1. INTRODUCTION

Land is the most important natural resource which embodies soil, water and associated flora and fauna involving the total ecosystem. The growing population and human activities are increasing the pressure on the limited land and soil resources for food, energy and several other needs. Comprehensive information on the spatial distribution of the land use/land cover categories and the pattern of their change is a prerequisite for planning, utilization and management of the land resources of the country.

Land use refers to the purpose the land serves, for example, recreation, wildlife habitat, or agriculture. Land use applications involve both baseline mapping and subsequent monitoring, since timely information is required to know what current quantity of land is in what type of use and to identify the land use changes from year to year.

Land cover refers to the surface cover on the ground, whether vegetation, urban infrastructure, water, bare soil or other. Identifying, delineating and mapping land cover is important for global monitoring studies, resource management, and planning activities. Identification of land cover establishes the baseline from which monitoring activities (change detection) can be performed, and provides the ground cover information for baseline thematic maps.

Land use and land cover (LULC) refers to the physical characteristics of earth surface, captured in the distribution of vegetation, water, soil and other physical features of the land, including those created solely by human activities. Information on the land use and land cover in the form of maps and data is very important for planning, management, and utilization of land for agriculture, forestry, urban, industrial, environmental studies and economic development.

It is important to distinguish this difference between land use and land cover, and the information that can be ascertained from each. The properties measured with remote sensing techniques relate to land cover, from which land use can be inferred, particularly with ancillary data or a prior knowledge.

Remote Sensing can provide an important source of data for land use/ land cover mapping and environmental monitoring Image classification, which is the systematic grouping of remote sensing and other geographically referenced data by categorical or increasingly, fuzzy decision rules is considered the best known and most widely used information extraction technique in remote sensing.

The usefulness and success of land use and land cover mapping depends on the choice of appropriate classification scheme for feature extraction. To determine the quality of information derived from the classification process, accuracy assessment of the classification is implemented. Error matrix, which is primarily used in remote sensing for accuracy assessment, is typically based on an evaluation of the derived classification against some 'ground truth' or reference dataset. This study also accomplishes accuracy assessment which helps to identify the accuracy of land use/land cover data.

Details of Satellite Image used for the study

S. No.	Satellite	Sensor	Launch	Date of RS data
1.	Landset 8	The Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS)	11 February 2013	5 june 2021
2	Landset 7	the Enhanced Thematic Mapper Plus (ETM+) sensor	April 15, 1999	June 2000

Geographical Location of jabalpur in India:

India

Geographical Coordinates of Jabalpur district

Latitude is **23.185884**

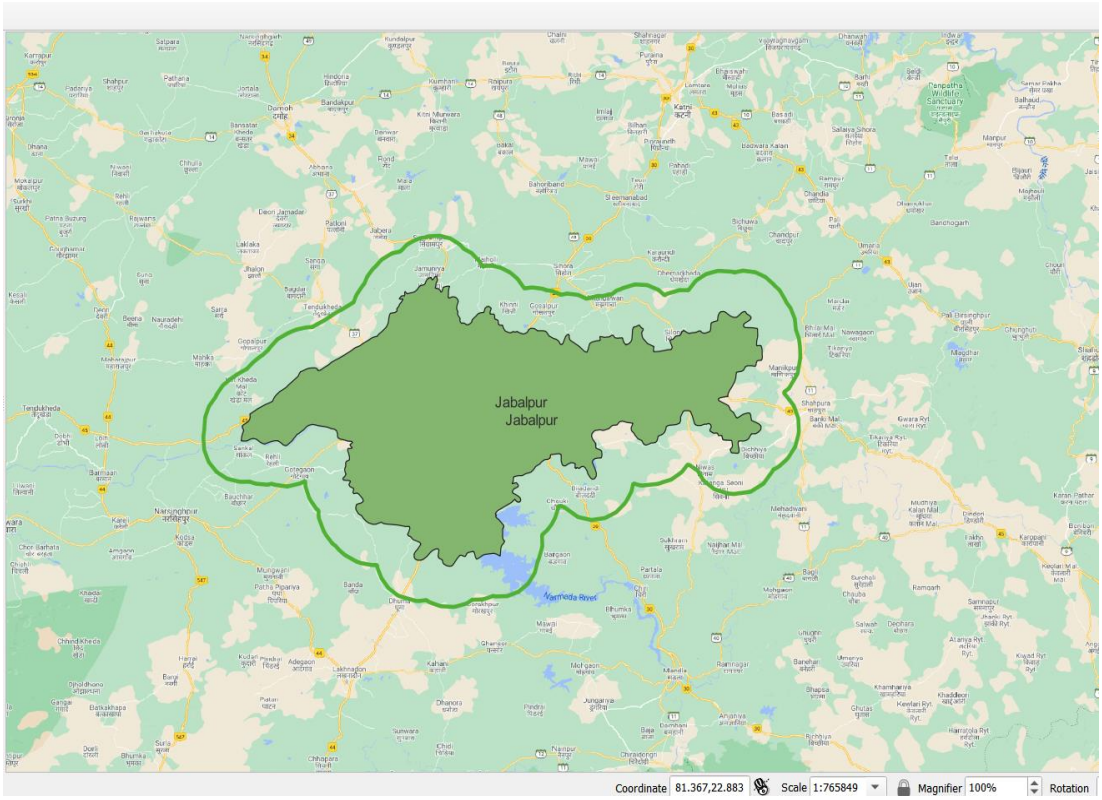
Longitude is **79.974380**.

GPS coordinates are **23° 11' 9.1824'' N and 79° 58' 27.7680'' E**.

Projection is **utm**

Zone is **44**

Datum is **WGS85**



About District

Jabalpur is a district in Madhya Pradesh, India's central state. The district's administrative seat is the city of Jabalpur.

The district covers 5,198 km2[1] and has a population of 2,463,289 people (2011 census). After Indore, it is Madhya Pradesh's second most populous district (out of 50) as of 2011.

The Jabalpur district is located in Madhya Pradesh's Mahakosha region, on the split between the Narmada and Son watersheds, though largely inside the Narmada valley, which passes through the famed Marble Rocks gorge and falls 30 feet over a rocky cliff (the Dhuan Dhar, or misty shoot). It is made up of a long, narrow plain that runs north to south and is surrounded on all sides by hills. This plain, which is an offshoot of the Narmada's major valley, is covered in its western and northern parts.

Jabalpur is the administrative headquarters of Jabalpur district. Historically, a center of the Kalachuri and Gond dynasties, the city developed a syncretic culture influenced by intermittent Mughal and Maratha rule. The High Court of Madhya Pradesh, headquarters of the West Central Railway and Army headquarters of five states (MP, Chhattisgarh, Orissa, Bihar and Jharkhand) are located here.

