INTERNATIONAL JOURNAL OF GEOMATICS AND GEOSCIENCES
Volume 4, No 3, 2014

Research article ISSN 0976 – 4380

Natural resource mapping using hybrid classification approach: Case study of Cooch Behar District, West Bengal

Aditi Sarkar[1], Pabitra Banik[1], Rana Dattagupta[2]

1- Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata-700108, India.

2- Computer Science and Engineering Department, Jadavpur University, Kolkata – 700032, India.

banikpabitra@gmail.com

ABSTRACT

An ever increasing population associated with various other factors has put tremendous pressure on the environment and its respective resources. Hence it has become necessary to study the human interactions with its surrounding environment to facilitate sustainable developmental plans. Land use and land cover is the most dominant factor to study these interactions. Remote Sensing technology has become an indispensable tool to evaluate the environmental processes and provide adequate information of available resources and to design strategies for the sustainable use of these resources. In the current study an attempt has been made to study the land use and land cover in the Cooch Behar district by mapping the natural resources of the area. The study has been made by integrating Geographical Information System (GIS) approach with Hybrid classification using District Planning Map and multispectral Indian Remote Sensing-P6 - LISS_III data for 2008 and 2009 based on an emerging digital classification technique. The land use land cover classification was accomplished on the basis of prior knowledge of the study area instead of the traditional classification schemes and the satellite imageries. A hybrid supervised/unsupervised classification approach coupled with GIS analyses was employed to generate land use/cover maps with eight classes; upper terrace, single crop, double crop, forest, degraded forest, rivers and water bodies, sandy area and settlements. The study highlights the usage of a non-conventional classification scheme purely based on researcher's in-depth knowledge of study area and application of a hybrid classification technique for far more accurate land use mapping.

Keywords: GIS, Remote Sensing, Landuse/ land cover, multispectral imagery, hybrid classification

1. Introduction

One of the most universal applications of Remote Sensing is image classification for generating land use and land cover maps (LULC). LULC is a critical aspect of the earth as it reflects the availability of resources at human disposal. Since long LULC classification methods purely relied on visual interpretation of satellite imageries by experts. However of late automatic classification methods have evolved. Automatic classification algorithms with improved accuracy have become more desirable to lessen the costs of photo-interpretation. Hybrid classification is one such algorithm that combines the unsupervised and supervised techniques to classify an image. It combines the benefits of - unsupervised (Iterative Self Organizing Data Analysis Technique) being non-biased, statistical method to separate clusters; followed by supervised (Maximum Likelihood Classifier) which utilizes the analyst's knowledge of area. There are more than one ways to execute a Hybrid classification.

A methodology for automatic classification of land cover from high resolution multi-spectral IKONOS imageries was developed by Zingaretti et al. (2009). A hybrid classification approach coupled with GIS analysis to generate land use and land cover pattern for Bindura District of Zimbabwe was employed Kamusoko and Aniya (2007). They monitored and analyzed land use land cover changes using Landsat satellite imageries. A vegetation mapping methodology was proposed to Landsat imagery, using a hybrid approach in classifying vegetation species diversity of the Lepini mountain chain, Italy by Caprioli et al. (2003). Monitoring and analyzing land use/ land cover changes become decisive as was established by Chapin et. al. (2000). They highlighted how the basic environmental processes were markedly affected by land cover and any ensuing changes in them from local to global scales. In the past the only way to study an area of interest was through in-situ observation and sampling. Defries and Townshend (1999) studied how Remote Sensing had fundamentally altered the approach of scientists in studying land, vegetation and other environmental aspects of the earth's surface. The current study aimed at mapping the land use and land cover of the Cooch Behar district through an exhaustive image classification technique to monitor the local natural resources. Such an attempt seems to be scarce in the area chosen for this particular study.

2. Study area

West Bengal has a vast diversity in terms of physiography, climate, soils, natural vegetation, agricultural practices and land uses. In India, the existing land administration system is a British legacy considering the block as an administrative boundary.

