Principal Component Analysis of multispectral remote sensing data for spectral characterization of hydrocarbon seepage induced alteration of subaerial regolith in a part of Son Valley Vindhyan Basin

Ashish Misra[1], Prem Chandra Maurya[2]*
[1]* KDMIPE, ONGC, Dehradun (misra_ashish@ongc.co.in); [2] GSI, Patna

Keywords
Multispectral remote sensing, VNIR-SWIR, spectral characterization, hydrocarbon seepage, Principal Component Transformation, FPCS

Summary
The main research objective of this work is to remotely detect and map the characteristic mineral assemblages and chemical alteration owing to seepage of gaseous hydrocarbons in a part of northern Son Valley sector of Vindhyan Basin contiguous to Damoh & Hatta in Madhya Pradesh, India. Active hydrocarbon exploration has already firmly established subsurface presence of hydrocarbons in the area.
Some of the characteristic alteration features include: bleaching of red beds, anomalous enrichment of ferrous-iron, clay minerals and carbonates. Many of these alteration minerals exhibit diagnostic spectral features in Visible through Shortwave Infrared range (VNIR-SWIR; 0.4µm-3.0µm) of the electromagnetic spectrum, which makes them amenable for detection using remote sensing techniques.
Landsat-8 Operational Land Imager (OLI) multispectral optical remote sensing data of the study area, processed to suppress the effects of surface vegetation has been utilized for the purpose. Spectral characterization involved application of first order Principal Component Transformation of the 6-band multispectral VNIR-SWIR dataset to compute the covariance eigenvectors among the multispectral bands to determine the principal components (PCs) with maximum opposite-sign loadings. Subsequently, four-band composites of multispectral bands with such highest covariance differences and highest spectral contrast for target mineral assemblages were subjected to second order feature-oriented principal component selection (FPCS) technique to derive PCs that reflect surface abundances of the target alteration mineral groups. Colour composites of these PCs clearly highlighted anomaly areas corresponding to sites of surface seepages.
The results have been field validated. Subaerial regolith samples have been collected from seepage sites (altered) as well as from unaltered locations. Distinct macroscopic differences in the soil properties could be observed. The results were further confirmed through spectral match among multispectral image and resampled library spectra which prove the efficacy of spectral characterization of HC seepage induced subaerial regolith alteration based on remote sensing techniques.

Introduction
The study area is located in Damoh district of Madhya Pradesh state, around Hatta in Central India (Fig. 1). The most prominent topographic feature of the Son Valley synclinorium in the study area is manifested by gently S-SE dipping (4° to 6°) homoclinal ridges of resistant sandstone units towards northern flank and flat alluvial surface dominantly underlain by incompetent beds of shale and dolomitic limestone. Stratigraphically, the outcropping rock units belong to the Rewa and Bhander Groups of Upper Vindhyan sequence.

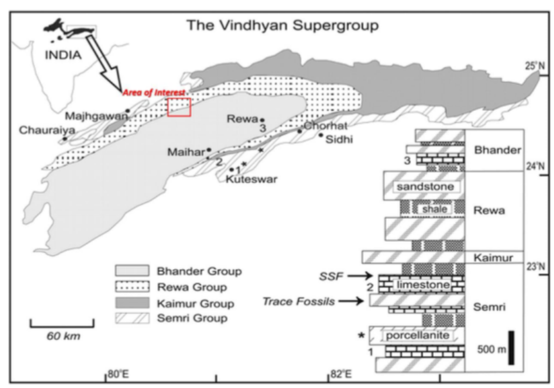

Figure-1 Regional setting and broad stratigraphic framework of Son Valley Vindhyan Basin; study area is indicated by a red box

The area is richly intersected by a network of non-perennial streams/ irrigation canals and interspersed by small irrigation ponds. In general the alluvial cover is in excess of ~20m thickness which is heavily cultivated. Whereas the land cover in rocky areas is sparsely forested.
Multiple macroseeps of flammable gaseous hydrocarbons have been reported intermittently from

agricultural fields in the area, with most of them are clustered around Mariyadu (Madiadoh) towards the base of the rocky ridge towards northern limits of the study area (e.g., Madiadoh/ Mariyadu, kaikheda, Devlai & Himmat-Patti). Structurally, the surface lineaments in the study area dominantly exhibit a NW-SE grain. A map showing surface geological units, lineaments and known seepage sites is presented in Fig. 2.