Country	India
State	Madhya Pradesh
Division	Jabalpur
Headquarters	Jabalpur
Government	
• Lok Sabha constituencies	Jabalpur
Area	
• Total	5,198 km2 (2,007 sq mi)
Demographics	
• Literacy	82.47%
• Sex ratio	925 females/1000 males

Language

Hindi (94.21%)

Urdu (1.93%)

Others (3.86%)

Population

Year	Pop.	±% p.a.
1901	430,810	—
1911	472,102	+0.92%
1921	471,951	−0.00%
1931	489,583	+0.37%
1941	576,167	+1.64%
1951	662,569	+1.41%
1961	821,579	+2.17%
1971	1,119,885	+3.15%
1981	1,484,235	+2.86%
1991	1,768,037	+1.77%
2001	2,151,203	+1.98%
2011	2,463,289	+1.36%

Neighbour District

Damoh district	Katni district	Umaria district
	↑ ←Jabalpur district→ ↓	Dindori district
Narsimhapur district	Seoni district	Mandla district

Religions in Jabalpur district (2011)

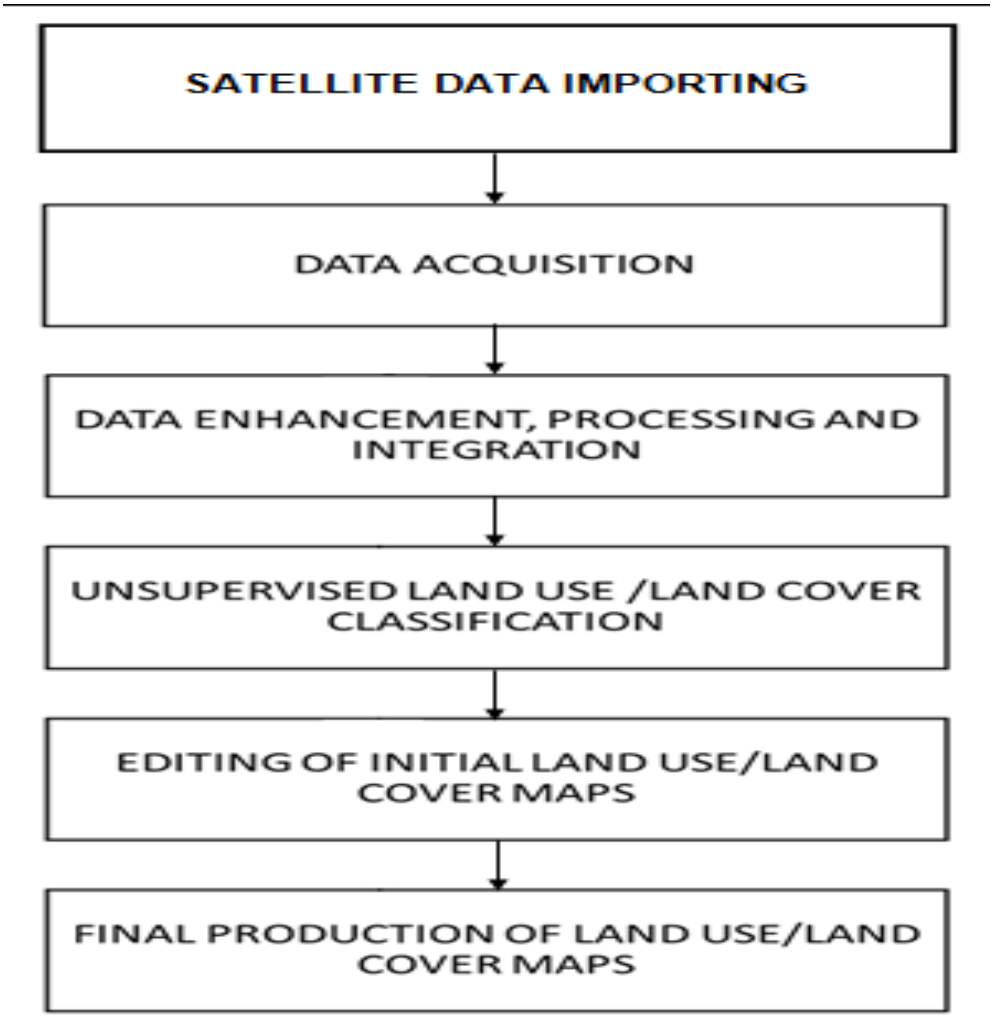
Religion	Percent
Hindus	87.65%
Muslims	8.27%
Jains	1.37%
Christians	0.94%
Other religions (tribal)	0.85%
Sikhs	0.54%
Other or not stated	0.38%

Main Tourism In Jabalpur District:-

1:- The Dhuandhar Falls are located in the Jabalpur district of Madhya Pradesh, India. The 10 metre high falls are found on the Narmada River in Bhedaghat. The Narmada River flows through the world-famous Marble Rocks, narrows, and then plunges into the Dhuandhaar waterfall. The noise of the plunge, which forms a bouncing mass of mist, can be heard from a long way. Additionally, the energy generated could be used in a hydroelectric power plant.

2:- Bargi Dam is one of the earliest dams built on the Narmada River in Madhya Pradesh, India, as part of a chain of 30 large dams. The Bargi Dam management has created two important irrigation projects, the Bargi Diversion Project and the Rani Avantibai Lodhi Sagar Project.

Methodology



Land use and land cover mapping of the Jabalpur District of Madhaya Pradesh was carried out by following standard methods of analysis of remotely sensed data. For this purpose, we are using different band data downloaded by USGS Explorer.

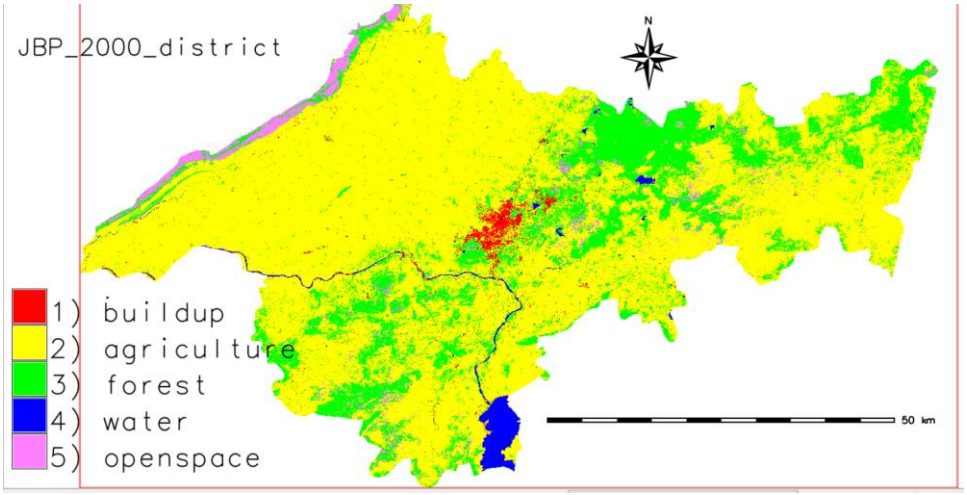
Before processing of any image for image enhancement, transformation or classification, pre-processing was done for band separation. Different bands were downloaded GRASS software.

Standard false colour composite (FCC) image of the study area was prepared using bands 2, 3,4 and 5 of Landsat 8 /Landsat 7 and discrimination of features was made by visual interpretation on this image. The interpretation key was based on the relationships between ground features and image elements like

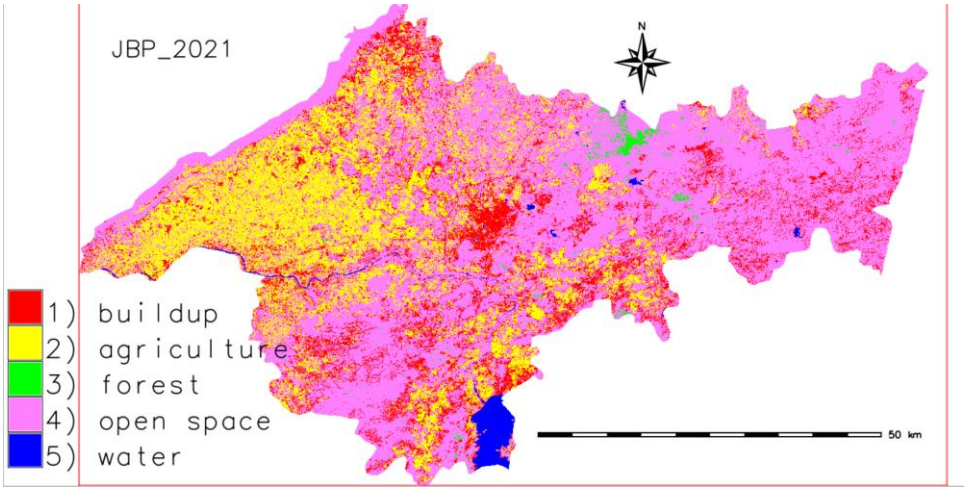
texture, tone, shape, location, and pattern analysis, image classification was done by unsupervised classification method. It helped in assigning the classification of the image into land use categories. The classified map was regrouped and merged. Finally need to calculate kappa and accuracy of train data collected from fcc image and test data from google earth.

