Soil classification using airborne hyperspectral data employing various approaches

Justin George K[1*], Vinay Kumar[1], Bhukya Tarun[2], Suresh Kumar[1], A Senthil Kumar[1]
[1] Indian Institute of Remote Sensing, ISRO, 4 Kalidas Road, Dehradun -248001, Uttarakhand, India
Email: justin@iirs.gov.in
[2] National Institute of Technology Karnataka, Surathkal, Mangalore - 575 025, Karnataka, India

KEY WORDS: AVIRIS-NG, Soil Classification, Spectral Angle Mapper (SAM), Support Vector Machine (SAM) and Minimum Noise Fraction (MNF)

ABSTRACT: Hyperspectral remote sensing technology is one of the advance technology for detailed land cover feature extraction. Hyperspectral datasets contain large number of contiguous spectral bands with a narrow spectral bandwidth which enables identification of peculiar absorption features for distinguishing different type of soils. The potential of Airborne Visible Infrared Imaging Spectrometer-Next Generation (AVIRIS-NG) data was tested for distinguishing black and red soils in the ICRISAT area near Hyderabad, Telangana. The AVIRIS-NG data captured in 432 narrow contiguous bands (346–2505 nm) with spectral sampling of 5 nm bandwidth and a 4m ground pixel size was used in this study. The dataset was first spectrally subsetted by identification and removal of bad bands and was atmospherically corrected by converting it to surface reflectance using FLAASH. The data was finally georeferenced using the Internal Geometry Module (IGM) parameters. Optimal spectral bands from the reflectance data were selected on the basis of different characteristics of various soils. Data dimensionality reduction technique Minimum Noise Fraction (MNF) was also performed to extract noise free components. Total five classes including red and black soils were considered for land cover classification. Pixel based classification techniques such as Spectral Angle Mapper(SAM) and Support Vector Machine (SVM) were performed on the reflectance as well as MNF transformed data. SVM was also performed on data containing noise free MNF components and the selected optimal spectral bands. In the resultant classified output of reflectance data, SVM classifier provided higher accuracy and was able to classify black and red soil in a better way than SAM technique. The results also suggested that use of MNF components and specific spectral bands altogether improvised the classification of black and red soil.

1. INTRODUCTION

Rapid and accurate detection of the characteristics and changes of the Earth's surface is a necessity for better understanding of the relationships and interactions between natural and human phenomena, which provide a foundation for the management and judicious use of natural resources (Lu et al., 2004). Hyperspectral remote sensing imagery enables identification and classification of different land cover features including different types of soils, mainly due to the information contained in large number of contiguous spectral bands with a narrow spectral bandwidth. The information present in the airborne hyperspectral dataset is coupled with improved spatial resolution which enables accurate identification of peculiar absorption features for distinguishing different land cover features. Airborne hyperspectral imagery not only gives abundance of spectral responses but also provides good ground pixel size that increase the classification accuracy.

To extract meaningful information from hyperspectral imagery, selection of optimum bands from the data is crucial for certain classes and also essential for better classification for that particular class. Band selection is one of the best way to reduce storage and processing time for hyperspectral remote sensing data (Landgrebe, 2003). Despite rich and fine spectral information of hyperspectral data, curse of dimensionality and Hughes phenomenon affect the land use/cover classification accuracy of such images. Hence it is required to reduce the high dimensionality and inherent multi-collinearity of the hyperspectral data, specifically for efficient and effective feature extraction using various techniques like Minimum Noise Fraction (MNF), Principal Component Analysis (PCA), Independent Component Analysis (ICA) etc. to increase the accuracy of the classification. Land cover classification using remote sensed data is challenging due to spectral and spatial complex features present. MNF components are used to improve the spectral separability of thematic classes in airborne hyperspectral imagery using any classifier (Green et al.,1988).

