Land-use and Land-cover Change Detection Using multi temporal data

DWDM PROJECT REPORT

by

Veena Nagar and Rounak Modi

Roll Number(s): 171CO151,
171CO236

Department of Computer Science and Engineering

National Institute Of Technology Karnataka, Surathkal (575025),
Mangalore, Karnataka

November 2020

Abstract

One of the primary uses of land in India is Agriculture. Information about the crop calendar is essential for the proper management of agriculture. Remote sensing sensors provide a regular, consistent, and reliable measurement of vegetation response at various crop growth stages. Vegetation indices are used effectively in crop monitoring and characterizing vegetation.

In the present study, a methodology is developed to show the vegetation density change, and extracts some key elements (such as date of peak vegetative stage, Greenery rate) . This paper also demonstrates the use of multi-temporal data to extract key phenological information calendar by interpreting temporal variations in spatial domain to generate crop cycle parameters(sowing and harvesting months and month of peak vegetative stage, Greenery rate). In this methodology the mean of difference images has been taken to detect the vegetation change. This observed mean data series has a considerable amount of local fluctuation in the time domain and needs to be smoothed to enhance dominant seasonal behavior. This smoothening is done based on a Savitzky-Golay filter. Based on the smoothen graph's temporal analysis, it is possible to extract the greenery rate(change in vegetation density), the number of crop cycles per year and their crop calendar.

This is illustrated by analysing multi-temporal data series of one agricultural year (from January 2017 to December 2018). Such an analysis is very useful for analysing dynamics of kharif and rabi crops.

Table of Contents

List of Figures

1. Introduction

1.1 Brief statement of the problem

Land use refers to the purpose the land serves, such as recreation, wildlife habitat, and agriculture. It is not the surface cover on the ground. And land cover change is a dynamic process taking place on the biophysical surfaces that have taken place over time, and space is of enormous importance in natural resource studies.

One of the primary uses of land in India is Agriculture. This paper has considered agricultural land use and found the land cover change over months to extract crop cycle parameters such as the number of crop cycles and greenery rate(i.e., rate of growth of vegetation in an area).

We can simulate crop growth cycles and generate reliable yields using the dynamic assessment of crop conditions. The crop calendar is formed by the chronological sequence of different physiological stages of a crop in its growth cycle.

The use of remote sensing data for estimating crop area has reached an operational level. But the dynamics of agriculture require monitoring the site repetitively in a year to map the crop regions, which calls for a time series analysis of satellite images.

This time series analysis can be done using moderate resolution satellite imagery with high revisit such as MODIS (Moderate-Resolution Image Spectroradiometer) data.

Satellite images are used to compute Vegetation indices, We are using these vegetation indices to detect vegetation change and as a result to detect the land-cover change using the multi temporal data.

1.2 Importance of the problem

One of the most critical sectors for India is Agriculture. Proper planning for this sector requires reliable and relevant information. India’s Planning commission recognizes the crucial role of remote sensing technology in generating quality crop statistics. To draw a development plan, we should first have an inventory of the resource.

Crop inventory includes knowing the area under a crop and its place, time, and quantity. Crop is affected by many biotic and abiotic factors, Since it is a very sensitive biological system. We can simulate crop growth cycles and generate reliable yields using the dynamic assessment of crop conditions.

1.3 Related Literature

Factors such as atmospheric noise conceal the actual variation in the land surface and thus hamper time series information extraction. It has already been reported in [1] a way that minimises the negative effects of noise. Firstly with the use of rule-based strategy all the key temporal points from the NDVI time series are identified. This segments the time series. Secondly the observed data points are fitted with a weighted double-logistic function, this emphasizes cloud free points. Finally the 3000 test points from the selected Sentinel-2 tiles are evaluated using the proposed method and a qualitative and quantitative comparison with the generally used Savitzky-Golay and harmonic analysis of time series is done. In other related work [2] the Normalized Difference Vegetation Index (NDVI) is used to provide a new dimension for describing vegetation growth cycle. This NDVI analysis at regular time intervals provides crop cycle stage information but this has some fluctuations and needs to be smoothed so that dominant season behaviour is enhanced. After successfully smoothening the NDVI series we can determine the crop cycles, various crop cycle parameters and the crop calendar. The algorithm proposed in this work considers the troughs and peaks in the graph and the trough-peak-trough are considered as a valid crop cycle if it satisfies both of the following conditions : (i) duration is greater than 60 days and (ii) NDVI value is greater than 0.3. The resultant graph obtained from this algorithm is smoothened using Gaussian Filter. It is known from the [3] that support vector machines (SVM) and artificial bee colony algorithms give better results compared to other traditional classification techniques.

