

Parametric Methods to Multispectral Image Classification using Normalized Difference Vegetation Index

Keerti Kulkarni, P. A. Vijaya,

Abstract: *The key to proper governance of the municipal bodies lies in knowing the geography of the region. The land cover of the region changes with respect to time. Also, there are seasonal variation in the layout of the waterbodies. Manual verification and surveying of these things becomes very difficult for want of resources. Remote Sensing Images play a very important role in mapping the land cover. In this paper, we consider such remotely sensed Multispectral Images, taken from Landsat-8. Parametric Machine learning algorithm like Maximum Likelihood Classifier has been used on those images to classify the land cover. Normalized Difference Vegetation Index (NDVI) has been calculated and integrates with the classification process. Four basic land covers have been identified for the purpose namely Water, Vegetation, Built-up and Barren soil. The area of study is Bangalore urban region where we find that the water bodies are decreasing day by day. An overall efficiency of 82% with a kappa hat of 0.67 has been achieved with the method. The user and the producer accuracies have also been tabulated in the Results part. The results show the land cover changes in a temporal manner.*

Keywords: *Land Cover Classification, Bangalore Urban, Multispectral Landsat Images, Maximum Likelihood Classifier, Normalized Difference Vegetation Index (NDVI).*

I. INTRODUCTION

The Mapping of the land cover plays a major role in a majority of the decisions taken by the civic bodies. In a rapidly changing urban landscape, keeping up with the changes in the land-cover by manual survey is very difficult. Hence the increasing reliance on the Remotely Sensed Images for the land cover Classification. The area under consideration is the Bangalore urban district which is very heterogenous in nature. The area is very densely packed with more number of built-up areas and less number of greenery and wetlands. The number of lakes and the vegetation is also decreasing day by day. Remotely sensed images have been used to estimate the land cover of the region. In this work, we have used the LANDSAT-8 images, freely downloadable from the glovis website, for the purpose of classification. The LANDSAT-8 Multispectral images consist of 11 bands, including the thermal bands. Temporal analysis requires certain types of corrections to be implemented on the images. The images need to be processed for Atmospheric Corrections and then used for Classification.

Revised Manuscript Received on December 15, 2019.

Keerti Kulkarni, Assistant Prof, ECE Dept, BNM Institute of Technology, Bangalore. (Email: keertikulkarni@bnmit.in)

Dr. P. A. Vijaya, Prof & Head, ECE Dept, BNM Institute of Technology, Bangalore. (Email: pavmkv@gmail.com)

A parametric machine learning model is the one which uses a limited set of parameters irrespective of the amount of training samples you provide to the model. Even if a large number of training pixels are provided to the model, the number of parameters used to summarize the data will still remain same.

II. LITERATURE SURVEY

The RESOURCESAT, LISS-3 images for the land cover classification has been done by the authors in [2]. Spectral angle mapping has been used for the Classification process. LANDSAT images have been traditionally used for the land cover classification. In [2] the authors have shown a comparison of SAR DATA and LANDSAT-8 images. They have compared the results of classification using Maximum Likelihood Classifier. Object based approaches have been increasingly preferred over pixel based approaches [3]. Land Cover Classification is done using a supervised object based approach. The results have been shown in [4], which gives a comparative accuracy of the different object based image analysis approaches. Geographic Information Systems (GIS) have been used in applications which involve applications where the exact geographic location is important, such as monitoring the growth of the cities or towns [5]. Markov chains are generally associated with predictions. Just getting the classification report is not enough, rather a future prediction of the land cover was proved by the authors [6] by using the prediction analysis. Supervised classification method has been integrated with the Global Land Cover Facility Site (GLCF) to increase the accuracy of classification [7]. An integration of one or more than one remotely sensed images obtained from different sensors can help in improving the overall accuracy as shown in [8]. The authors have used the Supervised- Maximum Likelihood classification to observe the extent of landcover changes in the Simly Watershed. American and French Satellite Images were used together. Degradation analysis is also possible by using thermal images. This was analysed in [9]. Maximum likelihood classification was used on the Landsat Thematic Mapper and Landsat 8 optical and Thermal images. The Land Surface Temperature is a function of Land cover, as shown by the authors [10]. The land surface temperature is generally higher for the built-up areas. Since the old ages, Cartography is used to map the land cover. Using the cartography of the given geographic area augments the classification process as shown by the authors [11].