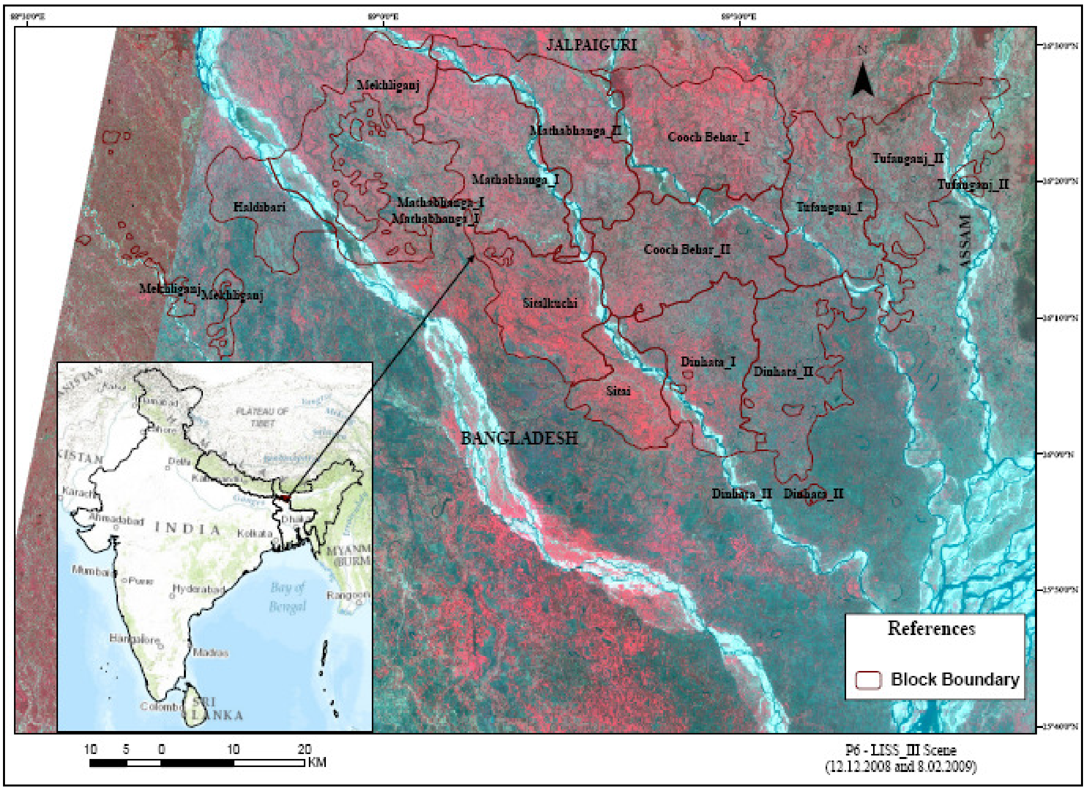

Figure 1: The map of the study area

Hence the analysis has been pre-conceived with the block boundary as the basic theme boundary. The choice of the area has been guided by the huge variability in terms of the natural as well as economic backdrop of the district. The area chosen for the study was Cooch Behar district (Figure 1). It is situated in the foothills of Eastern Himalayas and forms a part of the Himalayan Tarai of West Bengal. It is bounded by Jalpaiguri in the north, Assam in the east and Bangladesh in the south and south west. The district occupies a total area of 335765.79 ha. The areal extent of the district is between 25^0 57' 47" to 26^0 36' 20" N and 88^0 47' 44" to 89^0 54' 35" E. Cooch Behar and its surrounding regions face deforestation due to increasing demand for fuel and timber, as well as air pollution from increasing vehicular traffic. In absence of large forest area in the district, except at Patlakhawa, not many species of animal are found. The district is mainly agricultural in nature, and the economy is dependent on the same. Among the major crops grown here are cereals like wheat, rice, pulses, tobacco and jute. The coverage of the crops is gradually increasing in the region.

3. Data and software used

The data sets obtained for the analysis consists of

3.1 Satellite images

IRS-P6 – Linear Imaging Self-Scanning-III raster data (**Table 1**) with spatial resolution of 23.5 meters, taken on 12th December 2008 and 8th February 2009 was considered for the study. To obtain the entire district coverage three scenes were used.

Table 1: Specification of satellite data used for land use/ land cover classification

District	Latitude	Longitude	IRS-P6 (LISS - III)		Date of Pass
			Path	Row	
Cooch Behar	25°55' to 26°40'	88°45' to 90°	108	53	8.02.2009
			106	55	12.12.2008
			106	56	12.12.2008

3.2 Ancillary data

Ancillary data used for the study are - a) National Atlas and Thematic Mapping Organization (NATMO) District Planning map (DPM) of Cooch Behar at 1:250, 000 scale; b) NATMO map of West Bengal water resources for geo-referencing the district planning map; c) The district map of Cooch Behar obtained from the Census, 2001.

3.3 Software used

ERDAS IMAGINE 9.0 software was used for image pre-processing. ArcMap 10 software was used to accomplish the spatial database creation. Finally Microsoft Office was used for the documentation and calculations pertaining to the study.