Frassy et al (2013) had used Multispectral Infrared Visible Imaging Spectrometer (MIVIS) data for classification and to distinguish between two different features. They reported that classification of MNF-transformed data succeeded to do that, whereas classification of original hyperspectral data failed. This study demonstrated the use of MNF as a pre-processing procedure in improving the capability to extract specific information from airborne hyperspectral datasets (Frassy et. al., 2013). Hyperspectral data is normally affected with noise and redundancy of information (too many channels) which can be taken care of by MNF transformation which primarily isolate the noise, followed by generation of components that describe most of the variance present in the dataset. This technique resulted in modifying the dataset, that provide more accurate classifications when compared with geologically calibrated check areas than when the raw data are used alone (Harris et al., 2006). In this situation, optimal bands selection based on optimization procedures has high potential to improve the accuracy of hyperspectral image feature extraction and classification (Samadzadegan, et al., 2012).

To classify reflectance bands and MNF transformed output of hyperspectral data, Hegde et al (2014) had used two classifiers Spectral Angle Mapper (SAM) and Support Vector Machine (SVM). The results expressed that SVM gave better accuracy than SAM for urban land cover classification. It was also observed that SVM classifier gave best accuracy for MNF components in comparison to classification using reflectance bands (Hegde et al, 2014). The study concluded that the MNF bands are giving better accuracy than original bands. SVM was found to give more accuracy compared to random forest (RF) and maximum likelihood classifiers. Moreover, SVM gives good results even if we have less number of training datasets (Burai et al., 2015).

2. STUDY AREA

The study area is located in and around Patancheru of Sangareddy District, in the state of Telangana, located at the southern portion of India (Fig 1). It is at a distance of nearly 30kms from the main city centre on the Hyderabad-Solapur highway. The study area is approximately 20km x 10km (Total swath covered by airborne survey). The swath covered includes large stretches of agricultural fields, where variety of crops are cultivated throughout the year and thus was identified for crop and soil related studies using AVIRIS data. ICRISAT, the international organization conducting agricultural research in semiarid tropics for rural development is located within our study area.

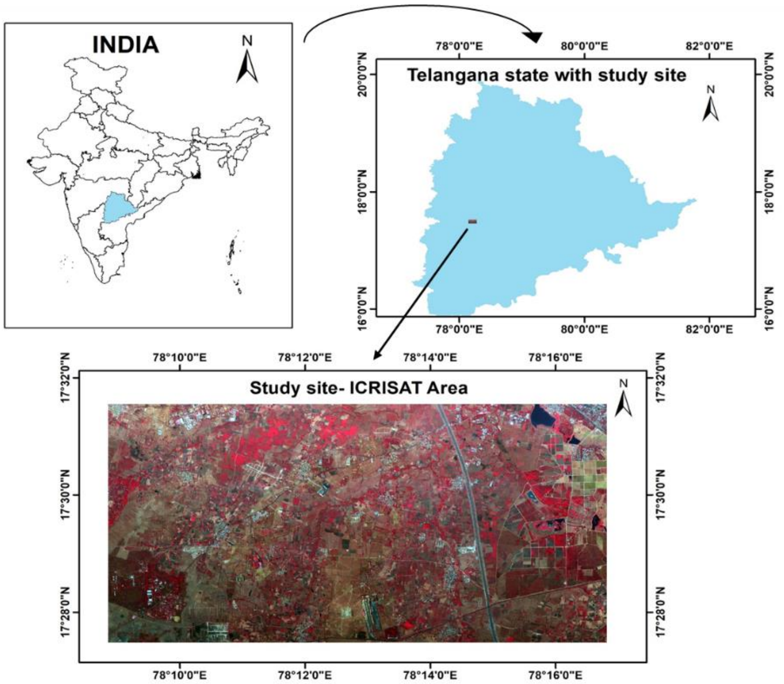

Fig 1: Location of study site within Telangana and FCC of ICRISAT site

The major crops cultivated and recorded in the area during field campaign are Rice, maize, tomato, Bengal gram, cotton, red gram, brinjal, niger etc at various locations. Majority of the area is under rice based cropping systems. The major soil types present in the study area are red soil, black soil and mixed soil.

3. DATA USED

The AVIRIS NG radiance data (L1) of study site was collected from National Remote Sensing Centre (NRSC), Hyderabad. The study area was covered as five strips. The details regarding the AVIRIS_NG data are given in the below.

