IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Satellite-based analysis of land use & land cover through several indices & K-means unsupervised Classification in Bokaro district, Jharkhand.

Utkarsha Rani* (Research Scholar, PG Dept. of Geography Ranchi University)
Rajeev Ranjan Shrivastava (Assistant Prof., St. Xavier's College, Ranchi)

Abstract:

The earth has transformed ages throughout the world to the service of humankind. In India explosive growth of the population and their growing need has steadily increased the demand to utilize the land resources at optimum for agricultural land, mining and industrial establishment, urban expansion, etc. with the expense of natural forest, grassland, and wetland that had been the valuable habitat of numerous species. Therefore, Land use and land cover change have become significant issues to study resource management and utilization in the present scenario. This study aims to detect the land-use land cover change in Bokaro District of Jharkhand, Eastern India from 2001 to 2021 using Landsat-7 ETM+ and Landsat 8 OLI satellite images. The DOS method has been used for atmospheric correction, and the classification NDVI, SAVI, MNDWI, and NDBI indices were first calculated. Afterward, K means unsupervised classification has been used to generate land use and land cover maps through the SCP plugin in QGIS. For the accuracy of classified land use and land cover maps, a confusion matrix & kappa coefficient were used to derive overall accuracy and the results were above the minimum and acceptable threshold level. The massive changes in land use land covers are noticed in vegetation, agricultural land, mining, allied activities, and waterbodies. The result reveals both the increase and decrease of the different land use/ land cover classes from 2001 to 2021. An appreciable increase in vegetation cover and a marginal decrease in the mining area was noticed.

Keywords: Landsat 7 ETM+, Landsat 8 OLI, DOS atmospheric correction, K means algorithm, Unsupervised classification, Kappa Coefficient

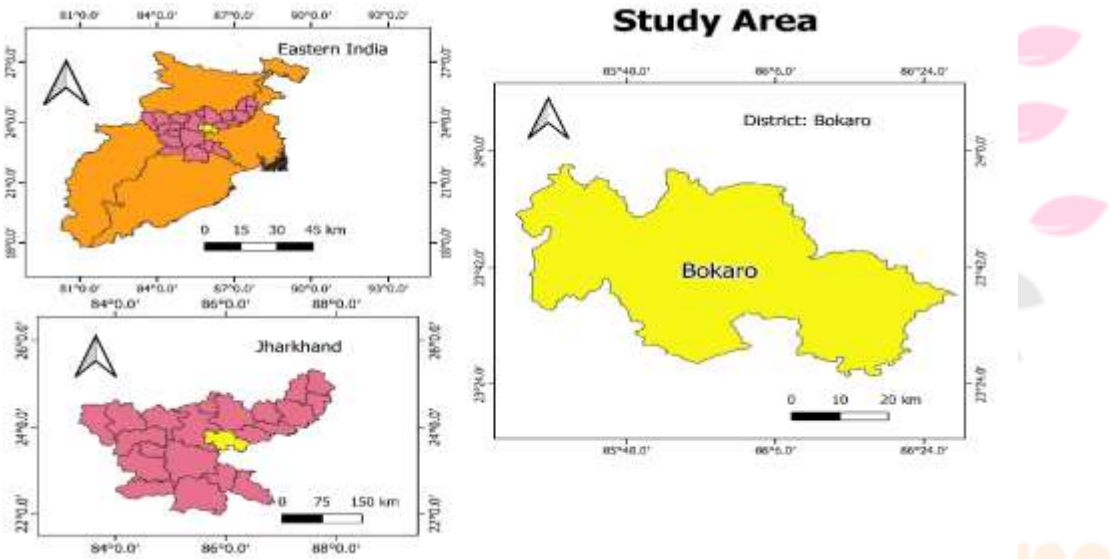
Introduction:

