Greener Journal of Agricultural Sciences

ISSN: 2276-7770

Vol. 15(2), pp. 76-84, 2025

https://gjournals.org/GJAS

DOI: https://doi.org/10.15580/gjas.2025.2.011723007

Mapping *Acacia senegal* (L.) Wild and other shrubs in North Kordofan State, Sudan using Remote Sensing Principal Component Analysis.

Hatim Abdalla M. ElKhidir[1]; Abdelaziz K. Gaiballa[2]; Nancy Ibrahim A. Khiralla[2]; Abdelrahman A. Khatir[1]

[1] Agricultural Research Corporation, ElObeid Research Station.
[2] Sudan University of Science and Technology, College of Forest and Range Science.

ABSTRACT

Mapping forest types in a natural heterogeneous forest environment using remote sensing data is a long-standing challenge due to similar spectral reflectance from different tree species and significant time and resources are required for acquiring and processing the remote sensing data. This study was conducted in North Kordofan State, Sudan to map *Acacia seneal* (L.) Wild, which is known as gum arabic tree "Hashab" that is found in many types of pure stands or associated with other tree species such as *Acacia tortilis "Seyal"* or *Leptadinia pyrotechnica* "Marakh". The objective was to understand means of identification and mapping of *A. senegal* trees (Hashab) in gum belt using remote sensing techniques. Three pure stands were selected to present the three tree species. The MOD13Q1v006 16-day and the MOD09A1 8-day products form MODIS were acquired for this study. Images provide Normalized Differences Index (NDVI) with 250 m resolution for 15 years while MOD09A1 product provide band 7 (soil) with 500 m resolution. Soil Adjusted Vegetation Index (SAVI) for the period of 15 years was measured from red and near infrared bands of MOD13Q1. Three variables of NDVI average of June, August and October and three variables of SAVI average of June, August and October with band 7 which present the soil factor were used to calculate the Principal Component Analysis. The colour composite image was made depending on PCA results to differentiate between the three tree species. The results of PCA in this study indicate that, PCA can be used to determine the effect of vegetation type and tree species. The results support using PCA and colour composite to distinguish tree species. The effect of soil as band7, NDVI and SAVI is significantly affecting vegetation cover and vegetation type.

ARTICLE'S INFO

Article No.: *011723007*
Type: *Research*
Full Text: PDF, PHP, HTML, EPUB, MP3
DOI: *10.15580/gjas.2025.2.011723007*

Accepted: *20/01/2023*
Published: *05/11/2025*

Keywords: *Acacia senegal, Principal component analysis, MODIS imagery, mapping.*

****Corresponding Author***

Hatim Abdalla M. ElKhidir

E-mail: *guzaia@ hotmail.com*

Article's QR code

LIST OF ABBREVIATIONS

CBS	Central Bureau of Statistics
FCC	False Colour Composite
MODIS	Moderate Resolution Imaging Spectroradiometer
NDVI	Normalized Differences Vegetation Index
PCA	Principal Component Analysis
SAVI	Soil Adjusted Vegetation Index

INTRODUCTION

Principal components analysis (PCA) is a method in which original data is transformed into a new coordinate system, which acts to condense the information found in the original inter-correlated variables into a few uncorrelated variables, called principal components. With respect to hyper spectral data, PCA transforms large data sets into relatively few meaningful uncorrelated orthogonal variables/dimensions (i.e., the principal components) that represent most of the information present in the original image. In any principal components rotation, the first component or dimension accounts for the maximum proportion of the variance of the original image, and subsequent components account for maximum proportion of the remaining variance (Holden and LeDrew, 1998; Zhao and Maclean, 2000).

Colour is widely used in remote sensing work. In many instances, the use of colour conveys additional information both visually and scientifically. Remote sensing satellites view the earth in different spectral bands, viz. near infrared (NIR), red, green, and blue bands, in a conventional multispectral imaging system. In the absence of a blue channel, colour images can be generated using near infrared, red, and green bands in what is known as a false colour composite (FCC) and does not look natural, like the image we see with the naked eye (Patra *et al.* 2006).

The color composite of the PCA provide a good option for visual interpretation. Jensen (2005) reveals that principal components can be used for the classification of remote sensing and false color compositing. It made up of PC1, PC2 and PC3 images. A false color composite produced by using the first principal component (PC1) as green, the second principal component (PC2) as blue and the third principal component (PC3) as red.

Although much effort has been being put into the development of automated extraction and recognition techniques, the manual interpretation and image visualization of remotely sensed data are still of basic importance for many areas of applications (Sabins 1987).

Objectives

The objectives of this study is to understand means of identification and mapping of *A. senegal* trees (Hashab) in gum belt using remote sensing techniques in North Kordofan Region, the Sudan, through understanding means of using multivariate analysis to rank different remote sensing-based variables expressing variations in *Acacia senegal* (Hashab), *Acacaia tortilis* (Seyal) and *Liptadenia pyrotechnica* (Marakh) covers.

