

Python Data Analysis with Google Earth Sentinel Imagery for the Area of Sundarbans West Bengal, India

Sumana Chatterjee*

Ph.D. Scholar (Computer Science), Nirwan University, Jaipur, Rajasthan, India.

*Corresponding Author: sumana.spssaha.chatterjee@gmail.com

Citation: Chatterjee, S. (2025). Python Data Analysis with Google Earth Sentinel Imagery for the Area of Sundarbans West Bengal, India. Journal of Commerce, Economics & Computer Science, 11(04), 23–28. <https://doi.org/10.62823/jcecs/11.04.8175>

Abstract

This paper is based on data analysis by python programming language on Google collaborator platform, data analysis for Google-Earth sentinel imagery, for the area of 'SUNDARBANS', situated in West Bengal, India, the world's largest mangrove forest, the UNESCO heritage property. The file for python code of analysis is linked with the 'Google Earth Engine', data pulled from this interface, using the coordinates of the area of interest, here which is the area of 'SUNDARBANS', the delta. In Google Earth Engine, there is to have a log-in credential and during the execution of this python programming to fetch data from this interface, it was required to create some project in 'Google Earth Engine', with a newly created project id, which was to incorporate in the programming file to initialize and authenticate the execution and also for pulling of data from that interface. In this way by 'Google Earth' initialization and providing project id, enabled with authentication from Google id, obtained the imagery data for sentinel-1 for visualization of VH(vertical-horizontal) backscatter and VV (vertical-vertical) backscatter on folium map for the area of interest. Then obtained result of NDVI value from sentinel-2 imagery along with visualization of NDVI value for a particular date range, incorporated in the program. VV is reflected wave received vertically to the radar sensor and VH is reflected wave received horizontally to the radar sensor. High VV back scatter indicates rough surface on earth such as structures etc and low VV back scatter indicates smooth surfaces like water. NDVI value, B4, B5 surface reflectance band are related for understanding vegetation on the earth surface. VV, VH values as well as visualization of backscatter time series graph and also NDVI values and visualization of NDVI help to understand changes in different geo-physical properties such as surface roughness, soil moisture, vegetation and other. From this type of data analysis obtaining result as well as visualization regarding change of such parameters, can provide valuable information to monitor the environment, natural resources, impact on human activity on earth's surface.

Keywords: Python, AI, Google Collaborator, Google Earth, VV, VH, Backscatter, NDVI, Time Series.

Introduction

Google earth engine is a platform for geospatial data from google. There are several data in this global interface including the Sentinel data . In this paper analysis has been done with the Sentinel 1 and Sentinel 2 data. The code snippet was extracted from 'Kaggle', the site of data science company where the syntax and several codes were actually imported snippets from google developer site . For this paper, application of the programme has been done with the area of interest Sundarbans, West Bengal, India. Sundarbans is the UNESCO, Heritage property, existing between the Ganges, Brahmaputra delta region.

The Sundarbans mangrove ecosystem covers approximately 10000 square kilometre with 40% in India and 60% in Bangladesh. This is a source of vital blue carbon storage system, land of biodiversity, the forest acts as a natural guard against cyclones and storm surges, protects millions of coastal inhabitants. Also the mangrove forest can protect the boundary of land with its special type of dense root system, protecting land from the sea water, sea water can not encroach into the land for mangroves, soil erosion is protected by this valuable system. So, the land, soil and vegetation of Sundarbans are considered as some unique, asset property acting as barrier for natural disaster and calamity. In this paper the analysis is subjected to the main focus area, latitude 21.9° N and longitude 89.2° E and the data extraction polygon under analysis was extended from $21^{\circ} 43' 2''$ – $21^{\circ} 55' 5''$ N latitude and between $88^{\circ} 42' 0''$ – $89^{\circ} 04' 0''$ E longitude. The objective is to execute the analysis of sentinel data obtained from google earth engine to understand the status of image, with the help of visualization technique of python code. Instead of manual measurement, the analysis of satellite imagery had been done to collect the imagery automatically from cloud. Sentinel 1 data and Sentinel 2 data from google earth engine and in addition with that land sat-8 from USGS earth explorer had been used for this research study. As Sundarbans is UNESCO property having great geographical significance, so to preserve and restore all its geo-physical entity should be our aim towards the goal of sustainable environment.