From an environmental point of view, the dynamic process of land use/land cover change is an indispensable concern all over the world, which indicates global environmental change (Ruiz-Luna and Berlanga-Robles, 2003), and this has been recounted as the most remarkable regional anthropogenic disruption of the environment (Lambin et al., 2003). The land use land cover change scenario in India has undergone a radical change since the onset of the economic revolution in the early 1990s. Along with it, the explosive growth of the population and their growing need have steadily increased the demand to utilize the land resources at optimum for agricultural land, mining and industrial establishment, urban expansion, etc. Therefore, Land use and land cover change have become significant issues to study resource management and utilization in the present scenario. Energy requirement in India is increasing day by day and the coal mining industries are eventually increasing their production to meet the requirement of energy production through thermal power plants, where coal is used for the generation of electricity (Provisional Coal Statistics, 2015-16). About 95% of the total coal reserves of India come from 44 known coal fields in Peninsular India out of which more than 20% are situated in Jharkhand. (Sachdev, 2007) In this context, it is essential to scrutinize the effect of mining on land use land cover change to minimize its impact on the environment as well as for proper land management and decision-making (Bocco et al., 2001; Laskar, 2003; Turner II et al., 2007). Mining and land use changes are correlated with each other, especially opencast/surface mining operations, which have a major impact on the LULC of the area during pre-mining and post-mining operations. (Kumar A, Pandey AC (2013) The current study is an attempt in this direction to evaluate the status and trend of land use/cover in a coal mining area using multi-temporal remote sensing and google satellite images for validation. Before coal mining

and industrialization, this region had a large forest cover with fertile cultivable patches. The main objective of the present study is to assess the spatial and temporal changes in land use and land cover of Bokaro District.

Study Area:

The East Bokaro coalfields are located in the Bokaro district of Jharkhand state in India. They lie between 23º45′ to 23º50′N latitude and 85º30′ to 86º03′E longitude. The coalfields covering an area of 259 sq km are a major source of medium coking coal in India and are third from, the east in the chain of coalfields in the Damodar Valley. The location map of the study area is shown below, where Coal is produced by underground as well as opencast mining methods by Central Coalfields Limited (CCL), a subsidiary of Coal India Limited. The important coal seams include the Jarangdih seam, Kargali seam, Bermo seam, and Karo seam. The area is dispersed with isolated hillocks which are the major geographic units. In general, the lower altitudes (230–300 m above mean sea level) have a sandstone type of landscape. The southern parts of the coalfield possess a gently undulating topography. In the northern part, the compact and massive sandstone of lower Barakar has given rise to comparatively rugged terrain. Lugu Hill is a prominent landmark and separates East from the West Bokaro coalfields. Bokaro Thermal Power Plant is situated in the central part on the banks of the Konar River. There are three important rivers in the coalfield, namely Bokaro in the central part, Konar in the east, and Damodar in the south. Tenughat dam/reservoir is situated in the southwestern parts and is an earth-fill dam with composite masonry and rock excavation in the foundation. The dam was constructed in 1974 (although the construction began much earlier) and has been developed into a popular tourist center.



Data and Methodological Procedure and Techniques:

1. The satellite data used for land use/land cover classification are derived from Landsat-7 Enhanced Thematic Mapper for five–year time periods of 2001, whereas Landsat-8 data sets are used for the year 2021. The Landsat-7 and 8 sensors have a spectral resolution of 30 m and a swath of 185 km. These satellite data sets were obtained from Earth Explorer (http://earthexplorer.usgs.gov/ The Landsat-7 & 8 data sets were projected to the reference system UTM (zone 45) and WGS 84 datum. The study area falls under path 144 and rows 43 and 44 in the Worldwide Reference System of Landsat which was merged for Landsat 7 imagery since this study area falls within two rows 43 and 44 respectively.