This study was conducted in North Kordofan State which situated in the central part of Sudan extending from 11.15ºN to 16.40ºN latitudes and from 27.30ºE to 32.25ºE longitudes (Map 1). The State covers an area of approximately 245,000 km^2, representing two third of the Greater Kordofan Region area. (CBS, 2008).

Annual rainfall ranges from less than 50 mm on the northern border to more than 600 mm on the southern border. The length of the rainy season varies from about one month or less in the north to about three months in the south. Rains occur between June-October with the peak in August. Within and between seasons variation in rainfall amount and distribution is common. The average daily temperature ranges between 10-35ºC with an annual variation of 15ºC. April, May and June are the hottest months of the year, and December, January and February are the coolest ones. Wind direction differs according to season; north-east in winter and south-west in summer (Van der Kevie, 1973).

MATERIALS AND METHODS

Site Selection

Acacia senegal (L. Willd) occurs in North Kordofan State in different stand's types. These types of occurrences can be pure Hashab stands or mixed with some other tree species like Marakh or Seyal. In this study pure Hashab stand were presented in ElHimaira Forest, Marakh presented in Umgalgy and Seyal in Mieliha stand. Table (1) showed the sites and their information.

Table (1). Study areas sites

Site	Tree Species	Area (km^2)	Village
1	*Acacia Senegal*	1.76	ElHimaira
2	*Acacia tortilis*	0.4	Mieliha
3	*Leptadinia pyrotechnica*	1.98	Umgalgy

Site selection was done using GPS logger and Arc GIS to make shapefiles for the selected stands. (Figure 1; Plates 1, 2 and 3).

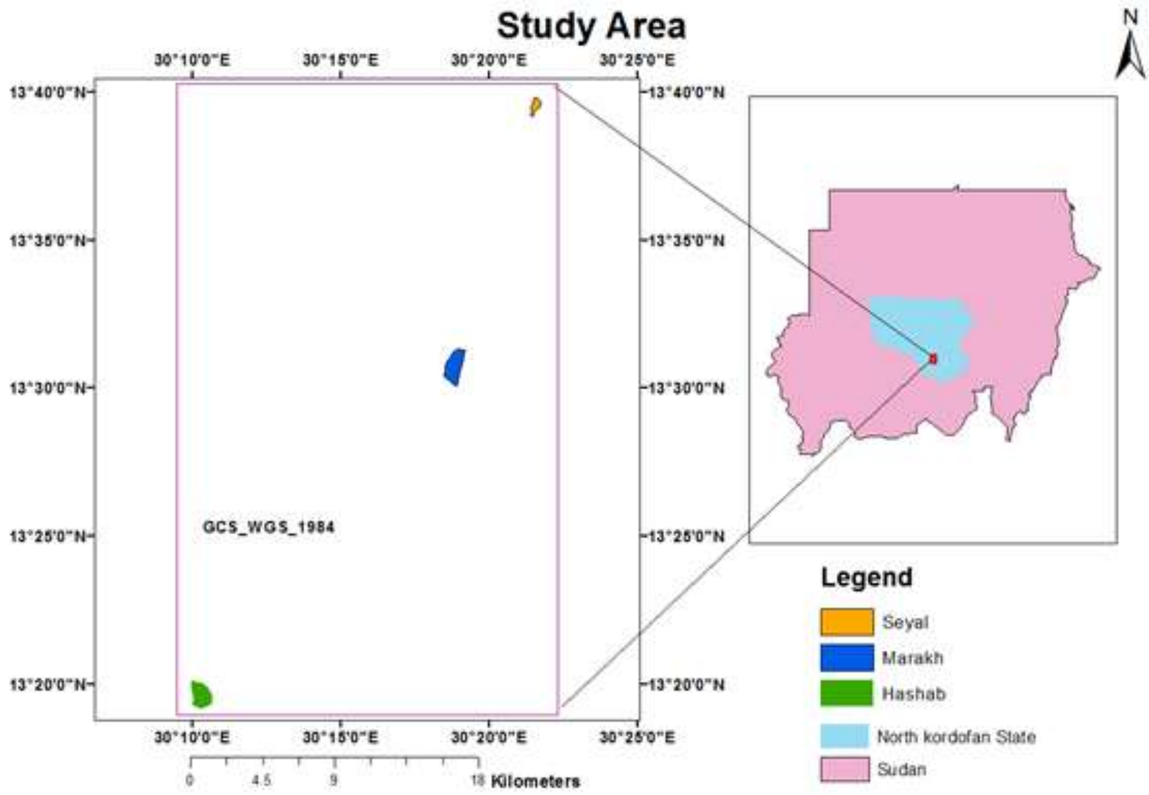

Figure 1: Study area

Plate (1). Hashab in Himaira Forest

Plate (2). Seyal in Meileha

Plate (3). Marakh in Umgalgy

The MOD13Q1v006 and the MOD09A1 products form MODIS were acquired for this study. The MOD13Q1 product provides a 250 m resolution of 16-day imagery including the red (620-670 nm) and infrared bands (841-876 nm) wavelength. This product provides Normalized Difference Index (NDVI).

