CO416

Data Warehousing and Data Mining

Research Paper Implementation - Presentation

Topic :

P23. Forecasting Land-use And Land-cover Change In Dakshina Kannada District, Karnataka

By:

Veena Nagar : 171CO151

Rounak Modi : 171CO236

Content

Introduction

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).

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.

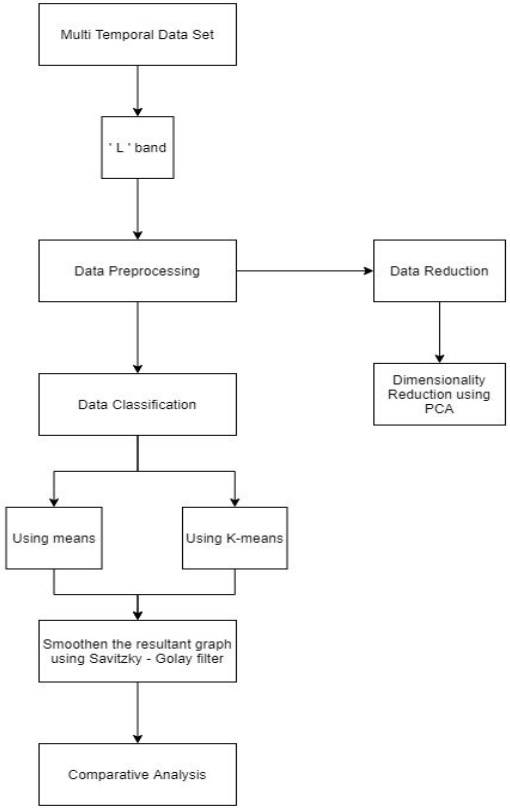
Introduction

Brief of our work : In this 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) .In this methodology, we have taken the mean of difference images to detect the vegetation change.Since this observed mean data series has a considerable amount of local fluctuation in the time domain, we have performed smoothening using Savitzky-Golay filter to enhance dominant seasonal behavior. Based on the temporal analysis of smoothen graph, We have extracted the greenery rate(change in vegetation density), the number of crop cycles per year and their crop calendar.

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

All steps of the methodology are shown in the following flowchart.



Flowchart of the methodology designed to extract the results

Description of Proposed Method

- The proposed methodology is based on the fact that the satellite dataset belonged to a single band.
- We have used multi temporal difference images.
- The difference images will show considerable change for land covered under vegetation.
- The change observed for buildings, rocks, water bodies, etc will be negligibly small.

Identification Of Bands

- It is useful to think of each pixel as having one value per band.
- **PIL.Image.Image.getbands()** used this to find the number of bands in our dataset image.
- Returns a tuple containing the name of each band in this image.

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

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

Bands Identification

Data Preprocessing

Multi temporal Data Set

For detecting the land-change over span of months we have taken a data set containing 48 .tif images. The images given are of the following form :

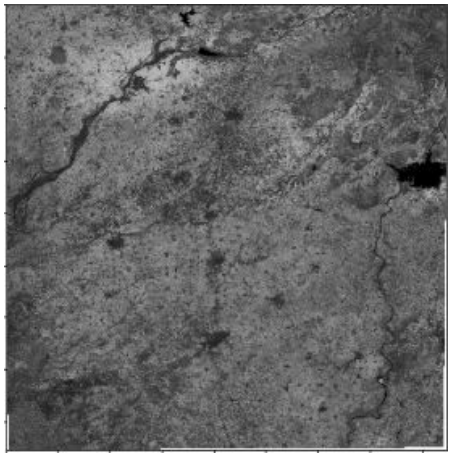


Fig.1 tif image of Jan_2017.

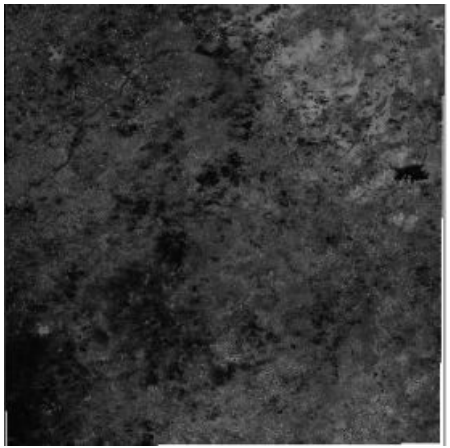
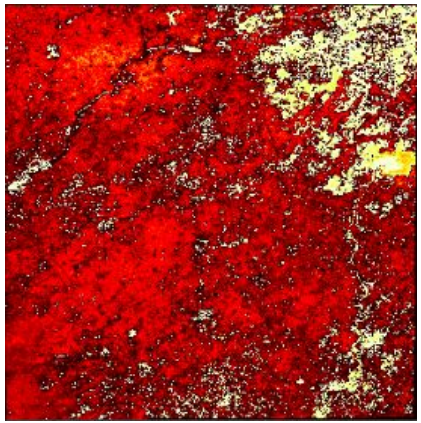


Fig.2 Tif image of Jan_2017 after interval of 15 days.

Data Preprocessing(contin...)