Table 1: AVIRIS NG Data Details of ICRISAT

Site Name	Date of Acquisition	Centre Lat& Long	No. of Bands	Wavelength (μm)	Resolution (m)
ICRISAT	19 /12/2015	17.438 & 78.134	432	0.3466 – 2.5054	4.0

4. METHODOLOGY

The overall methodology adopted in the study is depicted in Fig 2.

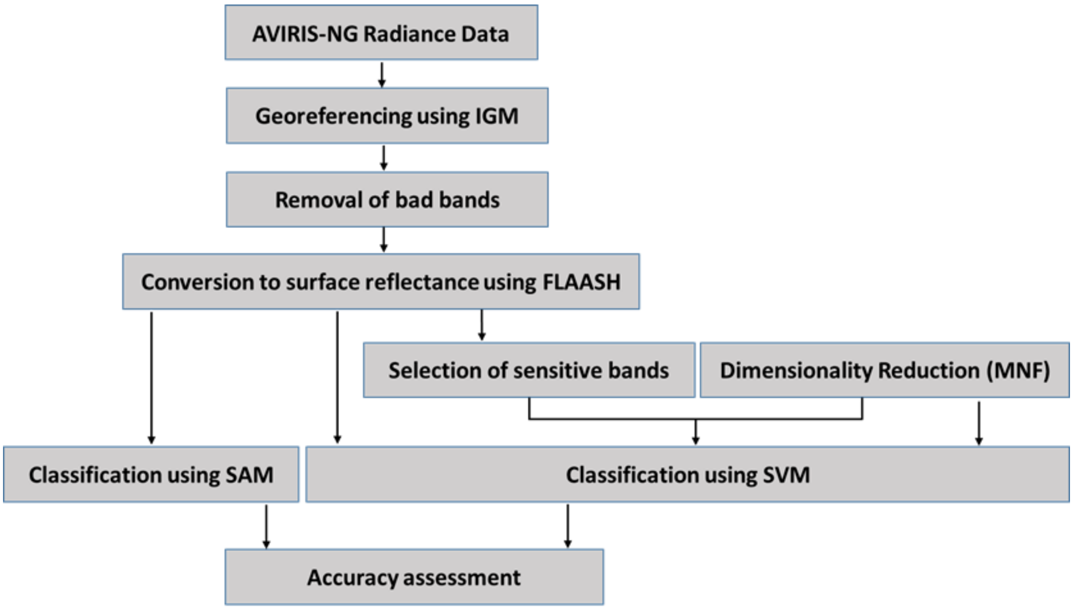

Fig 2: Overall methodology adopted in the study

4.1 Pre-Processing of Data

AVIRIS-NG data for ICRISAT site was obtained as 05 strips in radiance form. All the strips were checked for identification of bad bands with high amount of noise. All bands having low information/high noise and bands with bad columns were removed from the data. Overall total 359 bands were left after the removal of all the bad bands. After that the datasets were atmospherically corrected and converted to surface reflectance using FLAASH. All the atmospheric corrected five strips were geo-referenced using the IGM module parameters which was provided along with the radiance data. All the five georeferenced data strips were joined together by mosaicking to get a continuous image of the ICRISAT study area.

4.2 Selection of Optimal Bands

For selecting optimal bands, spectral profiles of various classes from atmospherically corrected data were plotted using training samples and their differences in characteristics were studied. Using these plots, optimal bands were selected by observing the band position where maximum differences in the reflectance characteristics of various classes were observed. This is clearly distinguishable as the spectral plots of different classes do not intersect at these wavelengths. The optimal bands which were selected based on their best discriminating ability were with band numbers 49, 82, 100, 130, 230, 280, and 304 corresponding to central wavelengths of 0.65, 0.817, 0.907, 1.057, 1.658, 2.069 and 2.189 respectively. These selected bands helped in better identification and classification of various classes because of their high separability at these corresponding wavelengths. Figure 3 shows spectral characteristics of different land cover features.