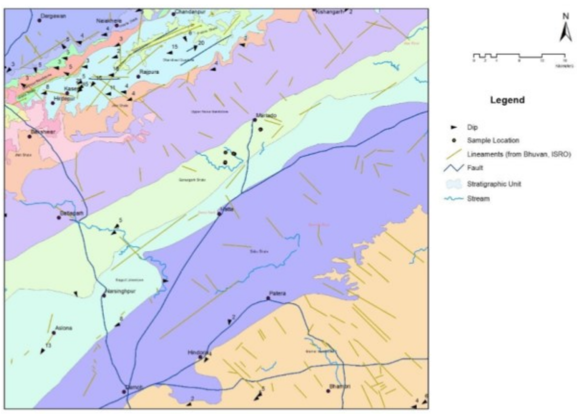

Figure-2 Geological map of the study area; mapped lineaments and surface HC seepage sites have been indicated (field sites for ground truthing and sample collection)

Theory

The assumption is that generated hydrocarbons migrate from source and trapped at depth may escape in small amounts to the surface. The amount of seepage is related to the lithology of overlying rock, pressure in oil/gas reservoirs, which is related to hydrostatic pressure and changes in lithospheric stress (Freek van der Meer, 2002). There are two types of hydrocarbon seepage – Macroseepage and Microseepage. The term macroseepage refers to visible oil and gas seeps. Hydrocarbon microseepages are defined as light, invisible, analytically detectable volatile and semi-volatile hydrocarbons moving vertically to the surface through seal and overburden rocks. Microbubble buoyancy or near-vertical migration is the dominant process in microseepage hydrocarbons (Freek van der Meer et al, 2002).

Direct evidence of hydrocarbon microseepages includes the presence of non-biogenic hydrocarbons, such as ethane or butane, in soils. Indirect evidence includes a suite of closely interrelated geomorphic, chemical, geobotanical, mineralogical, and geophysical anomalies, etc. (Fig. 3a). Mineralogical changes may promote geomorphic and physical alterations of the surface and near-surface strata that may be detected using remote sensing techniques. These changes may include: bleaching of red beds, radioactive anomalies, clay mineralization, visual tonal and textural anomalies, carbonate cementation, etc. A generalized model of hydrocarbon induced geochemical and geophysical alteration of soils and sediments has been suggested by Schumacher (1996) (Fig. 3b).

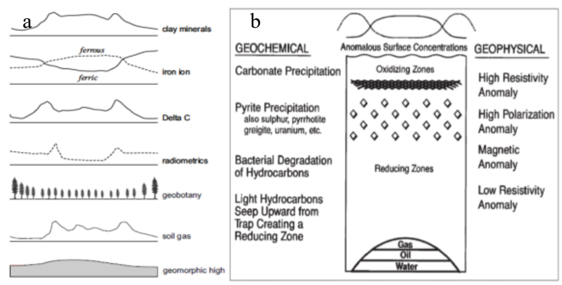

Figure-3 (a) Anomalies that can result from seeping hydrocarbons (after Werff, 2006); (b) Generalized model of hydrocarbon induced geochemical and geophysical alteration of soils and sediments (Schumacher, 1996)

In the region of reflected solar EMR (0.3-2.5μm), many minerals have diagnostic absorption features due to vibrational overtones, electronic transition, charge transfer and conduction. Altered rocks/ soils due to seeping hydrocarbons typically contain higher concentrations of clay minerals, such as kaolinite, montmorillonite, illite, alunite with a distinct Al-OH/ Mg-OH absorption feature at 2.2 μm and a less intense absorption feature at 2.35 μm (Fig. 4A). Iron oxide and hydroxide minerals such as limonite, jarosite and hematite tend to have spectral absorption features in the visible to middle infrared from 0.4-1.1 μm of the electromagnetic spectrum (Fig. 4B). Additionally, the increased percentage of Delta C leads to precipitation of calcite, with Fe, MgOH and CO_3 exhibiting absorption features in 2.31-2.33 μm range of reflected EMR.

