

# CoRE stack: Commoning for Resilience and Equality - Geospatial data layers

Draft version: April 2024

Shivani A Mehta<sup>1</sup>, Aaditya Sinha<sup>1</sup>, Aila Dutt<sup>2</sup>, Akshay P Sarashetti<sup>1</sup>, Akshay Pratap Singh<sup>1</sup>, Aman Gupta<sup>1</sup>, Aman Verma<sup>2</sup>, Ajay Tannirkulam<sup>4</sup>, Ananda Sreenidhi<sup>2</sup>, Ananjan Nandi<sup>1</sup>, Ankit<sup>1</sup>, Ankit Sharma<sup>2</sup>, Anamitra Singha<sup>1</sup>, Anant Gulgulia<sup>1</sup>, Ashima Mittal<sup>1</sup>, Atharv Dabli<sup>1</sup>, Athira P<sup>3</sup>, Balakumaran Ramachandran<sup>5</sup>, Chahat Bansal<sup>1</sup>, Chintan Sanjaybhai Sheth<sup>1</sup>, Craig Dsouza<sup>5</sup>, Deepak Kumar<sup>2</sup>, Dharmisha Sharma<sup>1</sup>, Harshita<sup>1</sup>, Jaskaran Singh<sup>1</sup>, Kapil Dadheech<sup>2</sup>, Ksheetiz Agrahari<sup>2</sup>, Om Krishna<sup>1</sup>, Pooja Prasad<sup>1</sup>, Priyadarshini Radhakrishnan<sup>1</sup>, Rittwick Bhabak<sup>1</sup>, Ruptirumal Sai Bodavula<sup>1</sup>, Ramita Sardana<sup>1</sup>, Saketh Vishnubhatla<sup>1</sup>, Samitha Haldar<sup>2</sup>, Sanjali Agrawal<sup>1</sup>, Shruti Kumari<sup>1</sup>, Siddharth S<sup>1</sup>, Sukriti Kumari<sup>2</sup>, Vishnu S<sup>2</sup>, and Aaditeshwar Seth<sup>1,2</sup>

<sup>1</sup>Indian Institute of Technology, Delhi

<sup>2</sup>Gram Vaani

<sup>3</sup>Indian Institute of Technology, Palakkad

<sup>4</sup>Magasool

<sup>5</sup>Well Labs

Contact: [contact@core-stack.org](mailto:contact@core-stack.org)



---

## Acknowledgements

None of this would have been possible without valuable insights and field support from our partners. A very special thanks to Subrata Singh, Chiranjit Guha, Yogesh Barola, Himani Sharma, Mahesh Jadav, Pratiti Priyadarshini and the wider FES team, for their constant support, encouragement, and thought partnership. Our sincere gratitude to the teams at SUPPORT in Dumka, 4S in Gaya, PRADAN in Sirohi, Jan Jagran Kendra in Giridih, CCD in Adilabad, SPWD and JJBA in Latehar, and IIT-IIT in Kolar, for field access. Finally, our heartfelt thanks to our donors: GIZ, Common Ground, Rainmatter Foundation, Tower Research, Hellermann Tyton and Tarides whose generous contributions have enabled us to build the CoRE stack so far. We would also like to especially acknowledge Sharat Singh from GIZ for believing in our vision and being the first one to mobilize support for the CoRE stack.



---

# Table of Contents

<b>1 Acknowledgements</b>	<b>3</b>
<b>2 Introduction</b>	<b>11</b>
<b>3 Hydrological boundaries</b>	<b>15</b>
3.0.1 Basins . . . . .	15
3.0.2 Subbasins . . . . .	15
3.0.3 Watersheds . . . . .	16
3.0.4 Microwatersheds . . . . .	16
<b>4 Hydrological variables</b>	<b>19</b>
4.0.1 Precipitation . . . . .	19
4.0.2 Runoff . . . . .	20
4.0.3 Evapotranspiration . . . . .	22
4.0.4 Change in groundwater . . . . .	23
4.0.5 Change in well depth . . . . .	24
4.0.6 Drainage lines with stream orders . . . . .	26
4.0.7 Aquifer systems in India . . . . .	28
<b>5 Climatic variables</b>	<b>31</b>
5.0.1 Drought Frequency and Intensity . . . . .	31
<b>6 Landscape variables</b>	<b>35</b>
6.0.1 Terrain classification . . . . .	35
6.0.2 Land use land cover . . . . .	36
6.0.3 Cropping intensity . . . . .	39
6.0.4 Water bodies . . . . .	40
6.0.5 First census of water bodies . . . . .	42
<b>7 Welfare variables</b>	<b>45</b>
7.0.1 NREGA assets categorization . . . . .	45
<b>8 Site assessment</b>	<b>49</b>
8.0.1 Lineament . . . . .	49
8.0.2 Lithology . . . . .	49
8.0.3 CLART . . . . .	50
<b>9 Administrative boundaries</b>	<b>57</b>
9.0.1 State . . . . .	57
9.0.2 District . . . . .	57

9.0.3	Block . . . . .	57
9.0.4	Panchayat . . . . .	57
9.0.5	Village . . . . .	58

---

## List of Tables

2.1	Hosting specification of vector layers in CoRE stack . . . . .	12
2.2	Hosting specification of raster layers in CoRE stack . . . . .	13
4.1	Categories of AMC conditions . . . . .	21
7.1	MGNREGA category and their corresponding keywords . . . . .	45
8.1	Recharge potential values and their corresponding scores for recharge . . . . .	52



