e.g. minerals) the scatter-plot of the first two principle components produced a triangle in which the ‘pure’ endmembers were located at the corners. Several studies have adapted this technique by analyzing different principal component pairs and have managed to successfully obtain image endmembers within different environments [35, 36, 37]. In this study a PCA was applied to Landsat images using ENVI to identify endmembers. The spectral mixing space as represented as orthogonal scatterplots of the first three PC bands were generated and the vertices of these plots were selected as endmembers after visualization in the original images. Endmember spectra were applied to SMA in order to produce the fraction images with associated the RMSE images.

**Change Detection:** Two approaches were conducted to evaluate the variation in land use and land cover (LULC). The first approach was visual interpretation of the land cover elements in the different years. Visual interpretation for each endmember was performed using a standard RGB composite by displaying fractions images of the three years 1987, 1999 and 2008 as blue green and red, respectively. The second approach consisted in Change Vector Analysis (CVA) [10]. CVA allows the direction and magnitude of change between two time periods to be evaluated. Vegetation and soil vulnerable to erosion fraction images were used to monitor the vegetation re-growing and desertification in three terms of time (1987-1999, 1999-2008 and 1987-2008). Change direction was measured as the angle of the change vector between two coincide pixels measurement in the two different years. Angles measured between 90° and 180° indicated an increase in sand and decrease in vegetation and therefore an increase in degraded area. On the contrary, angles measured between 270° and 360° indicated a decrease in sand and an increase in vegetation and therefore represented re-growth conditions. Angles measured between 0°-90° and 180°-270° indicated either increase or decrease in both vegetation and sand and consequently persistence in the conditions (Figure 2) [13]. Change of magnitude is measured as the Euclidean distance or length of the change vector between two coincide pixels measured in the two different years equation 3. Four classes of magnitude were represented for either degradation or re-growing according to Kuzera et al. [38].

\[ R = \sqrt{(x1-x2)^2 + (y1-y2)^2} \]  

(3)

Where:
- \( R \) : The magnitude of the vector change;
- \( x1 \) : Sand in T1;
- \( x2 \) : Sand in T2;
- \( y1 \) : Vegetation covers in time T1 and
- \( y2 \) : Vegetation cover in time T2
Field Survey: A 2-weeks field survey was conducted in October 2008 in order to test the accuracy of SMA using ground vegetation data as references. A total of 16 mixed ground cover plots (size 60 x 60m for each plot) were selected. Vegetation was composed of a mixture of acacia trees, bushes, grass and shrubs. Trees and bushes were georeferenced with a GPS and the crown diameters were measured and orthogonally projected to the ground surface to estimate the percentage cover. The percentage cover of grass and shrubs was estimated using the line point intersect sampling method [22]. Measurements of the grass and shrubs were taken along 30 60-m long transects, oriented in N-S direction, every 2 m. Measurement points were selected at 60 cm intervals along the transect. The grass and shrubs under the trees and bushes were ignored. The accuracy of SMA was estimated by scatter plot correlation comparing total percentage of live cover in each plot with the live cover (vegetation) fraction image.

RESULTS AND DISCUSSION

Endmember Spectra and SMA Applications: The PC analysis of TM5 Sep 15 data found that the first three components explained over 99% of the variance and that simulated data were mean-corrected and projected onto the space determined by those components. In this PC-reduced space five endmembers were manually selected (Figure 3): bright vegetation (BV), dark vegetation (DV), non-photosynthetic material (NPM), bright soil (BS) and dark soil (DS). BV consisted of all types of natural vegetation (e.g. dense shrubs, grass) and cultivated crops with higher leaf chlorophyll and water content. DV consisted of natural vegetation with lower leaf chlorophyll/water content (senescing vegetation). NPM identified villages (e.g. straw houses), dormant trees and senesced grass and shrubs. BS and DS represented coarse sandy soils and fine sandy soils with higher organic matter in the top layer, respectively. Both BS and DS absorbed more energy at 2.215 micrometer (band 7) than at 1.65 micrometer (band 5) which indicates availability of moisture content [39] (Figure 4). A higher soil organic matter content usually also implies a higher soil water holding capacity and subsequently higher water content. The effect of shadow was ignored since it is reduced for the sparse canopies typical of many semiarid bush species [40].