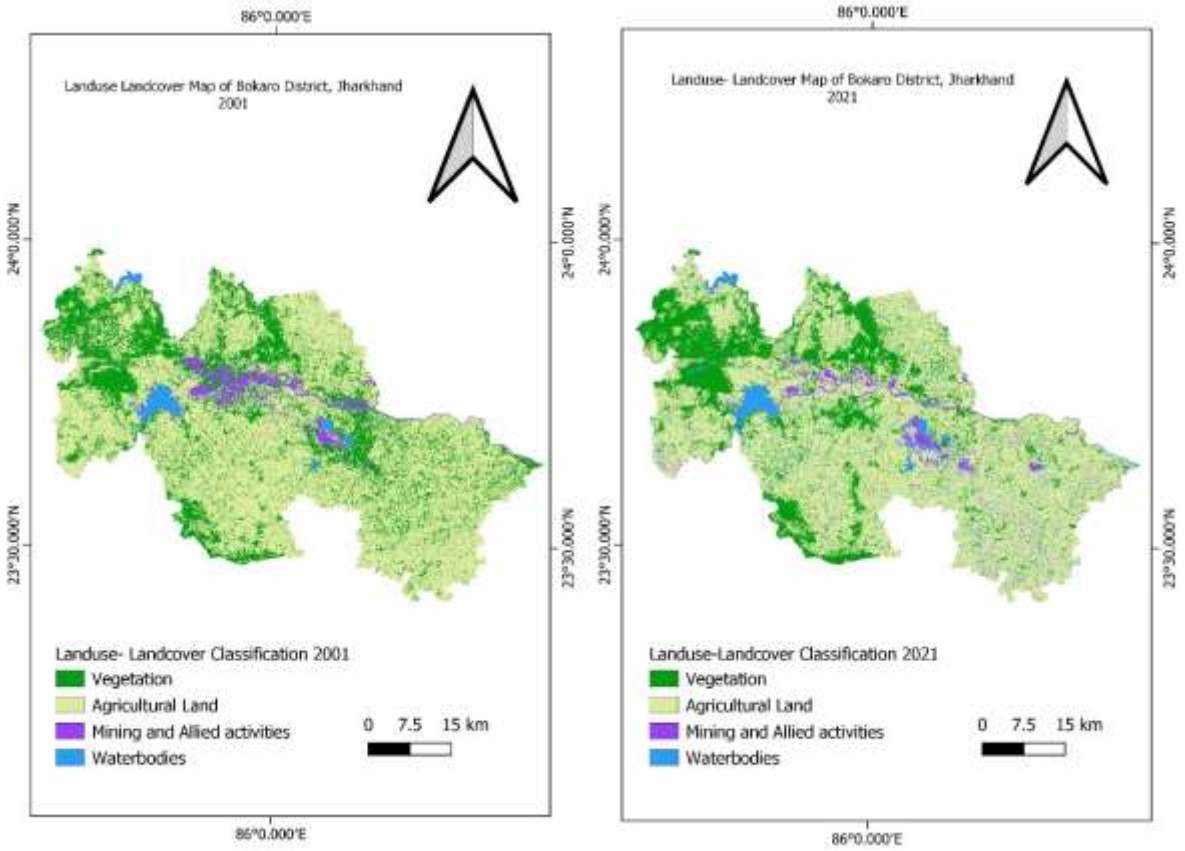
Landsat (7 ETM+) Year 2001	Landsat 8 (OLI) Year 2021 (micrometers)	Resolution
B1 (0.45–0.52) Blue	B1 (0.43–0.45) Coastal Aerosol	30 Meter
B2 (0.52–0.60) Green	B2 (0.45–0.51) Blue	30 Meter
B3 (0.63–0.69) Red	B3 (0.53–0.59) Green	30 Meter
B4 (0.76–0.90) NIR	B4 (0.64–0.67) Red	30 Meter
B5 (1.55–1.75) SWIR 1	B5 (1.85–1.88) NIR	30 Meter
B7 (2.09–2.35) SWIR 2	B6 (2.08–2.35) SWIR 1	30 Meter
	B7 (2.11-2.29) SWIR 2	30 Meter
	B9 (1.36-1.38) Cirrus	30 Meter

Table 1.1

Digital numbers (DN) from Landsat Level 1 products are suitable for single image analysis but not effective when comparing spectral values across images of different periods, therefore image pre-processing steps like radiometric and atmospheric calibration and corrections are required to bring the digital numbers to values that are of comparable scales. The present analysis requires the comparison of data over multiple periods and therefore needs pre-processing to obtain the actual surface reflectance values. The acquired data were corrected for radiometric and atmospheric errors. The procedure adopted for atmospheric correction is presented. All Landsat images have a metadata text file associated with them for Landsat 2001 and 2021. The raw digital numbers were first converted into radiance values based on the rescaling factors provided in the metadata file. The radiance values were then converted into top-of-atmosphere reflectance values using the solar zenith angle, Earth-Sun distance, and solar exo-atmospheric irradiances which vary with latitude, time, and date of the image acquisition.