---

# List of Figures

2.1	CoRE stack: Modular approach to build a digital public infrastructure . . . . .	12
3.1	Basin (left) and subbasin (right) boundaries across India . . . . .	16
3.2	Microwatershed (left) and watershed (right) boundaries across Mohanpur block, in Gaya district of Bihar . . . . .	17
4.1	Methodology to estimate runoff . . . . .	22
4.2	Spatial variation in precipitation, runoff, evapotranspiration and change in groundwater during 2022-23 across microwatersheds of Mohanpur block, Gaya district, Bihar . . . . .	24
4.3	Time series of water balance variables: precipitation, runoff, evapotranspiration and change in groundwater (Deltag) on fortnight basis between 2022-2023 for a microwatershed in Mohanpur. Time series of groundwater (G) is derived by assuming the groundwater in the first fortnight as zero in the worst case. . . . .	25
4.4	Change in well depth across microwatersheds of Mohanpur during 2017-22 (left) and 2018-23 (right) . . . . .	26
4.5	Snapshot of drainage lines with stream orders over Jamui district in Bihar, overlaid on open street map (OSM). . . . .	27
4.6	Principal aquifer systems of India . . . . .	29
4.7	Major aquifer systems of India . . . . .	29
5.1	Drought intensity is determined for each monthly window (rows) in terms of mild, moderate, severe and no drought. Weekly drought intensity (columns) is determined by the maximum drought intensity that has occurred across the corresponding monthly windows in which the week is present. . . . .	34
5.2	Flow chart of determining monthly drought intensity . . . . .	34
6.1	Scale dependency of TPI . . . . .	36
6.2	Landforms classification using slope and thresholds (denoted using $\lambda$ ) on both small-scale and large scale TPIs . . . . .	37
6.3	Landform classification of Masalia with micro-watershed boundaries in black.	38
6.4	Land use land cover classification of Raichur district in Karnataka for 2022 .	40
6.5	For a micro-watershed in Mohanpur, we analyze the percentage of area under single, double and triple cropping (left), and corresponding cropping intensity over the years. . . . .	41
6.6	Water bodies as seen from satellite (first column) in Masalia block of Dumka district in Jharkhand and their detection using the classifier (second column)	42

6.7	Bar graph depicting the surface area of a water body (in percentage) in Mohanpur block of Gaya district in Bihar . . . . .	43
6.8	The map shows geo-coordinate of waterbodies from the waterbody census data of Rajasthan, specifically zoomed into Mandalgarh block of Bhilwara district for visualization. . . . .	44
7.1	MGNREGA assets in Pindwara block of Sirohi district in Rajasthan . . . . .	47
7.2	Metadata of a MGNREGA asset in Masalia block of Dumka district in Jharkhand . . . . .	48
8.1	Lithology shapefile from Bhukosh for Bhilwara district of Rajasthan . . . . .	50
8.2	The recharge potential scores are combined with different slope percentages and land cover categories to derive treatment codes in CLART. . . . .	53
8.3	Flowchart of methodology to produce reprogrammable CLART using score assignment . . . . .	54
8.4	The figure shows the component layers for generating CLART output for Mandalgarh block in Bhilwara district of Rajasthan . . . . .	55
9.1	State, district and block boundaries (from left to right) of India . . . . .	58
9.2	Panchayat (top) and village (bottom) boundaries of Mohanpur block, in Gaya district of Bihar . . . . .	59

---

# Introduction

The document outlines water security related datasets in the Commoning for Resilience and Equality (CoRE) stack to build a water security plans. The CoRE stack is a collaborative effort with several academic research institutions, civil society organizations working on ecological sustainability and rural livelihood, innovation catalysts, and technology-led social enterprises. Being designed as a Digital Public Good, the CoRE stack enables an open-access co-creation network to innovate and scale digital technology solutions for ecosystem sustainability. Figure 2.1 shows the hierarchical modules of CoRE stack.

Datasets are essential to quantify and assess the state of water security. Water security requires tracking of supply-side variables such as precipitation and incidence of droughts, demand-side variables such as increasing number of cropping cycles being practiced and intervention variables such as the available surface runoff for harvesting. We leveraged remote-sensing and machine learning to temporally estimate various variables at the micro-watershed scale. A few initiatives are already attempting to provide water security related datasets [31, 29], but they tend to use out-dated or lower resolution data products and sub-optimal algorithms.

We are guided by the principles of open data and algorithms to promote transparency and foster innovation through collaborations. We publish temporally static datasets such as hydrological and administrative boundaries using drive folders, and publish block-wise temporally dynamic datasets such as hydrological and climate variables on the [Landscape explorer](#). To generate temporally dynamic datasets for your own block, we publish our GitHub repositories for reproducibility, collaboration and contribution.

Figure 2.1: CoRE stack: Modular approach to build a digital public infrastructure