The endmember set was selected to maximize the model performance for BV and BS which is more vulnerable to wind erosion than DS. Not all image components can be effectively modelled using simple endmember models [37]. To find the best quality of fraction images, three combinations of endmembers were tested [41]. The combinations were: 1) all five endmembers; 2) four endmembers with BV, NPM, BS and DS; 3) three endmembers with BV, NPM and BS. Fraction images derived from different combinations of endmembers were evaluated with visual interpretation, error extent and distribution in the error fraction image. The combination with four endmembers (BV, NPM, BS and DS) was chosen since it provided the best distinction of LULC types and lowest errors.

This set of endmember spectra was therefore used across the three selected images. Using the identical endmembers to analyze multitemporal images strengthened the change analysis [22]. Similar to using reference endmembers from a spectral library, using identical image endmembers for different images allows
a direct comparison of resulting endmember proportions [37]. The RMS error images for SMA process ranged from 0% to 3% for TM5 1987, from 0% to 1.5% for TM+7 1999 and from 0% to 2.8% for ETM+7 2008.

Figure 5 shows the scatter plot correlation between the percentage of vegetation determined with SMA (ETM+7 Oct 18) and field data. In general, the correlation between them is good with an $R^2$ of 0.91 but with a slight overestimation, especially at lower SMA values. There are three main sources of error that could have affected the comparison. The first one can be due to the misregistration of multidate scene and location of the field sites. This is potentially the largest source of error [37], especially in our case where the geometric rectification was done with 13 ground control points for all the scenes before subsetting of the study area. This was done because the study area had no fixed sharp points that could be used as control points. Moreover, most of the sites were characterized by a higher degree of scene heterogeneity that could have increased the uncertainty in location [37]. Other sources of error can be related to the application LLHM method error to fill the gaps in ETM+7 Oct 18 [15] and accuracy of the field survey, especially in the estimation of grass and bushes. Considering the comparative approach of the present work, the overestimation errors were considered acceptable to evaluate the LULC change.

**Change Detection:** Average estimation of endmember fractions is given in Figure 6. The effect of the rainfall was very clear for BV whose fraction increased of 16% from 1987 to 1999 (the rainiest year) and then decreased in 2008 to the initial level. The fraction of bright soil increased over the 21 yrs (+14%). However, average estimation is not sufficient to provide a clear representation of spatial change at landscape scale [42, 43].

Tempo-spatial variations of the endmember fractions are visually interpreted by displaying fractions for year 1987 in blue, year 1999 in green and year 2008 in red for each endmember (Figure 7). The visual interpretation of color composite shows that the major changes have
the most saturated colors while the minor changes have less saturated colors. White tones indicate no temporal change and grey tones indicated no existence for that endmember.

Vegetation fraction in general was higher in 1987 and 1999, mainly located around the villages, indicating lost of vegetation in 2008. Figure 7 shows an increase in vegetation fraction due to the high rainfall in 1999 in the north-east part (saturated green color) although the date of the image in November when the rain was less than in August and September. In 2008 vegetation fraction increased in small scattered areas (red color).
Fig. 8: Degradation and re-growth areas calculated by applying CVA, a) from 1987 to 1999, b) from 1999 to 2008; c) from 1999 to 2008

NPM was high in 1999 represented by saturated green most likely represented the seasonal growth of non-perennial vegetation (e.g. herbaceous species) due to high rain fall from August to October. The white color in NPM composite shows the stable area of the villages and the red color around them indicates clearly the extension of these villages. A drastic change of BS fraction can be observed in 2008 (Figure 7). It was spatially distributed around the villages and across the site from the northern to the eastern part.