Here the land cover changes on 2 study areas in the coastal area of Dakshina Kannada district in the state of Karnataka, South India are studied. The study was conducted during the years 2004-2008. In these 2 study areas, classification was carried out using artificial bee colony algorithms and support vector machine algorithms. These classification algorithms helped in finding the OCA values for the 2 study areas for the year 2004 and 2008. The results obtained from post-classification of study area 1 indicated that urbanisation in the study area 1 has almost increased twice.

1.4 Scope of the project

Final result of this project should detect the change in vegetation density(Greenery rate) and the number of crop cycles in two years(From January 2017 to December 2018), which as a result will detect the Land-use and Land-cover change if the changing parameter is considered as vegetation of the given area.

Materials and Methods

Overview of the proposed method

The main steps for the proposed algorithm are shown in the flowchart (Figure 3).

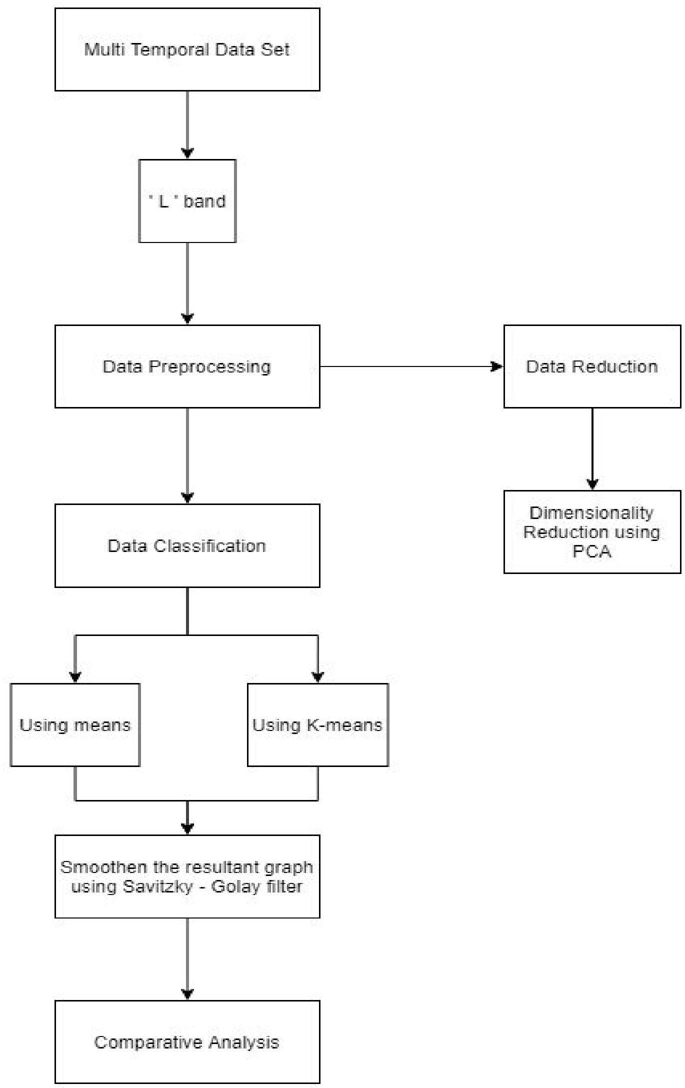

Fig 1.a. Flowchart of the methodology designed to extract the results

Overview of the K-means Clustering method

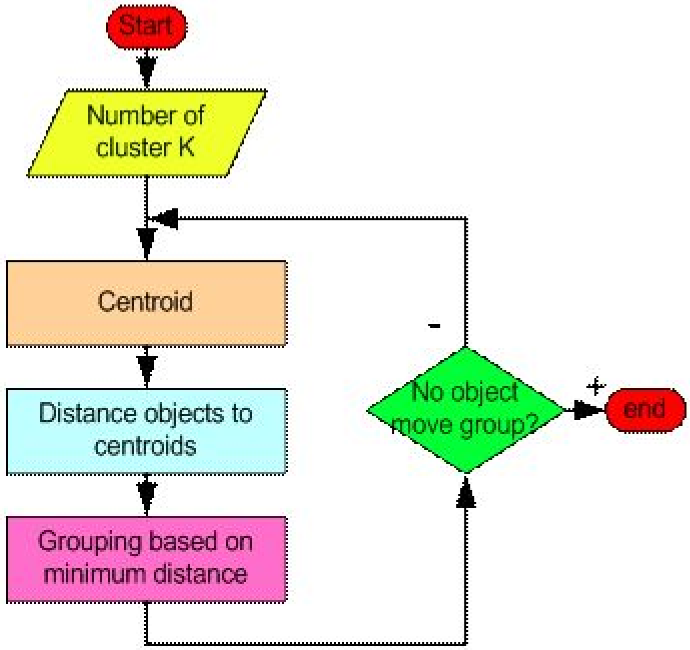

Fig 1.b. Step by step K-mean clustering algorithm (copyright from https://people.revoledu.com/kardi/tutorial/kMean/Algorithm.htm)

Description of the Proposed Method

Our proposed methodology is based on the fact that the satellite dataset belonged to a single band Fig 1. and hence this dataset was not helpful for calculating the NDVI values as NDVI values need red and infrared band images. In the proposed work we have used multi temporal difference images, these difference images will show considerable change for land covered under vegetation while the change observed for buildings, rocks, water bodies, etc will be negligibly small. Taking January as a sowing month we have found difference images Fig 3 and then performed data classification.