The authors have integrated Remote Sensing images, GIS and Cartography to map the land cover. Normalized difference indices were used to validate the results. Object based approaches along with GIS in remotely sensed images resulted in a classification accuracy of 80% [12]. Classification results cannot be held completely valid unless it is substantiated by the ground reality. The authors have stressed upon the obtaining the ground reality along with the remotely sensed images in the complete mapping of LULC [13]. Atmospheric Corrections, geometric corrections and orthorectification are very important if the classification results have to be temporally analysed. Atmospheric Corrections and calculating the accuracies have to go hand-in hand if the classification results have to be meaningfully used [14]. Maximum likelihood classification was used for the process of classification in the paper. A comparative study of all the methods used in the classification of Land Cover was done by the authors in [16].

III. STUDY REGION

The study region is the complete Bangalore Urban District with the range 77.5E, 13N to 77.75E, 13N. Bangalore, being a technical hub of India, the city has developed in leaps and bounds by the way of urban infrastructure. Remotely sensed Satellite images are collected from the glovis website, Landsat-8 has 8 bands having a resolution of 30m, one panchromatic band having a resolution of 15m, and two thermal bands having a resolution of 100m. Two datasets have been used one collected on 7th Nov 2013 and the other on 5th April 2017. Both the datasets have been orthorectified.

IV. MATERIALS AND METHODS

A. Images used in the work

The following Table shows the details of the Landsat images used.

Table 1 : Images used

Data Set	Year of Acquisition	Spatial Resolution	Satellite	File Format	Cloud Coverage
1	07 th Nov 2013	30m	Landsat-8	GeoTiff (Raster)	0-10%
2	5 th April 2017				

B. Pre-processing

The LANDSAT-8 images are represented in terms of Digital Numbers (DN Values) which is nothing but the pixel values. But for processing we need the values which have been actually reflected from the object on earth in the absence of atmospheric effects. Hence the need for preprocessing.

C. Feature Selection

Before we actually go to the classification process, creating a Regions of Interest (ROIs) is very important. Regions of Interest are created based on the spectral signatures and the spectral distances between each class. The spectral signatures of the different land cover classes used for

classification is shown in Figure 4.1. Spectral Distances are calculated using the Jeffries-Matusita Distance which calculates the separability of a pair of probability distribution. It is given by the following equation.

$$J_{ab} = 2(1 - e^{-B}) \tag{1}$$

Where

$$B = \frac{1}{8} ((a - b)^T (\frac{\Sigma a + \Sigma b}{2})^{-1} (a - b) + \frac{1}{2} \ln(\frac{\Sigma a + \Sigma b}{|\Sigma a| |\Sigma b|}) \tag{2}$$

where:

- a = spectral signature vector of the first class;
- b = spectral signature vector of the second class;
- Σ a = covariance matrix of sample a;
- Σ b = covariance matrix of sample b;

The distance value is 0 if signatures are same.

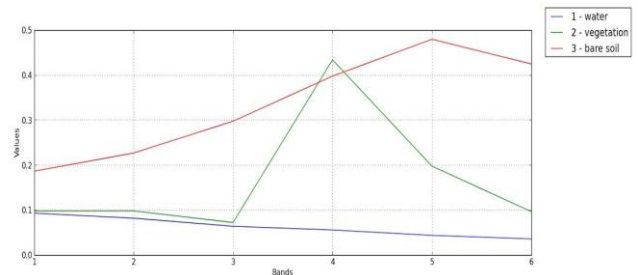


Figure 4.1 : Spectral Signatures of different Land cover Classes.

The study area is very heterogenous in nature. Hence the spectral signatures are not always very different. Someplace, water-bodies are covered in marsh hence the spectral signatures are between that of water and vegetation and hence difficult to evaluate. In such cases the near infrared band and the red bands are used to calculate the various Vegetation Indices. Popularly used is the Normalized Difference (NDVI), which aids in the distinction of the classes. NDVI is defined as:

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)} \tag{3}$$

NDVI values range from -1 to 1. Greenery shows values nearer to 1, while built-up areas have low NDVI values. For water, the NDVI values are nearer to zero.