4. Methodology

The study has made use of multi-sourced primary and secondary data. The LISS - III satellite data was digitally interpreted to study the LULC pattern. The following flow chart (**Figure 2**), gives an elaborate insight into the methodology adopted to accomplish the study.

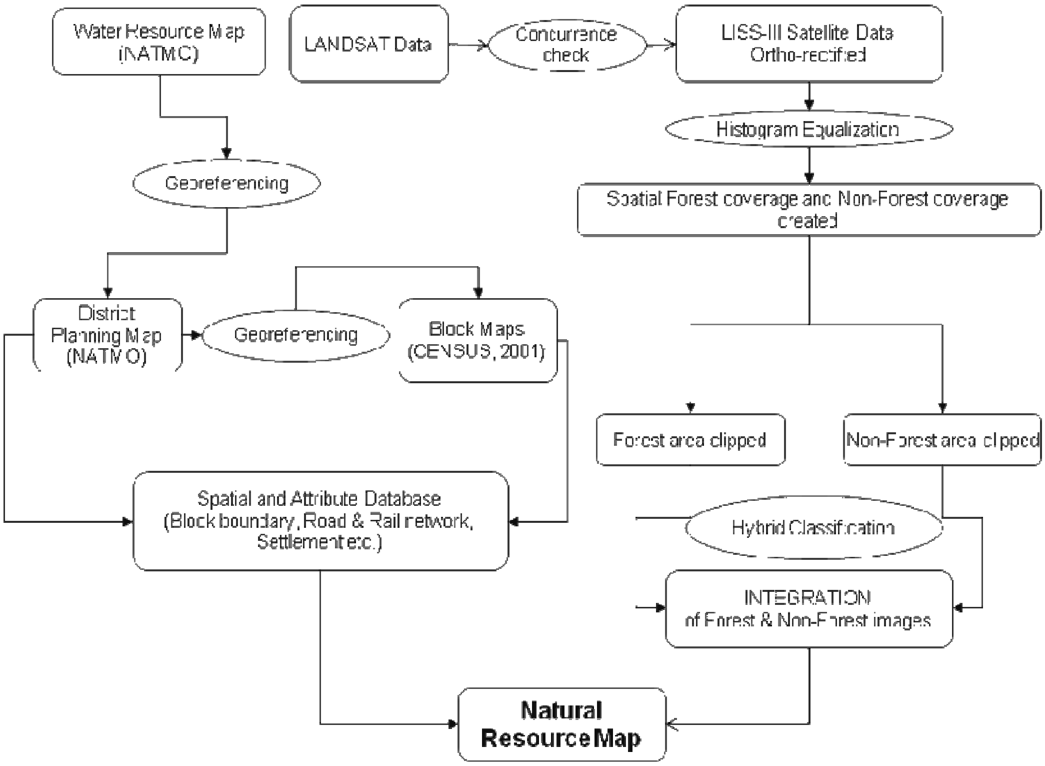

Figure 2: Schematic representation of methodology used to prepare natural resource map

4.1 Creation of geo-spatial database for the district

The Cooch Behar district being restricted area for security reasons, the relevant topographical map sheets could not be obtained. The West Bengal Water Resource map published by NATMO (1:1,000,000) was obtained and geo-referenced and then the DPM obtained from NATMO was geo-referenced using the Water Resource Map with root mean square error (RMSE) of 0.367. The district map of Cooch Behar obtained from the Census, 2001 was geo-referenced using the geo-referenced district planning map (RMSE: 0.51). The individual blocks from the Census, 2001 were georeferenced using this previously mentioned district map (total RMSE: 0.73). These blocks were individually handled during thematic map preparation. The thematic maps of block boundary, road & rail network map, stream map, settlement map etc. were generated for each block.

4.1.1 Administrative database

The district map obtained from Census of India, 2001 was used to create digital administrative boundary (district & blocks) map coverage of the study area in ArcMap 10. Once the district boundary feature class was vectorized, the individual blocks were eventually digitized too. Along with them the settlements were also digitized from the district planning maps (NATMO).

4.1.2 Communication network database

The district planning map in coherence with Google Earth was used to create digital road and railway network map coverages of the entire district.

4.2 Preparation of Land use/ Land Cover

The following stages were involved in the mapping of land use and land cover of Cooch Behar district.