Identification of bands

The identification of bands is an essential step to check the number of bands in the images of the data-set. Fig 2 shows that the images in the data-set are single-band(L-band) images. Single band images in the data-set eliminated all other methods such as band averaging and NDVI calculation methods to estimating crop cycle parameters.

Assumption : January is a sowing period.

```
print("Multiband image", im1x)
```

```
Multiband image ('L',)
Multiband image ('L',)
```

Fig 2. Bands Identification

Data Preprocessing

The data preprocessing is a fundamental step to transform the raw data in a useful and efficient format. Data reduction is a part of data preprocessing, firstly we have detected the change for all images with respect to the first image. And all difference images with respect to the first image from the given data-set have been generated. The difference image shown in Fig 3 was the difference of the image_2 and image_1 of January.

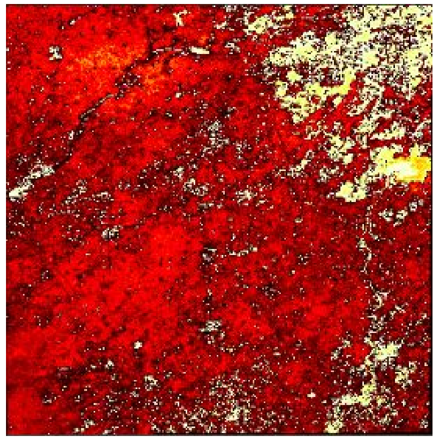

Fig 3. diff_image in hot scale

Then the reduction of the difference image is done by performing Principal Component Analysis (PCA) dimensionality reduction. PCA is used for linear dimensionality reduction of the data to project it to a lower dimensional space. The following 2 PCA methods were performed :

1. Fit_transform(X[,y] :- to apply the dimensionality reduction on X by fitting the model with X Fig 4.a.
2. inverse_transform(X) :- to transform data back to its original space Fig 4.b.