D. Classification

In Accordance to the Bayes' Theorem, classification of a pixel to any one of the four classes, i.e., vegetation, water, built-up or soil, depending on the probability distribution for the classes. We also need to calculate the spectral distance (or Separability) between the training samples or pixels. This helps us to know the similarity between the chosen training samples. If the samples are too similar, it may cause errors in the classification process. The discriminant function, is calculated for every pixel and is defined as:

$$g_k(a) = \ln p(C_k) - \frac{1}{2} \ln |\Sigma_k| - \frac{1}{2} (a - b_k)^T \Sigma_k^{-1} (a - b_k) \tag{4}$$

k takes the value 1,2 3, and 4 for the four land cover classes respectively

Where C_k = any one of the four classes;

a = spectral signature vector of a image pixel;

Therefore:

$$a \in C_k \Leftrightarrow g_k(a) > g_j(x) \forall k \neq j \quad (5)$$

E. Post Processing

The Classification process is not usually, the end result. Accuracy assessment is needed to know the classification errors. It is done by calculating the error matrix. The Classification reports for the two different time durations are generated which further help is generating the bar graph. The Kappa value is a metric that compares an Observed Accuracy with an Expected Accuracy (random chance). Kappa is calculated by the following equation.

$$kappa = \frac{totalAccuracy - randomAccuracy}{1 - randomAccuracy}$$

Where

$$totalAccuracy = \frac{A + B}{A + B + C + D}$$

Where A= True Positive, B=True Negative, C=False Positive, D= False Negative

$$randomAccuracy = \frac{ActualFalse * PredictedFalse + ActualTrue * PredictedTrue}{Total * Total}$$

$$randomAccuracy = \frac{(A + C) * (B + D) + (D + A) * (C + A)}{Total * Total}$$

The above values are calculated from the confusion matrix or the error matrix. The error matrix for the algorithm is shown in the Table 2.

V. METHODOLOGY

The flowchart in Figure 5.1 gives a brief idea of the classification process.

After the preprocessing, the Thermal band (Band 10) is separated out for further processing. Band numbers 2,3,4,6 are used mapping the land cover. Band 4 and Band 5 are used to calculate the NDVI values. The NDVI raster further helps in the classification process. The classification process consists of the following steps.

1. Decide the number of Classes used for Classification. Here we use the following 4 classes.

- Water (lakes)
- Vegetation (parks and forests)
- Built-up (buildings, roads etc)
- Soil (barren, with some minor vegetation)

2. Choose the correct band combination. Here we choose bands 4-3-2 and bands 3-2-1 for both the datasets.

3. Data Preparation using the Spectral Distances and Regions of Interest (ROIs).

4. Calculate and display the NDVI Raster.

5. Applying the various Parametric Classification Algorithms using the NDVI Raster as the Metadata, as described in the previous sections.

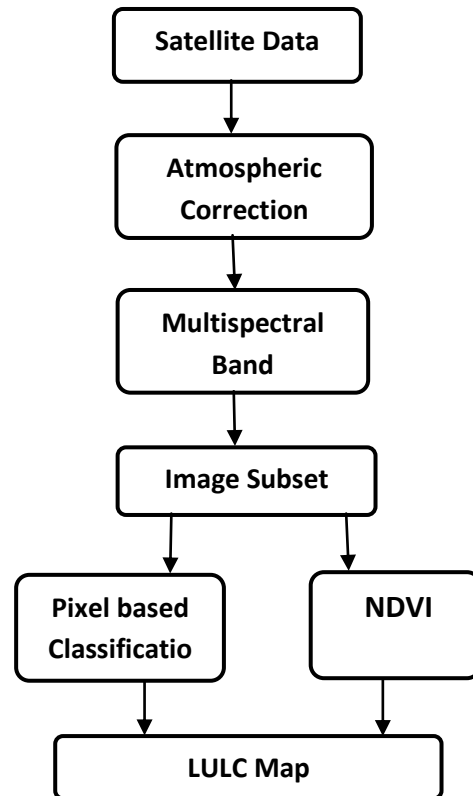


Figure 5.1 : Flowchart for classification using machine learning models

VI. RESULTS AND DISCUSSIONS

The study region used for classification is shown in Figure 6.1. The image is a raster of Bangalore Urban region downloaded from the usgs website [15].