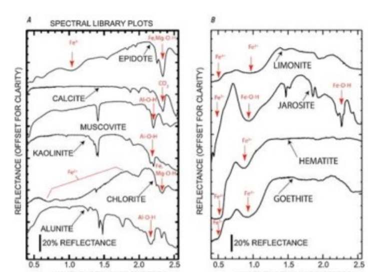

Figure-4 USGS spectral library spectra of (A) clay minerals – epidote, calcite, muscovite, kaolinite, chlorite and alunite; and (B) iron oxides – limonite, jarosite, hematite and goethite (after Clark et al., 2003)

Principal Component Analysis of multispectral remote sensing data for spectral characterization of hydrocarbon seepage induced alteration of subaerial regolith in a part of Son Valley Vindhyan Basin

Data, Methodology & Image Processing Results

Landsat-8 OLI 6-band multispectral satellite image data corresponding to VNIR-SWIR channels acquired on 21.05.2015 forms the primary dataset. Ancillary data include compiled and mosaicked ASTER GDEM surface elevation data (30m spatial resolution), public domain 1:50000 scale geological map and location information of known HC macroseepage sites. All geospatial data were appropriately georeferenced and co-registered and a project GIS was created for final interpretation.

The step-by-step approach to address the study objectives is shown in Fig. 5.

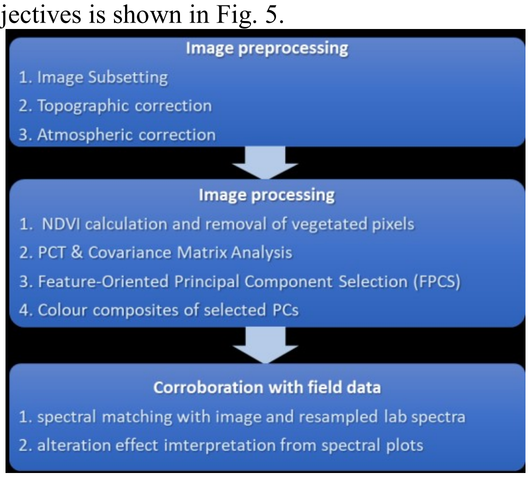

Figure-5 Methodology adopted for the study

Surface reflectance image subset corresponding to the study area derived from radiative transfer model-based atmospheric correction of the raw Landsat Top-of-Atmosphere (TOA) radiance data is shown in Fig. 6a and the topographic layout of the study area based on ASTER GDEM based surface elevation model is depicted in Fig. 6b.

To suppress the effects of surface vegetation on spectral response of bare earth, a normalized difference vegetation index (NDVI) image derived using NIR and red bands was processed using iterative thresholding (NDVI range -0.077 to -0.073) to map vegetation-dominated pixels (Fig. 6c). These pixels were excluded from the multispectral surface reflectance data to obtain the final vegetation-suppressed multispectral surface reflectance dataset (Fig. 6d).

Based on the logic that multispectral image data can be enhanced digitally to detect the presence of subtle variations in the concentration of target mineral species provided the remote sensing data are collected in spectral bands where the diagnostic spectral features lie, lab spectra of the expected minerals, soils and rock types expected in the study area, as indicated in the published maps, were compiled. The full resolution spectral library spectra were resampled corresponding to the band centres of the Landsat-8 OLI VNIR-SWIR image bands for the soils and rocks (containing target alteration minerals) (Fig. 7a,b for rocks and Fig. 8a,b for soils).