To further enhance the interpretability of the images, the atmospheric correction was carried out to get the surface reflectance values by removing the effects of the atmosphere. Atmospheric correction requires information about the atmospheric conditions and aerosol properties at the time of image acquisition which is difficult to implement. Thus, the simplest approach using dark object subtraction (DOS) which is based on the information contained within the image was employed. DOS method although of lower accuracy is useful when there is no available data on atmospheric measurements. This method assumes that the radiance/reflectance values of the image pixels in complete shadow are entirely due to atmospheric scattering/path radiance and this value is subtracted from each pixel in the image to get the actual land surface reflectance. Most image-processing software packages have tools to calibrate the data. In QGIS (open source), the Semi-automatic Classification Plugin Tool converts Landsat digital numbers into surface reflectance values using dark-object subtraction. Therefore, correction of the Landsat 2001 and 2021 datasets was carried out in QGIS using the parameter specifications from the [.MTL]text file. NDVI MNDWI, NDBI, and SAVI were prepared for the respective years. These indices were then stacked together to create a false colour composite (FCC) for all the images. Visual image interpretation is the major tool for getting information about land use/land cover from satellite data. However, K means classification was run to unsupervised classification into 10 groups, and re-classification of those group outcomes was done in the desired class. For this paper, 4 classes were generated namely Vegetation, agricultural land, Mining, and Allied activities & Waterbodies.
An unsupervised method using the K-means classifier was used to perform the classification and generate the land use/land cover maps. But pre-processing of images was done by calculating several indices and stacking them together. The entire process for finding indices through different band calculations is written below.

1. NDVI = Green- NIR / Green +NIR
2. MNDWI = Green- SWIR / Green + SWIR
3. NDBI = SWIR- NIR/ SWIR+ NIR
4. SAVI = (NIR- R)/ (NIR+ R+L) * (1+L)
   where L = Soil brightness correctness factor= 0.5



Selection for validation Samples: For each classification, the area under coverage was generated through a classification report in which a percentage of the land cover of each class is also given.

The main requirement is to provide an adequate number of samples for each class, even if the class area proportion (Wi) is low. The number of samples for both raster images is calculated as:

$$N = (\sum i=1(Wi*Si)/So)^2$$

where:

- Wi = mapped area proportion of class I;
- Si = standard deviation of stratum i;
- So = expected standard deviation of overall accuracy;
- c = total number of classes;

In this study, we hypothesize that the user's accuracy is lower and standard deviations Si is higher for classes having low area proportion. The classification report table contains the percentage of each class, which we divide by 100 to get the required Wi. we assumed So=0.01 and conjectured Si values reported in the following table.
Therefore,

$$N_{1991} = (0.1141/0.01)^2 = 130$$

$$N_{2020} = (0.1215/0.01)^2 = 148$$

are the number of samples that we distributed among classes.

A rough approximation is considering the mean value between equal distribution ($Ni=N/c$) and weighted distribution ($Ni=N*Wi$), which is $Ni=(N/c+N*Wi)/2$ as illustrated in the following Table 1 and 2 for the year 2001 and 2021 respectively.

Macro C	Class Name	Area(m2)	Percentage	Weight (Wi)	Standard Dev (Si)	Wi*Si	N*Wi	N/c	Total	x=Total/2	No. of sar
1	Vegetation	606380000	10.32259163	0.103	0.3	0.0309	13.39	32.5	45.89	22.945	22
2	Agriculture	2116060000	36.02233458	0.36	0.2	0.072	46.8	32.5	79.3	39.65	39
3	Mining and Allied	108630000	1.849241612	0.018	0.4	0.0072	2.34	32.5	34.84	17.42	17
4	Waterbodies	49230000	0.8380573	0.008	0.5	0.004	1.04	32.5	33.54	16.77	17
						0.1141				96.785	
						11.41					
						130.1881					

