Assignment 1: Report

Sajja Patel - 2021101107, Aishani Pandey - 2022121009

26 September 2024

The Kaveri River Basin, spanning parts of Karnataka and Tamil Nadu in southern India, is one of the most significant river systems in the region. The river originates in the Western Ghats and flows eastward into the Bay of Bengal. The basin supports diverse ecosystems and plays a crucial role in agriculture, drinking water supply, and hydroelectric power generation for millions of people. However, the basin is also prone to varying climatic conditions, leading to periods of drought that can severely impact agricultural productivity and water availability.

1 Drought characterisation

1.1 Introduction to Drought

A drought is a prolonged period of abnormally low precipitation that leads to a shortage of water, affecting ecosystems, agriculture, and human activities. Droughts are characterized by several key features:

- **Meteorological Drought**: Occurs due to a sustained lack of rainfall compared to normal conditions.
- **Agricultural Drought**: Affects crop production and soil moisture, leading to food shortages.
- **Hydrological Drought**: Involves reduced water levels in rivers, lakes, and groundwater.
- **Socioeconomic Drought**: Impacts the economy and society, often leading to water restrictions, food scarcity, and migration.

The following are factors by which droughts are charcterised-

- **Intensity**: The degree of water deficiency relative to the normal climate conditions of a region, typically measured using drought indices such as the *Standardized Precipitation Index (SPI)* or *Standardized Precipitation Evapotranspiration Index (SPEI)*. These indices quantify deviations from normal precipitation and evapotranspiration conditions over different time scales, indicating how severe a drought is.
- **Frequency**: The number of drought occurrences over a specific period, often expressed as the return period (e.g., a drought that occurs once every 10 years). It indicates how often a drought of a certain intensity is likely to happen in a given region.
- **Areal Extent**: The geographical scope or spatial coverage of the drought, often expressed as the total area affected by drought conditions. This characteristic helps determine how widespread the impact is on a local, regional, or national scale.
- **Severity & Duration**: The cumulative impact of drought, often defined as the product of intensity and duration. It reflects the overall effect of the drought on water availability, ecosystems, and human activities, incorporating both the deficit in water resources and the length of time that the deficit persists.

1.2 Introduction to SPI

The Standardized Precipitation Index (SPI) is a widely used statistical measure that quantifies precipitation deficits over varying time scales. It helps assess the severity of drought conditions and is a useful tool for monitoring climate variability.
The SPI is calculated using the following steps:

1. **Calculate the Precipitation Total**: Determine the total precipitation for the desired time scale.
2. **Fit a Probability Distribution**: Fit a probability distribution (often the Gamma distribution) to the precipitation data for the time scale of interest.
3. **Compute the Cumulative Probability**: Calculate the cumulative probability of the observed precipitation.
4. **Standardize**: Use the cumulative probability to derive the SPI value.

SPI values are standardized to have a mean of zero and a standard deviation of one for a specific location and time period. This allows for comparisons across different regions and time periods.

SPI	Drought category
$SPI \geq 2.00$	Extreme wet
$2.00 > SPI \geq 1.50$	Very wet
$1.50 > SPI \geq 1.00$	Moderate wet
$1.00 > SPI \geq -1.00$	Normal
$-1.00 \geq SPI > -1.50$	Moderate drought
$-1.50 \geq SPI > -2.00$	Severe drought
$-2.00 \geq SPI$	Extreme drought

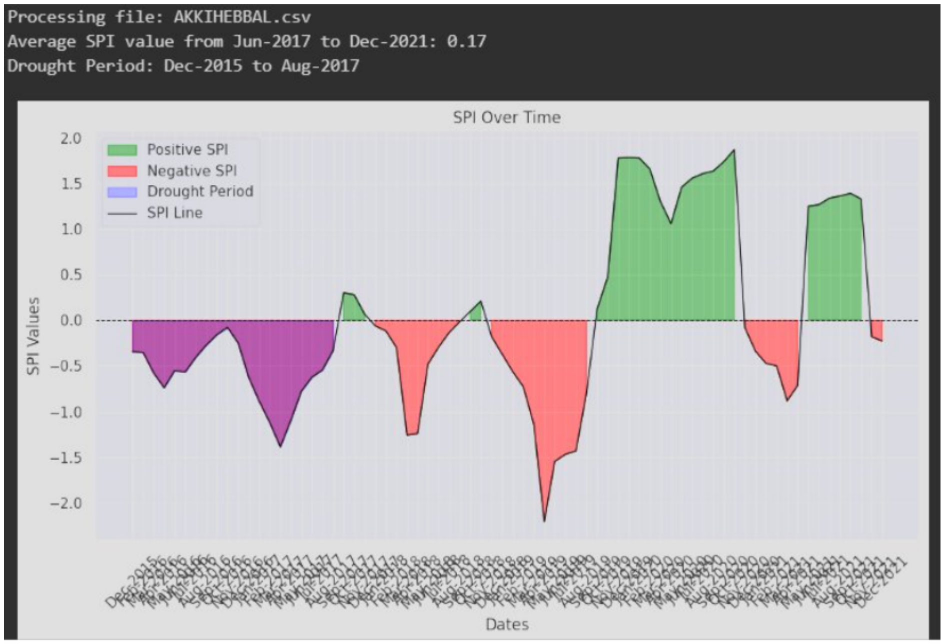
1.3 Data Collection & Pre-processing

Uniform Rainfall Data: We will gather uniform rainfall data across the Kaveri Basin from multiple meteorological stations from this link. The data will be formatted in CSV files, containing columns for "Dates" and "ACTUAL (mm)" of rainfall to maintain uniformity across years and stations. This uniformity is crucial for accurate SPI calculation and analysis.