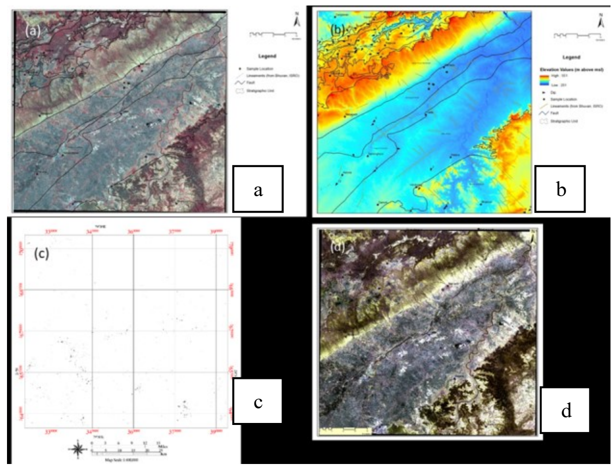

Figure-6 a) False Colour Composite (FCC) surface reflectance image subset (bands 5-4-3 (RGB)) corresponding to the study area; b) Digital Elevation Model extracted from ASTER GDEM degree tile for the study area; c) NDVI threshold image; d) vegetation-suppressed surface reflectance image data FCC; absence of red pixels indicates successful vegetation suppression. Lithological boundaries, dips and lineaments have been overlain on the images for clarity.

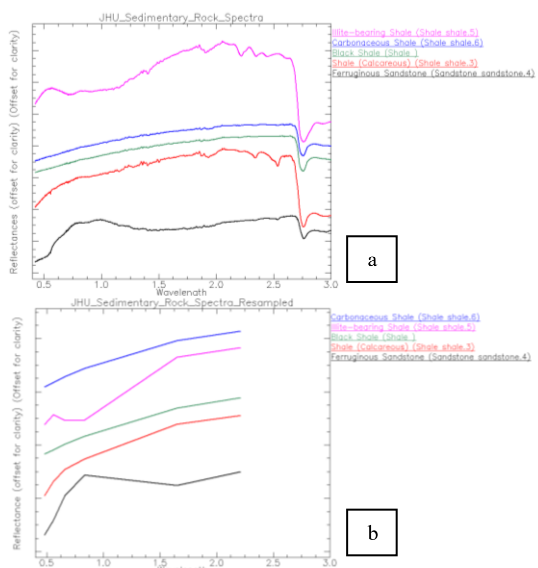

Figure-7 JHU spectral library spectra of sedimentary rocks expected in the area: a) full feature, b) resampled to Landsat-8 image channels

Principal Component Analysis of multispectral remote sensing data for spectral characterization of hydrocarbon seepage induced alteration of subaerial regolith in a part of Son Valley Vindhyan Basin

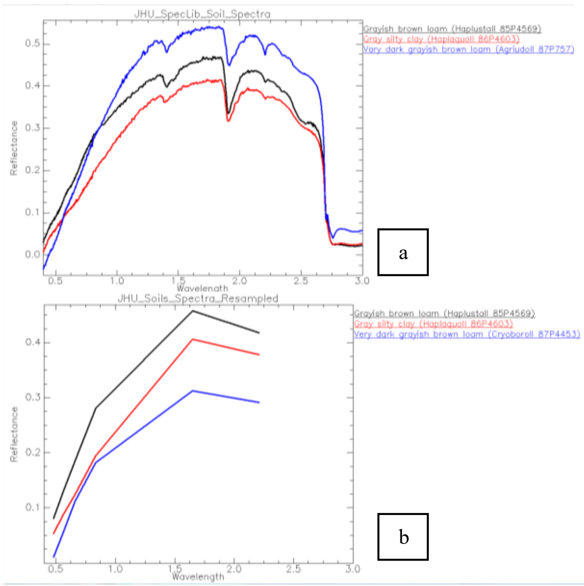

Figure-8 JHU spectral library spectra of soil types present in the study area: a) full feature, b) resampled to Landsat-8 image channels

From resampled spectra it can be observed that spectral details present in the full resolution laboratory spectra are lost but important absorption features remain well preserved.

Principal Component Transformation (PCT) has been implemented in ENVI image processing software, which involves a two-step process. In the first step *n* histograms of the scene are used as inputs to the principal component algorithm. This algorithm calculates *n* principal components, which are actually orthogonal vectors in n-dimensional space that are oriented along directions of maximum remaining variance. The variances of the principal components are called eigenvalues. The total variance of all principal components sum up to 100% of the total variance of the data; with the first three PCs usually accounting the majority (50% - 95%) of the variance, when $n > 3$. The second step consists of the transformation of the image to principal component space. Using the equation, *n* principal component values for each pixel are calculated. The resulting PCs include information from all other bands. The first principal component displays the terrain features and incorporates highest contribution from all input channels. The *n*th principal component image usually gives the homogeneous image, interrupted by a few bright and dark pixels that are spectrally unique for that image scene, based on relative contributions (separation) among input image bands.