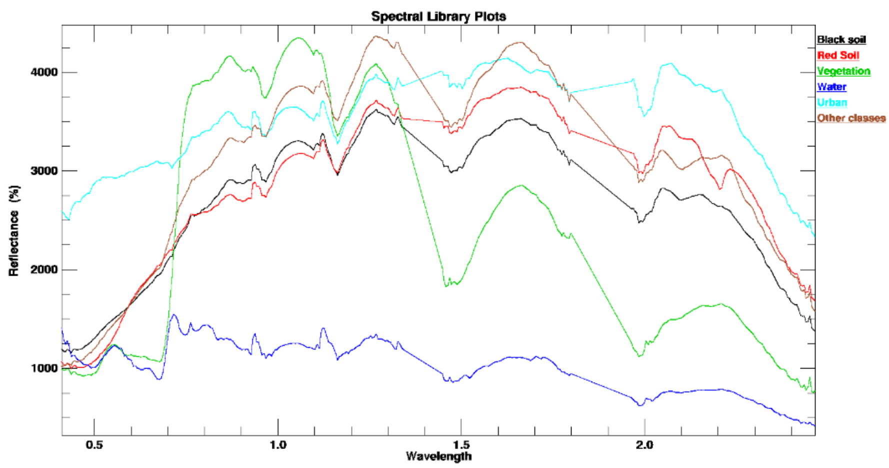

Fig 3: Spectral profiles of various land cover classes

4.3 Data Dimensionality Reduction Using MNF

Hyperspectral data in comparison to multispectral, contain huge amount of spectral information because of the continuous nature of spectral bands and their narrow band widths. This voluminous data may result in redundancy of information across different bands. Data dimensionality reduction technique like Minimum Noise Fraction (MNF) is widely used to bypass the redundancy issue and for feature extraction from these voluminous remote sensing datasets such as AVIRIS-NG. MNF components were generated in order to explore the high information content present in hyperspectral remote sensing data, by segregating the noise from data. MNF transformation helps to achieve high signal to noise ratio. In this study a forward MNF transformation was applied to entire reflectance image in order to take care of the problems due to high data dimensionality.

The first 20 MNF components, with high signal to noise ratio (SNR), were selected for further processing and classification. These bands accounted for most part (>99%) of variance present in the original data. The output of MNF are uncorrelated bands ranked based on their eigen values, from highest to lowest.

4.4 Image Classification

For the identification and classification of various classes/features, two supervised classification methods namely Spectral Angle mapper (SAM) and Support Vector Machine (SVM) were used in this study. First SAM and SVM was performed on the reflectance data. For SAM classification multi values of maximum angle in radians ranging from 0.1 to 0.3 are given for classifying five different classes. SVM which is an advanced technique for classification was performed with Radial Basis Function kernel. SVM parameters (C = 100 and = 0.05) are selected to perform SVM classification on reflectance and MNF bands. SVM was also carried out for classifying data which contained optimal spectral bands and MNF components altogether.

4.5 Accuracy Assessment

The accuracy of the obtained classified datasets was assessed using ground truth data collected during field survey. The accuracy was estimated from the datasets by deriving various parameters such as overall accuracy, kappa coefficient, producer's and users' accuracy in addition to visual inspection for identification of potential errors and misclassification.

5. RESULTS AND DISCUSSION

The classification results obtained by using various approaches are shown in Fig 4. The results revealed that SAM classifier produced lowest overall accuracy of 79.32% whereas the SVM method outperformed it with an overall accuracy of 81.12% when only reflectance image was used as input for classification (Table 2). The high accuracy of SVM classifier can be attributed to the consideration of nonlinearity by the classification technique. The overall accuracy and kappa have shown significant improvement when the SVM classification was carried out on the MNF transformed output as well as combining data of optimal spectral bands with MNF components. This is mainly due to the segregation and removal of noise from useful data, during MNF transformation. However, the increase in overall classification accuracy with combined optimal bands with MNF components was very marginal in comparison to classification using only the MNF components (increase from 86.64% to 86.87%).