4.2.1 Image Selection and preprocessing

Three cloud-free IRS-P6, LISS-III scenes over a period of December 2008 to February 2009 were obtained from NRSC data center, Hyderabad. The imageries were obtained as GeoTIFF files in four bands (2, 3, 4, 5). The layers were stacked in order of 5, 2, 3, and 4 to obtain a standard FCC for each scene. To generate the FCC for individual district as a separate study area, the individual FCCs were mosaiced following the path-row covering the entire district. The images had a Polyconic projection with Everest spheroid and Indian 1975 datum. The Landsat images of Coochbehar area were used to cross-check the referencing of already geocoded LISS-III images. Once the concurrence between the two data sets were checked and found to be satisfactory, they were reprojected to Universal Transverse Mercator projection, with WGS-84 spheroid and datums to match with the other datasets incorporated through GIS. To clip out the exact district FCC, the district boundary vector was used as clipping features on the satellite imagery. To ensure that the imagery was well interpreted in terms of its visual nature, histogram equalization enhancement was exercised over the imagery. This resulted in a better visual output. Since the study area was only a part of the entire scene, a subset of the district was clipped from the imagery to obtain the area of interest using the district coverage vector. The district FCC was then used to generate land use pattern.

4.2.2 Land use classification, modification, area calculation and accuracy assessment

Once the district FCC was clipped out, the forests were separated out at the initial level, using the forest coverage file. The forest vector map was prepared by visual interpretation and on-screen digitization from the satellite imageries. The forest areas were then clipped out of the satellite data using the forest vector files, to be left with two image datasets of the non-forest area and of the forest area. This was to make sure that the existing natural vegetation did not conflict with the cropped areas during classification, as they tend to reflect more or less similar spectral signatures. So finally, a forest raster and a non-forest raster were generated for separate classifications. For the study a general LULC scheme was adopted that was based on the a priori knowledge of the study area. Overall 8 LULC classes were considered: 1) upper terrace; 2) single cropped area; 3) double cropped area; 4) forest; 5) degraded forest; 6) rivers and water bodies; 7) sandy areas; 8) settlements.

There are several ways to perform a hybrid classification, of which the most commonly used technique has been adopted for this study. Here the signature file from an unsupervised classification is used in a supervised classification. This hybrid approach to classification involves three segments. Firstly to reduce the spectral conflict, the forest and non-forest areas were clipped out from the original LISS - III data before the unsupervised classification. For the unsupervised classification of the non-forest raster, the ISODATA clustering algorithm was performed using 75 clusters, which uses the minimum spectral distance formula to form clusters and saving the signature file. The clusters were then assigned to one of the five classes (upper terrace, single crop, double crop, rivers and water bodies and sandy areas). The process was repeated for the forest raster separately with 20 clusters, where the clusters were assigned to one of the four classes (forest, non forest, rivers and water bodies and sandy areas). This cluster assignment to any of the classes was aided by the original LISS - III data and the knowledge of the study area. In the second segment the supervised classification was performed on the original forest raster data and non-forest raster data, using the signature files saved at the previous segment. In the final segment, to obtain the complete land use/ land cover, GIS analysis such as recode and overlay functions were used to fuse the forest and non-forest classified images. To refine the final classified image, a median statistical filter was applied with a window size of [3×3]. This resulted in a smooth output for a better visual interpretation. Once the classification process was complete, the area statistics for the land use and land cover was generated. An accuracy assessment of the land-cover classification map was performed. The basic idea was to compare the assigned classification of each pixel with the actual classification. For the study, the ground data could not be collected for error estimation due to inaccessibility to certain areas for security purposes. Hence random point selection was the only option left. For the accuracy assessment a random 50 points were generated. These points were plotted on each LULC and the class values were automatically generated. The class values were then cross checked with the corresponding reference values. This technique enabled the analysis of each pixel which was then entered into the error table as the reference value. Then an error matrix calculating the overall accuracy and kappa coefficient was created.