Compare b/w Jabalpur district map of year 2000 vs 2021

2000 map



2021 map



Geometric Correction :-

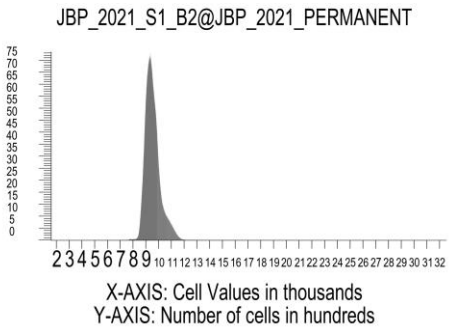
We are continuously checking coordinates of RS data with map in google earth to remove geometric correction. We take coordinates of min 15 different point in RS data (like in road, lakes, pond, forest, agriculture field train track etc) by using query option in Grass Gis map and checked it with google earth one by one. If any coordinate is not matched with main coordinates system, we check, is coordinate system in UTM or not. If it is okey, then It is mistake during uploading boundaries data in grass software

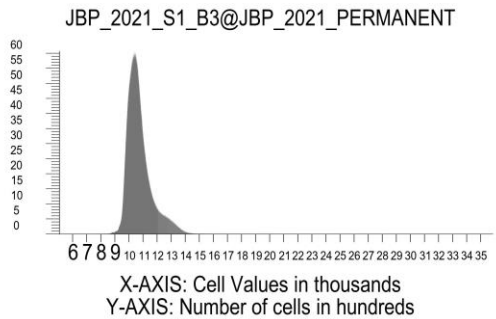
Radiometric correction

We are rescaling all band data of RS data one by one. If RS data is of year 2021, we are rescaling from 0 to 65535. And If RS data is of year 2000 , we are rescaling data from 0 to 255. We are rescaling data because to reduce radiometric error.

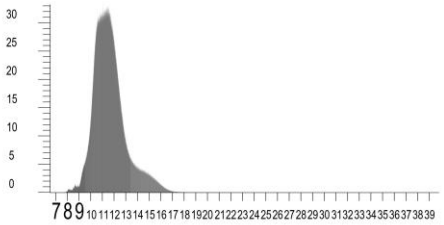
--> Histogram of 2000 and 2021 data before removing radiometric correct (before rescaling)

From 2021 data set –before resize



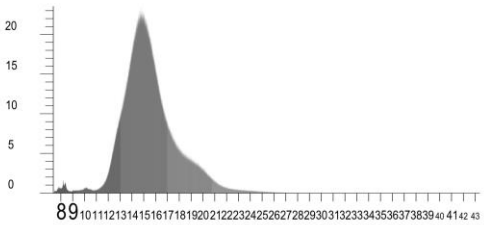


JBP_2021_S1_B4@JBP_2021_PERMANENT



X-AXIS: Cell Values in thousands
Y-AXIS: Number of cells in hundreds

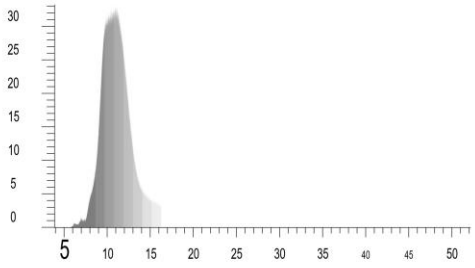
JBP_2021_S1_B5@JBP_2021_PERMANENT



X-AXIS: Cell Values in thousands
Y-AXIS: Number of cells in hundreds

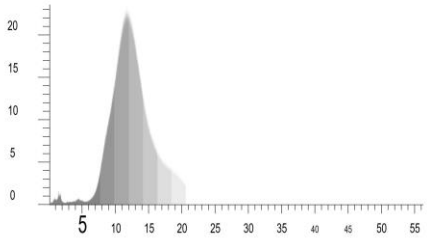
From 2021 data set –after resize

JBP_2021_S1_B4_RESIZE@JBP_2021_PERMANENT



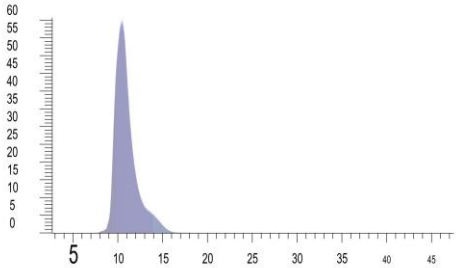
X-AXIS: Cell Values in thousands
Y-AXIS: Number of cells in hundreds

JBP_2021_S1_B5_RESIZE@JBP_2021_PERMANENT



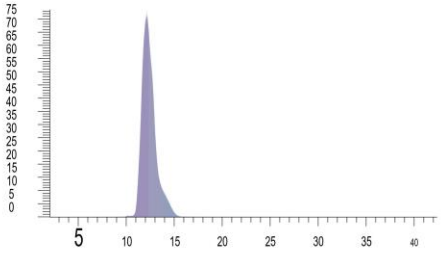
X-AXIS: Cell Values in thousands
Y-AXIS: Number of cells in hundreds

JBP_2021_S1_B3_RESIZE@JBP_2021_PERMANENT



X-AXIS: Cell Values in thousands
Y-AXIS: Number of cells in hundreds

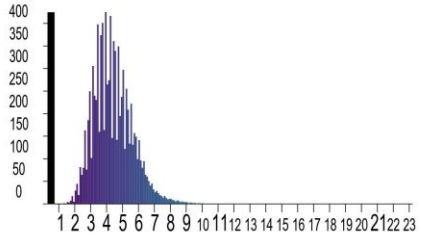
JBP_2021_S1_B2_RESIZE@JBP_2021_PERMANENT



X-AXIS: Cell Values in thousands
Y-AXIS: Number of cells in hundreds

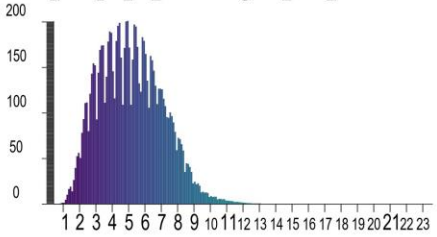
From 2000 data set –after resize

JBP_2000_S1_B2_RESCALE@JBP_2000_PERMANENT



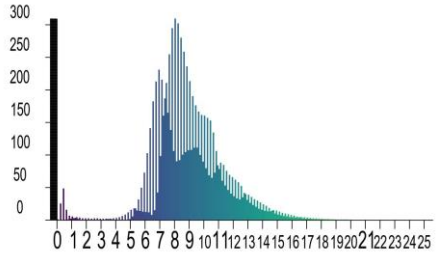
X-AXIS: Cell Values in tens
Y-AXIS: Number of cells in thousands

