Water Body Identification using Satellite Data

1 Aim

The objective is to identify water bodies within a designated region during a hydrological year, extract them as polygons, and subsequently store various data pertaining to these bodies.

2 Methodology

1. **Water body identification**

   Input:

   - Region of interest and the hydrological year.
   - Sentinel-1 and Sentinel-2 data.
   - Dynamic World data.

   Output:

   - Raster layer with class 2 representing water in Kharif, class 3 representing water in Kharif and Rabi, class 4 representing water in all the seasons.

   Algorithm:

   (a) Segment the hydrological year into three seasons: Kharif, Rabi, and Zaid.

   (b) Water Detection in Kharif Season:

      i. Use Sentinel-1 (SAR) data to detect water pixels during the Kharif season.

      ii. Apply a thresholding method on VV band:

         - If VV value is less than -16 for a pixel, classify it as water; otherwise, classify it as non-water.

      iii. Aggregate and classify all images within a month using the threshold value.

      iv. Determine the presence of water in each Kharif month using mode.

v. If water is detected in at least 3 out of 4 months, classify it as present for the entire season.

(c) Water Detection in Rabi and Zaid Seasons:

i. Use Dynamic World data for detecting water pixels during Rabi and Zaid seasons.

ii. Aggregate images for each month and calculate mode to classify water presence.

iii. If a pixel is classified as water in at least 2 out of 4 months, classify it as water for that season.

(d) Combine outputs for all seasons and classify them into classes:

- Class 2: Water presence in Kharif.
- Class 3: Water presence in both Kharif and Rabi.
- Class 4: Water presence in Kharif, Rabi, and Zaid.
- Adjust outputs to backfill for other combination. For instance, if later seasons feature water, then preceding seasons should also reflect this presence of water. (Order - Kharif,Rabi,Zaid)

(e) Refine the output using Sentinel-2 band NDWI:

- If maximum NDWI band value for a pixel over the year is less than 0.15 and it was classified as water, flip it to non-water.

Figure 1: Water body identification

2. **Extracting water bodies**

   Input:

   - Raster layer output from water body identification.

   Output:

   - Feature collection of water body polygons.

   Algorithm:

   (a) Generate output for each hydrological year from 2017 to 2022 using the water body identification method.

   (b) Convert output images to binary images:

   - Assign pixel value as 1 if the class value is 2, 3, or 4; otherwise, assign it as 0.

   (c) Perform logical OR operation on all binary images to produce the final image output.

   (d) Extract water bodies from the final image as a feature collection:

   - Delineate each body by grouping connected pixels with a value of 1 in 8 directions.

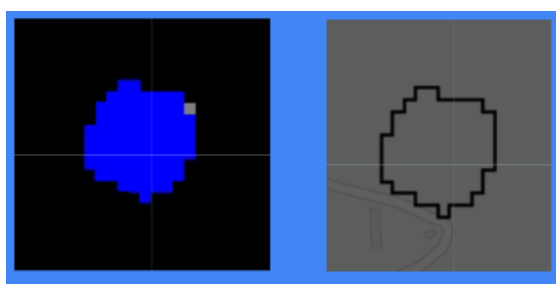

Figure 2: Left side image shows the water body raster image output and right side shows the extracted water body.

3. **Adding data of water bodies**

- For each feature, diverse metadata is stored including size-range, area-ored, unique ID, area for each year, presence of water during Kharif for each year, presence of water during both Kharif and Rabi for each year, and presence of water during Kharif, Rabi, and Zaid for each year, census_id.
- Size-range - It tells the size range in which the water body belongs. There are 4 size ranges(sqm) - (0-500),(500-1000),(1000-5000),(5000 and above).
- Area_ored - It tells the size of the water body in sqm after OR of all the years.
- Unique_id - It is a unique id generated for every water body using the micro watersheds intersecting with the water body.
- Census_id - It signifies the unique mapping of the water body with the water bodies census data.

  WBC data:

  - The objective of the 1st Census of Water Bodies is to have a comprehensive national database of all water bodies by collecting information on all important aspects. It has columns like state name, district name, unique id, capacity, depth, surface area of water body, lat lon and many more. Capacity, depth, surface area, unique id and the coordinates of the location are the columns used for generating this unique mapping.
- Following details are added for each hydrological year :-
  - Area for each year - It tells the area(sqm) of the water body in that year.
  - Presence of water during Kharif for each year - It tells the percentage of pixels that have water in kharif season in that year.
  - Presence of water during both Kharif and Rabi for each year - It tells the percentage of pixels that have water in kharif and rabi season in that year.
  - Presence of water during Kharif, Rabi, and Zaid for each year - It tells the percentage of pixels that have water in all seasons in that year.

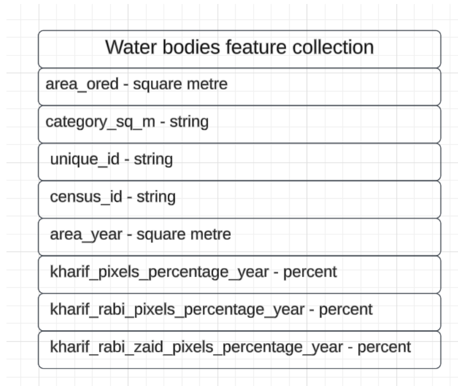

Figure 3: Schema

3 How to execute?

To generate the surface water bodies,

1. Run the provided Google Colab script 1 for years 2017-2022. **Input:**
    - Region of interest.
    - Start and end date.

   Output:
    - Raster layer for each hydrological year between then start and end date with class 2 representing water in Kharif, class 3 representing water in Kharif and Rabi, class 4 representing water in all the seasons.

2. Subsequently, run GEE script 1 to create the vector layer containing combined bodies and their associated properties.

   Input:
    - Region of interest (table).
    - the images generated from above script (image to image5).

   Output:
    - Feature collection of water bodies containing properties like area in different seasons, category-size, ored-area etc.

3. Then, execute the GEE script 2 to generate unique IDs for each water body.

   Input:
    - micro watershed layer of required region (aoi).

- water bodies feature collection generated in previous script. (water-bodies).

Output:

- Feature collection of water bodies with unique id column added.

4. Finally, execute the GEE script 3 to get the census ID for the water body which signifies the unique mapping of the water body to the census point from the water body census data.

   Input:

   - WBC census state data (census state).
   - water bodies feature collection generated in previous script. (water bodies).
   - region of interest. (geometry).

   Output:

   - Feature collection of water bodies with census id column added.

4 Github repository link

- Link - `https://github.com/vatsal3812/water_bodies`

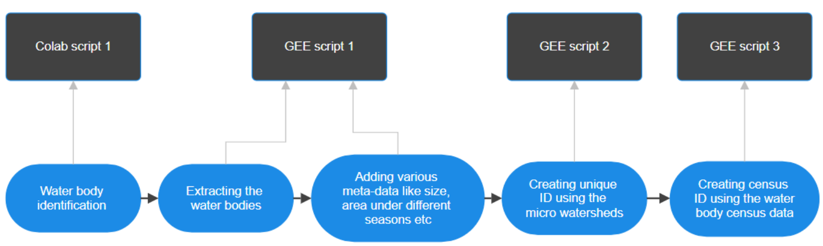

Figure 4: Execution process