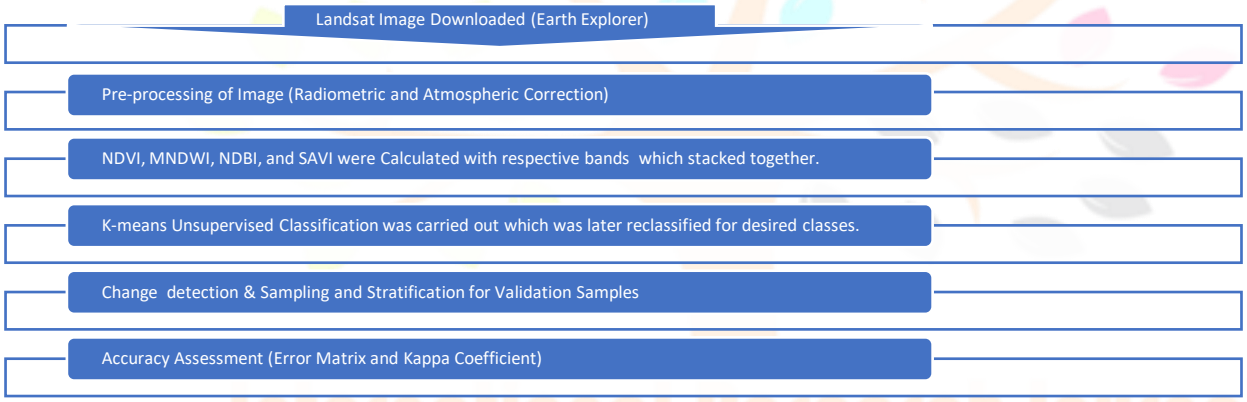
Table: 2.1

Macro C	Class Name	Area(m2)	Percentage	Weight (Wi)	Standard Dev (Si)	Wi*Si	N*Wi	N/c	Total	x=Total/2	No. of sar
1	Vegetation	689005800	11.74783139	0.117	0.3	0.0351	17.316	37	54.316	27.158	27
2	Agriculture	1902503700	32.43846814	0.324	0.2	0.0648	47.952	37	84.952	42.476	42
3	Mining and Allied	115716600	1.973015475	0.0197	0.5	0.00985	2.9156	37	39.9156	19.9578	20
4	Waterbodies	173225700	2.95356921	0.0295	0.4	0.0118	4.366	37	41.366	20.683	20
				0.12255		0.12155			220.5496		
						12.155					
					No. of sample(N)=	147.744					

Table: 2.2

Afterward, the accuracy of image classification was assessed with the help of collected google Satellite Images. An area proportion sample design was used, out of the classification report. Google Satellite verification was done by visual inspection of the map to verify the interpreted land cover features wherever accessible. For each point, the appropriate land cover type was identified from the classified maps and google Satellite Imagery and compiled into an error matrix. Kappa statistics (x), a measure of agreement between producer and user accuracies, were also calculated.

Flow Chart of Process under carried:



Results and Discussion:

Classified Groups	**Area of land 2001**	**Area of land 2021**	**Change in % of the area, 2001-2021**
Vegetation	**10.32**	**11.75**	**1.43**
Agricultural Land	**36.02**	**32.44**	**-3.58**
Mining and Allied activities	**1.85**	**1.97**	**0.12**
Waterbodies	**0.84**	**2.95**	**2.11**
Unclassified Area	50.96	50.88	

Table: 3.1

The above table shows the changes in land use and land cover area from 2001 to 2021 where the growth of vegetation area is much appreciated and mining and allied activities have also grown in the area and the agricultural land area has reduced by 3.58%. The reason for the change in waterbodies is erroneous as mining

area greatly influences visibility. For the valid justification of these results, an accuracy assessment was carried out with the help of manual error matrix calculation on an excel spreadsheet which is as follows.

ERROR MATRIX, 2001 & 2021

		References						
		Vegetation	Agricultural Land	Mining & Allied Activities	Waterbodies	Total	Commision Error	User's Accuracy %
Classification	Vegetation	824	41	32	1	898	8.24%	91.76%
	Agricultural Land	128	847	23	5	1003	15.55%	84.45%
	Mining & Allied Activities	0	0	824	16	840	1.90%	98.10%
	Waterbodies	0	0	8	1299	1307	0.61%	99.39%
	Total	**952**	**888**	**887**	**1321**	**4048**		
	Ommision Error	**13.45%**	**4.62%**	**7.10%**	**1.67%**			
	Producer's Accuracy %	**86.55%**	**95.38%**	**92.90%**	**98.33%**			
	Overall Classification Accuracy %	**93.73%**					overall accuracy p(o)=	**0.937252964**
	Kappa Coefficient p(o)-p (r) /1-p(r)=	**0.9155081**					**random accuracy p(r)=**	**0.257360476**

Table:4.1

		References						
		Vegetation	Agricultural Land	Mining & Allied Activities	Waterbodies	Total	Commision Error	User's Accuracy %
Classification	Vegetation	21270	1	8	0	21279	0.04%	99.96%
	Agricultural Land	5	4698	106	0	4809	2.31%	97.69%
	Mining & Allied Activities	0	4	4714	101	4819	2.18%	97.82%
	Waterbodies	4	13	129	15252	15398	0.95%	99.05%
	Total	**21279**	**4716**	**4957**	**15353**	**46305**		
	Ommision Error	**0.04%**	**0.38%**	**4.90%**	**0.66%**			
	Producer's Accuracy %	**99.96%**	**99.62%**	**95.10%**	**99.34%**			
	Overall Classification Accuracy	**99.20%**					overall accuracy p(o)=	**0.991987906**
	Kappa Coefficient p(o)-p (r) /1-p(r)=	**0.9878022**					**random accuracy p(r)=**	**0.34315105**