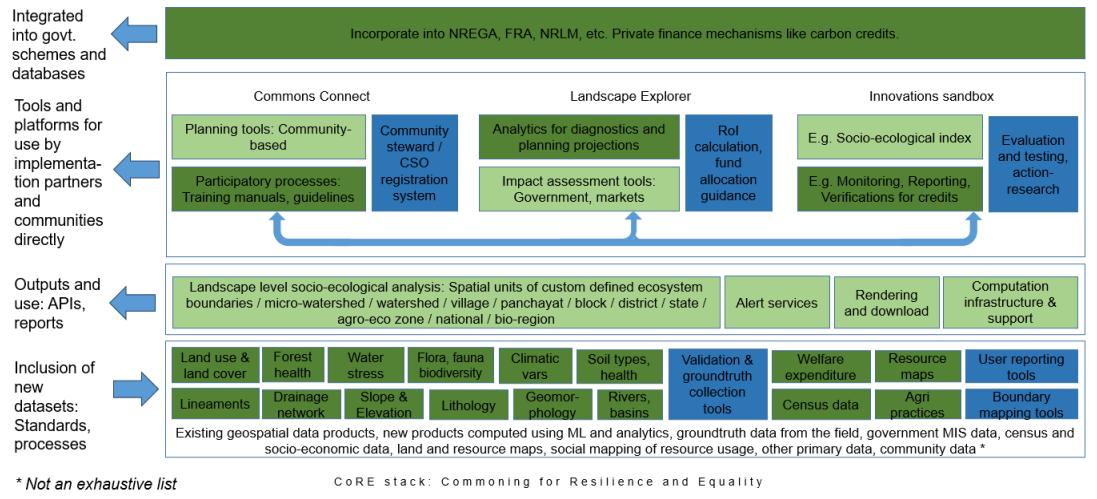


Table 2.1: Hosting specification of vector layers in CoRE stack

Layer	Type	Spatial resolution	Temporal resolution	Dataset	Codebase
Basin	vector	NA	static	Google drive folder	-
Sub-basin	vector	NA	static	Google drive folder	-
Watershed	vector	NA	static	Google drive folder	-
Micro-watershed	vector	NA	static	Google drive folder	-
Precipitation	vector	NA	fortnightly	-	GitHub repository
Runoff	vector	NA	fortnightly	-	GitHub repository
Evapotranspiration	vector	NA	fortnightly	-	GitHub repository
Change in groundwater	vector	NA	fortnightly	-	GitHub repository
Change in well depth	vector	NA	five years	-	GitHub repository
Drainage lines with stream orders	vector	NA	static	Google drive folder	-
Aquifer systems in India	vector	NA	static	Google drive folder	-
Drought intensity and frequency	vector	NA	yearly	-	GitHub repository

Water bodies	vector	NA	yearly	-	Google drive folder	GitHub repository
First census of water bodies	vector	NA	static	-	Google drive folder	-
Cropping intensity	vector	NA	yearly	-	Google drive folder	GitHub repository
NREGA assets categorization	vector	NA	yearly	Google drive folder	Google drive folder	GitHub repository
Lithology	vector	NA	static	Google drive folder	Google drive folder	-
State	vector	NA	static	Google drive folder	Google drive folder	-
District	vector	NA	static	Google drive folder	Google drive folder	-
Block	vector	NA	static	Google drive folder	Google drive folder	-
Panchayat	vector	NA	static	Google drive folder	Google drive folder	-
Village	vector	NA	static	Google drive folder	Google drive folder	-

Table 2.2: Hosting specification of raster layers in CoRE stack

Layer	Type	Spatial resolution	Temporal resolution	Dataset	Codebase
LULC	raster	10m	yearly	-	Google drive folder
Terrain	raster	30m	static	-	Google drive folder
Lineament	raster	194m	static	Google drive folder	-
CLART	raster	30m	static	-	Google drive folder



---

# Hydrological boundaries

## 3.0.1 Basins

A watershed, also known as a drainage basin or catchment area, is an area of land where all the water that falls or drains into it ultimately drains into a common outlet, such as a river, lake, or ocean. This includes surface water runoff, precipitation, and even underground water flow. Watersheds play a crucial role in the hydrological cycle, influencing the distribution and movement of water across the landscape.

Based on the size, the hydrological units are termed as basin, sub-basin, catchment, sub-catchment, watershed, sub-watershed and micro-watershed. India Water Resources Information System (WRIS) divides the Indian subcontinent into 25 major river basins with major river basin of Ganga-Brahmaputra-Meghna followed by Indus, Mahanadi, Godavari and Krishna.

Basin boundaries for India were obtained from WRIS ArcGIS portal. The layers were fetched on QGIS by using the url of WRIS ArcGIS portal in the QGIS's ArcGIS rest server. Basin boundaries across India are shown in figure [3.1](#)

### Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: static
- Dataset: [Google drive folder](#)

## 3.0.2 Subbasins

The subbasin boundary layer obtained from WRIS ArcGIS portal had attribute 'sbcode' which refers to the sub-basin code to which the watershed belongs. The watershed layer was dissolved on the 'sbcode' attribute to combine all the watersheds belonging to the same sub-basin. Subbasin boundaries across India are shown in figure [3.1](#)

### Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: static
- Dataset: [Google drive folder](#)

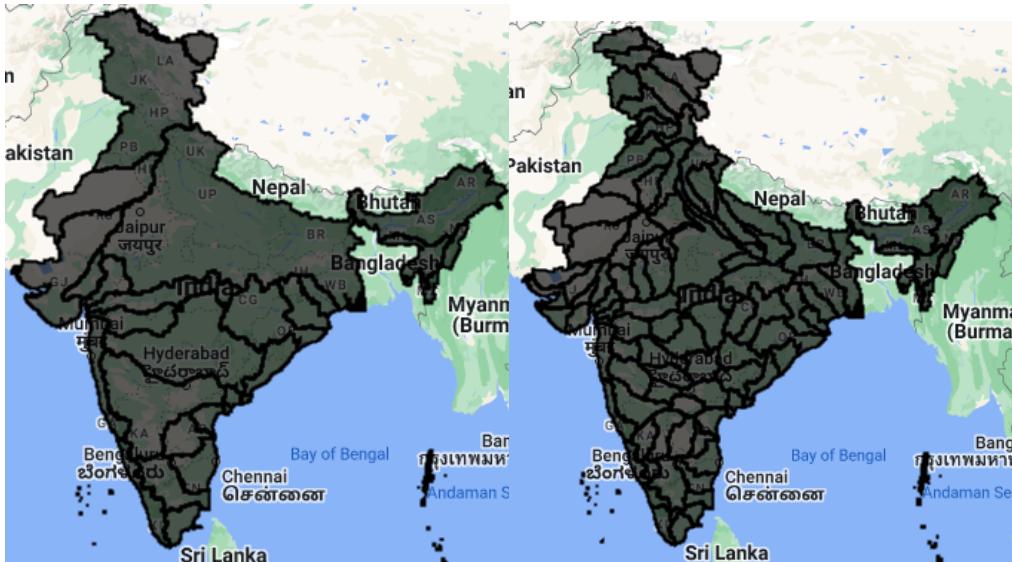


Figure 3.1: Basin (left) and subbasin (right) boundaries across India

### 3.0.3 Watersheds

Watershed boundaries for India were obtained from WRIS ArcGIS portal. The layers were fetched on QGIS by using the url of WRIS ArcGIS portal in the QGIS's ArcGIS rest server. Watershed boundaries across Mohanpur block, in Gaya district of Bihar are shown in the figure 3.2

#### Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: static
- Dataset: [Google drive folder](#)

### 3.0.4 Microwatersheds

Within each watershed, micro watersheds were delineated using the land elevation data. We used digital elevation model (DEM) data of Shuttle Radar Topography Mission (SRTM) for land elevation. SRTM was an international research effort that obtained digital elevation models with a spatial resolution of 1 arc-second (30 m) for global coverage. The DEM data can be obtained for a region of interest using the QGIS SRTM Downloader plugin and can be clipped for every watershed.