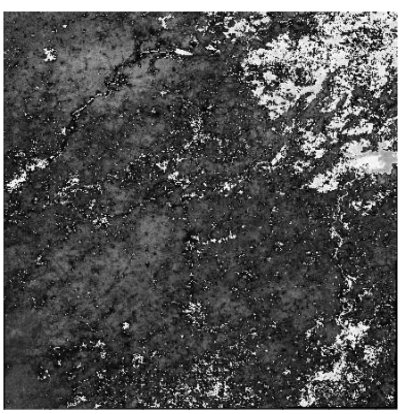

Fig 4.a. Jan_2017-After fit_transform

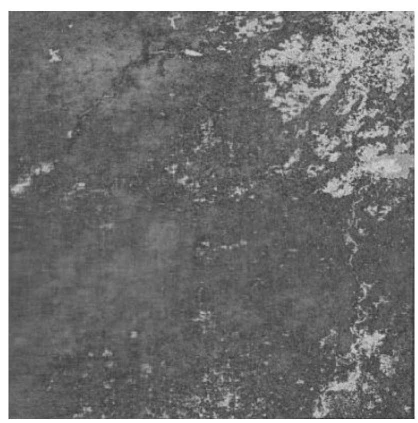

Fig 4.b. Jan_2017-After inverse_transform

Data Classification

1. Using Mean(Average) : In this method the crop cycle parameters are estimated by directly taking the mean of preprocessed data. Fig 5 shows the graph between the mean values of the preprocessed data in the span of 2 years.

```
In [9]: plt.plot(means)
Out[9]: [<matplotlib.lines.Line2D at 0x7f4b1aa398e0>]
```

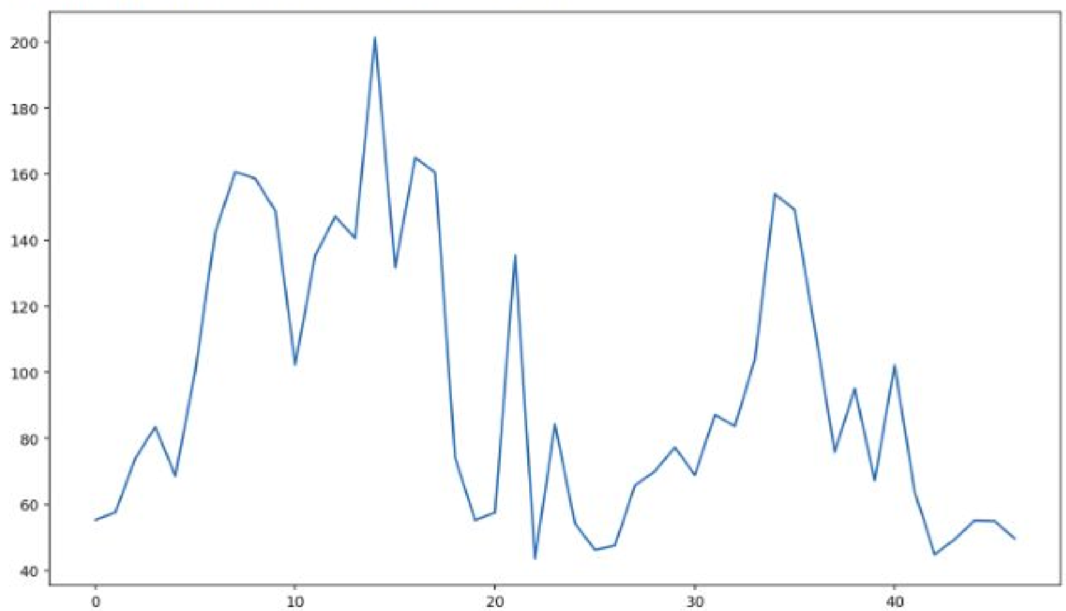

Fig 5. Vegetation density graph using normal Mean method

2. Using K-Means Clustering: In this method, Firstly the pre-processed images have been clustered using KMeans clustering algorithm then the mean values of those clustered images has been calculated. Fig 6 shows the graph between the calculated mean values in the span of 2 years.