JBP_2000_S1_B3_RESCALE@JBP_2000_PERMANENT



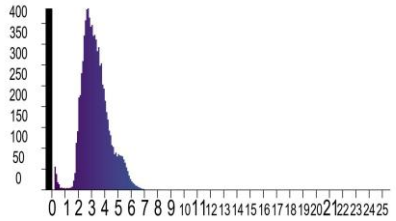
X-AXIS: Cell Values in tens
Y-AXIS: Number of cells in thousands

JBP_2000_S1_B4_RESCALE@JBP_2000_PERMANENT



X-AXIS: Cell Values in tens
Y-AXIS: Number of cells in thousands

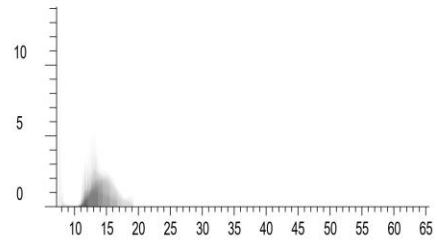
JBP_2000_S1_B5_RESCALE@JBP_2000_PERMANENT



X-AXIS: Cell Values in tens
Y-AXIS: Number of cells in thousands

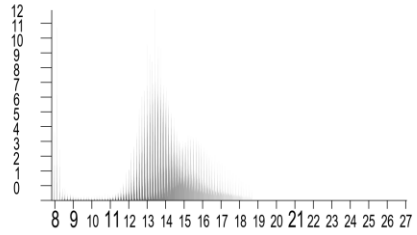
From 2000 data set –before resize

JBP_2000_S1_B5@JBP_2000_PERMANENT



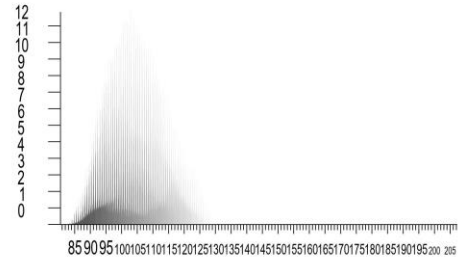
X-AXIS: Cell Values in thousands
Y-AXIS: Number of cells in thousands

JBP_2000_S1_B4@JBP_2000_PERMANENT



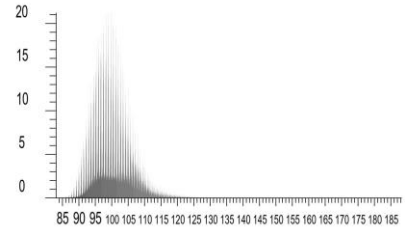
X-AXIS: Cell Values in thousands
Y-AXIS: Number of cells in thousands