Watershed usually covers an area of about 20,000 to 1,50,000 ha that contains many micro-watersheds (500 to 1500 ha size). Therefore, micro-watersheds were generated within each watershed with a minimum area threshold of 5555 pixels (500 ha divided by  $900 \text{ m}^2$ , 1 pixel = 30m by 30m) as an input parameter to the r.watershed function of GRASS library in QGIS. Microwatershed boundaries across Mohanpur block, in Gaya district of Bihar are shown in the figure 3.2

## Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: static
- Dataset: [Google drive folder](#)

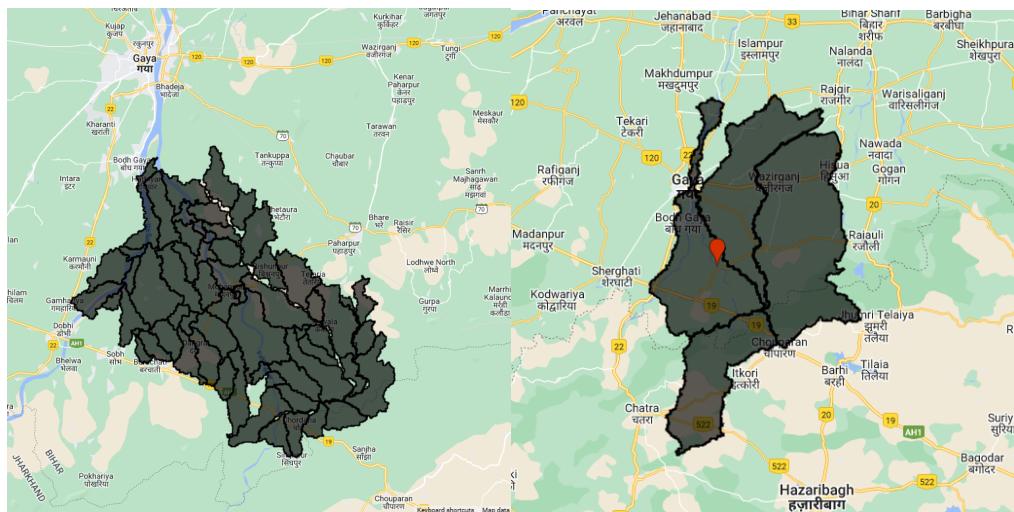


Figure 3.2: Microwatershed (left) and watershed (right) boundaries across Mohanpur block, in Gaya district of Bihar



---

# Hydrological variables

## 4.0.1 Precipitation

### Introduction

In the context of water balance, precipitation is the key input that contributes to the positive side of the water balance, representing the input of water to a particular region. Global precipitation data is available from satellite based products for last two decades which is critical for water resource management, agricultural production and addressing issues related to drought or flood assessment in a given area.

### Input layers

Precipitation can be calculated during an interval using the Global Satellite Mapping of Precipitation (GSMP) dataset available on Google Earth Engine's data catalogue [13]. GSMP provides a global precipitation in mm/hr at spatial resolution of approximately 11km. GSMP is a product of the Global Precipitation Measurement (GPM) mission, which uses multi-band passive microwave and infrared radiometers from a constellation of satellites to estimate hourly precipitation.

### Methodology

Given a time duration and a region of interest, we filter the daily images of dataset between the start and end dates of the duration and add these images to get total precipitation of the duration on Google Earth Engine. The total precipitation image is clipped for the region of interest and a mean is performed across all the pixels in the region of interest to derive total precipitation of the region of interest for that duration. We derive fortnightly precipitation for each microwatershed across the five blocks. The spatial variation of precipitation is shown figure 4.2.

### Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: fortnightly
- Codebase: [Github repository](#)

## 4.0.2 Runoff

### Introduction

Surface runoff refers to the flow of water that occurs when excess rainwater cannot be absorbed by the soil or vegetation and, as a result, flows over the Earth's surface. It is a major component of water cycle and is also known as overland flow. The spatio-temporal variation of runoff generation plays an important role in planning rainwater harvesting structures for effective water resource management, agricultural production and flood control. Remote sensing and GIS based layers such as land cover and soil type are very useful input data to runoff modelling. Runoff modelling can be grouped into three categories: empirical black-box, lumped conceptual and distributed physically based modelling [38]. Empirical models lack well defined representation of physical processes that converts a part of rainfall into runoff and distributed physically based modelling are highly data intensive. Therefore, we focus on conceptual modelling of surface runoff by using Soil Conservation Service Curve Number (SCS-CN) method. The SCS-CN method is widely used to convert rainfall to surface runoff using the curve number and involves the relationship of land cover, hydrologic soil groups, slope and 5-day antecedent rainfall.

### Input layers

- **Precipitation data:** using the GSMP dataset [13] at a spatial resolution of 11Km.
- **Soil type:** using the HYSOGs250m dataset [12] that represents a global gridded dataset of hydrologic soil groups (HSGs) with a spatial resolution of 250m, developed by USDA to support curve-number runoff modeling. Hydrologic soil groups A, B, C, and D correspond to low, moderately low, moderately high, and high runoff potential, respectively.
- **Slope:** using the NASA SRTM DEM dataset [25] at a spatial resolution of 30m.
- **Land cover:** using the Dynamic World dataset [9] which has a spatial resolution of 10m and includes nine categories of land covers such as water, trees, grass, flooded vegetation, crops, shrub and scrub, built-up, bare and snow .