Literature Review

The analysis of data in this study is based on analysis in python on google collaborator platform, with the sentinel 1 and sentinel 2 satellite imagery, pulled from google earth engine interface and land sat-8, satellite imagery, pulled from USGS google earth explorer. The source of python code and preliminary idea of such analysis was 'Kaggle', the application was done with different location and the analysis and code was appended and modified accordingly as per the compatibility of execution with recent time period. Besides this source, different research papers as mentioned later herewith as reference part, had been studied to acquire the knowledge behind such type of data analysis to have valuable insights. Knowledge and idea about how the research study with satellite image data for any type of change in earth surface such as land erosion, shoreline change, insights of NDVI and vegetation density, NDWI or water salinity, effect of cyclone on NDVI, Flood mapping, land use/land cover study can be done, all these ideas were gained from Kaggle, google earth engine interface, USGS google earth explorer as well as research papers related to this particular subject. In this paper, analysis had been considered for some years, 2019 to 2025 (till August 2025 as available till 20.8.2025), the time series trend of VV, VH back scatter values as well as output of other analysis as obtained there, helped to gain valuable insights regarding change of various geo-physical properties of earth surface, occurred if any, all of which output and result helped to take decisions towards suitable measure to restore sustainable environment of the study area Sundarbans.

Research Gap

In this research study, the data analysis had been done with sentinel data of satellite imagery pulled from google earth engine using python code. Two types of sentinel imagery were pulled using python codes to fetch different folium maps, gee maps on google collaborator platform. These two types are sentinel 1 and sentinel 2. Other source of satellite imagery data in addition to these to enrich this research study was land sat data -8 from USGS google earth explorer. The other similar research papers with the analysis of satellite imagery related to sentinel imagery from google earth engine or USGS earth explorer had been studied also but in this present research study satellite image data had been obtained only by python script from both interface simultaneously i.e. from google earth engine and google earth explorer exclusively by python code snippet. The analysis was with the study of the change of vegetation and change of various geo-physical properties on earth surface and also study with after-effect of heavy rainfall or cyclone or natural calamity like storm surge as well. In this research no manual effort was given to collect data and to get result, only the effort of automatic data pulling by python code snippet from GIS interface had been taken under consideration. For analysis of effect of cyclone, in this research study, the cyclone 'REMAL', May 2024, has been taken under consideration.

Research Questions /Hypothesis

For this research study, execution was done with python script, on google collaborator platform. The code was applied accordingly to pull the required imagery data for analysis and visualization of satellite imagery as obtained by python code from google earth engine and google earth explorer. The objective was to analyse the plot of visualization to understand the change, the present picture of land

portion, shore-line, salinity of water and also vegetation of Sundarbans, whether any significant changes have been occurred on account of natural disaster like heavy rainfall, cyclones, storm surge etc. To understand the change, two time periods have been considered for analysis of change, one time span of early period and another for later period than that, depending upon the data available in google cloud. For all cases as later period, till present period was taken under consideration except shore line change where the early period taken was 1990 and the later period was 2020. In this manner gee map as well as folium map provided with geo coordinate, were obtained for visualization of change of NDVI for change of mangrove health, NDWI for change of salinity, flood mapping shore line change for a certain period and also for understanding of land use/land cover. This was the main objective and research question related to that was collection of visual plot necessary to study, whether any significant changes occurred or not.