A color composite of PC6:PC5:PC3 in RGB for PCA performed on bands 2, 3, 4, 5, 6 and 7 of the Landsat-8 OLI image of the study area showing maximum variance for the image data is shown in Fig. 9a.

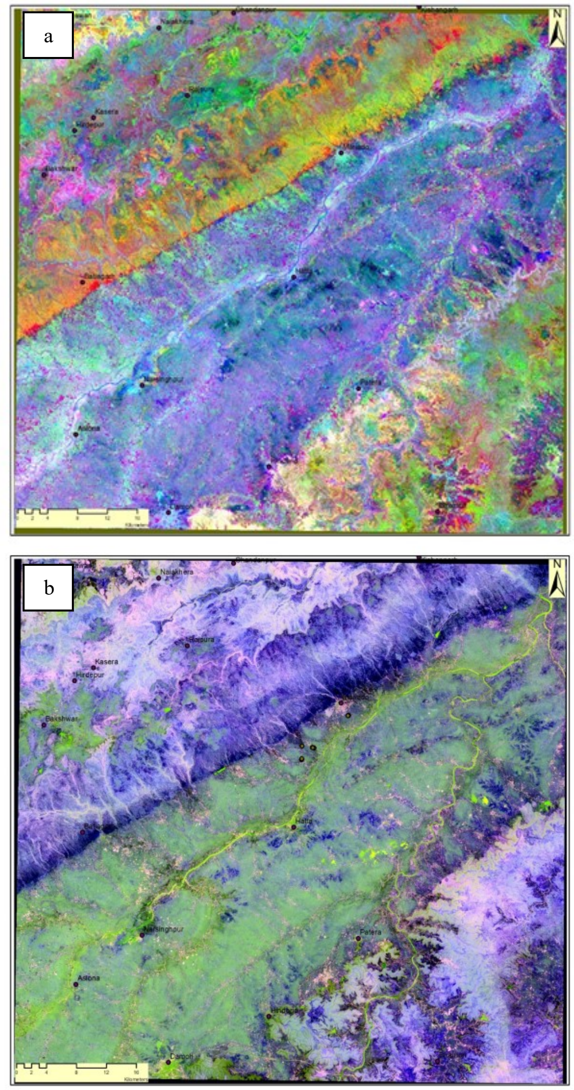

Figure-9 a) Result of PCT: FCC of PC6:PC5:PC3 in RGB, showing highest spectral variance within the study area; b) Result of FPCS analysis: Color composite of 2456-PC3 (red), 2467-PC3 (green) and 3467-PC3 (blue), respectively. Darker pinkish-purple shades indicate areas of iron oxide and/or clay/ carbonate mineral anomalies.

Feature-oriented Principal Component Selection (FPCS) (Crosta & McM-Moore, 1989) is based on the examination of eigenvector values to determine the Principle Components (PC) that concentrate spectral information about specific target materials. Besides, it predicts whether the target materials are represented by bright or dark pixels in the PC images according with the magnitude and sign of the eigenvectors. In FPCS the PC transformation is performed for a select combination of bands such that a relationship between the spectral responses of target materials and numeric value extracted from eigenvalue matrix used to calculate the PCs is established. In this study, the spectral subsets were selected according to the position of characteristic spectral features of target alteration minerals exhibited in their resampled spectra. Based on the detailed analysis of eigenvectors and eigenvalues derived from different combinations of PCA, four principal components (PCs) including 2456-PC3 (the third component of the PCA on band 2, 4, 5, 6), 2467-PC3 and 3467-PC2 showed steady spectral enhancement for the hydrocarbon-induced target alteration minerals. Fig. 9b shows the colour composite of 2456-PC3 (red), 2467-PC3 (green) and 3467-PC3 (blue), respectively. Eigenvector matrices for these indices have been summarized in Table-1.