MOD09A1 provides Bands 1-7 at 500-meter resolution in an 8-day gridded level-3 product in the Sinusoidal projection. This product provide the band 7 with a wave length of (2105-2155 nm) which is an area of soil reflectance that is used as an additional parameter to be analyzed using principal component analysis (PCA) that usually used to reduce the number of variables to be investigated which use in this study.

MODIS data were downloaded from MODIS data site (http://daacmois.orl.gov/cgi-bin/MODIS/GLBVIZI_GIb/Modis_subset_order_global_col5.pl). Fifteen years from 2000 to 2014 MODIS data were studied. MODIS NDVI data of years 2000 to 2012 was already calculated and ready when downloaded but data of years 2013 and 2014 were downloaded in bands and NDVI calculated using below formula with remote sensing software applying the following equation:

NDVI = (float(b2)-(b1))/(float(b2)+(b1))

The analysis of MODIS data was done and all NDVI images were resized to fit the study area frame. After that NDVI data were layer stacked for all years using all months to study the NDVI patterns in dry and rainy seasons. All stands were overlapped in study area frame to get the NDVI mean of all Hashab, Seyal and Marakh in selected sites.

Formula below was used to calculate NDVI means over all fifteen years.

NDVI average = (float(b1)+float(b2)+float(b3)+float(b4)+float(b5)+float(b6)+float(b7)+float(b8)+float(b9)+float(b10)+ float(b11)+float(b12)+float(b13)+float(b14)+float(b15)) /15

Where: b1= red and b2= near infrared

Principal component Analysis (PCA)

Standardized Principal Component Analysis was applied as an additional mean of analysis to reduce a large data matrix into some important data (Principal Components). According to Eastman and Fulk (1993) the standardized PCA approach is more effective than unstandardized PCA in the analysis of change in multi-temporal image data sets. (PCs) are supposed to minimize different sources of noise. Standardized principal Component analysis was used in this study to find the most effective variable data can be used to classify the study area based on the most important data derived from PCA.

PCA was done to identify patterns in data, and expressing the data in such a way as to highlight their similarities and differences. NDVI means of June, August and October in all fifteen years of the study were used beside SAVI means of these months and Band 7, which presented the soil factor. Subset included Hashab, Seyal and Marakh was used.

Preparing PCA data

NDVI means

The formula of NDVI was used to make NDVI from MODIS image when NDVI was not available as in 2013 and 2014 years.

NDVI = (float(b2)-(b1))/(float(b2)+(b1))

Where: b1= red and b2= near infrared

The below formula used to get NDVI means of fifteen years in three months presented the beginning of rainy season (June), Mid of rainy season (August) and end of rainy season (October).

NDVI average: (float(b1)+float(b2)+float(b3)+float(b4)+float(b5)+ float(b6)+float(b7)+float(b8)+float(b9)+float(b10)+ float(b11)+float(b12)+float(b13)+float(b14)+float(b15)) /15

Where: b1= red and b2= near infrared

The soil-adjusted vegetation index (SAVI) was used based on the idea that, depending on the vegetation cover, the NDVI for different cover conditions does not converge at the same location. Therefore, an L-factor is used to adjust the NDVI so that different vegetation densities will intersect the soil line at the same location.

The soil adjusted vegetation index (SAVI) is defined as:

SAVI = (NIR-R)/NIR+R+L)*(1+L)

Where NIR stands for near-infrared, R stands for red and L stands for soil cover. An L factor of 0.5 was used, as suggested by Huete (1988).

MODIS images were used to make SAVI image using below formula:

SAVI = (1.5*(float(b2)-float(b1)))/(float(b2)+float(b1)+0.5)

Where the constant 0.5 has been adjusted to account for first order soil background variation.

SAVI average of June, August and October of fifteen years were done using ENVI 4.7 software by the previous formula in NDVI average.

Bad 7 was extracted from MODIS image with 500 x 500 m resolution to present soil factor in Principal Component Analysis (PCA). This extracted image was resized to fill the frame of study area as NDVI and SAVI images. Band 7 was resampled using ENVI 4.7 software to match NDVI and SAVI images which 250 x 250 m resolution.

PCA calculation

Principal Component Analysis used to sort all these seven factors; NDVI "June, August and October", SAVI "June, August and October" and B7). The basic statistics for each image, mean, variance, standard deviation, variance-covariance, and correlation matrices, eigenvector, eigenvalue, in addition to ratio of contribution of principal components were obtained. The first PC band contains the largest percentage of data variance and the last PC bands appear noisy because they contain very little variance, much of which is due to noise in the original spectral data or the potential to express the variability.