JBP_2000_S1_B3@JBP_2000_PERMANENT



X-AXIS: Cell Values in hundreds
Y-AXIS: Number of cells in thousands

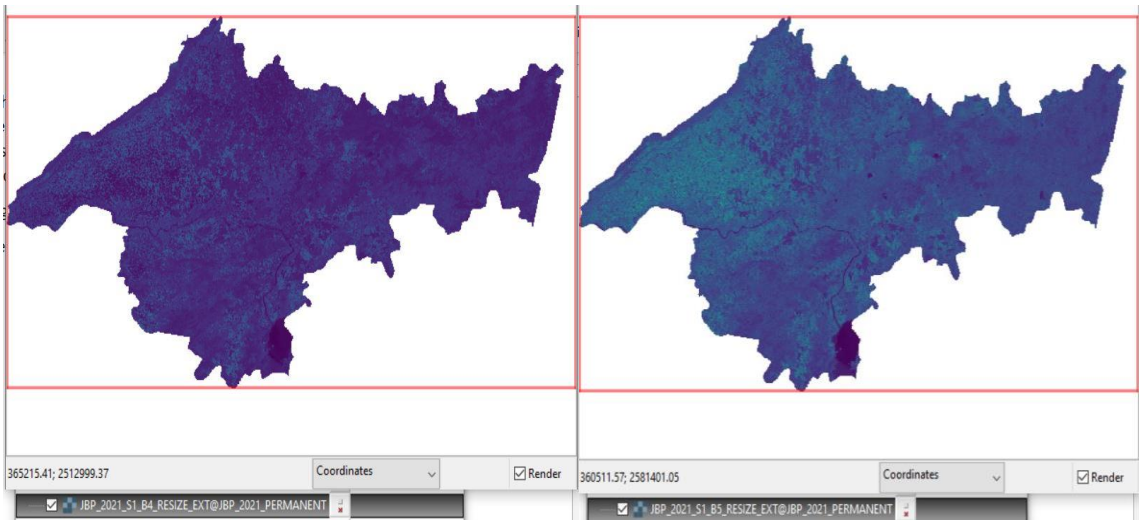
JBP_2000_S1_B2@JBP_2000_PERMANENT

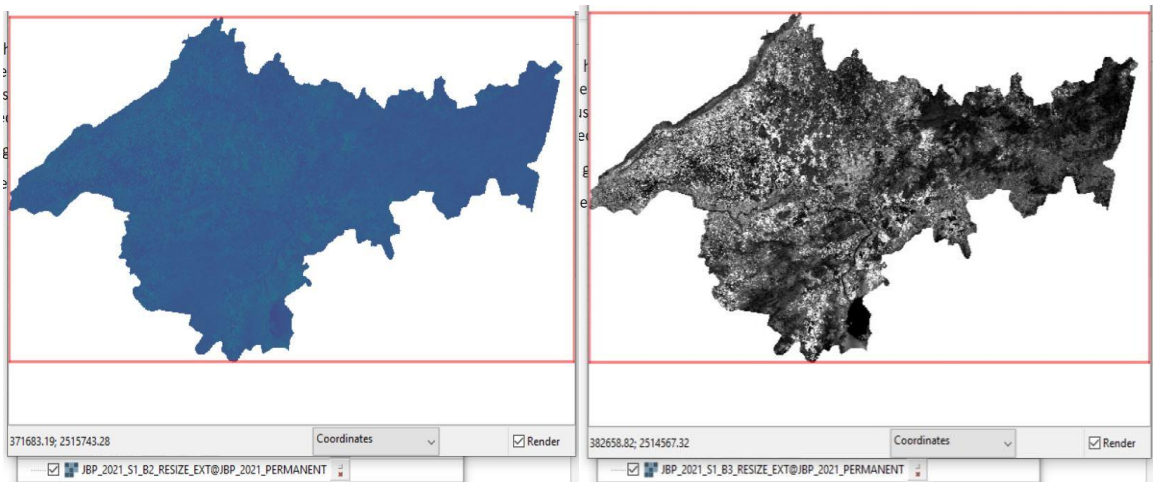


X-AXIS: Cell Values in hundreds
Y-AXIS: Number of cells in thousands

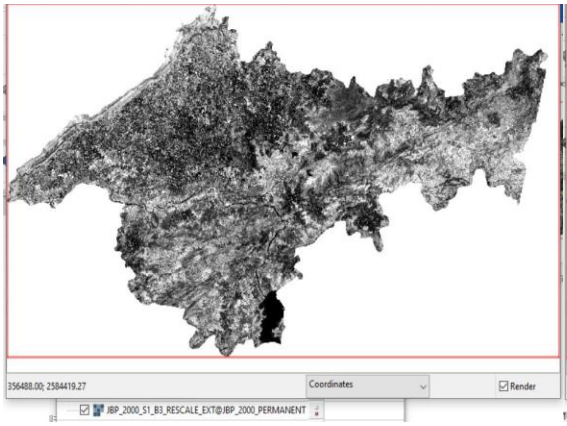
Extracted Image of band is given below-

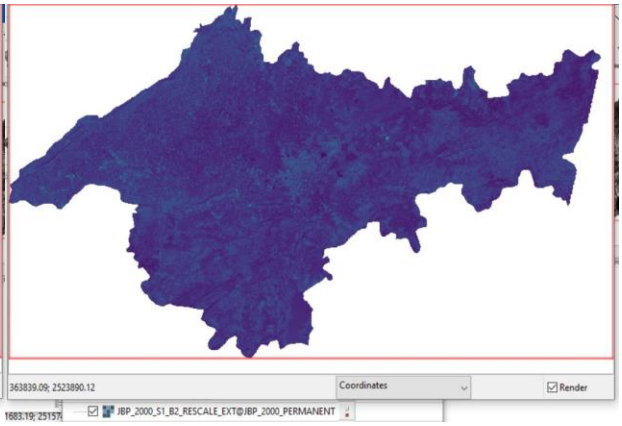
Data set of 2021

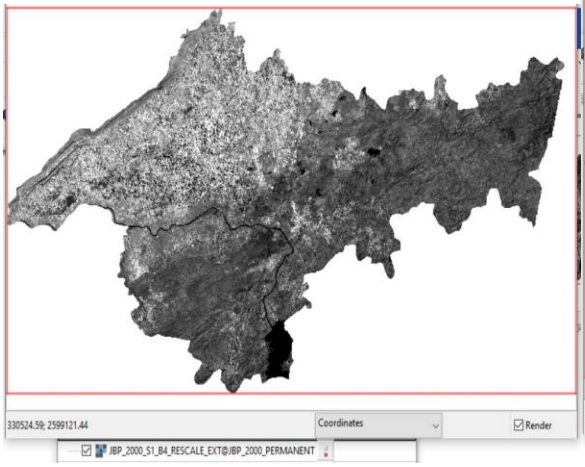


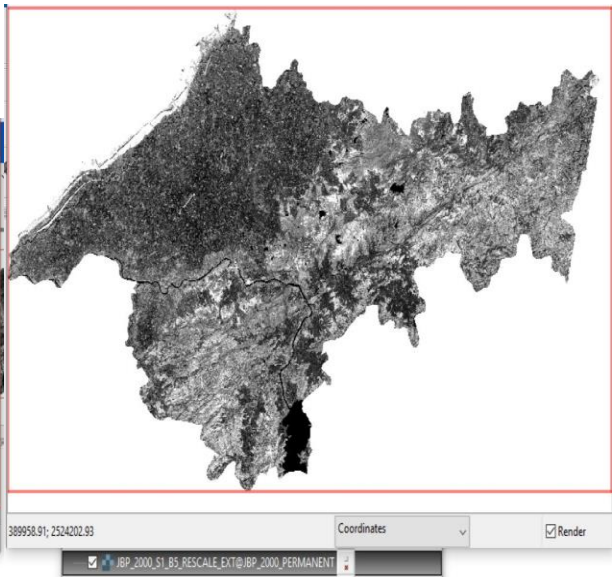


Data set of 2000


356488.00; 2584419.27
Coordinates
Render
JBP_2000_S1_B3_RESCALE_EXT@JBP_2000_PERMANENT


363839.09; 2523890.12
Coordinates
Render
JBP_2000_S1_B2_RESCALE_EXT@JBP_2000_PERMANENT


330524.59; 2599121.44
Coordinates
Render
JBP_2000_S1_B4_RESCALE_EXT@JBP_2000_PERMANENT


389958.91; 2524202.93
Coordinates
Render
JBP_2000_S1_B5_RESCALE_EXT@JBP_2000_PERMANENT

Jabalpur area distribution and percent cover of

data set of 2021

Sno	description	acres	%
1	buildup	148,765	6.75
2	agriculture	211,065	9.58
3	forest	6374	0.29
4	open space	589,246	26.74
5	water	14,574	0.66

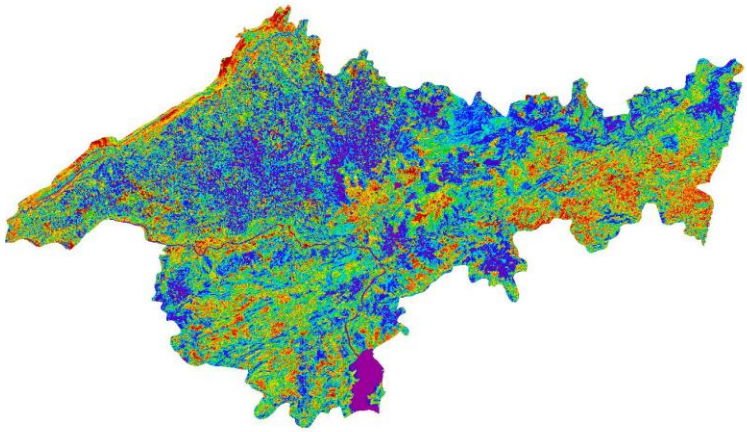
Data set of 2000

Sno	Area	acres	cover
1	buildup.	75,348	3.42
2	agriculture	1,03,542	4.7
3	forest	2,15,451	9.78
4	water.	3,54,512	16.09
5	openspace	2,22,446	10.1

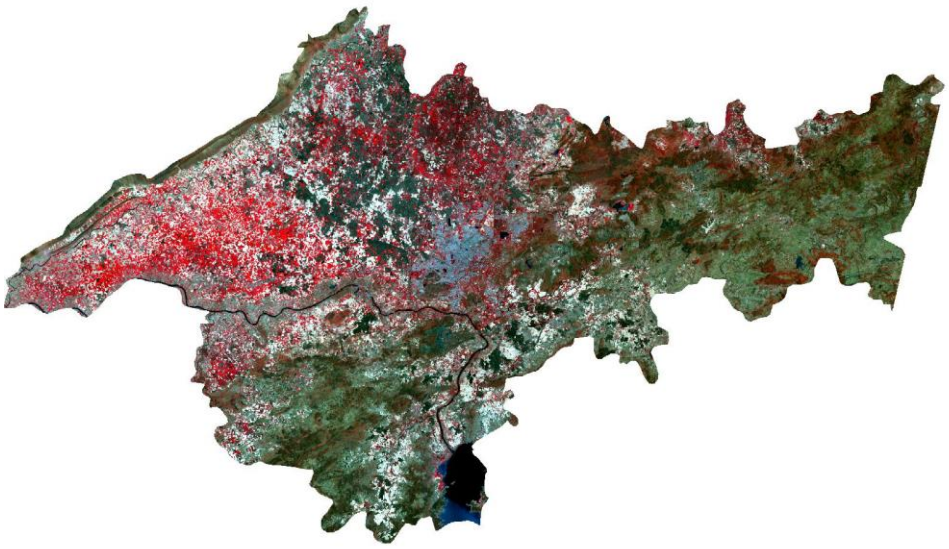
FCC

FCC false color composition is used to assign false colour to different color band to identify feature uniquely. Ex I am assigning red color in place of green, green in place of blue and infrared in place of red. B2 is of blue band, B3 is for green color band, B4 is for red color band and B5 is for infrared color band .