1.4 Data Analysis & visualisation

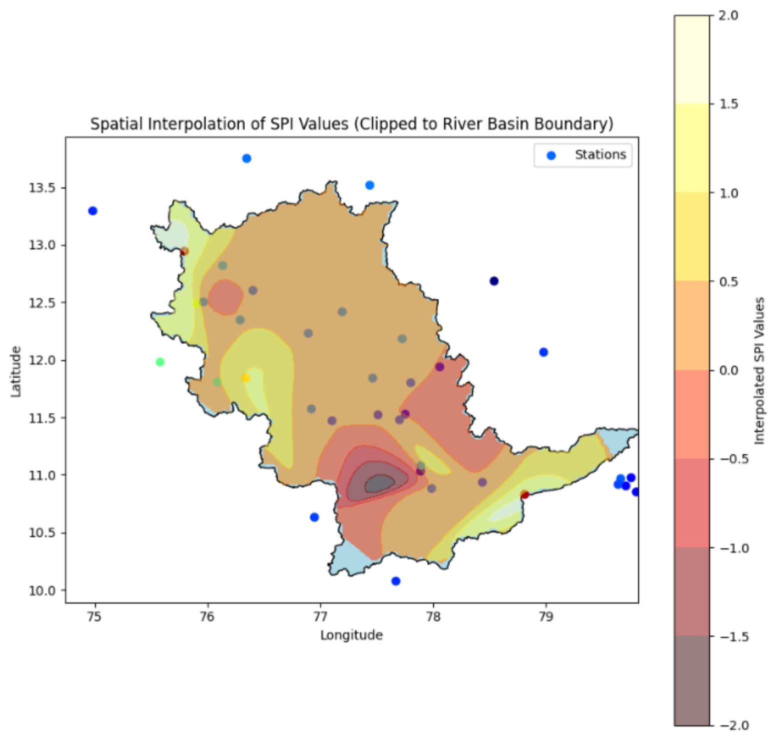
1.4.1 SPI over time for all stations

This plot will show the temporal variations in SPI across different locations in a user specified time frame. Within this, for drought identification we have used the longest period of negative SPI of every station.



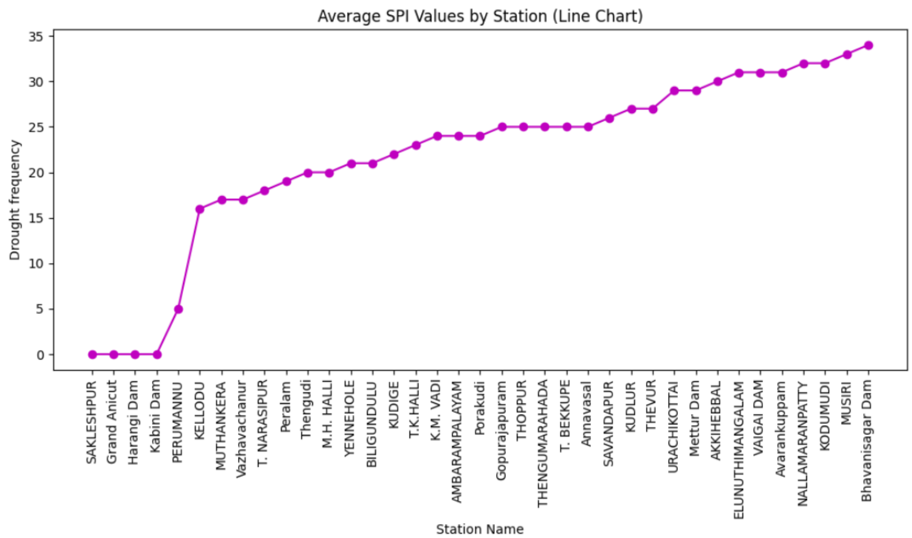
1.4.2 Contour Plot of SPI Over Time

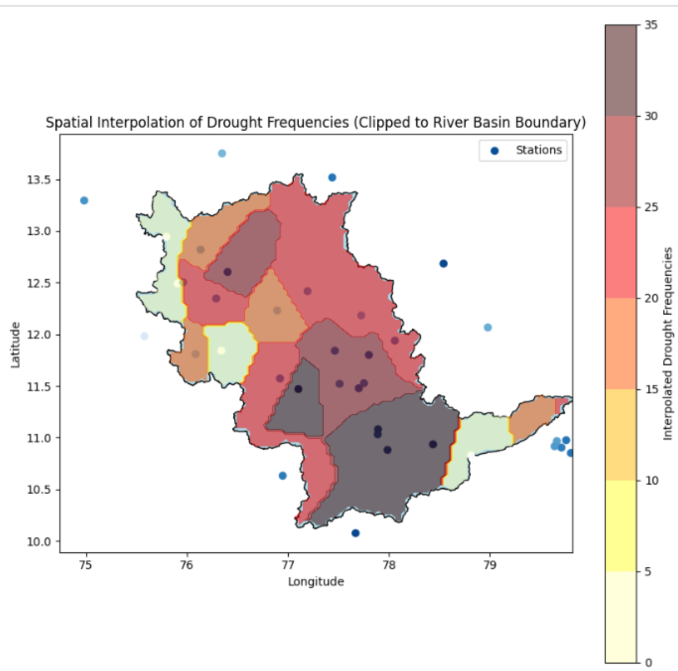
By visualizing SPI values over time, we can illustrate the spatial distribution of drought conditions across the basin. In this context, the griddata function from scipy.interpolate is being applied to interpolate the SPI values over a grid defined by longitude (X) and latitude (Y). The method='cubic' parameter specifies that cubic interpolation is used.