4.3 Results

The road network map (Figure 3) shows a general road distribution throughout the district. It was also possible to categorize the roads into four major classes viz, metalled road, unmetalled road, cart track and pack track. Their length segmentation clearly reflects that the unmetalled roads are dominant over the metalled roads. Approximately 650 km of the district roads were found to be unmetalled, whereas only 545 km of roads were metalled. A distance of about 250 km exists in the form of cart track. This is an indication to negligent developmental programmes in the district. As a result it is easy to assume that the communication system too is in quite a pitiful situation. From the final classification (Figure 4) it was observed that Cooch Behar is a dominantly double cropped area occupying 50 percent of the total area (Table 2). Nearly 30 percent of the district practices single cropping. The district has about 3 percent area under upper terrace. They are mainly the upland areas characterized by light soils with dominant sand and gravel concentration. As a result the upper terrace has very poor water retention capacity. These are the areas of concern that need to be addressed for better resource management. The most significant realization is the near absence of forested areas. Combining forest and degraded forest lands, a mere 0.9 percent of the district is covered. This is quite an alarming fact keeping in mind the global warming and may be attributed to the rising urbanization and even increase in agricultural practices. The district known for its very frequently scattered water bodies has about 4.68 percent area under rivers and water bodies. The alarming 4.41 percent sandy area however indicates that a substantial water body area has undergone severe changes and has condensed. It is only during the peak monsoon periods that they are flushed with upstream flow. Nearly 5.30 percent of the study area is occupied by settlements.