Colour Composite

The colour composite image was made depending on PCA results to differentiate between the three tree species using ENVI 4.7 and ArcGIS software.

RESULTS AND DISCUSSIONS

Principal Component Analysis (PCA) is widely used as dimensionality reduction technique in literature (Jolliffe, 2005, Gonzalez and Woods, 2002). It condenses most of the information spread across many channels into fewer number of channels. Variance of a data set, denoted by eigen values, decreases from first principal component to last component i.e. first principal component contains maximum amount of total variance of the data set (Byrne *et al*., 1980, Gonzalez and Woods, 2002). Last few components contain less variance and hence dropped in classification process. Therefore, by condensing higher number of channels into fewer number it reduces computational demand and possibly improves performance.

In this study, Principal component analysis (PCA) was performed on selected types of data including band 7 which represent soil effects and two vegetation indices (NDVI and SAVI) for three periods in season (June, August and October). PCA was computed to order and sort the effects of these seven components on Hashab, Seyal and Marakh vegetation patterns from 2000 - 2014. Figures (1) shows the seven components before computed to PCA.

The results of the eigenvalues which computed for each of these components are shown in table (2) and figure (2). First three components are chosen for further analysis (Colour Composite) based on the fact that these three components contain 99.83 percentage of total information and the rest of the components were discarded. Components that contain about 95 percent are considered fair for analysis requirements (Eeti *et al.*, 2014).

The results obtained from Table (2) provide a visualization of the relationships between the components and input trends. The eigenvectors indicates that the first component (PC1) displays a strong positive relation with three variables band7, NDVI-August and SAVI-August with loading values of 0.71, 0.37 and 0.32, respectively, while the most values are negative in the rest of components. These three factors presents positive coefficients in the eigenvectors of the variance-covariance matrixes and gave highest values correlated to Component one (PC1). Estomell, J. *et al.* (2013) successfully applied this analysis in Landsat images, they showed that the three principal components may contain over 90 percentage of the information in the original seven bands. These calculations have been widely used in remote sensing to classify the land surface (Jia, and Richards, 1999) and detect changes (Eastman, and Filk, 1993).