Methods

The whole analysis, for the purpose of compatible execution, had been done in three python files created on google collaborator platform. The script of the file was to fetch the satellite imagery from google earth engine and from USGS earth explorer as well. The code of analysis was similar in pattern to some python file as obtained from 'Kaggle', but the area of interest and analysis was different, in this case the area of Sundarbans was subjected to analysis. In the python file, first step was importing necessary libraries and packages for all types of execution and uploading of visual plot for satellite imageries. The necessary libraries were folium for interactive maps, libraries for authentication and initialization, also importing of matplotlib, seaborn for visualization, lpython for interactive python code, numpy for numerical calculation of python and also importing Branca, which is directly linked with folium map used for the satellite imagery obtained from google earth engine. In this paper, the main focus area of consideration was with coordinate 21.9° N, 89.2° E. Using the GeoJSON file, the data extraction polygon was between $21^{\circ} 43' - 21^{\circ} 55'$ N latitude and $88^{\circ} 42' - 89^{\circ} 04'$ E longitude. Before execution of the python file, it was required to authenticate and initialize by google id, after which the code snippet from google earth engine could be executed. In google earth engine there was required to create a project with a new generated project id which was to copy in the python code to authenticate and initialize the python script. Then with the particular coordinate, fetched the imagery data as polarization available there. In this case the whole process was completed. On this basis, vertical-vertical (VV) and vertical-horizontal (VH) sentinel images were obtained, number of collected images were 389. Thus VV backscatter and VH backscatter on map of folium could be obtained. VV and VH backscatter values tell us how much radar energy is returned from the surface. Different surfaces on earth interact with radar in different ways depending on roughness, moisture and structure. VV is when the radar energy is reflected back vertically and VH is when the radar energy is reflected back horizontally. Very low VV and VH often indicates very smooth surface, that means water bodies. For soil /bare land, VV low to moderate and VH, very low. For rough or moist soil, VV increases significantly because moisture improves conductivity. VH is still having weak value for this type of surface. For soil moisture detection, generally VV is used. For vegetation or forest or crops, VV with moderate values, VH much stronger, due to volume scattering inside canopy. For dense forest, VH is close to VV. For crops VH rises during growth of crops and drops after harvest. VH/VV ratio is a strong vegetation indicator. For urban structures, high rise, higher values of VV and VH. For urban structures, VV high values due to double bounce of scattering between walls and grounds, VH lower values than VV. After obtaining VV and VH backscatter, the sentinel 2 imagery from google earth engine was obtained using python code snippet. Total 982 data were obtained for NDVI analysis from sentinel 2 imageries. Both for sentinel 1 and sentinel 2 imageries, the satellite imageries were obtained from Copernicus interface under ESA (European space agency). For the NDVI calculation $NDVI = (NIR - Red) / (NIR + Red)$, 'NIR' is reflected images of 'near infra red' images and 'Red' is reflected 'Red images'. For this purpose, the B4 and B8 bands had been collected for red images and for near infra red images respectively. Along with calculation of NDVI images, the folium map of NDVI image for understanding of vegetation pattern was also obtained. Other than sentinel imageries, the land sat imageries, with B4 and B5 bands from USGS google earth explorer, were also obtained to get information about NDVI frequency. The coefficient of back scatters for different time period shown as time series graph was also plotted to understand the variation of earth surface as mentioned earlier. Changes in the backscatter values can indicate changes in the properties of surfaces, vegetation, soil moistures and other geo-physical properties. So from the backscatter time series, can have the valuable idea about changes of environment, natural resources etc. Based on NDVI categories, plot the pie chart

with different categorization of NDVI values such as healthy vegetation (0.2%), barren or sparse vegetation 39.6%, dry vegetation(60.2%). Plot the pie chart with separate type of category on earth surface, such as water surface 22.4%, land surface 1.3%,dry vegetation 1.2%, vegetation 75.0% . After this step, to understand deforestation changes, LANDSAT -8 images were collected from USGS google earth explorer by python snippet with filtering date from 2015 to 2023. Then filtering data for recent period, August 2025, determined the median NDVI values with normalized differences between B4, B5 bands to monitor mangrove health and similarly determined the median NDWI values with normalized differences between B3,B8 bands to monitor salinity of water, both by gee map. Filtering data for 1.1.2020 to 31.8.2025 determined median VV to understand flood mapping area for 2020 to 2025. Filtering two date range 1990 and 2020, taking median value with respect to land sat 8 image data with bands SR_B2,SR_B3,SR_B4, extracted gee map of shore line change between these time period. Extracted gee map for land use/land cover. Lastly taking median value of NDVI for before and after period of cyclone REMAL, may 2024, before period as April 2024 and after period as June 2024, taking normalized difference of B4 and B8 bands as mapped function of NDVI extracted gee map of effect of cyclone REMAL on NDVI change.

Significance of the Study

The google earth engine satellite imagery analysis for the area of interest Sundarbans is most significant as Sundarbans is world heritage UNESCO property, so any change of this heritage property should be monitored very well. There is two types of methods to monitor the changes, either by manual measurements, field based surveys of the mangrove environment or by online process of fetching satellite imageries from google earth engine and or from USGS google earth explorer . In this research paper, the automatic data pulling from google earth explorer could be done by python code snippet .Instead of manual survey the code snippet helped to have valuable insights to monitor various geo-physical properties of the area of interest. The valuable insights about the deforestation of mangrove, were obtained from the sentinel-1 and sentinel 2 imageries extracted from the Copernicus (ESA)HUB .In addition to sentinel imagery insights, the land sat -8 imageries, were also obtained from the USGS google earth explorer which also could help to gain idea about NDVI frequencies. Moreover soil and salinity map with NDWI values, rainfall analysis, global surface water monitoring could help to monitor all the changes to take necessary steps in favour of sustainable environment. Sundarbans is like a natural guard against storm surge, cyclones, so to monitor land erosion, damage of vegetation, increase in salinity all are very essential by the process of gaining insights through the study of satellite images as stated here.