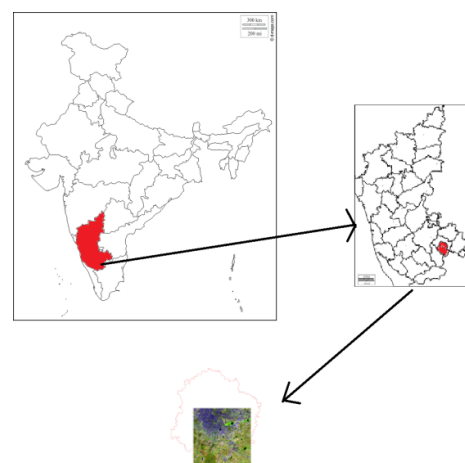


Figure 6.1: Study Area

Results using Maximum Likelihood Estimation

Two different data sets, one for the year 2013 and the other for the year 2017 have been downloaded. Each data set consists of 11 bands. The data set also contains one additional Metadata file.

Figure 6.2 shows the Band-4 of the LANDSAT-8 image captured in 2013. The Band-4 is a raster and the Bangalore Urban clip is a vector. Both have been tiled to verify that the work is actually carried on at the specified study area. Figure 6.3 shows the band combination raster for the bands 4-3-2. Here the vegetation is shown in red and the urban areas or built-up areas are shown in blue. This band combination is particularly useful in identifying the vegetation. Figure 6.4 shows the band combination raster for the bands 4-3-2. In this combination vegetation is shown in green and the built-up areas are bluish. Hence this band combination is also called the Natural Color Composite and the band combination 3-2-1 is called the False Color Composite. Figure 6.5 shows the NDVI raster derived from the formula discussed previously. NDVI aids in the classification process, since the Regions Of Interest (ROIs) can be mapped in a better manner with the knowledge of the NDVI values. The Maximum Likelihood algorithm has the spectral distance set to 0.01 and the threshold value to 0. Finally, the change in the land cover is displayed in the form of a bar graph (Figure 6.8). This has been derived from the Classification reports of both the years. The overall Accuracy is 82% with a kappa hat of 0.67.

Overall accuracy, user’s accuracy, producer’s accuracy and Kappa hat, are the general accuracies used to evaluate the calculation algorithm. Overall Accuracy gives us an idea as to how the classification was accurate geographically. Classification accuracy of 100% is generally desired, which in turn shows that all the land cover classes were classified perfectly.