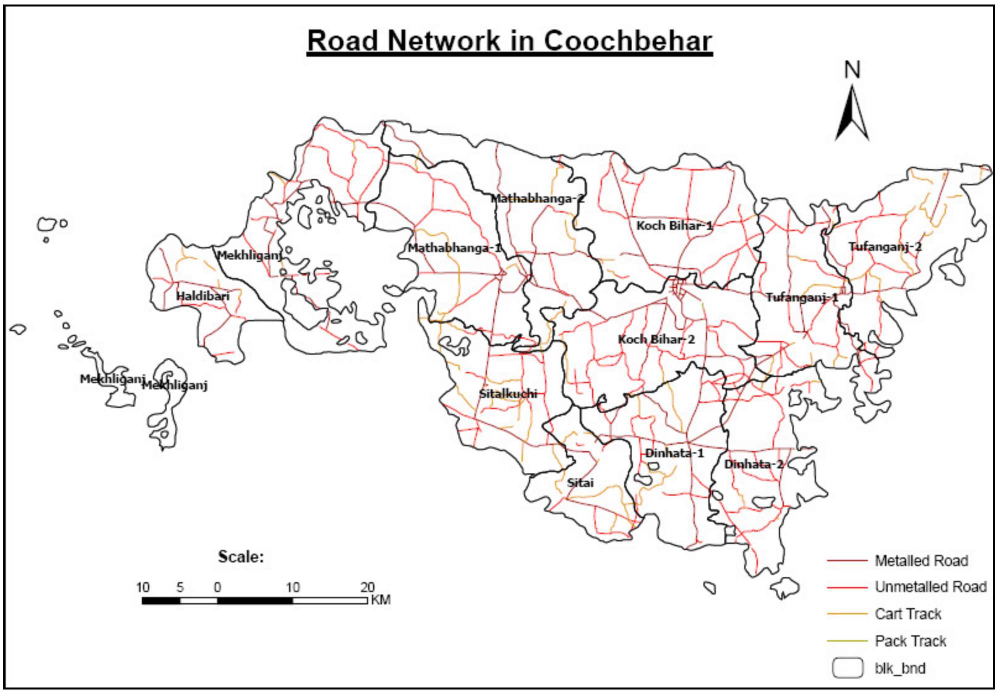

Figure 3: Road network of Cooch Behar district

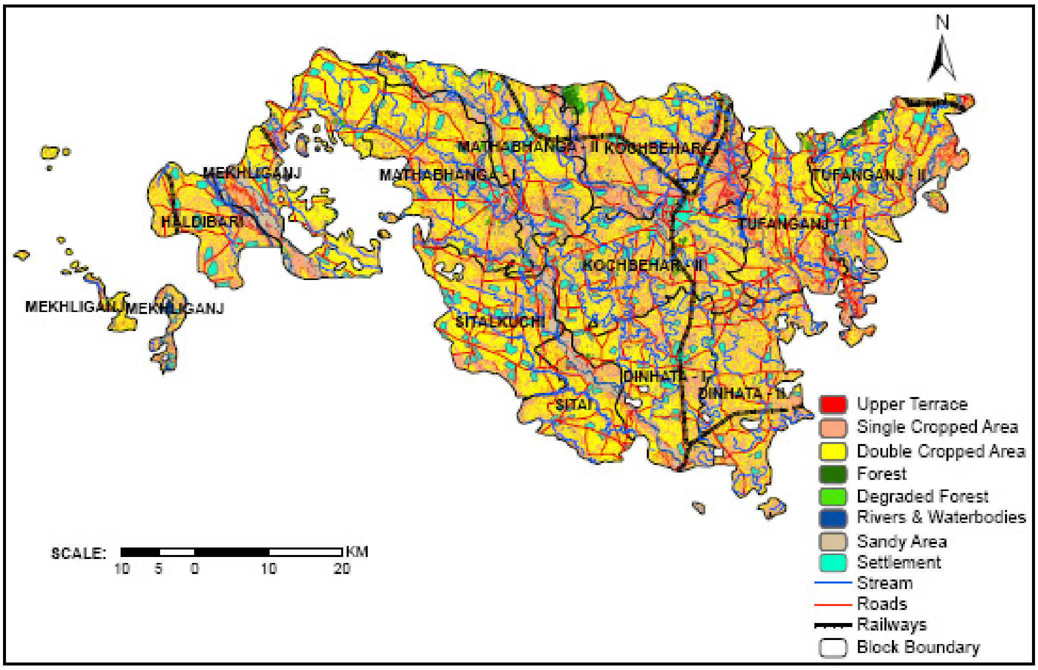

Figure 4: Land use and land cover map of Cooch Behar district

Table 2: Land use and land cover area statistics

Classes	Area (ha)	Area (%)
Upper Terrace	10898.72	3.25
Single Cropped Area	103996.25	30.97
Double Cropped Area	169539.17	50.49
Forest	1809.79	0.54

Degraded Forest	1193.41	0.36
Rivers and Water bodies	15697.69	4.68
Sandy Area	14823.42	4.41
Settlement	17807.34	5.30
TOTAL	**335765.79**	

Given that enough pixels were checked, the percentage of accurate pixels was found to be 88.4 percent and it gave a good estimate of the accuracy of the whole classification. The Kappa Coefficient was found to be 8.1. These accuracy levels are ample for the area of interest since it corroborates the stipulated minimum of 85 percent accuracy level (Anderson et al. 1976). Moreover since the spectral conflict of forest and non-forest areas was well reduced, it was effectively reflected in the improvement of the accuracy of per pixel land use classifications.

5. Conclusion

This study demonstrates a land use/land cover mapping methodology that attempts to incorporate the richness and goodness of both the traditional classification techniques. Since times immemorial Remote Sensing and GIS have proved to be critical contributors to mapping the natural resources and aiding in sustainable resource management planning. Updated natural resource information can prove to be a significant tool for an efficient and sustainable land use planning and natural resource management. The primary goal of the study is to establish an integrated approach of GIS and hybrid classification technique as a systematic procedure in land use mapping thereby increasing the accuracy levels by incorporating the advantages of both the traditional classification algorithms. It is also an attempt to establish a contemporary classification scheme based on the observers acute knowledge of the study area instead of the empirical schemes followed till date. As a result a unique distribution of classes within the study area is observed during the study. The study supports our aim to develop an amalgamated classification technique to extract complex yet accurate outcomes.

Acknowledgements

The authors express their gratitude to West Bengal State Council of Science and Technology for funding the project work and providing the necessary logistics during the tenure of the project.

6. References

1. Kamusoko, C., and Aniya, M., (2009), Hybrid classification of Landsat data and GIS for land use/cover change analysis of the Bindura district, Zimbabwe, International Journal of Remote Sensing, 30(1), pp 97-115.

2. Zingaretti, P., Frontoni, E., Malinverni, E.S., and Mancini, A., (2009), A Hybrid Approach to Land Cover Classification from Multi Spectral Images', in Proceedings of Image Analysis and Processing – ICIAP: 15th International Conference Vietrisul Mare, September 8-11, 2009, Italy, pp 500-508.

3. Caprioli, M., Leone, A., Ripa, M.N., and Tarantino, E., (2003), A hybrid land cover classification of Landsat 7 ETM+ data for an efficient vegetation mapping, in

Proceedings of Camarda D. (ed.), Grassini L. (ed.). Local resources and global trades: Environments and agriculture in the Mediterranean region, Bari, CIHEAM, pp 451 - 460.

4. Chapin, F.S. III., Zavaleta, E.S., and Eviner, V.T., (2000), Consequences of changing biodiversity, Nature, 4(5), pp 234-242.

5. DeFries, R.S., and Townshend, J.R.G., (1999), Global land cover characterization from satellite data: from research to operational implementation, Global Ecology and Biogeography 8(5), pp 367-379.

6. Anderson, J.R., Hardy, E.E., Roach, J.T., and Witmer, R.E., (1976), ‘ land use and land cover classification system for use with remote sensor data, US Geological Survey Professional paper 964, Sioux Falls, Washington, DC, USGS.