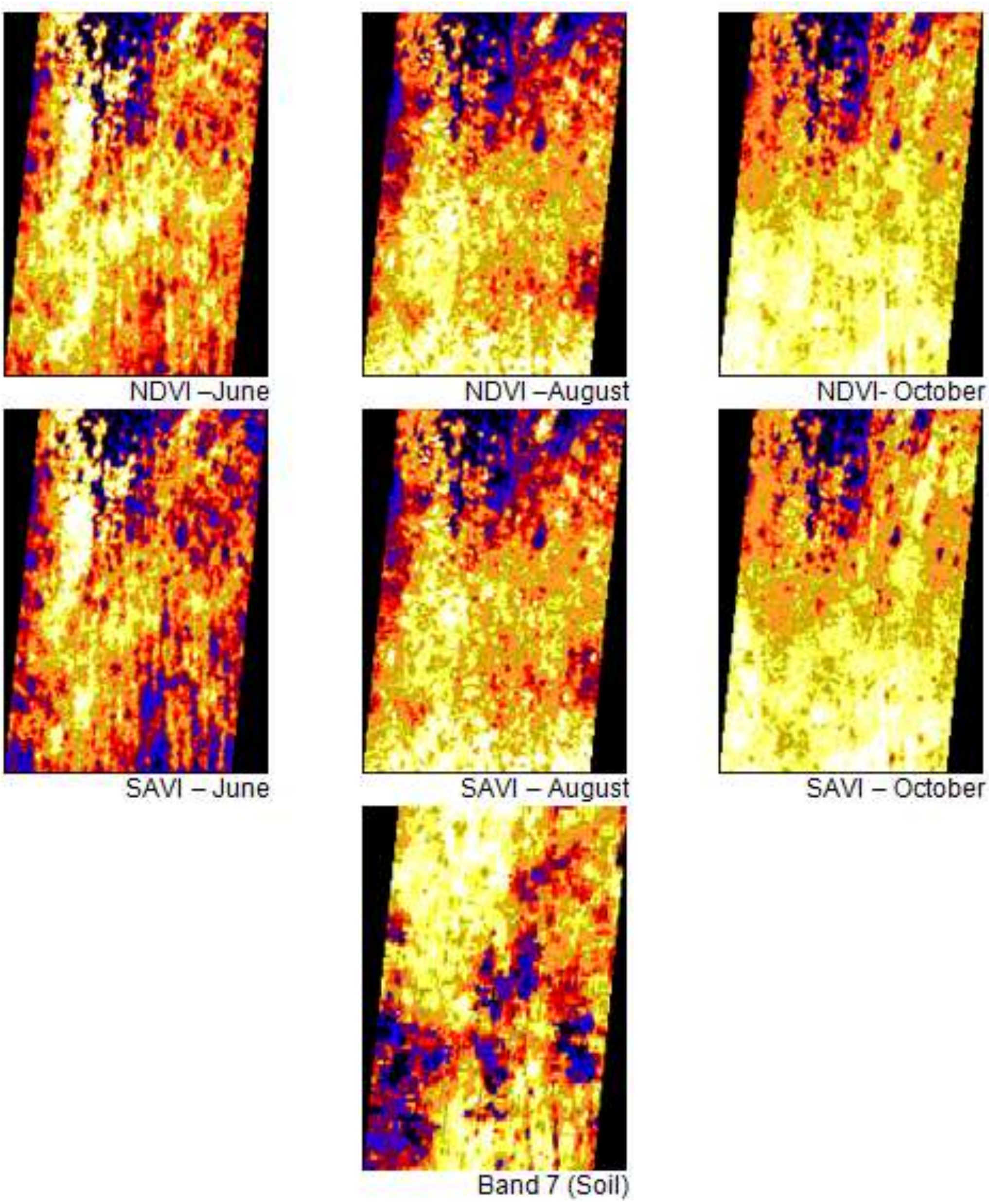

Figure (2). Band 7, NDVI and SAVI images before computed to PCA

This study agreed with (Khairalla, 2013) who revealed that largest contribution of variance is attributed to the first three principal component that are more important for representing the variation of selected rangeland sites. These three components (figure, 4) were used as the second input to the color composite images and classification algorithms to distinguish between Hashab, Marakh and Seyal in study area. Khairalla, (2013), used colour composite images to classify algorithms to distinguish between rangelands sites.

Table (2). Eginvalues and Associated Percentages

ncipal Components	Eigenvalues	Percentages	Cumulative
PC1	0.0618	92.9599	92.9599
PC2	0.0043	6.4222	99.3820
PC3	0.0003	0.4461	99.8281
PC4	0.0001	0.1453	99.9734
PC5	0.0000	0.0191	99.9925
PC6	0.0000	0.0064	99.9989
PC7	0.0000	0.0011	100.000

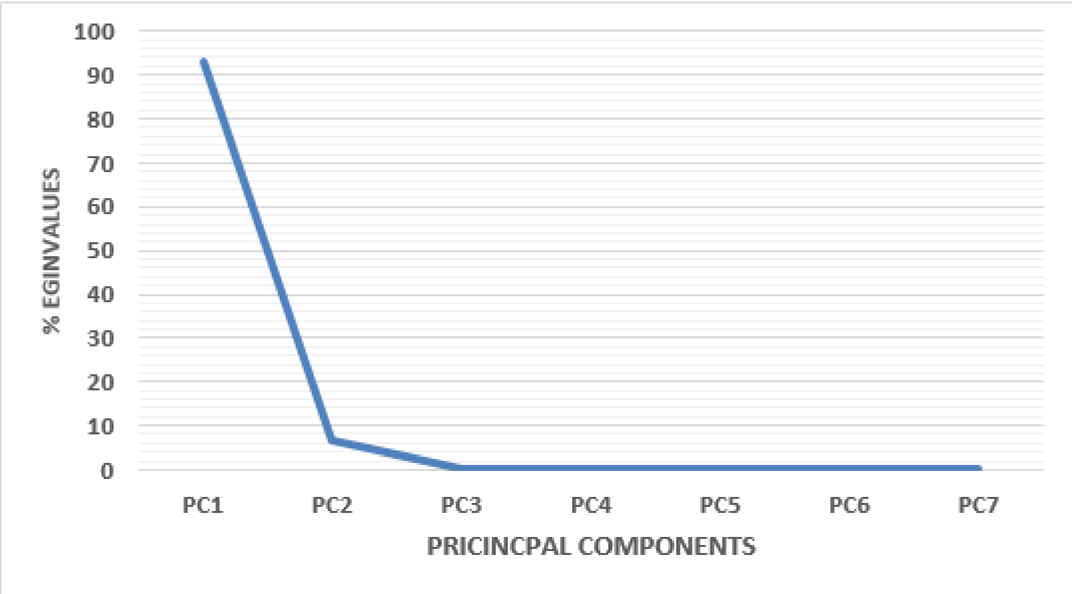

Figure (3). Eginvlues of Principal Component

Table (3). Eginvectors

Principal Components	NDVI Average June	NDVI Average August	NDVI Average October	SAVI Average June	SAVI Average August	SAVI Average October	Band 7 (Soil)
PC1	0.2250	**0.3658**	0.3091	0.1984	**0.3220**	0.2600	**0.7137**
PC2	-0.1699	-0.4322	-0.3295	-0.1305	-0.3481	-0.2319	0.6955
PC3	-0.3757	0.5478	-0.3257	-0.3631	0.3680	-0.4251	0.0685
PC4	0.5155	0.0436	-0.5625	0.5429	0.1452	-0.3134	-0.0435
PC5	0.3130	0.4091	0.3489	-0.0794	-0.6143	-0.4785	0.0141
PC6	0.2190	-0.4583	0.4262	-0.1716	0.4904	-0.5389	0.0041
PC7	-0.6064	0.0002	0.2646	0.6937	-0.0276	-0.2833	-0.0007