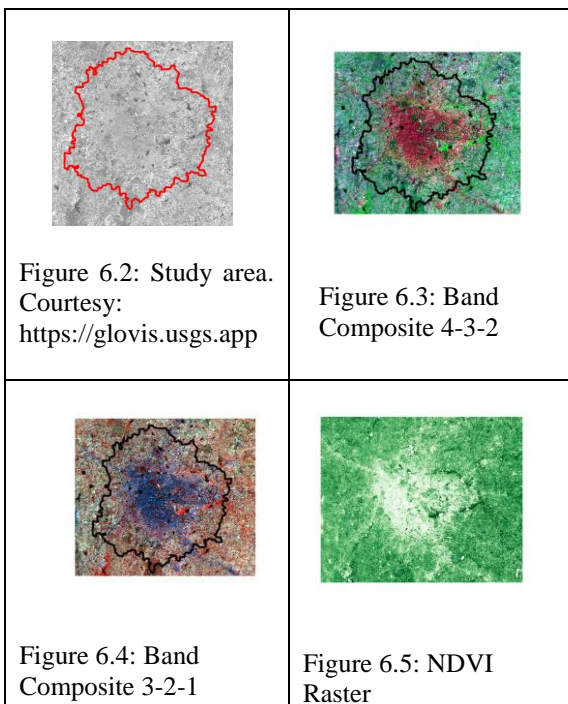


Figure 6.2: Study area. Courtesy: <https://glovis.usgs.app>

Figure 6.3: Band Composite 4-3-2

Figure 6.4: Band Composite 3-2-1

Figure 6.5: NDVI Raster

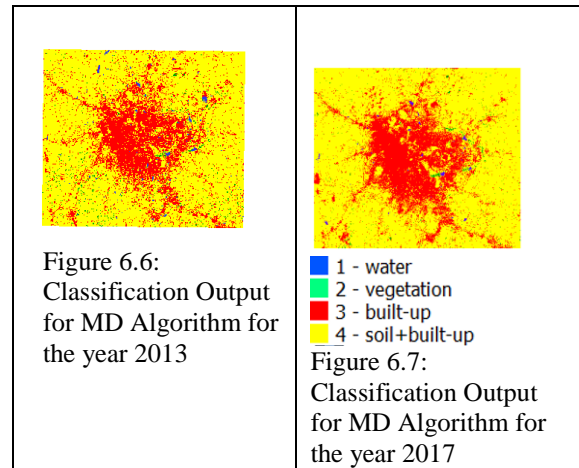


Figure 6.6: Classification Output for MD Algorithm for the year 2013

Figure 6.7: Classification Output for MD Algorithm for the year 2017

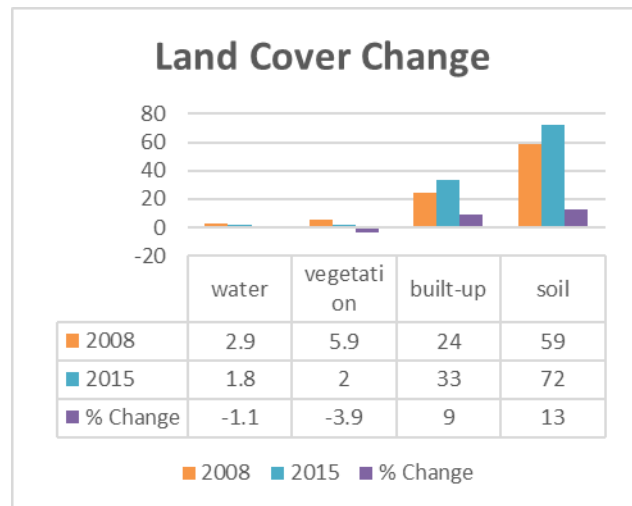


Figure 6.8 Land Cover Classification

Table 2: Area based Error Matrix

		Reference			
		Class-1	Class-2	Class-3	Class-4
Classified	Class-1	0.0022	0	0.0009	0
	Class-2	0	0.3004	0.0619	0
	Class-3	0	0.0998	0.4242	0.001
	Class-4	0	0.009	0.1004	0.001
	Total	0.0022	0.4009	0.5904	0.002

Table 3: Producer’s and user’s accuracy

	Class-1	Class-2	Class-3	Class-4
Producer’s accuracy	98	75.8	86.2	95.2
User’s accuracy	72.7	81.7	84	70.4

VII. CONCLUSION

The Maximum Likelihood Classifier gives a good accuracy for the Land Cover Mapping of the area under consideration, given the fact that the area is very heterogenous in nature. The accuracy can be further improved by using the non parametric methods such as Random Forest and Support Vector Machines

ACKNOWLEDGEMENT

We extend my heartfelt thanks to the Visvesvaraya Technological Institute (VTU) for giving us a proper platform for doing our research work. We also sincerely thank the Management of BNMIT for providing us the resources required to carry out our research work.