Table 2: Accuracy assessment results for classified outputs using various approaches

Classification approach	Overall accuracy (%)	Kappa coefficient
SAM using reflectance image	79.32	0.748
SVM using reflectance image	81.12	0.772
SVM using MNF components	86.64	0.837
SVM using optimal bands and MNF components	86.87	0.84

The classification results with respect to black and red soils were also compared and red soils have shown very good accuracies in all the different classification approaches used in the study (Table 3). Whereas the black soil has low producer's and user's accuracy in comparison to red soil classification in all the approaches. This may be because of the distinct colour of red soils due to the presence of iron oxides, which show much more distinct characteristics effects on the overall spectra in comparison to black soil. The characteristic features of black soil spectra may be obliterated by effect of soil organic matter as well as increased moisture conditions occurring at the time of data acquisition. In case of black soil, the accuracies when reflectance image was classified using SAM classifier increased from less than 80% to more than 80% when SVM classifier was employed. In case of black soil, both the producer's and user's accuracies increased drastically when MNF components were used alone or in combination with optimal bands instead of using only reflectance image as input for classification. Whereas in case of red soil, there was only a very slight increase in user’s accuracy, as we move from the SAM classification of reflectance image approach towards different SVM classification approaches using different input layers. This was accompanied by a decrease in producer’s accuracy in case of red soil classification.

Table 3: Accuracy assessment results for black and red soil using various approaches

Classification approach	**Black soil**		**Red Soil**	
	User’s accuracy	**Producer’s accuracy**	**User’s accuracy**	**Producer’s accuracy**
SAM using reflectance image	77.52	72.92	98.10	92.20
SVM using reflectance image	81.96	54.9	98.28	80.56
SVM using MNF components	89.13	77.71	98.86	77.62
SVM using optimal bands and MNF components	90.17	77.40	98.86	77.62