Table: 4.2

Based on the results obtained from the accuracy assessment, the table was generated and the error matrices were prepared on an excel sheet individually for the years 2001 and 2021. This indicates that the classification done for the year 2001 is 93.73% accurate and its kappa coefficient value is much higher than the 75% limit in this case 91.5%. Whereas, the next classification is 98.78% accurate. The reason for better classification accuracy cannot be ignored since the Landsat satellite data in OLI carrier have better resolution than the previous one which is an enhanced thematic mapper in Landsat-7.

Conclusion:

Summing up, although we are only in the early stages of analyzing the future state of land use and land cover in Bokaro. In the area where mining activities are prevalent, it is difficult to decipher another land cover. The area of Study has a broad coverage of Mining zones and the prevalence of ashes obstructs to segregation of the Mining area from the Settlement and Agricultural land adjacent to it. This is the limitation of classification as spectral signatures are quite similar in those areas. Although some improvement has been noted in vegetation, which implies better policy implications. However, a clear message of the scenarios of the particular importance of local change is that current land-use/cover patterns are not static the trend could reverse in demand of the growing population. Indeed, major changes in the land cover over the next several decades, including trend reversals, are not implausible. The fact that some scenarios only begin to show distinctive trends after two or three decades also implies that a long-term view is needed to better anticipate the future of the land.

References:

1. Abou Samra R. M.& El-Barbary S.M. A. (2018) The use of remote sensing indices for detecting environmental changes: a case study of North Sinai, Egypt, Korean Spatial Information Society,
2. Congedo, Luca(2012) Adapting to Climate Change in Coastal Dar es Salaam Rome, 03 December 2012 https://www.academia.edu/8097334/Development_of_a_Methodology_for_Land_Cover_Classification_Validation
3. Garai D. & Narayan A. C. (2018) Land use/land cover changes in the mining area of Godavari coal fields of southern India, The Egyptian Journal of Remote sensing and Space Science Vol, 21, Pages 375-381
4. Global Strategy to improve Agricultural and Rural Statistics (GSARS). 2017. Handbook on Remote Sensing for Agricultural Statistics. GSARS Handbook: Rome.
5. Hongmei ZHAO, Xiaoling CHEN (2005) Use of Normalized Difference Bareness Index in Quickly Mapping Bare Areas from TM/ETM+, IEEE
6. Kumar et al. (2018) Satellite Image Based Land Use Land Cover Change Analysis of Ranchi District, Jharkhand, SGVU J CLIM CHANGE WATER Vol. 5, pp. 1-8

7. Lambin E. F. & et.al, (2003) Dynamics of land use Land Cover change is tropical Region, Annual Review Environment Resources.
8. Madasa A. & et. Al (2021) Application of geospatial indices for mapping land cover/use change detection in a mining area, Journal of African Earth Sciences. Vol.175
9. Maury C. & Sharma V.N. (2020) Land use/ Land cover Change Detection in Auranga River Basin, Jharkhand, NGJI- BHU vol 66, pp. 1-8
10. Millennium Ecosystem Assessment (2003) Ecosystems and human well-being: A framework for assessment. Island Press, Washington D.C., pp. 245
11. Roy P.C. & Giriraj A. (2008) Land-use landcover analysis in the Indian context, Journal of Applied Science vol.8, pp.1346-1353
12. Ruiz-Luna & César Alejandro Berlanga-Robles (2003) Land use, land cover changes & coastal lagoon surface reduction associated with urban growth in northwest Mexico, Landscape Ecology vol.18, p.p.159-171
13. Shibendu S. Ray & et.al, (2005) Derivation of indices using remote sensing data to evaluate cropping systems Journal of the Indian Society of Remote Sensing, vol. 8
14. Upgupta S. & Singh P.K., (2019) Quantifying the Dynamics and Drivers of Landscape Change in an Opencast Coal Mining Area of Central India (East Bokaro, Jharkhand) The National Academy of Sciences