Generated the consecutive difference images from the given data-set. The difference image shown below was the difference of the image_2 and image_1 of January.

Formula Used: diff_image1 = jan_2017_2.tif - jan_2017_1.tif

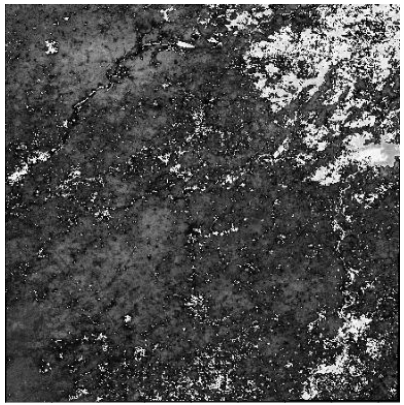


diff_image in hot scale

Data Preprocessing(contin...)

Performed 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.



Jan_2017-After fit_transform

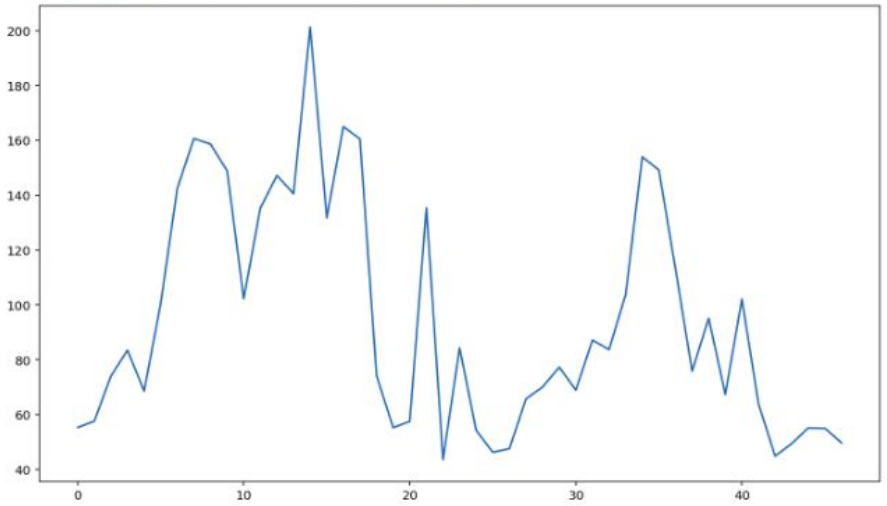
Jan_2017-After inverse_transform

Data Classification

1. **Using Mean(Average) :**

In this method the crop cycle parameters are estimated by directly taking the mean of preprocessed data.
The figure 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>]
```

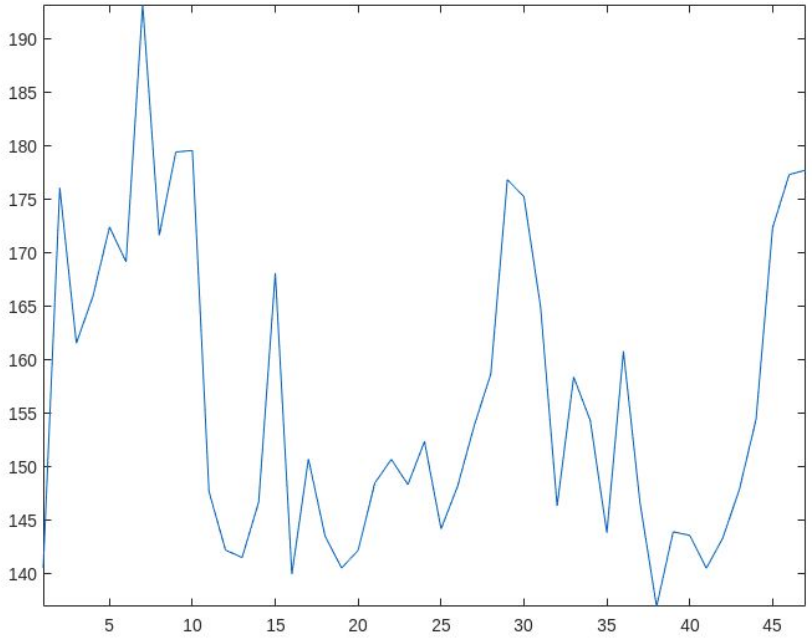


Vegetation density graph using normal Mean method

Data Classification

2. Using K-Means Clustering:

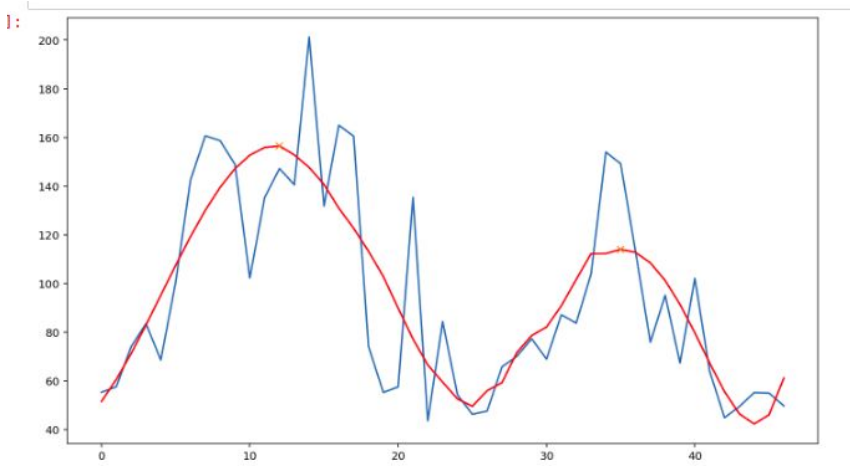
In this method, Firstly the pre-processed images have been clustered using K-Means clustering algorithm then the mean values of those clustered images has been calculated. Fig 6 shows the graph between the calculated mean values in the span of 2 years.



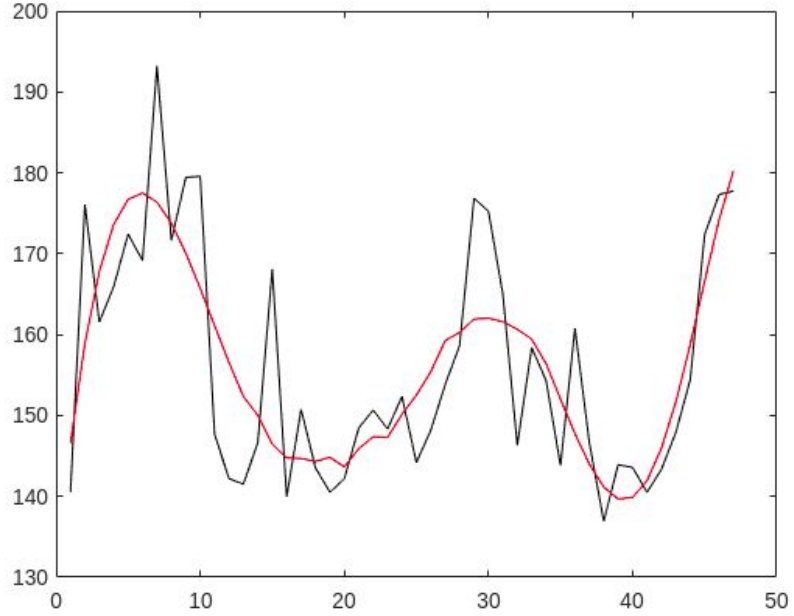
Vegetation density graph using K-Mean clustering method

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.



Smoothen Vegetation density graph Using Normal Mean Method