Timeline

To get the insights about various geo-physical properties of Sundarbans, west Bengal, India, mainly the initial part was to prepare the code snippet to enable successful execution for analysis, compatible with google collaborator platform, related with collection of suitable images of satellite imageries from google earth engine and USGS google earth explorer . Sentinel 1 and sentinel 2 images were obtained from the first one and land sat -8 data were obtained from earth explorer, the second one. The analysis was done in three python files . The first one was to obtain image of sentinel 1 and 2 from Copernicus (ESA) HUB to get VV, VH backscatter time series, land sat-8 imageries for NDVI analysis, imageries pulled from google earth explorer, pie plot on the basis of various types of vegetation and on the basis of categorization on earth surface,such as vegetation, land, water etc. The second and third python file was used to understand NDVI (mangrove health), NDWI (water salinity),Flood mapping(sentinel-1 SAR),shoreline change, Land-use/land cover(LULC), effect of cyclon Remal, span 24.5.2024 to 28.5.2024, all based on extraction of gee map with the help of python code snippet .The total time to fetch the satellite data, successful execution and analysis with visual plots, was approximately two months. No manual field survey was applied for the whole process. The screenshots of images of output as obtained from several analysis are given herewith.

Conclusion and Future Work

From this study it was to predict the present situation as well as status of change for different geo-physical properties of earth surface. In future like this area of interest, analysis from google earth engine and land sat -8 imageries, for other area of interest also can be done similarly to gain insights of that area of interest for the purpose of taking necessary decisions towards environmental sustainability.

Sumana Chatterjee: Python Data Analysis with Google Earth Sentinel Imagery for the Area of....

Image of Models and few Screenshots of Plots as Obtained

VV backscatter (SENTINEL 1 from Copernicus)

VH backscatter (SENTINEL2 from Copernicus)

NDVI MAP (1.1.2019 to 20.8.2025)

NDVI frequency (1.1.2019-20.8.2025) (Landsat-8 USGS earth explorer)

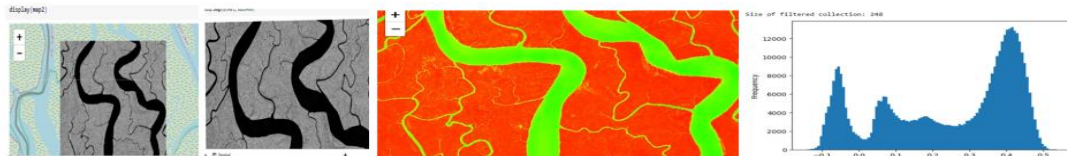


Figure 1: Folium map of sentinel data and NDVI frequency of Landsat-8 data

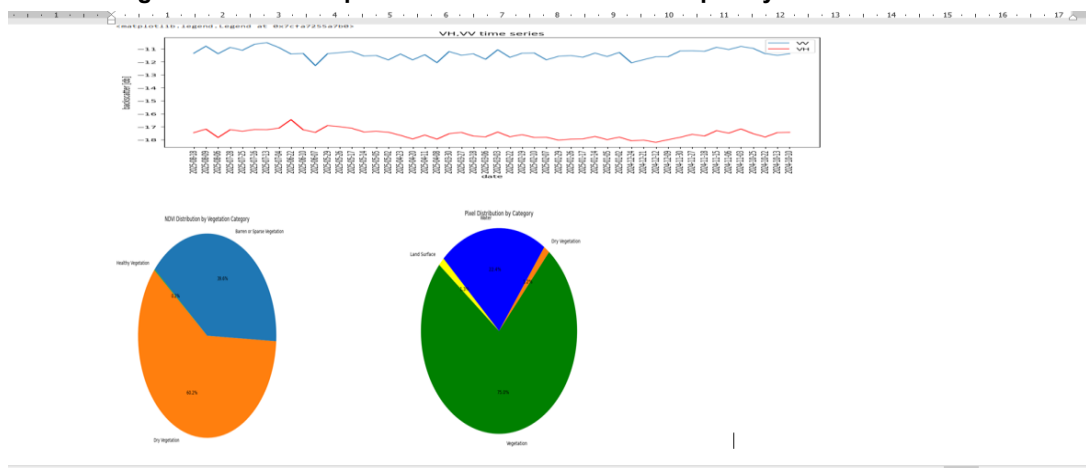


Figure 2: VH,VV TIME SERIES just below . Below that, left image is for NDVI vegetation category and on the right, category of different portions on earth surface (land, water, dry vegetation, vegetation)