1.4.3 Drought Frequency Over Time

Drought frequency is the number of months with negative frequency in the user specified time frame. First, it plots the river basin's boundaries and marks the meteorological stations, with each station's color representing its drought frequency. The code then performs spatial interpolation using griddata to estimate drought frequencies over a grid of points across the river basin. The interpolated values are clipped to ensure they only cover areas within the basin's boundaries. Finally, the interpolated drought frequencies are displayed as a contour plot, providing a visual representation of how drought intensity varies spatially across the basin. This approach helps analyze areas affected by different drought frequencies in the river basin.



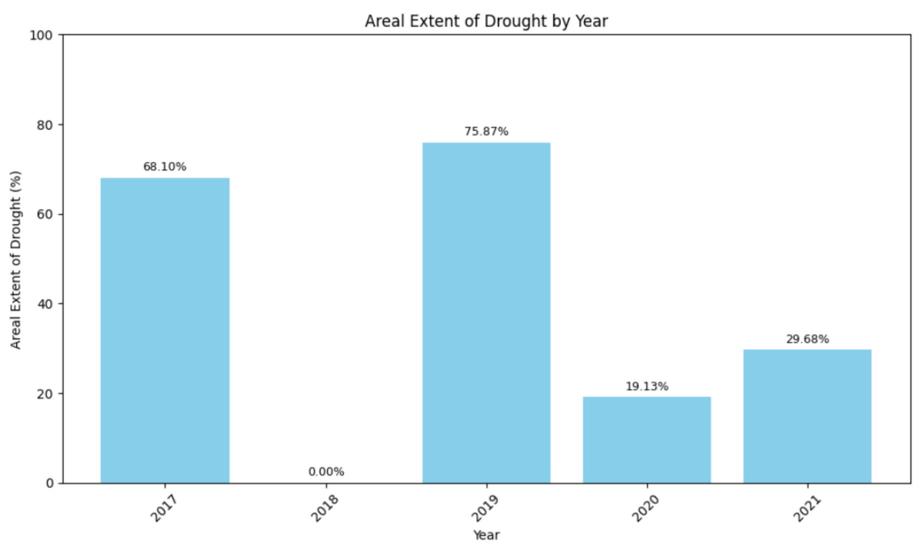


1.4.4 Areal extent

The code snippet calculates the percentage of the areal extent of drought within the clipped_SPI data array. It counts the total number of valid SPI values (excluding NaN values) and the number of those values that are negative (indicating drought conditions). Finally, it computes the percentage of the area affected by drought by dividing the count of negative values by the total count of valid values and multiplying by 100. In this case, the output indicates that approximately 18.35% of the area experiences drought conditions.

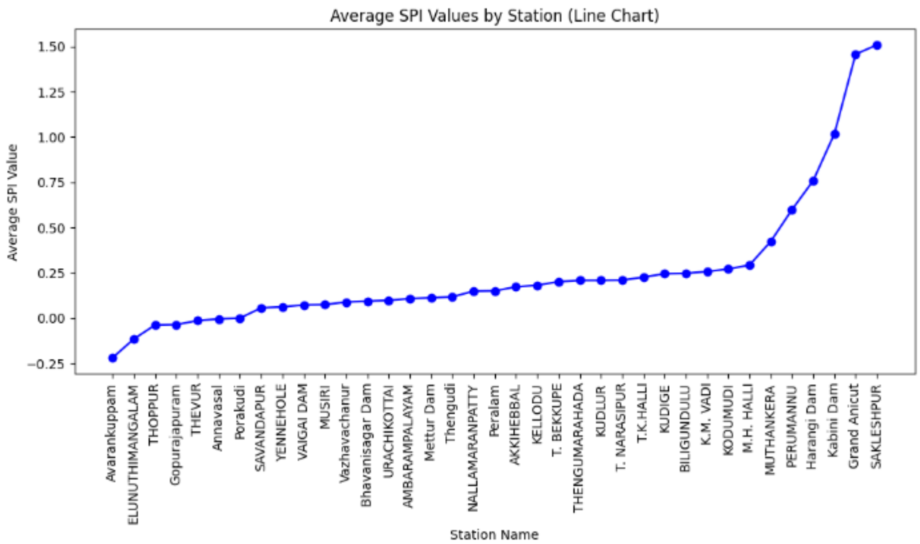
```python
cnt = 0
dro = 0
for i in range(clipped_SPI.shape[0]):
    for j in range(clipped_SPI.shape[1]):
        if (np.isnan(clipped_SPI[i][j])) == False:
            cnt += 1
            if(clipped_SPI[i][j] < 0):
                dro += 1
print("Percentage of areal extent of drought: ", (dro/cnt)*100)
```

```
[39]
... Percentage of areal extent of drought:  18.346870520783565
```