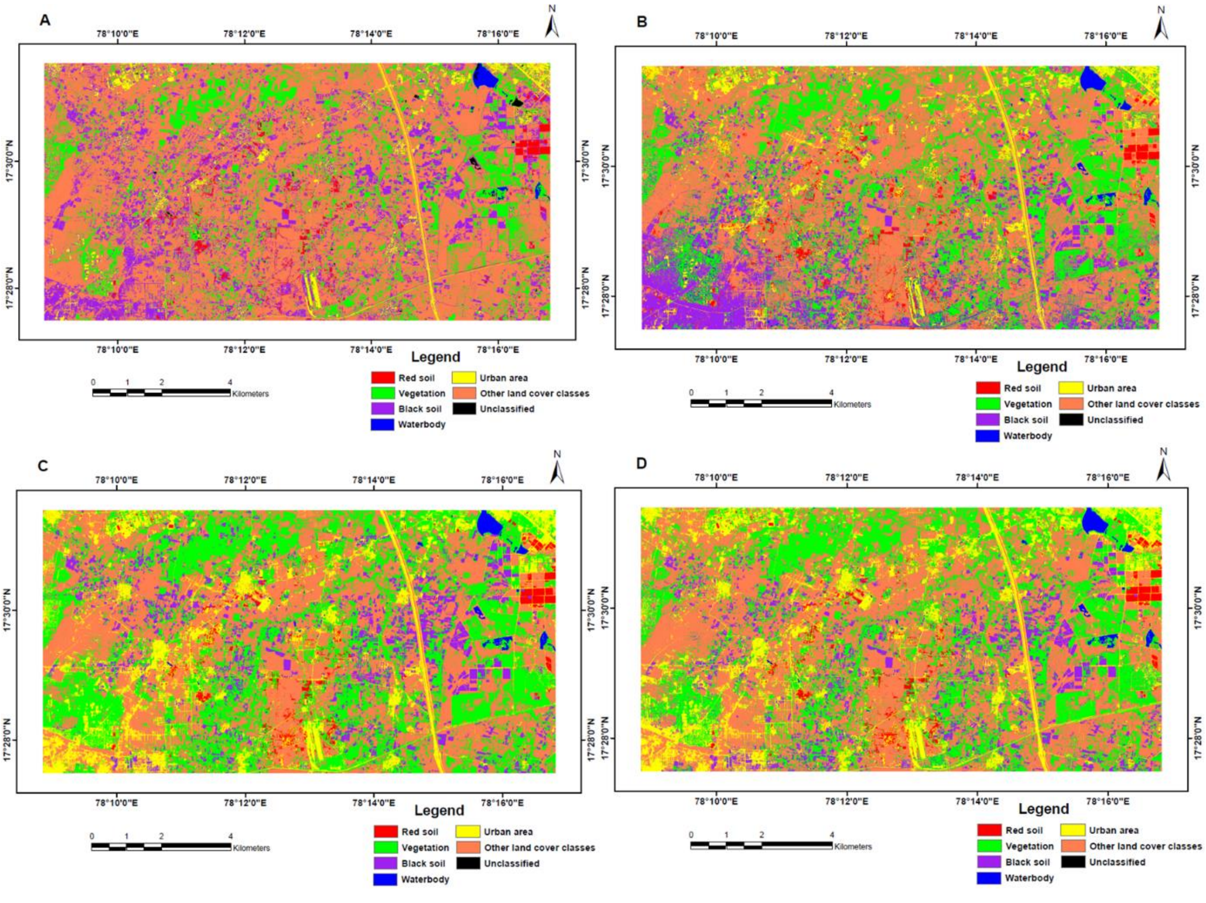

Fig 4: Classification results using various approaches such as (A) SAM using reflectance image, (B) SVM using reflectance image, (C) SVM using MNF components and (D) SVM using MNF components along with optimal bands.

6. CONCLUSIONS AND RECOMMENDATIONS

This study was aimed at accurate identification and classification of various soil classes using airborne hyperspectral remote sensing data(AVIRIS-NG) employing various approaches. The study revealed the improved performance of SVM classifier in comparison to SAM classifier for hyperspectral data classification. The MNF transformation performed in the study helped us to reduce redundancy and obtain noise free components. These noise free MNF components significantly improved the overall classification accuracies when used as input layers alone or in combination with selected optimal spectral bands, in comparison to the original reflectance image with large number of bands. The classification of red soil could be accomplished with higher accuracies when compared to black soil. Also the computing or processing was faster in case of SVM when compared to SAM and it again reduced when less number of bands were used for classification in the form of MNF components and optimal bands instead of original reflectance image.

7. ACKNOWLEDGEMENT

The authors would like to acknowledge the AVIRIS-NG coordinators from SAC, Ahmedabad and field campaign team from NRSC, Hyderabad for giving us the opportunity to be a part of AVIRIS-NG campaign and for providing necessary support during field data collection.

8. REFERENCES

Burai, P., De´ak, B., Valk´o, O., and Tomor., T., 2015 Classification of Herbaceous Vegetation Using Airborne Hyperspectral Imagery. Remote Sensing, 7, 2046-2066.

Frassy, F., Dalla Via, G., Maianti, P., Marchesi, A., Rota Nodari, F., & Gianinetto, M., 2013. Minimum noise fraction transform for improving the classification of airborne hyperspectral data: two case studies. 5th Workshop on Hyperstectral Image and Signal Processing: Evolution in Remote Sensing.

Green, A. A., Berman, M., Switzer, P., and Craig, M. D., 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions on Geoscience and Remote Sensing, 26(1), pp. 65-74.

Harris, J. R., Ponomarev, P., Shang, J., and Rogge, D., 2006. Noise reduction and best band selection techniques for improving classification results using hyperspectral data: application to lithological mapping in Canada's Arctic. Canadian Journal of Remote Sensing, 32(5), pp. 341–354.

Hegde, G. J., Ahamed, M., Hebbar, R., and Uday Raj., 2014. Urban Land Cover Classification Using Hyperspectral Data The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-8, pp. 751-754.

Landgrebe, D.A., 2003. Signal Theory Methods in Multispectral Remote Sensing, John Wiley & Sons, New York

Lu, D., Mausel, P., Brondizio, E. and Moran, E., 2004. Change Detection Techniques. International Journal of Remote sensing, 25(12), pp. 2365–2407.

Samadzadegan, F., Hasani, H., and Schenk, T., 2012. Simultaneous feature selection and SVM parameter determination in classification of hyperspectral imagery using ant colony optimization. Canadian Journal of Remote Sensing, 38, pp. 139–156.