### Methodology

The SCS-CN method uses curve number along with previous 5-day rainfall for surface runoff estimation. Curve number is dependent on the type of land cover, hydrologic soil group and slope as shown in the flowchart 4.1. Antecedent moisture condition (AMC) is calculated based on the total previous 5-day rainfall. AMC is categorized into three categories: AMC-I that refers to dry condition, AMC-II that refers to normal condition and AMC-III refers to wet condition. Curve number for AMC-II (CN2) is determined for each combination of land cover type and HSG [50]. Slope is also an important factor for determining the movement of runoff and therefore slope adjusted CN2 values using Sharpley and Williams method [51] are derived, namely  $CN_{2s}$ . Curve numbers for AMC-I (CN1) and AMC-II (CN3) can be calculated using CN2 values. The higher the AMC, the higher the CN and therefore the surface runoff.

The NRCS-CN method modified by antecedent moisture (M) to estimate runoff [43] is defined as:

$$Q = \frac{(P - I_a)(P - I_a + M)}{(P + I_a + S_r + M)} \quad (4.1)$$

$$M = 0.5 * (-S_r + \sqrt{S_r^2 + 4P_5S_r}) \quad (4.2)$$

$$I_a = \lambda * S_r \quad (4.3)$$

$$S_r = \left( \frac{25400}{CN_s} - 254 \right) \quad (4.4)$$

$$CN_{2s} = \frac{1}{3} * (CN_3 - CN_2)(1 - 2e^{-13.86s}) + CN_2 \quad (4.5)$$

$$CN_3 = CN_2 e^{0.00673(100 - CN_2)} \quad (4.6)$$

$$CN_{1s} = \frac{4.2CN_{2s}}{10 - 0.058CN_{2s}} \quad (4.7)$$

$$CN_{3s} = \frac{23CN_{2s}}{10 + 0.13CN_{2s}} \quad (4.8)$$

where  $P_5$  is the antecedent 5-day total rainfall,  $P$  is the total precipitation,  $I_a$  is the initial abstraction,  $Q$  is the direct run-off,  $S_r$  is the potential maximum retention, and  $\lambda$  is the initial abstraction coefficient ( $\lambda = 0.2$ ). Here,  $CN_2$  and  $CN_3$  are the SCS curve numbers for AMC-II and AMC-III conditions, and  $s(mm^{-1})$  is the soil slope.  $CN_{1s}$  and  $CN_{3s}$  are slope adjusted curve numbers for AMC-I and AMC-III conditions.

The above datasets with different spatial resolutions were scaled up and scaled down to compute runoff at 30m spatial resolution on daily basis. Given a duration and a region of interest, we generate runoff rasters between the start and end dates of the duration and add these rasters to get total runoff of the duration. The total runoff raster is clipped for the region of interest and a mean is performed across all the pixels in the region of interest to derive total runoff of the region of interest for that duration. We derive fortnightly runoff for each microwatershed across the five blocks. The spatial variation of runoff is shown figure 4.2.

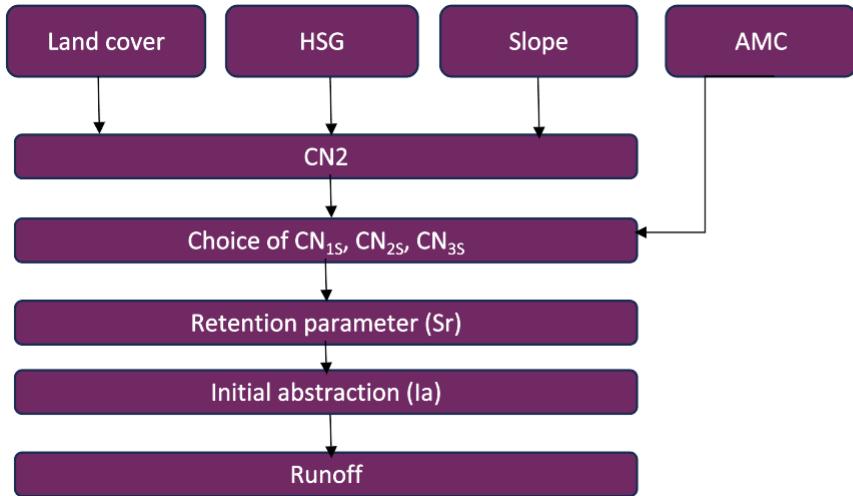
Table 4.1: Categories of AMC conditions

Antecedent rainfall	AMC condition
$0 < \text{rainfall previous 5days} \leq 35 \text{ mm}$	AMC-I
$35 < \text{rainfall previous 5 days} \leq 52.5 \text{ mm}$	AMC-II
$\text{rainfall previous 5days} > 52.5 \text{ mm}$	AMC-III

### Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: fortnightly
- Codebase: [Github repository](#)

Figure 4.1: Methodology to estimate runoff



### 4.0.3 Evapotranspiration

#### Introduction

Actual evapotranspiration plays a crucial role in the water balance of a region, contributing to the movement of water from the earth's surface to the atmosphere. It returns more than 60% of the total precipitation from land to atmosphere [36]. Evapotranspiration (ET) is the combination of transpiration from vegetation (Et), evaporation from bare soil (Eb), evaporation of intercepted precipitation from vegetated canopy (Ec), and evaporation from open water surfaces (Eo) [32]. Evapotranspiration is estimated using potential/reference evapotranspiration which provides the evaporative demand of atmosphere based on meteorological parameters and is independent of crop type, crop growing phase and management practices [8].

There are numerous remote sensing-based ET products that are available in the public domain and updated regularly for the user community. The products can be broadly classified into three categories [49] based on the methodology of computing ET: Penman-Monteith equation[44], Priestley-Taylor [47] equation and models that performs surface energy balance to compute evapotranspiration [53]. Remotely sensed meteorological parameters such as air temperature, net radiation and LULC are commonly used as inputs/forcings to land surface models to estimate evapotranspiration over large areas [40].

We used ET product from FLDAS Noah land Surface model [10] developed by NASA and USGS. The choice of product was determined by its spatio-temporal resolution, spatial coverage and its update frequency for data access. Noah ET is the sum of Ec, Et and Eb weighted by respective land surface coverage fractions. Noah's potential evapotranspiration (PET) is computed using Penman approach where PET is scaled by variables of canopy, transpiration and soil soil moisture availability to compute Ec, Et and Eb respectively.