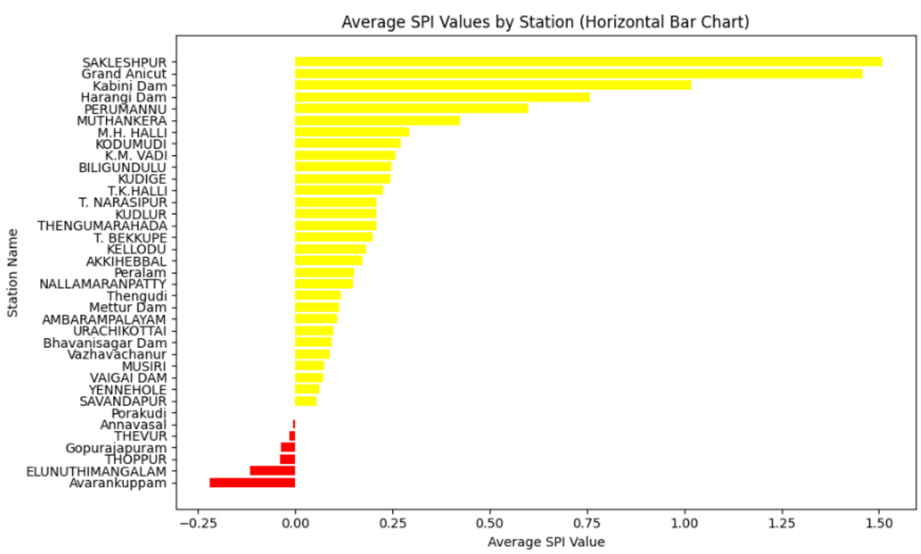


1.4.5 SPI values vs Stations

The code generates a line chart that visualizes the relationship between average SPI (Standardized Precipitation Index) values and various meteorological stations. It loads data from a CSV file, sorts the stations by their average SPI values, and then plots these values against the station names.



For better comparison, we have also made a horizontal bar chart which shows deviation on both sides.



1.5 Conclusion

In conclusion, the assessment of drought conditions in the Kaveri River Basin highlights the significance of understanding the various characteristics and impacts of droughts on the ecosystem and human activities. Through the application of the Standardized Precipitation Index (SPI), we have quantified and visualized drought intensity, frequency, and areal extent, providing valuable insights into the temporal and spatial dynamics of drought in the region. The findings emphasize the need for effective water management strategies and timely interventions to mitigate the adverse effects of drought, ensuring the sustainability of water resources for agriculture, drinking supply, and overall community resilience. By integrating data collection, analysis, and visualization, this study contributes to a deeper understanding of drought patterns and supports informed decision-making for the future.

2 Remote Sensing Analysis of NDVI for Drought Characterization

2.1 Introduction to NDVI and Its Relevance

The Normalized Difference Vegetation Index (NDVI) is a crucial metric for assessing vegetation health and density across landscapes. NDVI values range from -1 to 1, where higher values (closer to 1) indicate healthier vegetation, and lower values (closer to -1) signify barren land, stressed vegetation, or water bodies.

- Low NDVI: Below 0.3 (represents barren areas or unhealthy vegetation).
- Medium NDVI: Between 0.3 and 0.5 (representing stressed vegetation).
- High NDVI: Above 0.5 (healthy and dense vegetation).

For our analysis, NDVI is used to evaluate the impact of drought conditions over different time periods in the Kaveri Basin.

2.2 Data Collection and Processing

We utilized Landsat 8 imagery (Collection 2, Tier 1, Level 2) for NDVI calculation from 2017 to 2021. The following key steps were undertaken:

- **Cloud Masking**: To ensure accuracy, we applied a cloud mask to filter out pixels obscured by clouds. Before Cloud Masking, our shapefile looked like:

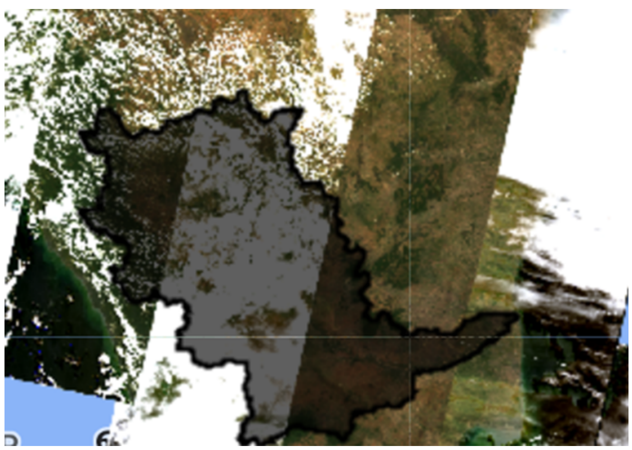


Figure 1: Initial shapefile

After cleaning the cloud cover, we got:

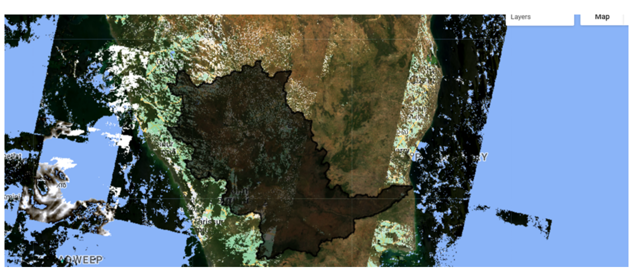


Figure 2: After cloud masking

- **Scaling Factors**: Landsat images require the application of scaling factors to convert raw pixel values to reflectance.