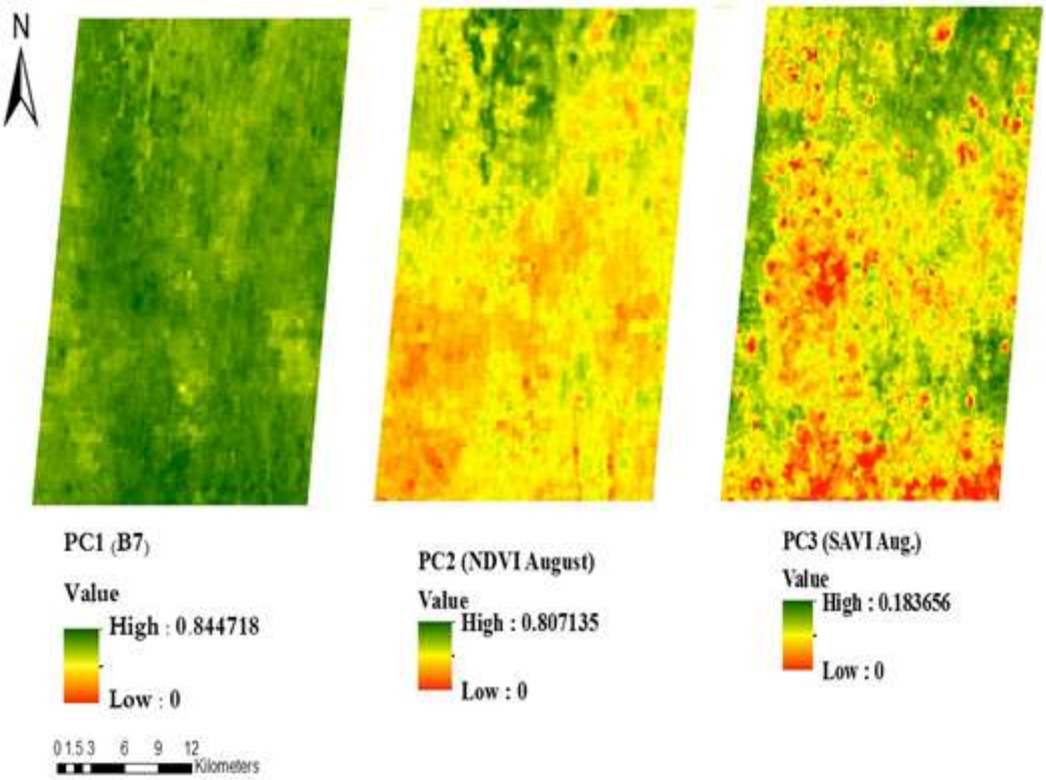

Figure (4): Principal Component Images from the Study area from 2000-2014.

Table (4) shows the correlation between each band and component ranges from +1 to –1. A value of (+1) represents a perfect relationship between the bands and a value of (-1) represents an inverse relationship. The first principal component has high degree of correlation with all seven bands ranging from 0.9722 to 0.9502 with SAVI-August and NDVI-June, respectively. The second PC has highest degree of correlation with band7 (0.2481) lowest and negitive degree of correlation in other bands. PC3 has possitive correlation with NDVI-August, SAVI-August and band7 (soil) with 0.0986, 0.0759 and 0.0064, respectively.

Very low correlation values were observed in the rest of four principal component with all bands which reachs zero value and almost negitive values.

Table (4).Correlation Between Input Rasters and Principal Components

Principal Components	NDVI Average June	NDVI Average August	NDVI Average October	SAVI Average June	SAVI Average August	SAVI Average October	Band 7 (Soil)
PC1	0.9706	0.9502	0.9581	0.9722	0.9586	0.9663	0.9687
PC2	-0.1927	-0.2951	-0.2685	-0.1681	-0.2724	-0.2265	0.2481
PC3	-0.1123	0.0986	-0.0699	-0.1233	0.0759	-0.1094	0.0064
PC4	0.0879	0.0045	-0.0689	0.1052	0.0171	-0.0460	-0.0023
PC5	0.0193	0.0152	0.0155	-0.0056	-0.0262	-0.0255	0.0003
PC6	0.0078	-0.0099	0.0109	-0.0070	0.0121	-0.0166	0.0000
PC7	-0.0092	0.0000	0.0029	0.0120	-0.0003	-0.0037	-0.0000

The results of PCA in this study indicate that, PCA can be used to determine the effect of vegetation type and tree species. Many studies have been used PCA for various purposes including detect geomorphologic features an sediment textural classes (Dewider and Frihy, 2003) as one of the index in decision tree classifier for land use classification (Ding, *et al*. 2011) and to distinguish between geologic features (Sadiq and Howari, 2009).