Table-1 Eigenvector matrices for FPCS analysis of Landsat-8 OLI 6-band multispectral image dataset for the study area

	Band 2	*Band 4*	*Band 5*	*Band 6*
PC1	0.532804	0.627358	0.543089	0.166120
PC2	0.650240	0.124155	-0.731816	-0.161924
PC3	-0.540508	0.760840	-0.357857	0.030184
PC4	0.034039	-0.110134	-0.203563	0.972252

	Band 2	*Band 4*	*Band 6*	*Band 7*
PC1	-0.641376	-0.742370	-0.186637	-0.051874
PC2	-0.766426	0.610857	0.181896	0.079740
PC3	-0.026368	0.273256	-0.940033	-0.202422
PC4	-0.023091	0.032736	0.220043	-0.974667

	Band 3	*Band 4*	*Band 6*	*Band 7*
PC1	0.659162	0.728002	0.181419	0.051045
PC2	0.711450	-0.522953	-0.441818	-0.158608
PC3	-0.239410	0.439681	-0.853006	-0.147457
PC4	0.045021	-0.056692	-0.210392	0.974933

Field Validation

Field data corresponding to similar dates as the date of scene acquisition was collected to bring closest comparison between the image analysis results and actual ground truth. Validation of image-based spectral characterization results was carried out through field sampling of regolith collected from three HC seepage sites, namely – Kaikheda, Deori Fatehpur, and Mariyadu/ Madiadoh.

All of these sites are located in agricultural fields. Samples were obtained by digging up to 2 ft. shallow ditches to avoid effects of diurnal weathering. Representative altered samples were collected from the reported seepage sites, and additional samples representing unaltered areas were collected to ascertain the petrographic characteristics of regolith (Fig. 10). Distinct differences in macroscopic properties (colour & texture) among altered and unaltered samples were observed. Altered samples exhibit visible effects of bleaching (reduction), and more nodular nature owing to enrichment of clays & carbonates.

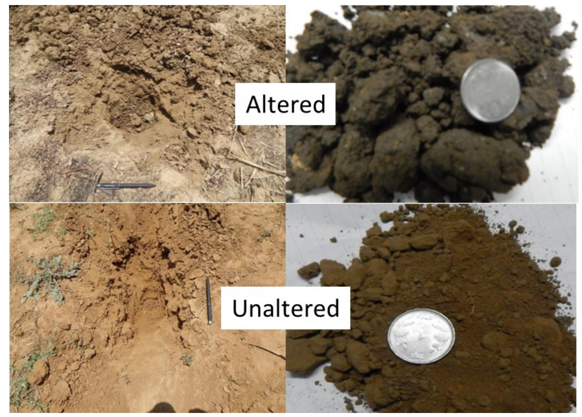

Figure-10 Representative field photographs of altered and unaltered regolith from one of seepage sites (Kaikheda)

Spectral Characterization

The results of spectral characterization are shown in Fig. 11. Image spectra for pixels covering 'altered' soil and 'unaltered' soil for the three seepage sites in the study area have been plotted along with lab spectra for anticipated surface sediment/ soil types from JHU spectral library.

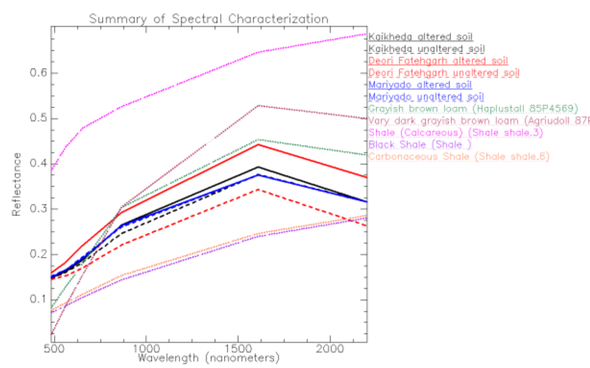

Figure-11 Plot of image spectra corresponding to pixels of field sample sites and JHU spectral library spectra of regolith type

It can be seen that distinct differentiation among altered and unaltered image pixel spectra is clearly brought out. Also, a close correspondence between soil type 'Very Dark Grayish Brown Loam' and most image pixel spectra has been achieved which exibits the power of remote sensing based surface type detection even with low resolution multispectral satellite data, such as Landsat-8 OLI.