FCC of year 2000



FCC of year 2021



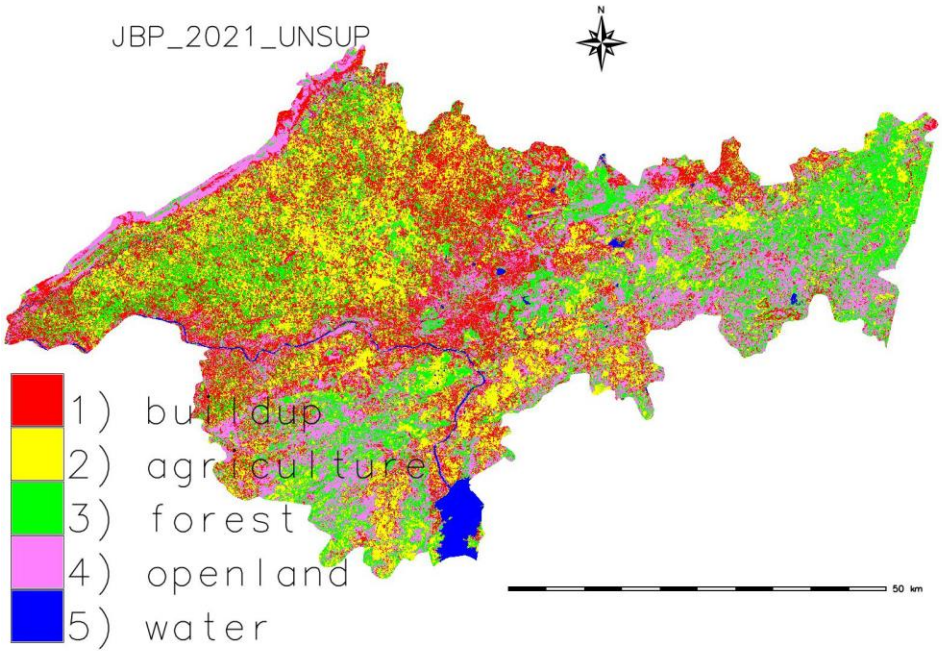
Unsupervised Learning

Unsupervised learning, also known as unsupervised machine learning, analyses and clusters unlabeled datasets using machine learning techniques. Without the need for human intervention, these algorithms uncover hidden patterns or data groupings.

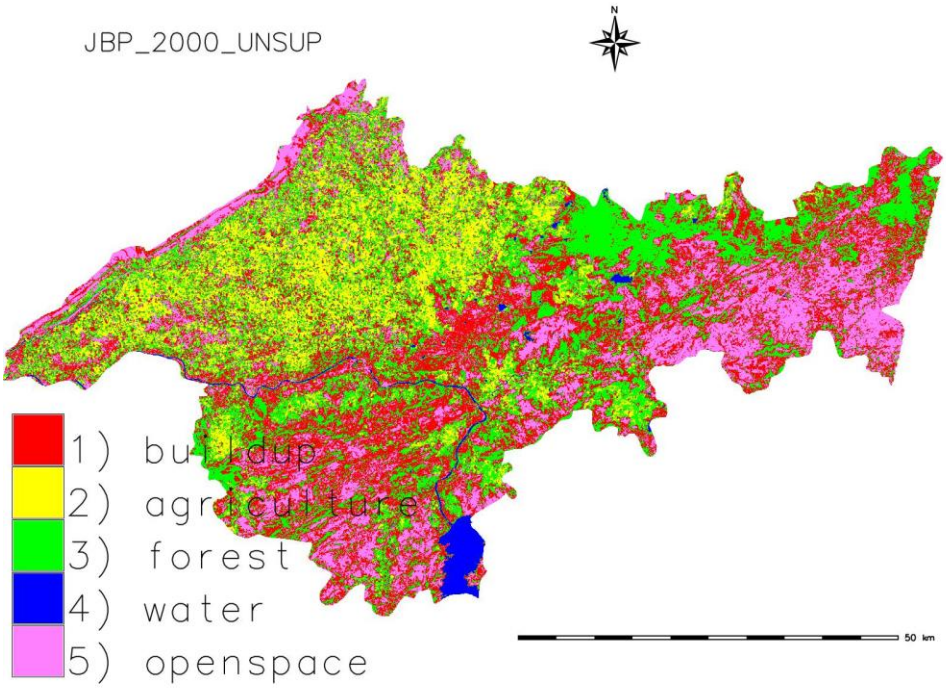
An AI system is supplied with unlabeled, uncategorized data in unsupervised learning, and the system's algorithms act on the data without any prior training. Compared to supervised learning systems, unsupervised learning algorithms can handle more complex processing tasks.

For unsupervised learning, we make 8 different class FCC map to distinguish different feature in FCC. In new data total 2^16 class is possible and in old data 2^8 class is possible, but it is huge data so it will become impossible to distinguish different class so we choose 2^3 to perform better classification.

Unsupervised classification of year 2021 data:-



Unsupervised classification of year 2000 data:-



Supervised Learning

Supervised learning, often known as supervised machine learning, is an artificial intelligence and machine learning subcategory. Its use of labelled datasets to train algorithms that accurately classify data or predict outcomes defines it.

Its use of labelled datasets to train algorithms that accurately classify data or predict outcomes defines it. As input data is fed into the model, the weights are adjusted until the model is properly fitted, which happens during the cross-validation phase.

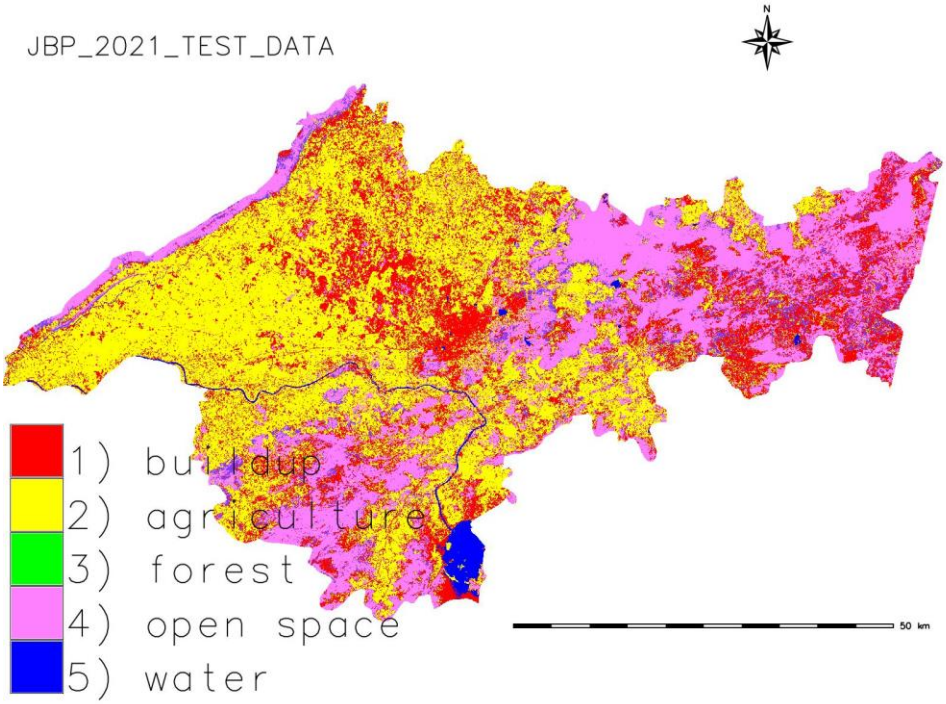
A training set is used in supervised learning to teach models to produce the desired output. This training dataset contains both correct and incorrect outputs, allowing the model to improve over time. The loss function is used to assess the algorithm's correctness, and it is adjusted until the error is suitably minimized.