#### Input layers

We use FLDAS NOAH dataset [10] that provides the ET on a daily basis at a spatial resolution of 0.01 degrees ( 1.31 km).

## Methodology

Given a time duration and a region of interest, we filter and download the daily images of dataset between the start and end dates of the duration by using the python APIs of NSIDC DAAC's Data Access [1]. We add these images to get total evapotranspiration of the duration. The total evapotranspiration image is clipped for the region of interest and a mean is performed across all the pixels in the region of interest to derive total evapotranspiration of the region of interest for that duration on Google Earth Engine. We derive fortnightly evapotranspiration for each microwatershed across the five blocks. The spatial variation of evapotranspiration is shown figure 4.2.

## Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: fortnightly
- Codebase: [Github repository](#)

### 4.0.4 Change in groundwater

#### Introduction

Water balance in the discipline of hydrology that aims to estimate the unknown water fluxes. It is an equation which is expressed in terms of water inputs, outputs and storage in a hydrological unit such as watershed [39]. We intend to estimate the net change in groundwater on a fortnightly basis for a micro watershed by solving the water balance equation.

With the current groundwater state in hand, the objective is to improve the groundwater state through interventions such as construction of rainwater harvesting structures, change in cropping patterns and growing plantations. The interventions will affect different components of the water balance equation and will allow us to project the groundwater state in future. For example, change in cropping patterns will affect evapotranspiration in the water balance equation. Groundwater projections will facilitate scientific and participatory planning within the community.

#### Input layers

We used the datasets of precipitation, runoff and evapotranspiration mentioned in section 4.1, 4.2 and 4.3 to compute change in groundwater as shown below for a micro-watershed in a fortnight.

## Methodology

Water balance equation will take precipitation ( $P$ ), outgoing runoff ( $Q_{out}$ ) and evapotranspiration ( $ET$ ) as inputs to output change in groundwater ( $\Delta G$ ) as shown in the equation 9.

$$P = Q_{out} + ET + \Delta G(\text{unknown}) \quad (4.9)$$

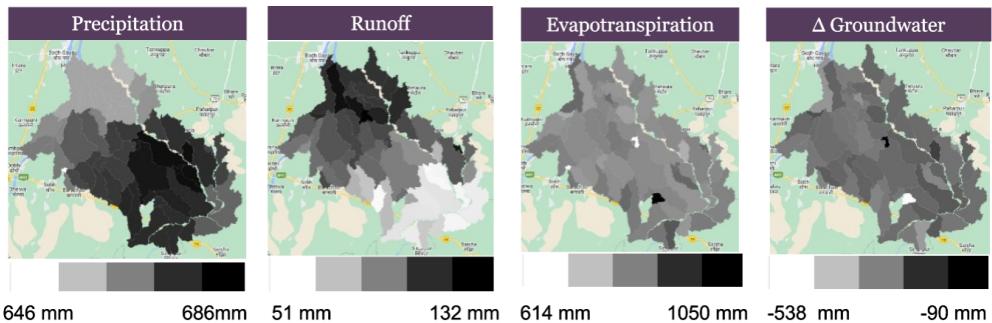
We assume incoming runoff, change in soil moisture and change in surface water storage as zero in the water balance and intend to model them in future. Each of these water balance inputs are derived using remote sensing products in order to diagnose the groundwater state

of a micro-watershed on fortnight basis as shown in the figure 4.3 and the spatial variation of change in groundwater is shown figure 4.2.

### Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: fortnightly
- Codebase: [Github repository](#)

Figure 4.2: Spatial variation in precipitation, runoff, evapotranspiration and change in groundwater during 2022-23 across microwatersheds of Mohanpur block, Gaya district, Bihar