The color composite used provided a good option for visual interpretation. Jensen (2005) reveals that principal components can be used for the classification of remote sensing and false color compositing. It made up of PC1, PC2 and PC3 images. A false color composite produced by using the first principal component (PC1) as blue, the second principal component (PC2) as red and the third principal component (PC3) as green. Figure (5) shows color composite images of the subset covered three selected tree species (Hashab, Marakh and Seyal). The Hashab site appears yellow, Marakh appears yellowish to oily and Seyal appears almost blue. By comparing this colour composite with NDVI time series groph, Hashab and Seyal are clearly different in NDVI mean values and RGB image while, Marakh is significantly different from Seyal and Hashab in NDVI values. Marakh has dark colour which can distinguish it from Hashab also and the different from Seyal is cleared.

The results support using PCA and colour composite to distinguish tree species. The effect of soil as band7, NDVI and SAVI is significantly affect vegetation cover and vegetation type. Siahaya, W.A. *et al.* 2015. Reported an accuracy estimation of change in land cover using NDVI (94.3%) and colour composite (84.7%). They found that the whole land cover types have experienced increased in both methods, except high density vegetation. Siahaya *et al*. (2015) stated that, for effective and sustainable land cover analysis, the use of vegetation index NDVI and Composite Colour should be encouraged.

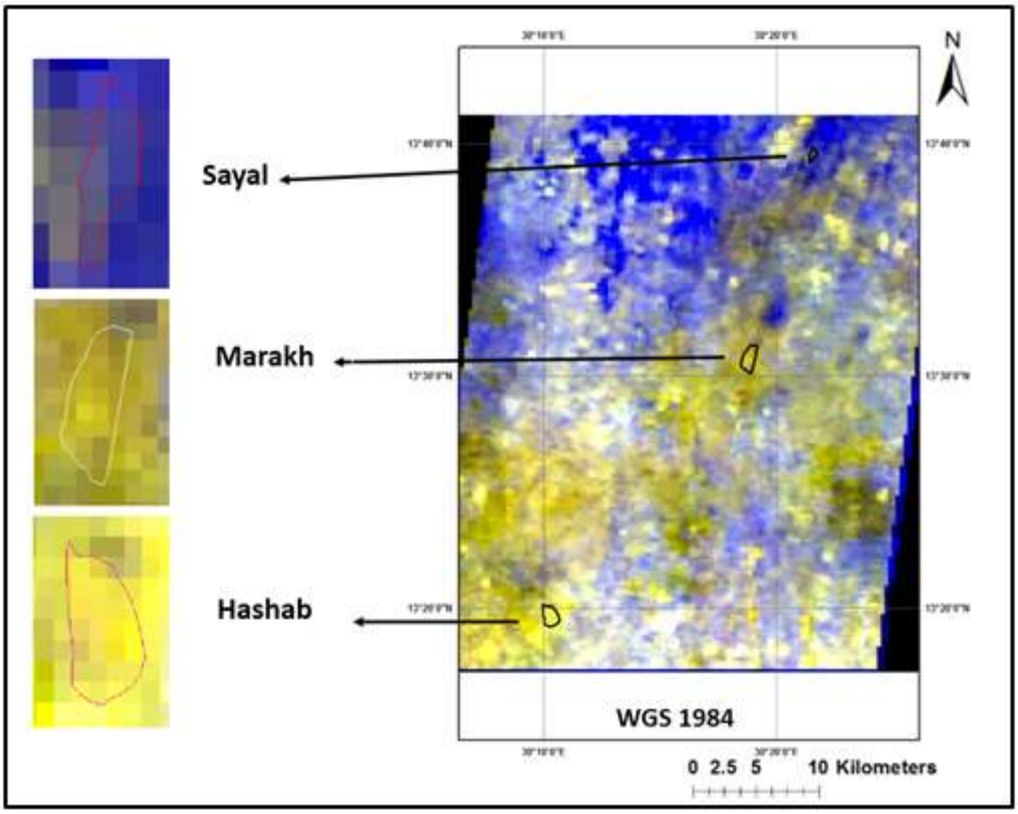

Figure (5). Colour composite of MODIS image (RGB) for better visualization of variation in Hashab, Marakh and Seyal during rainy season from 2000-2014 in study area.

CONCLUSIONS

The findings of this study indicated that the first three principal component "PC" (band7, NDVI August and SAVI August) represent 99.83% of the data and showed high contribution to the difference variance component and eigen-values and could use for differentiate between tree species.

The results also showed that the use of three variable of PC for making colour composite 3, 1, 2>RGB for the period of fifteen years enabled differentiation between three tree species stands in the enhanced images.

AUTHORS CONTRIBUTIONS

Hatim Abdalla M. ElKhidir. The main and corresponding author who write the draft.
Abdelaziz K. Gaiballa. Supervising the study and edited the draft.
Nancy Ibrahim Khiralla. Help in analysis during the study.
Abelrahman A. Khatir. Help in analysis and edited the draft.