For supervised learning, we need train data and test data. For making train data, initially we make vector data of different class (agriculture, buildup, forest, water, open land) by using polygon. Initially we make polygon on FCC data (at least 5 polygon in each and every sector). After making polygon of each class, I have converted vector data into raster data and then I have used input for supervised.

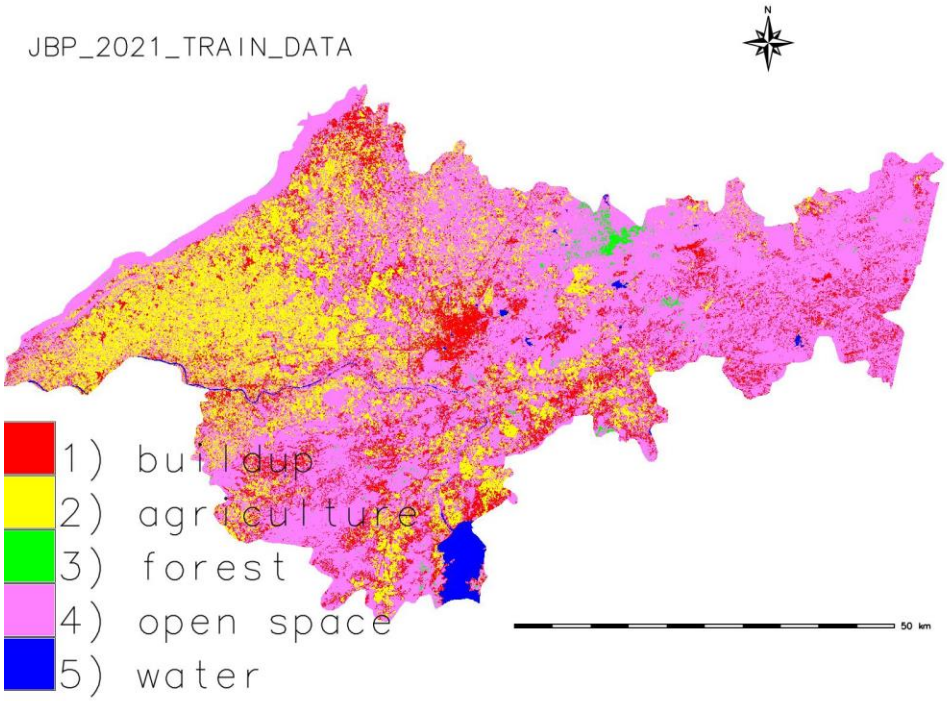
For test data, I make polygon on google earth map with same concept as for FCC data. After making polygon of each class. I have downloaded data in kml file and imported it into grass gis . I have converted vector data into raster data and then I have used input for supervised.

After input for supervised, I have assigned name to each polygon in grass location folder using notepad in grass location folder. After combining, I have used MLC classifier to classify data after importing combine data of test data and train data.

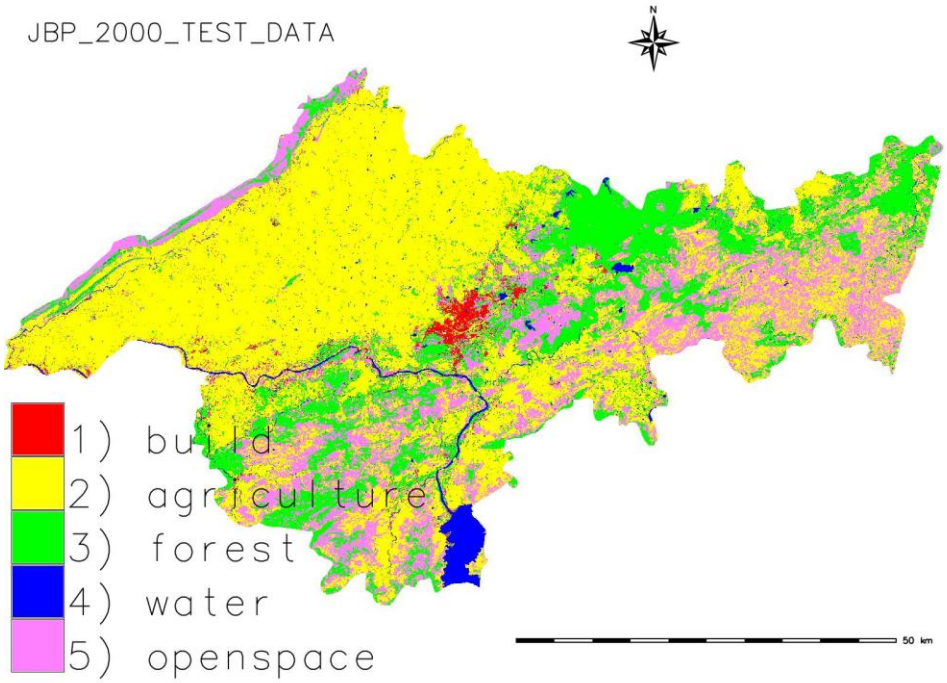
Test Data of year 2021



Train Data of year 2021



Test Data of year 2000



Kappa and Accuracy

kappa is a metric often used to assess the agreement between two raters. It can also be used to assess the performance of a classification model.

Cohen's kappa coefficient () is a statistic that is used to assess qualitative (categorical) item inter-rater reliability (and also intra-rater reliability). [1] It is typically believed to be a more reliable statistic than a simple % agreement estimate since it considers the potential of agreement occurring by chance. Due to the difficulties in understanding indices of agreement, Cohen's kappa has sparked debate. According to some studies, evaluating disagreement between items is conceptually simpler.

One parameter for evaluating classification models is accuracy. Informally, accuracy refers to the percentage of correct predictions made by our model.

Formula used to calculate accuracy

$$\text{Overall accuracy} = \frac{\text{total number of correctly classified pixels}}{\text{total number of reference pixel}}$$

Formula used to calculate kappa value

$$\hat{\kappa} = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}\left(x_{i+} \cdot x_{+i}\right)}{N^2 - \sum_{i=1}^{r}\left(x_{i+} \cdot x_{+i}\right)} = \frac{\text{overall accuracy} - \text{expected accuracy}}{1 - \text{expected accuracy}}$$

Value of kappa and accuracy on year 2000

```
LOCATION: JBP_2000_DATA_SET                              Mon Nov 27 22:01:54 2021
MASK: none
MAPS: MAP1 = Reclass of GE_TEST_DATA_NEW_MLC@JBP_2000_PERMANENT in JBP_2000_PERMA
      MAP2 = Reclass of Final_seg_train_out@JBP_2000_PERMANENT in JBP_2000_PERMAN
Error Matrix (MAP1: reference, MAP2: classification)
Panel #1 of 1
                        MAP1
     cat#       1       2       3       4       5
 M     1        49130   128777  29359   959     130596
 A     2        0       465602  0       0       0
 P     3        0       607933  360538  281     81
 2     4        27165   1142983 205154  129428  89424
       5        72      55943   294480  8       649782
Col Sum         76367   2401238 889531  130676  869883
cat#    Row Sum
F     1     49130      128777       29359         959      130596     338821
i     2         0      465602           0           0           0    1143244
n     3         0      607933      360538         281          81    2916500
a     4     27165     1142983      205154      129428       89424    6283910
l     5        72       55943      294480           8      649782   10651605
 21334080
Cats    % Comission     % Omission      Estimated Kappa
1       85.499718       35.665929       0.129788
2       0.000000        80.609919       1.000000
3       62.786363       59.468754       0.211562
4       91.881086       0.955034        0.052852
5       35.040314       25.302368       0.562454
Kappa           Kappa Variance
0.55721         0.000000
Obs Correct     Total Obs       % Observed Correct
2746403         4367695         62.879934
```