### 4.0.5 Change in well depth

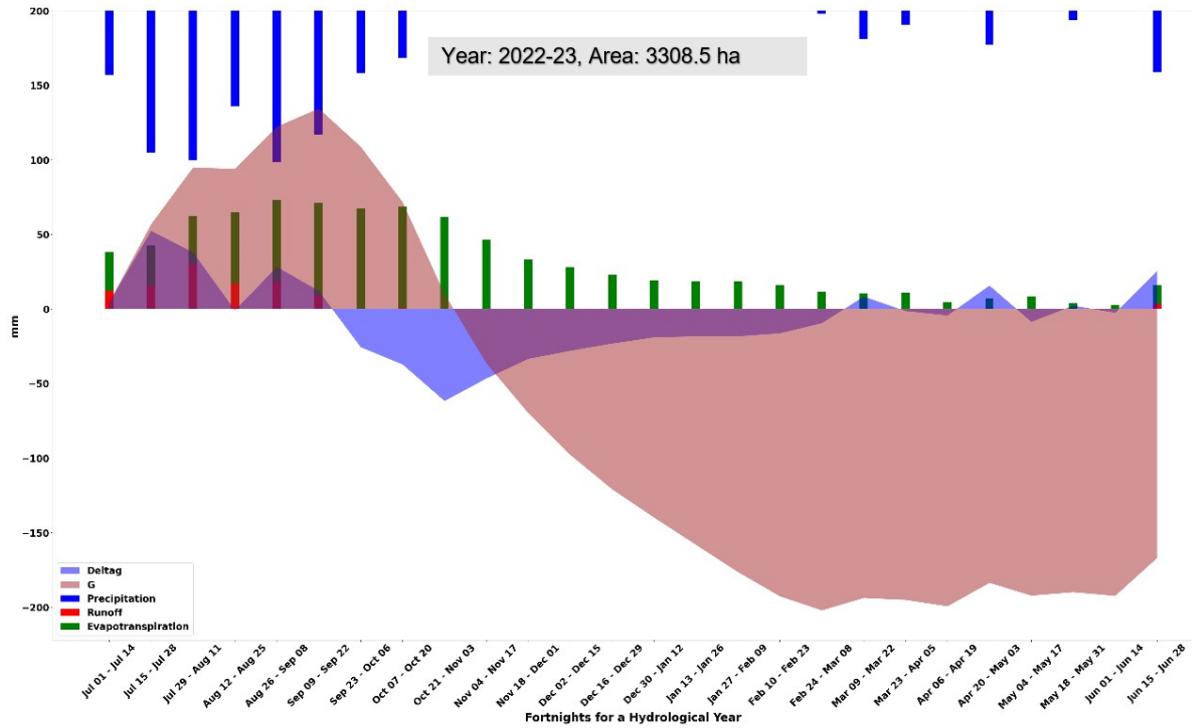
#### Introduction

The change in groundwater in an aquifer is related to the change in well depth through the specific yield of the aquifer [20]. Specific yield is a key parameter in hydrogeology that represents the fraction or percentage of groundwater that a saturated rock or soil can release under the influence of gravity. It is a measure of the storage and release capacity of an aquifer or a formation. Mathematically, specific yield is defined as the ratio of the volume of water drained from an aquifer by gravity to the total volume of the aquifer.

#### Input layers

We used the change in groundwater layer mentioned in section 4.4 and aquifer layer to compute change in well depth for each micro-watershed. Aquifer mapping was performed in the entire country by Ministry of Jal Shakti in India under Aquifer Mapping and Management Programme (NAQUIM) [4]. NAQUIM was taken up by Central ground Water board (CGWB) to take appropriate water resource management measures and understand the underlying ground water potential and recharge. Aquifer layer for India was obtained from India Water Resources Information System (WRIS) [ArcGIS portal](#). The CGWB has divided India across 14 principal aquifers and 42 major aquifers [4].

Figure 4.3: Time series of water balance variables: precipitation, runoff, evapotranspiration and change in groundwater (Deltag) on fortnight basis between 2022-2023 for a microwatershed in Mohanpur. Time series of groundwater (G) is derived by assuming the groundwater in the first fortnight as zero in the worst case.



## Methodology

The equation 10 describes the relationship between change in ground water storage of an aquifer and change in groundwater level [20].

$$\Delta G = \Delta h * A * S_y \quad (4.10)$$

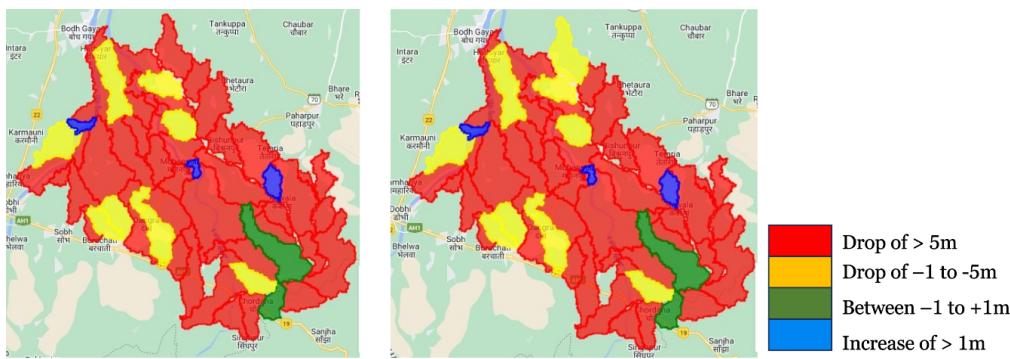
where,  $\Delta G$  is change in groundwater volume,  $\Delta h$  is change in well depth,  $A$  is the cross-sectional area of the aquifer and  $S_y$  is the specific yield of aquifer. The  $\Delta G$  in the equation assumes the change in the groundwater volume due to vertical fluxes and horizontal fluxes such as lateral flows across the boundaries of the aquifer is assumed to be zero. For example, alluvium aquifer shows fluctuation in groundwater by both lateral and vertical fluxes [20] while basalt aquifer which has medium to low permeability shows groundwater fluctuations due to vertical fluxes [20]. Therefore, to bring more accuracy for  $\Delta G$ , we will try to fine-tune the water balance equation as per different aquifer types.

Currently, the aquifer layer is clipped for each microwatershed to determine its specific yield. The change in well depth is computed over five year window for each microwatershed using the change in groundwater over 5 years and its specific yield. The figure 4.4 shows the variation of change in well depth across the micro-watersheds of Mohanpur between 2017-22 and 2018-23.

### Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: five years
- Codebase: [Github repository](#)

Figure 4.4: Change in well depth across microwatersheds of Mohanpur during 2017-22 (left) and 2018-23 (right)



### 4.0.6 Drainage lines with stream orders

#### Introduction

Drainage lines in hydrology refer to the visible channels through which water flows from higher elevations to lower elevations on the surface of earth such as streams and rivers. These lines play a crucial role in the movement of surface water. They collect runoff from precipitation events and transport it downstream, eventually flowing out from the outlet of the watershed.

#### Input layers

Drainage lines uses digital elevation model (DEM) raster dataset as input. The DEM provides pixel level information on the elevation of the terrain, which is used to determine the flow of water across the landscape. We use Shuttle Radar Topography Mission (SRTM) digital elevation model [25] generated by NASA and the National Geospatial-Intelligence Agency (NGA) in February 2000.