Discussion & Conclusions

Mapping of hydrocarbon microseepages and their surface manifestations became an issue of consideration for petroleum exploration in the last few decades since they can be direct indicators of possible subsurface hydrocarbon reservoirs. Remote sensing techniques offer a powerful way to detect such subtle changes in the surface mineralogy/ soil character/ geobotanical anomalies owing to their cost & time effective synoptic coverage of large regions. While the spectral feature of solid/ liquid hydrocarbons are too subtle (a faint absorption near 1.7 µm) to be sensed from a spaceborne system, its effect on the surface mineralogy leads to increased concentration of carbonates, clays and ferrous oxides, while reduction in ferric iron compounds. Minerals containing these chemical compositions have long been known to exhibit strong and distinctive spectral characteristics (absorption), particularly in the visible through infrared region of the EM spectrum (0.4 to 3.0 µm) which have been the basis of their remote detection even from spaceborne imaging systems, such as the popular Landsat satellites.

Through established image processing methods, the area has been assessed for distinct alteration anomalies resulting from hydrocarbon seepage. PCT of 6-band subset of Landsat-8 OLI scene led to creation of a color composite of PC6:PC5:PC3 with highest spectral variance. Similarly, in order to highlight the distribution of iron oxides and clays/ carbonates, the image was processed using Crosta technique which focuses on the PC selection based on spectral nature of the material of interest. Thus, selectively, PCT on bands 2456, 2467 and 3467 was performed and the most diagnostic PCs from these FPCS were combined to produce a color composite image showing relative abundances of iron oxides, clays & carbonates. Overlaying known seepage site locations on spectral anomalies helped in correlating alteration/ anomalous areas with regolith properties. Subtle enrichment of the expected mineralogical compounds was observed for all sites, as exhibited in spectral slope variations in corresponding pixels, however, any clear distinction could not be observed between seepage site spectra and background spectra, possibly for two reasons:

1. The area is highly cultivated, and soil cover is regularly reworked, with adding of nutrients which obliterates the natural soil properties and effects of seepage
2. The dips of the strata in the region are gentle, which could mean that the background seepage anomalies are very widespread, instead of being localized

It is possible that a more exhaustive spectral classification of the data might reveal even more subtle features, but it was beyond the scope of this study. The study, however, was highly successful in revealing the direct detection/ distinction of surface sediments/ soils using multispectral remote sensing data as a very close match of the spectral curve shapes between the image derived spectra over field sites and lab spectra for the type rock/ soil were achieved.

References

Clark R.N., Swayze, G.A., Wise, R., Livo, K.E., Hoefen, T.M., Kokaly, R.F., and Sutley, S.J. 2003. USGS Digital Spectral Library splib05a, U.S. Geological Survey, Open File Report 03395.

Crosta, A. P., and J. McM.Moore, 1989. Enhancement of Landsat Thematic Mapper Imagery for Residual Soil Mapping in SW Minais Gerais State, Brazil: A Prospecting Case History in Greenstone Belt Terrain. Proceedings of the 7th (ERIM) Thematic Conference: Remote Sensing for Exploration Geology. Calgary, 2-6 Oct., p. 1173-1187.

Meer, F.V.D., Dijk, P.V., Werff, H.V.D. and Yang H. 2002. Remote sensing and petroleum seepage: a review and case study; Terra Nova, 14, No.1, 1–17.

Schumacher, D. 1996. Hydrocarbon-induced alteration of soils and sediments, AAPG Memoir 66, p. 71-89.

Werff, H.V.D. 2006. Knowledge-based remote sensing of complex objects: recognition of spectral and spatial patterns resulting from natural hydrocarbon seepages. Unpub. PhD Thesis.

Acknowledgements

Ashish Misra expresses his gratitude to Director (E), ONGC for his kind permission to publish this work. He is further grateful to HoI, KDMIPE Dr. Harilal, Head Geology Group, KDMIPE Shri Shekhar Srivastava and I/c Remote Sensing & Geomatics, Geology Group, KDMIPE Dr. Subhobroto Mazumder for their encouragement and support.

P.C. Maurya is grateful to DG, GSI for permission to publish the work which was undertaken as part of his M.Tech. Dissertation at IIT Roorkee. His is grateful to his supervisors for their encouragement and support.