Value of kappa and accuracy on year 2021

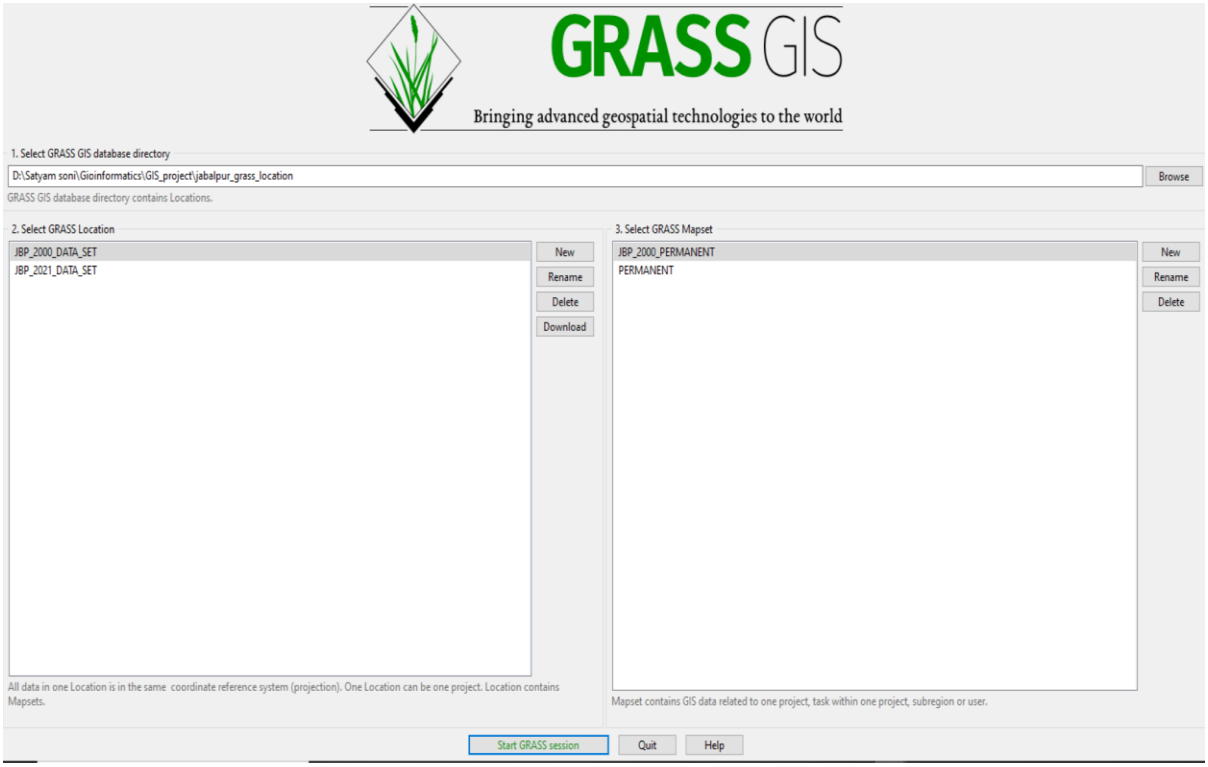
```
LOCATION: JBP_2021_DATA_SET                          Tue Nov 28 06:52:03 2021
MASK: none
MAPS: MAP1 = Reclass of GE_NEW_TEST_DATA_MLC@JBP_2021_PERMANENT in JBP_2021_PERMANEN
      MAP2 = Reclass of FINAL_SG_T1@JBP_2021_PERMANENT in JBP_2021_PERMANENT (FINAL_
Error Matrix (MAP1: reference, MAP2: classification)
Panel #1 of 1
                          MAP1
     cat#       1       2       3       4       5
 M      1       489398  162195  7268    10012   88
 A      2       25226   921736  470     0       1678
 P      3       588     23      28045   7       0
 2      4       546063  835522  1071411 188712  7991
        5       10963   1709    0       0       52865
Col Sum         1072238 1921185 1107194 198731  62622
cat#    Row Sum
F      1    489398      162195        7268       10012          88     668961
I      2     25226      921736         470           0        1678    2287032
N      3       588          23       28045           7           0    3933766
A      4    546063      835522     1071411      188712        7991    8230199
L      5     10963        1709           0           0       52865   12592169
 27712127
Cats    % Comission     % Omission      Estimated Kappa
1       26.842073       54.357335       0.644092
2       2.884176        52.022528       0.948456
3       2.156090        97.467020       0.971105
4       92.877983       5.041488        0.026885
5       19.335642       15.580786       0.803827
Kappa           Kappa Variance
0.51727         0.000000
Obs Correct     Total Obs       % Observed Correct
2596769         4361970         59.532039
```

Procedure

1 :- Initially I have downloaded India district boundaries and Remote sensing data from Deva Gis and earth explorer respectively. And imported it in QGis to extract area of interest (Jabalpur District). And I have created 10 km buffer of Jabalpur district boundaries to understand district information correctly. Projection of Coordinates is UTM. It is saved in particular folder of GIS project that's why we can get this data any instant of time during performing project.

2 :- Now I have uploaded boundaries data in Grass software using read projection and datum terms from a georeferencing option. I have created special folder for grass software where all processing data will store. I have created two locations of grass where two different year data will store.

Some image is given below-



3 :- I have given JBP_2000_DATA_SET and JBP_2021_DATA_SET, name of folder where all data processing of 2000 year and 2021 year respectively will store.

4 :- Then I have imported boundaries in grass software then I have imported remote sensing data in software of selected band that are B2, B3, B4, B5.then I have used raster map calculator to get area of interest of remote sensing rescaled data.

5 :- After importing data, I have removed Geometric correction using google earth cordinates then I have removed radiometric error by rescaling new data from 0 to 65535 and rescaling old data from 0 to 255.

6 :- After removing error , I have extracted area of intrest .I have converted all vector data into raster data and choosen cat option . As Raster data stored in discrete number called pixel or category. And vector data is stored in attribute.

7 :- then I have formed FCC (false colour composition) to understand different class.

8 :- then I have created group and subgroup with all selected band data.

9 :- then for unsupervised , I have made 8 cluster .then I have used Maximum likelihood classifier to classify data. After mapping 8 cluster class to five class (

buildup,agriculture,crop,forest,water,openland). I have saved area in acre and cover percentage.

10 :- then add raster legend ,scale bar , north arrow and text to map display.

11 :- then for supervised learning, we need to create test and train data . So add vecter legand of buildup,agriculture,forest,water,openland to map display .then create area polygon in each section. then convert vector data into raster data And store in signature.
12 : I have done same in google earth explorer and store it into specific folder. Import data it into grass software then Proceed same step as given in step no. 11.

14 :- I have Open folder where test data and train data was stored. Then I have assigned specific name to each data of buildup,agriculture,forest,water,openland raster data. And store it in single file of test data and train data.

13 :- then import use MLC to classify test data and train data and thenI have reclassify it according to report . finally I have calculated Kappa and accuracy.