2.3 NDVI Calculation

For each year, NDVI was calculated using the standard formula:

$$\text{NDVI} = \frac{(NIR - Red)}{(NIR + Red)}$$

Where NIR corresponds to the near-infrared band and Red to the visible red band. This calculation was applied to each pixel in the images covering the Kaveri Basin.

2.4 NDVI Calculation for the Whole Basin (2021)

1. We began by calculating the NDVI for the entire Kaveri Basin for the year 2021. We specifically took for year 2021 as due to computational complexity, we could not calculate across 2017-2021 for the whole basin, and our code was crashing. Using Landsat 8 data, we computed both the mean and median NDVI values. This provided insights into the overall vegetation health across the basin.

    - **Mean NDVI (2021)**: This provided an overall quantitative measure of vegetation health in the basin, where higher values indicated healthy vegetation and lower values indicated stressed or barren land.
    - **Median NDVI (2021)**: The median value was calculated and visualized over the whole basin. This helped eliminate the effect of extreme NDVI values and provided a balanced understanding of overall vegetation health.

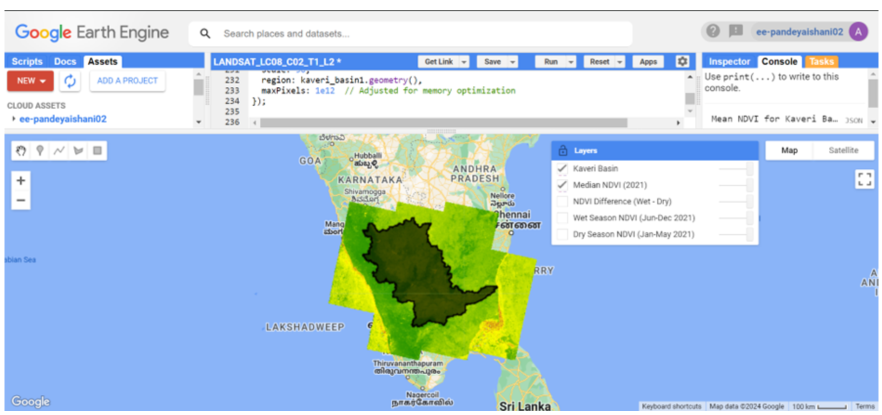


Figure 3: NDVI for the Kaveri Basin (2021)

    Mean NDVI for 2021 = 0.5170405744673834. This indicates **"Healthy"** vegetation in Kaveru basin

2. **Seasonal Analysis:** We also calculated the NDVI separately for the dry (Jan-May) and wet (Jun-Dec) seasons in 2021. The difference between these values provided valuable insights into the areas of vegetation growth or degradation.

    The NDVI was calculated separately for dry and wet seasons to observe vegetation changes between seasons:

    - **Dry Season (Jan-May)**: The dry season typically experiences lower NDVI values due to reduced rainfall. Lower NDVI values during this period indicated regions with reduced vegetation growth or stress due to limited rainfall.

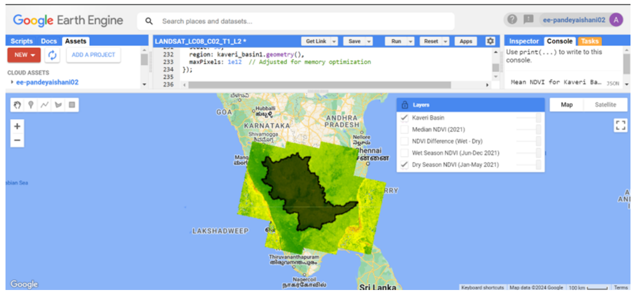


Figure 4: Dry Season Analysis - With Basin layer

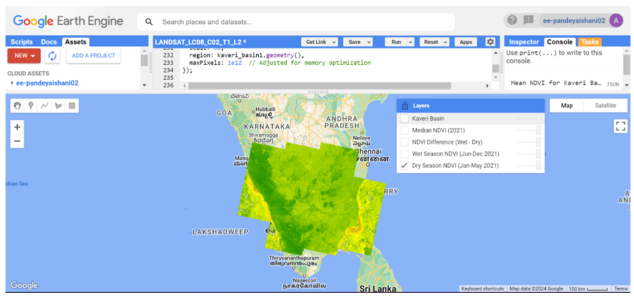


Figure 5: Dry Season Analysis - Without Basin layer

- **Wet Season (Jun-Dec)**: The wet season sees higher NDVI values due to increased water availability for vegetation. Higher NDVI values in the wet season reflected healthier vegetation due to abundant rainfall.