CVA quantified land cover degradation in degree of severity in (Figure 8). The difference in BS fraction for was provided in CVA equation for soil vulnerable to erosion parameter while the difference in BV fraction was substituted for vegetation fractions. According to CVA (Figure 8), the magnitude of desertification ranges from low to extreme, with a prevalence of severe degradation conditions (high or extreme) in the eastern part. Change detection analysis also shows the existence of re-growth conditions, mostly spread in the south-western part.

Fig. 9: Rainfall Anomaly Index (RAI) from 1973 to 2008.
Estimating degradation and re-growth areas in separate terms of time (1987-1999 and 1999-2008) make sense in observing the effect of rain in long run. While degradation in land cover affects small areas from 1987 to 1999 shown in Figure 8 very large areas were affected in extreme to high condition of degradation in land cover from 1999 to 2008. High rain in the 1990s and fluctuating rain in 2000s according to the Rainfall Anomaly Index (RAI) [44] (Figure 9) was one of the factors affecting the degradation in the study site.

Considering map of degradation and re-growth from 1987 to 2008 is the most reliable estimation since the amount of rain in the two years was comparable and the also the acquisition time of the two images, degradation prevailed over re-growth, (Figure 10) affecting an area of 153,867 km², with a prevalence of medium (70,944 km²) and high (48,578 km²) classes of magnitude. Re-growth was estimated on an area of 35,313 km², mainly classified as medium (17,005 km²) and low (13,708 km²). However, average estimation is not sufficient to provide a clear representation of driving factors of change at landscape scale [42, 43].

Re-growth conditions observed in the SW part appears in the map from 1987 to 2008 were mainly due to Government reforestation projects in last decade and sustained by higher rainfall in the last years in the study area. RAI time series (Figure 10) confirmed the existence of favourable conditions for vegetation growth from the 1990s to 2008, with higher frequency of positive anomalies than in the 1970s and 1980s.

The degradation appears in the map from 1987 to 2008 was driven by various factors, which operated with different intensity in the areas. In the eastern part the expansion of villages triggered the change in land use and mismanagement of the natural resource, mainly caused by deforestation to supply wood for domestic uses i.e. building, cooking, etc. and overgrazing [45].
The recent satellite and model based studies of the Sahel [46, 43, 47] which demonstrated that vegetation has recovered from the peak drought conditions suffered in the region in the 1980s are approved this result. For example, Anyamba and Tucker [43], monitoring the Sahelian vegetation dynamics using NDVI in the period 1981-2003, observed the prevalence of greener than normal conditions from the 1990s to 2003. Indeed, NVDI time series followed a similar increase in rainfall over the region during the last decade and indicates a gradual slow but persistent recovery from the 1980s.

CONCLUSION

The complication in providing information on vegetation types and structure from images acquired during the rainy season and to discuss it with the variation of soil has a number of important implications regarding the availability of the images free of cloud and the role of satellite imagery in increasing our understanding of vegetation structure and soil type in semiarid areas. Particularly, it highlights the need for long-term ground-based monitoring of changes in vegetation and soil characteristics and data from new satellite sensors in order to help in developing different reflectance models and images analysis for the vegetation community and soil type in arid and semi-arid areas.

Site-specific strategies according to accurate information in estimation the driving factors of land cover degradation at local scale are thus necessary to combat desertification, avoiding the implementation of untargeted measures. In order to identify the soundest strategies, more study in the interpretation techniques and high-resolution tools must be applied. In this study, the application of spectral mixture analysis with change vector analysis to Landsat data appeared to be a consistent and low-cost technique to obtain information on vegetation cover, soil surface type and identify risk areas.

REFERENCES


17. Markham, B.L. and J.L. Barker, 1986. Landsat MSS and TM post-calibration dynamic ranges, exoatmospheric reflectances and at-satellite temperatures. EOSAT Landsat Technical Notes, No. 1.