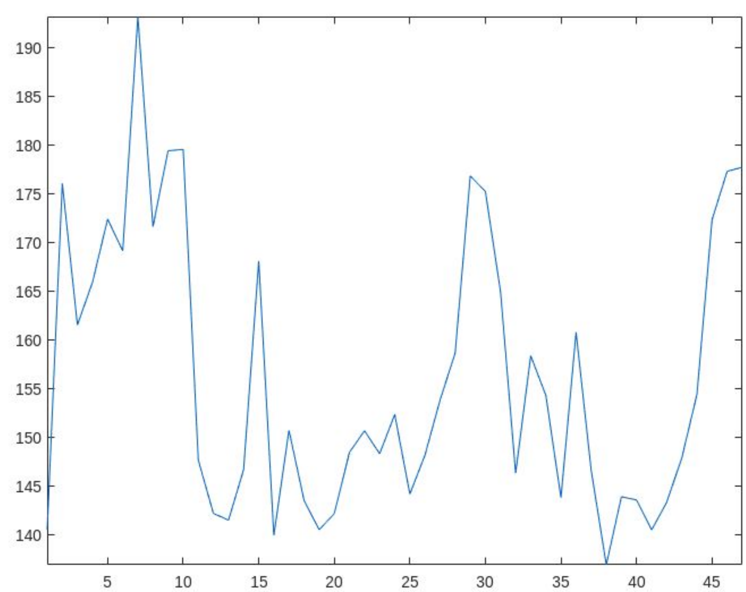

Fig 6. Vegetation density graph using KMeans clustering Method

Applying Savitzky-golay Filter

This smoothen graph has increased our performance by giving us as the approximated extrema of the graph. We have used the MATLAB extrema function to find the local maxima and minima of the graph.

1. For Mean(Average) :

   Fig 7 is the obtained smoothen graph after applying a savitzky-golay filter in Fig.5

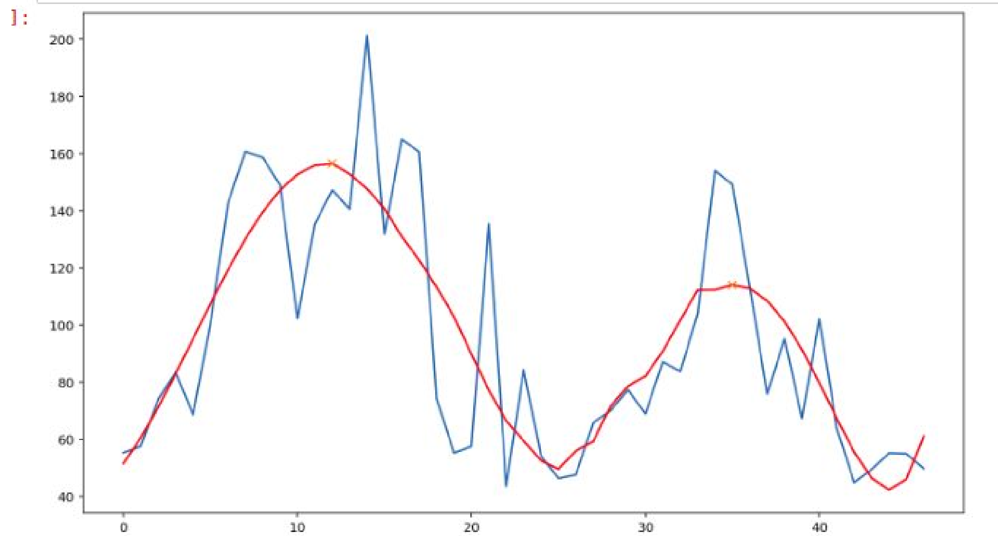

Fig.7 Smoothen Vegetation density graph Using Normal Mean Method

2. For K-Means Clustering :

   Red colored graph shown in Fig 9 is the obtained smoothen graph after applying savitzky-golay filter in Fig 6.