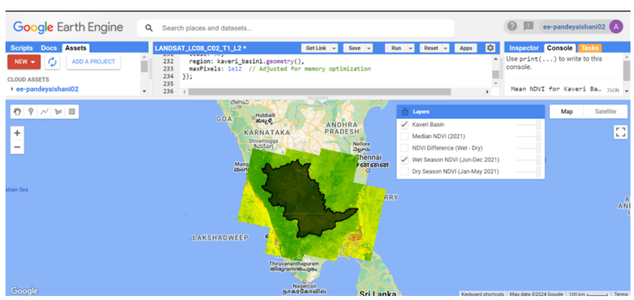


Figure 6: Wet Season Analysis - With Basin layer

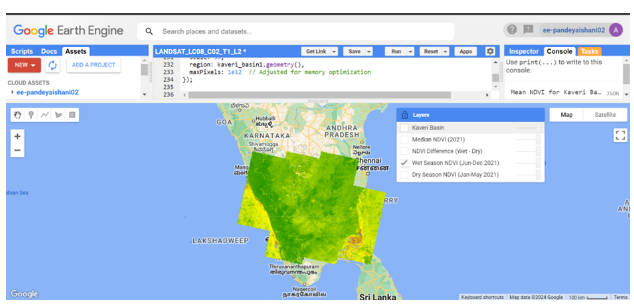


Figure 7: Wet Season Analysis- Without Basin layer

3. **Color Palette Interpretation:**
   - **Red/Brown:** Indicates vegetation stress, typically in drought-prone regions.
   - **Yellow:** Moderate vegetation health, some stress may be present.
   - **Green:** Healthy vegetation, indicating no significant stress.

2.5 NDVI Difference Analysis

To further understand the impact of seasonal drought, the difference between the wet and dry season NDVI values was calculated. This highlighted areas where vegetation either improved or worsened. The difference between the wet and dry season NDVI was calculated to highlight areas that saw improvement in vegetation (positive NDVI difference) or degradation (negative NDVI difference). Areas where vegetation worsened during the dry season were highlighted in red, and areas where it improved were highlighted in green.

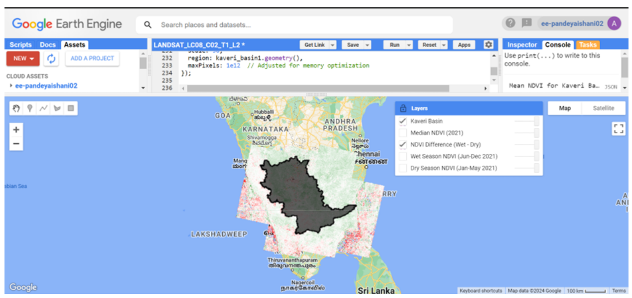


Figure 8: NDVI Difference (Wet - Dry) 2021

Analysis of NDVI Difference without the basin layer:

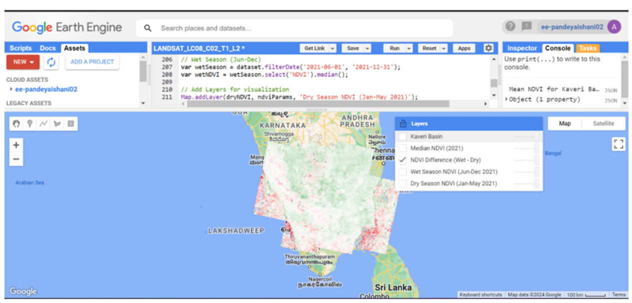


Figure 9: NDVI Difference (Wet - Dry) 2021 - without basin layer

The difference layer clearly shows the spatial variation, with red indicating a loss of vegetation (or worse conditions) and green indicating improvement. This provides insight into vegetation recovery or drought conditions between seasons.

2.6 NDVI Calculation for AOI (2019-2021)

1. After calculating for the whole basin, we focused on a smaller Area of Interest (AOI).

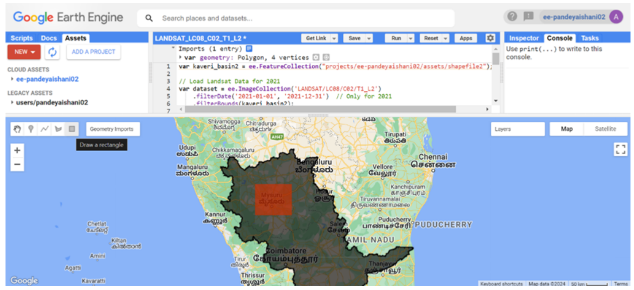


Figure 10: Area Of Interest

Initially, we tried calculating the NDVI from 2017 to 2021 and 2018 to 2021, but we obtained null values due to insufficient data, likely due to data gaps or insufficient coverage in the Landsat dataset for certain years.

Possible Causes:

- **No Data for the Time Range**: It's possible that the Landsat dataset for the selected AOI has missing data in certain areas.

- **Low Image Count**: Since we are working with NDVI data over multiple years (2017-2021), if there are not enough cloud-free images during that period within the AOI, the median or mean calculation may not be meaningful, leading to a null result.

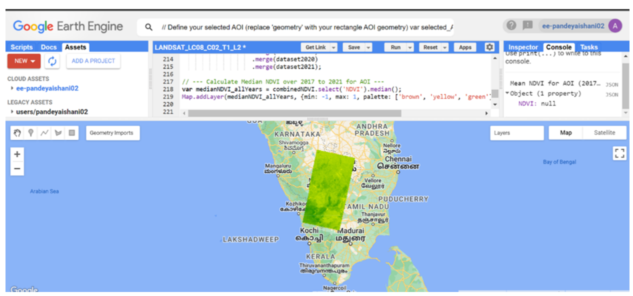


Figure 11: Mean and Median NDVI for 2017-2021 giving Null Value

2. Finally, we calculated NDVI successfully for 2019-2021 for both the AOI and the whole basin. The median NDVI was displayed on the map, while the mean NDVI was calculated and printed in the console.