REFERENCES

1. Keerti Kulkarni, Dr. P. A. Vijaya, ' Experiment of Multispectral Images using Spectral Angle Mapper Algorithm for Land Cover Classification', International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8, Issue-6S4, April 2019, pp 96-99.
2. Ali, M. Z., Qazi, W., & Aslam, N. (2018). The Egyptian Journal of Remote Sensing and Space Sciences A comparative study of ALOS-2 PALSAR and landsat-8 imagery for land cover classification using maximum likelihood classifier. The Egyptian Journal of Remote Sensing and Space Sciences, 21, S29-S35. <https://doi.org/10.1016/j.ejrs.2018.03.003>
3. Ma, L., Li, M., Ma, X., Cheng, L., Du, P., & Liu, Y. (2017). ISPRS Journal of Photogrammetry and Remote Sensing A review of supervised object-based land-cover image classification. ISPRS Journal of Photogrammetry and Remote Sensing, 130, 277-293. <https://doi.org/10.1016/j.isprsjprs.2017.06.001>
4. Ye, S., Gilmore, R., Jr, P., & Rakshit, R. (2018). ISPRS Journal of Photogrammetry and Remote Sensing A review of accuracy assessment for object-based image analysis: From per-pixel to per-polygon approaches. 141,137-147.
5. Monitoring urban growth and land use change detection with GIS and remote sensing techniques in Daqahlia governorate Egypt, Ibrahim Rizk Hegazya, Mosbeh, Rashed Kaloopb, International Journal of Sustainable Built Environment, Volume 4, Issue 1, June 2015, Pages 117-124
6. Monitoring land use/cover change using remote sensing and GIS techniques: A case study of Hawalbagh block, district Almora, Uttarakhand, India, J.S.Rawat, Manish Kumar, The Egyptian Journal of Remote Sensing and Space Science, Volume 18, Issue 1, June 2015, Pages 77-84
7. Land use change mapping and analysis using Remote Sensing and GIS: A case study of Simly watershed, Islamabad, Pakistan, Amna Butt, Rabia Shabbir Sheikh, Saeed Ahmad, Neelam Aziz, *The Egyptian Journal of Remote Sensing and Space Science, Volume 18, Issue 2, December 2015, Pages 251-259*
8. Land use classification and change detection by using multi-temporal remotely sensed imagery: The case of Chunati wildlife sanctuary, Bangladesh, Kamrul Islam^a Mohammed Jashimuddin^a Biswajit Nath^b Tapan Kumar Nath^c, *The Egyptian Journal of Remote Sensing and Space Science, Volume 21, Issue 1, April 2018, Pages 37-47*
9. Detection of land use and land cover change and land surface temperature in English Bazar urban centre, Swades Pal, Sk Ziaul *The Egyptian Journal of Remote Sensing and Space Science, Volume 20, Issue 1, June 2017, Pages 125-145*
10. Mapping Land Cover Using Remote Sensing Data and GIS Techniques: A Case Study of Prahova Subcarpathians, Marina-Ramona Rujoiu-Mare Bogdan-Andrei Mihai, *Procedia Environmental Sciences, Volume 32, 2016, Pages 244-255*
11. Changes in land use/cover using geospatial techniques: A case study of Ramnagar town area, district Nainital, Uttarakhand, India, J.S.Rawat, Vivekanand Biswas, Manish Kumar, The Egyptian Journal

of Remote Sensing and Space Science, Volume 16, Issue 1, June 2013, Pages 111-117

12. Global land cover mapping at 30 m resolution: A POK-based operational approach, JunChen^a JinChen^b AnpingLiao^a XinCao^b LijunChen^a Xuehong Chen^b ChaoyingHe^a GangHan^a ShuPeng^a MiaoLu^a WeiweiZhang^a Xiaohu aTong^a JonMills^d, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 103, May 2015, Pages 7-27
13. Land Use-Land Cover dynamics of Huluka watershed, Central Rift Valley, Ethiopia, HagosGebreslassie¹, International Soil and Water Conservation Research, Volume 2, Issue 4, December 2014, Pages 25-33
14. Land cover change detection using GIS and remote sensing techniques: A spatio-temporal study on Tanguar Haor, Sunamganj, Bangladesh, Md. Inzamal Haque, Rony Basak, The Egyptian Journal of Remote Sensing and Space Science, Volume 20, Issue 2, December 2017, Pages 251-263
15. Kulkarni, K., & Vijaya, P. A. (2018). A comparative study of land classification using remotely sensed data. Proceedings of the International Conference on Computing Methodologies and Communication, ICCMC 2017, <https://doi.org/10.1109/ICCMC.2017.8282720>
16. <https://glovis.usgs.app>

AUTHORS PROFILE



Keerti Kulkarni has around 15 years of teaching experience. She has obtained her BE degree from the University of Bombay, M.tech from DSCE, VTU. She is currently working towards her Ph.D degree in the area of Satellite Image Classification.



Dr. P. A. Vijaya is presently Professor and Head of the Department, Department of ECE, B N M Institute of Technology. Dr. P A Vijaya obtained BE degree from MCE Hassan, ME & Ph.D from IISc Bengaluru. She incorporates a wealthy teaching experience of thirty-two years and more than a hundred publications to her credit. Three students have completed their doctorate under her guidance.