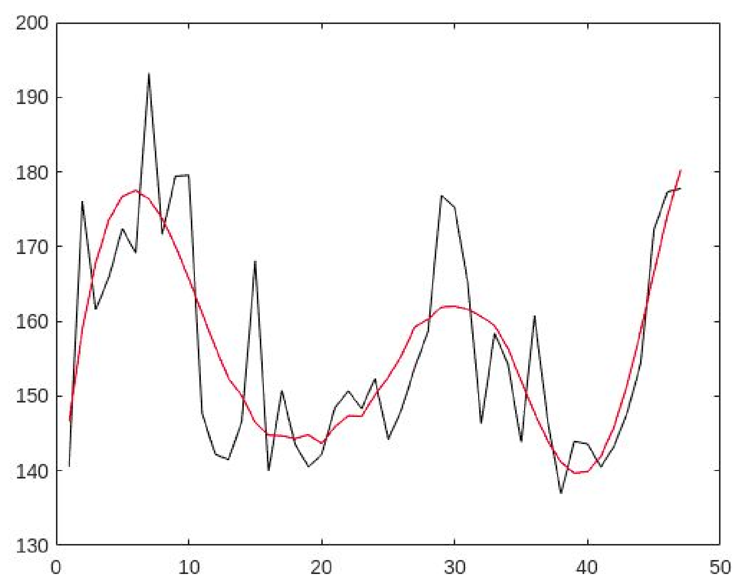

Fig.8 Smoothen Vegetation density graph using KMeans clustering Method

Results And Discussions

a. First method used in our methodology is the normal mean method where we have taken the mean of our pre-processed data. And plotted a vegetation density graph with the observed mean values.

b. Second method used is K-mean clustering, which is used to partition the every given image of data-set k clusters in which each observation belongs to the cluster with the nearest mean (i.e., cluster centroid), serving as a prototype of the cluster.
From the analysis, we can say that
   - if the value of k is a big number, then the result will shift towards the normal mean methodology.
   - And if the value of k is very small then the result obtained using k-mean clustering will have the most variance from the result obtained using normal mean method but the runtime of the program will simultaneously increase.

   We have analysed at different values of k and obtained the most appropriate change can be detected at k=10 and which is why it is chosen for performing the experiment.
   This method will cluster the related points to its centroid (center point). This will cluster each water body, building, and vegetation into different clusters.

c. This observed mean data series using both methodologies have considerable amount of local fluctuation in the time domain and needs to be smoothed to enhance dominant seasonal behavior. This smoothening is done based on a Savitzky-Golay filter.