Smoothen Vegetation density graph using KMeans clustering Method

Results and Discussions

a. First method used in our methodology is the normal mean method.
b. Second method used is K-mean clustering, from the analysis for k-mean clustering, we can say that :
    i. if the value of k is a big number, then the result will shift towards the normal mean methodology.
    ii. 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.
c. Based on the temporal analysis of smoothen graph, we detected the changes in land cover using vegetation density as a changing parameter in two years.

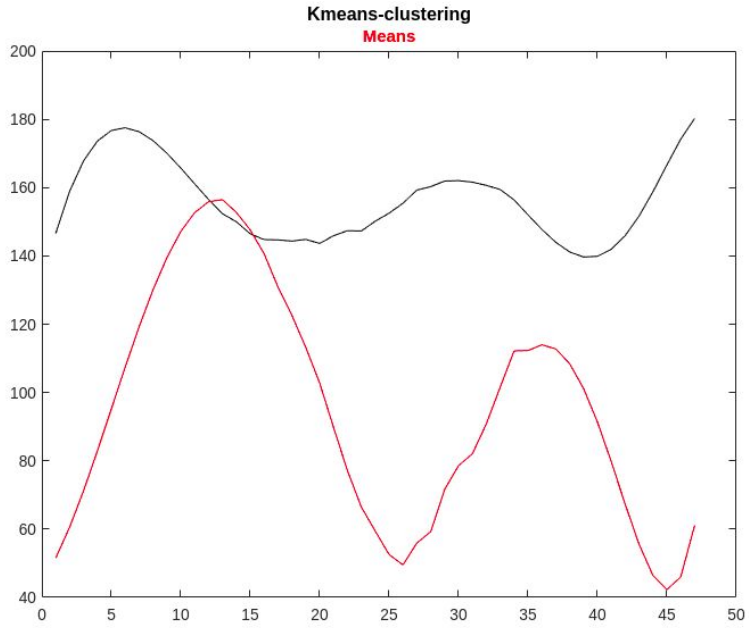
Results and Discussions(conti..)

d. Our Estimated crop cycle parameters were as follows:

1. crop cycle duration
2. Greenary rate
3. Local maxima - highest vegetation density and harvesting months of the year.
4. Local minima - least vegetation density and the sowing months of the year.

e.

- 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.



Comparison of both the resultant graphs means and K means

Conclusions

1. Since from the results obtained we can see a significant amount of difference in the vegetation density of a Land-cover at harvesting and sowing months of the year. This shows the land-use and land-cover change over a span of two years.
2. 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.

Thank You