NDVI GEE MAP(1.1.2025-31.8.2025)

NDWI,SALINITY GEE MAP(1.1.2025-31.8.2025)

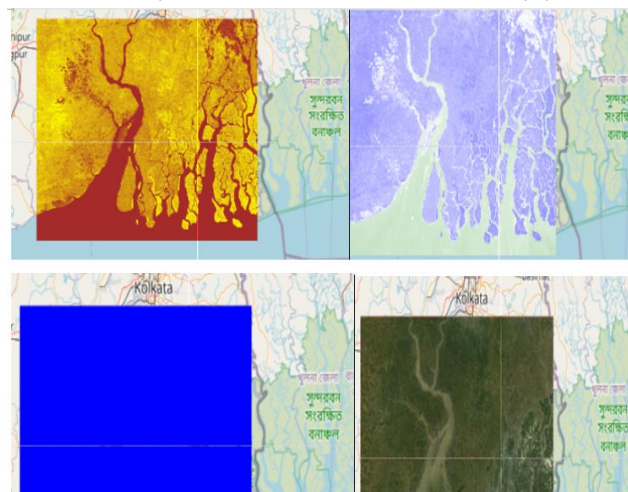


Figure 3: Datewise various gee map

FLOOD MAPPING(1.1.2020-31.8.2025) SHORELINE CHANGE(1990-2020)

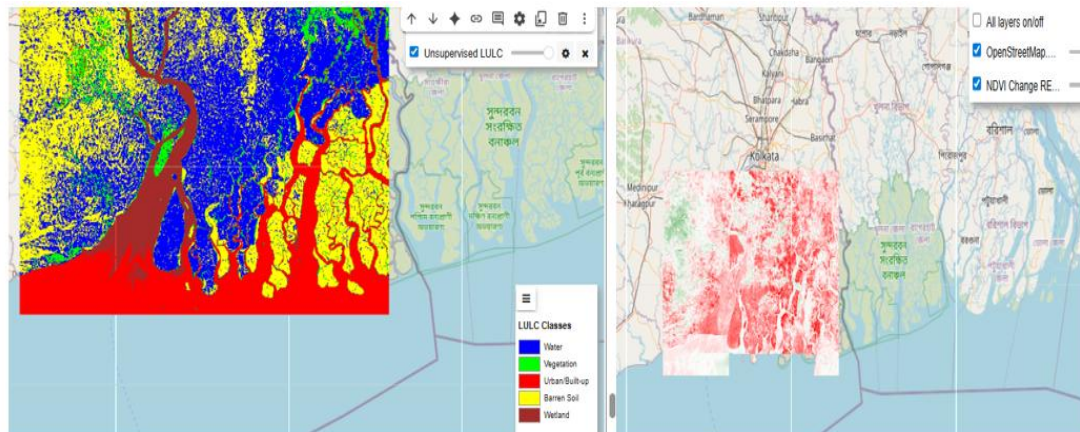


Figure 4: (Left) RECENT LAND USE, LAND COVER (Sundarbans) (Right) EFFECT OF REMAL, 2024

References

1. B. H. (2024). Google Earth Engine and Sentinel 1/2 data-based forest degradation. *ScienceDirect*, 9, 1-10.
2. Karsch, G. (2023, 3). Annual Mangrove Vegetation Cover Changes (2014–2020) in Indian Sundarbans National Park Using Landsat 8 and Google Earth Engine. *Frontiers in Wetland Ecology and Environmental Sustainability*, 1-9.
3. Hosaka, M. S. (2020). Assessing cyclone disturbances (1988–2016) in the Sundarbans mangrove forests using Landsat and Google Earth Engine. *Springer Nature*, 102, 133-150.
4. Santra, M. (2024). Quantifying shoreline dynamics in the Indian Sundarban delta with Google Earth Engine (GEE)-based automatic extraction approach. *Springer Nature*, 65, 426-442.
5. Hossain, K. A. (2024). Land cover change across 45 years in the world's largest mangrove forest (Sundarbans): the contribution of remote sensing in forest monitoring. *EUROPEAN JOURNAL OF REMOTE SENSING*, 57(1), 1-17.