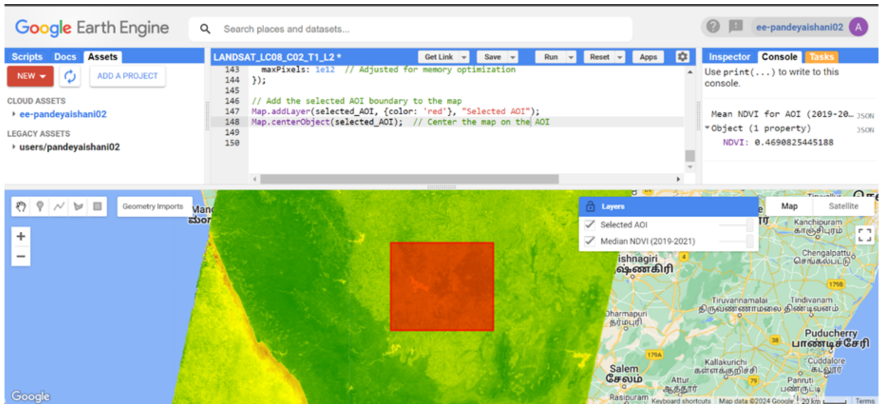


Figure 12: NDVI for AOI (2019-2021)

Mean NDVI Output: In the console, the mean NDVI for the selected AOI (from 2019-2021) is printed, with a value of **0.469082544518**8. This value reflects the **average vegetation health** in that region over the three years.

For 2019-2021, we even calculated NDVI for whole basin:

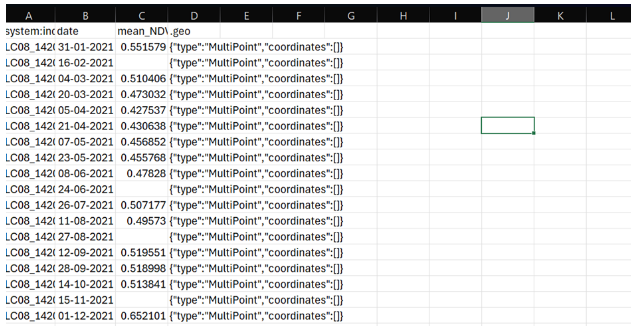
Figure 13: NDVI for Kaveri Basin (2019-2021)

We got the **Mean NDVI value = 0.501719440184364**1. This indicates **Healthy vegetation** during 2019-2021 across whole Kaveri Basin

3. **Color Palette Explanation:** The same color palette was applied to identify drought-prone areas, with red and brown highlighting areas with low vegetation health and green areas representing healthy vegetation.

2.7 NDVI Time Series

The NDVI time series was exported to a CSV file for each month - date-wise. This enabled temporal analysis of NDVI values over the study period, helping us track the progression of drought conditions and their effects on vegetation health.



A	B	C	D
system:inc	date	mean_NDV	.geo
LC08_142(	31-01-2021	0.551579	{"type":"MultiPoint","coordinates":[]}
LC08_142(	16-02-2021		{"type":"MultiPoint","coordinates":[]}
LC08_142(	04-03-2021	0.510406	{"type":"MultiPoint","coordinates":[]}
LC08_142(	20-03-2021	0.473032	{"type":"MultiPoint","coordinates":[]}
LC08_142(	05-04-2021	0.427537	{"type":"MultiPoint","coordinates":[]}
LC08_142(	21-04-2021	0.430638	{"type":"MultiPoint","coordinates":[]}
LC08_142(	07-05-2021	0.456852	{"type":"MultiPoint","coordinates":[]}
LC08_142(	23-05-2021	0.455768	{"type":"MultiPoint","coordinates":[]}
LC08_142(	08-06-2021	0.47828	{"type":"MultiPoint","coordinates":[]}
LC08_142(	24-06-2021		{"type":"MultiPoint","coordinates":[]}
LC08_142(	26-07-2021	0.507177	{"type":"MultiPoint","coordinates":[]}
LC08_142(	11-08-2021	0.49573	{"type":"MultiPoint","coordinates":[]}
LC08_142(	27-08-2021		{"type":"MultiPoint","coordinates":[]}
LC08_142(	12-09-2021	0.519551	{"type":"MultiPoint","coordinates":[]}
LC08_142(	28-09-2021	0.518998	{"type":"MultiPoint","coordinates":[]}
LC08_142(	14-10-2021	0.513841	{"type":"MultiPoint","coordinates":[]}
LC08_142(	15-11-2021		{"type":"MultiPoint","coordinates":[]}
LC08_142(	01-12-2021	0.652101	{"type":"MultiPoint","coordinates":[]}

Figure 14: NDVI Time Series

3 Drought Mapping & comparison

In this section, we provide a visual comparison between two key indicators of drought:

- A contour plot showing the spatial distribution of SPI values across the Kaveri Basin for the period 2017–2021.
- NDVI values in raster format, representing vegetation health, for the year 2021.