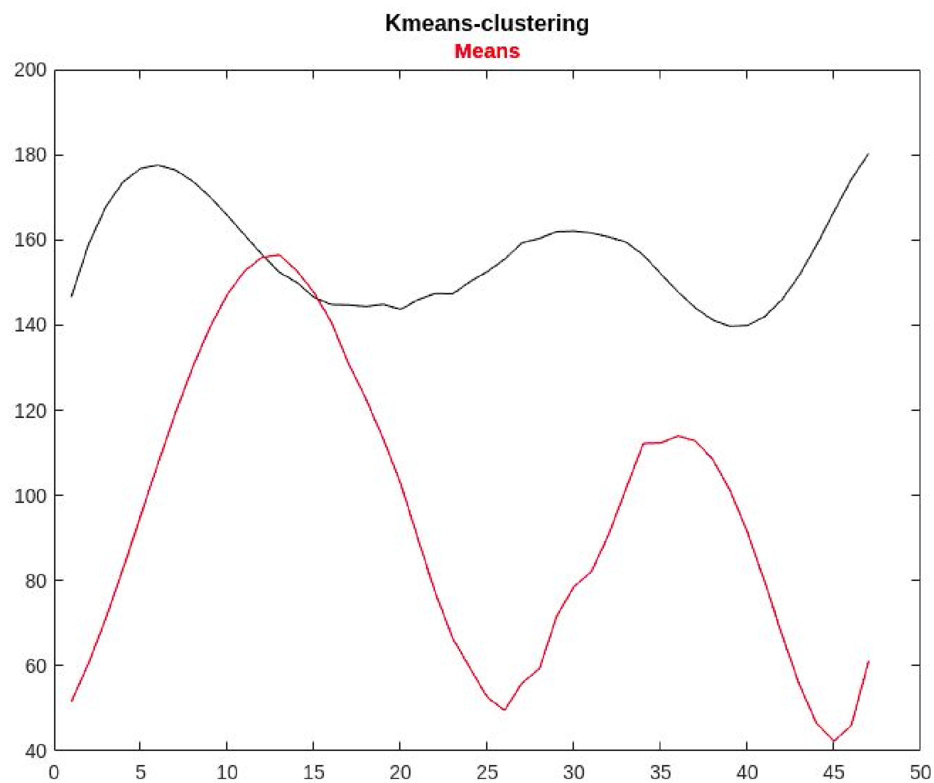

Fig.9. Comparison of both the resultant graphs shown in Fig7 and Fig8 .

Based on the smoothen graph's temporal analysis, We can detect the changes in land cover using vegetation density as a changing parameter in two years.
The greenery rate obtained from the methods tells us about the change of vegetation density of an area in two years. And the crop cycle duration obtained in the result referred to the number of months after which it will repeat the process of plantation(sowing and harvesting).

Estimated crop cycle parameters as follows:

a. crop_cycle_duration = (2ndlocalmax+1)/2-(1stlocalmax+1)/2
b. Greenary_rate = (yhat1(1stlocalmax) - yhat1(1)) / (1stlocalmax - 0)
c. We can also say that the local maxima in the vegetation density graph has the highest vegetation density and thus are the harvesting months of the year.
d. Similarly, the local minima in the vegetation density graph has the least vegetation density and thus are the sowing months of the year.

The crop cycle duration obtained using both methods are 12 months.
Whereas, Greenery rate obtained from Normal Means Method is 8.7462.
Greenery rate obtained from KMeans Clustering Method is 5.1529.

As shown in Fig.9. there would be a significant amount difference in the vegetation density of a Land-cover at harvesting and sowing months of the year. Which shows the land-use and land-cover change over a span of two years.

Conclusion

Land cover change can be used to monitor the changes on the land over a period of time. We have used land cover change to study the change in the vegetation over a duration of 2 years and extract the crop calendar features. Means and K-means (with k=10) algorithms were used for classification of data. The resultant of means and K-means was smoothed and the local troughs and peaks were identified. The crop cycle obtained from both of the algorithms was found to be of 12 months approximately but the greenery rate obtained from means was 8.7462 and from K-means was 5.1529.
In the study demonstrated the work is done in studying the crop calendar parameters that have occurred over a duration of 2 years and not on forecasting or predicting the crop calendar parameters. Crop calendar parameter forecasting can be done using different predictive analytics models. Also different data classification algorithms can be used to study the crop cycle parameters and a comparison from the results obtained by means and k-means can be done.

References

1. Yang, Yingpin, Jiancheng Luo, Qiting Huang, Wei Wu, and Yingwei Sun., "Weighted double-logistic function fitting method for reconstructing the high-quality sentinel-2 NDVI time series data set." *Remote Sensing* 11, no. 20 (2019): 2342.
2. Patel, Jayesh H., and Markand P. Oza. "Deriving crop calendar using NDVI time-series." *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 40, no. 8 (2014): 869.
3. Jayanth, J., T. Ashok Kumar, Shivaprakash Koliwad, and Sri Krishnashastry. "Identification of land cover changes in the coastal area of Dakshina Kannada district, South India during the year 2004–2008." *The Egyptian Journal of Remote Sensing and Space Science* 19, no. 1 (2016): 73-93.

4. How to make K-means CLustering work. [K-Mean Clustering Tutorial: Algorithm]