REFERENCES

Byrne, G.; Carapper, P. and Mayo, K. (1980). Monitoring Land Cover Change by Principal Component Analysis of Multitemporal Landsat Data. Remote Sensing of Environment 10 (3), pp. 175-184.

CBS. (2008). Central Bureau of Statistics. Ministry of Councils of Ministries. Annual Report Statistical Book. Sudan.

Dewider, K.H.M. and Frihy, O.E. (2003). Thematic Mapper analysis to identify geomorphologic and sediment texture of Eltineh Plain, North-West Coast of Sinai, Egypt. Int.J. Remote Sens., 24: 2377-2385.

Eastman, J. R. and Filk, M. (1993). Long sequence time series evaluation using standardized principal components. Photogrammetric Engineering and Remote Sensing. 59(6) 991-996.

Eeti, L.N.; Buddhiraju, K.M. and Bhattacharyg, A. (2014). A single Classifier using Principal Components vs Multi-Classifier System in Landuse –Landcover Classification of World View -2 Sensor Data. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume II.8. 2014. ISPRS Technical Commision VIII Symposium, 09-12. 2014, Hyderabad, India.

Estomell, J.; Jesus, M.; Mart-Gavila, M.; Teresa Sebastia, M. and Mangual, J. (2013). Principal Component Analysis Applied to Remote Sensing. Modelling in Science Education and Learning. Volume 6 (2). No. 7. Instituo Univestitario de Matematica Pura y Aplicada. Valencia, Spain.

Gonzalez, R. C. and Woods, R. E. (2002). Digital image processing Using Matlab, Pearson Prentice Hall, Upper Saddle River. doi:10.1117/1.3115362.

Gonzalez, R. C. and Woods, R. E. (2002). Digital image processing Using Matlab, Pearson Prentice Hall, Upper Saddle River. doi:10.1117/1.3115362.

Holden, H.; LeDrew, E. (1998). Spectral discrimination of healthy and non-healthy corals based on cluster analysis, principal components analysis and derivative spectroscopy. Remote Sens. Environ. 1998, 65, 217-224.

Huete, A.R. (1988). A soil-adjusted vegetation index (SAVI)' Remote Sensing of Environment, vol. 25, issue 3, pp. 259-309. DOI: 10.1016/0034-4257(88)90106-X

Jensen, J. R. (2005). Introductory Digital Image Processing: a Remote sensing perspective. 3rd Ed. Pearson Education, ISBN 0-13-145361-0, Prentice Hall, New Jersey.

Jensen, J. R. (2005). Introductory Digital Image Processing: a Remote sensing perspective. 3rd Ed. Pearson Education, ISBN 0-13-145361-0, Prentice Hall, New Jersey.

Jia, X. and Richards, A. (1999). Segmented Principal Components Transformation for E cient Hyperspectral Remote-Sensing Image Display and Classi cation. IEEE Transactions on Geoscience and Remote Sensing. 37(1). pp. 538{542.

Jolliffe, I. (2005). Principal component analysis. Wiley Online Library.

Khairalla, N. I. A. (2013). Monitoring and Assessment of Rangeland in Semi Arid of Sudan using Remote Sensing Techniques: North Kordofan State. Sudan University of Science and Technology. Ph.D. Thesis. 2013

Patra, S. K. Manish Shekher, S. S. Solanki, R. Ramachandran & R. Krishnan (2006). A technique for generating natural colour images from false colour composite images. International Journal of Remote Sensing, 27:14, 2977-2989.

Sabins, F.F. (1987). Remote Sensing: Principles and Interpretation, 2nd edition (W.H.Freeman and Company).

Sadiq, A. and Howari, F. (2009). Remote sensing and spectral characteristics of desert and sand from Qatar Peninsula, Arabian/Persian Gulf. Remote Sens, 1: 915-933

Siahaya, W.A.; Danoedoro, P.; Khakhim, N. and Baiquni, M. (2015). The comparison analysis of land cover change based on vegetation index and multispectral classification (a case study at Leihitu Peninsula of Ambon City District). Journal of Degraded and Mining Lalnds Management. ISSN: 2339-076X, Volume 2, Number 4 (July 2015): 415 – 422

SSA. (1975). Sudan Survey Authority. Annual Report. Sudan.

Zhao, G.; Maclean, L. (2000). A comparison of canonical discriminant analysis and principal component analysis for spectral transformation. Photogramm. Eng. Remote Sens. 66, 841-847.

Cite this Article: Abdelrahman, AK;Nancy, IAK;Abdelaziz, KG;Hatim, AME (2025). Mapping *Acacia senegal* (L.) Willd and other shrubs in North Kordofan State, Sudan using Remote Sensing Principal Component Analysis. *Greener Journal of Agricultural Sciences*, 15(1): 76-84, https://doi.org/10.15580/gjas.2025.2.011723007.