The SPI plot highlights areas affected by drought over a multi-year period, while the NDVI data offers insight into the basin's vegetation conditions during 2021. Due to computational constraints, we were unable to analyze NDVI trends over the same four-year period as SPI. However, by comparing these plots, we can clearly observe how areas with low SPI values, indicative of drought, align with regions of low NDVI, signifying stressed vegetation. This correlation emphasizes the close relationship between precipitation deficits and the health of ecosystems in the basin.

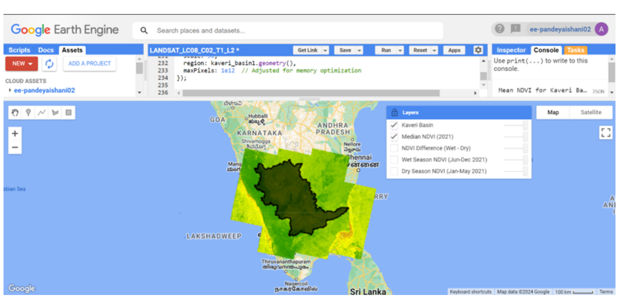


Figure 15: NDVI for the Kaveri Basin (2021)

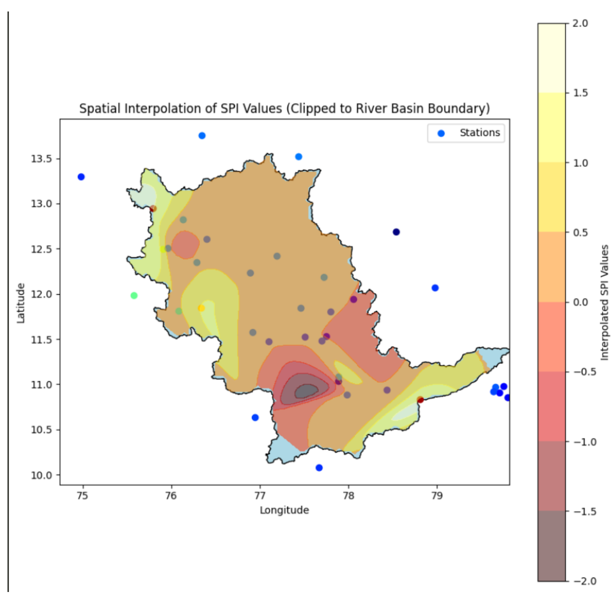


Figure 16: Spatial interpolation of SPI values

The additional plot of drought frequencies complements the first two visualizations by providing a long-term perspective on how often drought conditions have occurred across the Kaveri Basin. While the contour plot of SPI values offers insights into the intensity and spatial distribution of drought for the 2017-2021 period, and the NDVI plot shows the vegetation's response in 2021, the drought frequency plot adds another dimension by showing how frequently different areas have experienced drought. This helps reinforce the relationship between SPI values (precipitation deficits) and NDVI (vegetation health) by showing that regions with more frequent droughts also tend to have lower SPI and NDVI values, underscoring the persistence of drought impacts on the landscape.

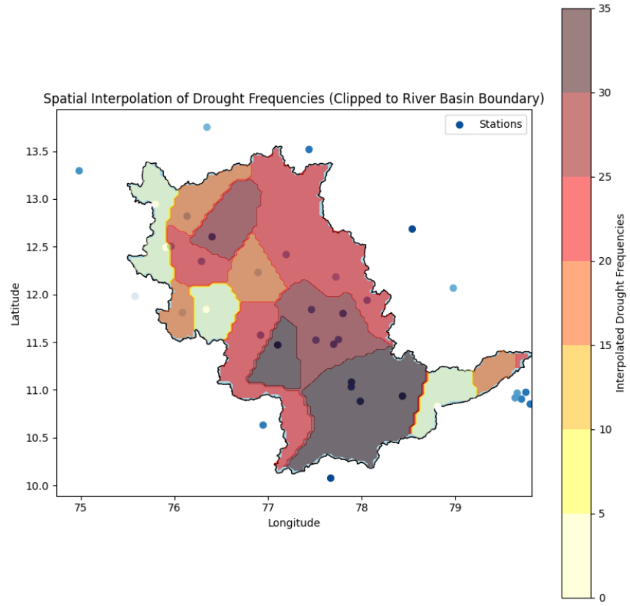


Figure 17: Spatial interpolation of drought frequencies

Conclusion

Through the analysis of SPI values, NDVI data, and spatial interpolation techniques, we successfully characterized drought conditions in the Kaveri River Basin from 2017 to 2021. By using SPI values from meteorological stations, we quantified the intensity and duration of droughts, while the contour maps helped visualize how drought conditions varied across the basin. The NDVI analysis for 2021 allowed us to assess the impact of these drought conditions on vegetation health, revealing a clear correlation between low SPI values and diminished vegetation activity. Although computational constraints prevented us from performing a multi-year NDVI analysis, the single-year data still provided meaningful insight into the basin's vegetation response to drought.

Additionally, the contour plot of drought frequencies added a temporal dimension to our study by highlighting regions where droughts have occurred more frequently. This, when combined with the SPI and NDVI analyses, painted a comprehensive picture of the drought landscape, showing the relationship between precipitation deficits, vegetation stress, and the persistence of drought conditions.

The combination of these tools and datasets provides a strong foundation for understanding drought patterns, which can be valuable for managing water resources, planning agricultural activities, and mitigating the long-term impacts of drought in the region. These insights are critical for improving resilience in drought-prone areas and developing strategies for sustainable resource management in the Kaveri River Basin.