#### Methodology

The drainage lines are generated using the *r.stream.extract* function of GRASS tools in QGIS. The function extract stream channels from the DEM based on flow accumulation threshold. The flow accumulation threshold determines the minimum accumulated flow required to consider a pixel as part of a stream. The function outputs a binary raster map representing the extracted stream channels where pixels are classified as streams with a value of 1, while

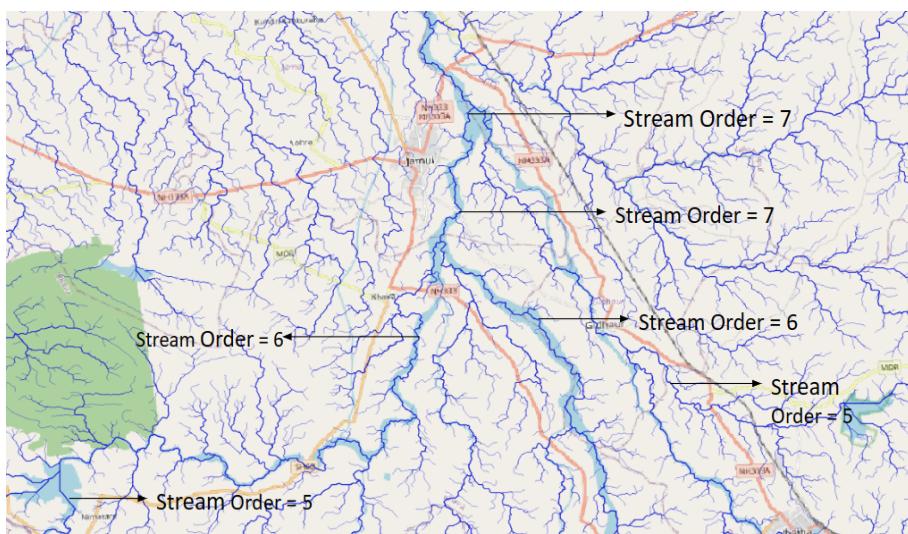
non-stream pixels are assigned a value of 0. We computed the drainage lines with a flow accumulation threshold of 100.

Stream order provides insight into the hierarchical structure of stream network. Each extracted stream channel is assigned stream order based on its relative hierarchy in the stream network. Streams without any tributaries are classified as first-order streams. When two first-order streams converge, they form a second-order stream. When two second-order streams meet, they create a third-order stream and so on. We assign stream orders based on the graph generated from stream network where the edges are streams and the nodes are intersection of streams. Each node has either degree of 1 (0 incoming – 1 outgoing edge) or 3 (2 incoming – 1 outgoing edge). Corresponding stream/edge of a node with a degree of 1 is assigned with a stream order of 1. While for the remaining nodes, the stream order of outgoing stream/edge is determined based on the stream orders of incoming streams/edges. If both the incoming edges have the same stream order, then the stream order of outgoing stream/edge is assigned as the stream order of the incoming edge + 1. And if the incoming edges have distinct stream orders, then the the stream order of outgoing stream/edge is assigned as the maximum of the two stream orders.

### Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: static
- Dataset: [Google drive folder](#)

Figure 4.5: Snapshot of drainage lines with stream orders over Jamui district in Bihar, overlaid on open street map (OSM).



## 4.0.7 Aquifer systems in India

### Introduction

For increasing economic growth and meeting its food security goals, India has relied heavily on groundwater by becoming largest consumer of groundwater in world. About more than 80% of rural drinking and more than 50% of agriculture's irrigation demand is met by groundwater [4]. Unplanned over exploitation of this valuable resource has led to drying of aquifers. Demarcation of aquifer systems across the country is essential to assess the ground water resource. In view of this, Central Ground Water Board (CGWB) of India carried out an exercise of mapping the lateral and vertical extent of aquifer systems across the nation and created a geospatial data that can be imported on GIS platforms. Based on the hydrogeological characteristics, the entire nation is divided into 14 principal aquifers which are in turn divided into 42 major aquifers [4] by rock types and their age as shown in the figures [4.6](#), [4.7](#) respectively.

### Methodology

The aquifer shapefile for India was obtained from India Water Resources Information System (WRIS) [ArcGIS portal](#). Each geometry in the shapefile represents a major aquifer uniquely identified by its region-specific aquifer properties such as aquifer thickness, specific yield, transmissivity and whether it is a unconfined and confined aquifer. Aquifer properties and their description can be found here [3].

### Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: static
- Dataset: [Google drive folder](#)

Figure 4.6: Principal aquifer systems of India

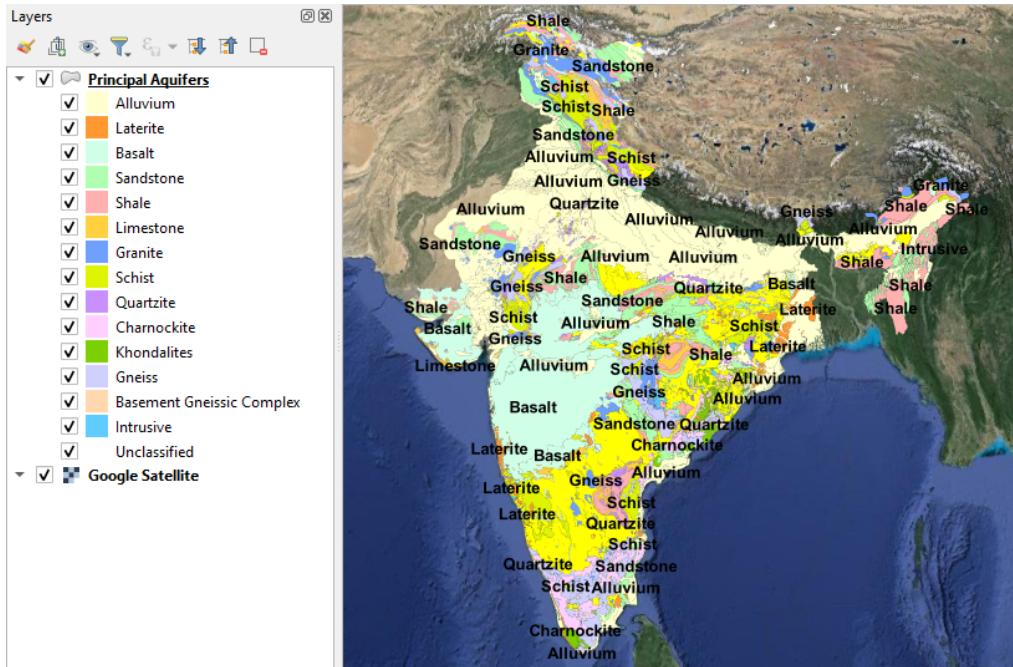
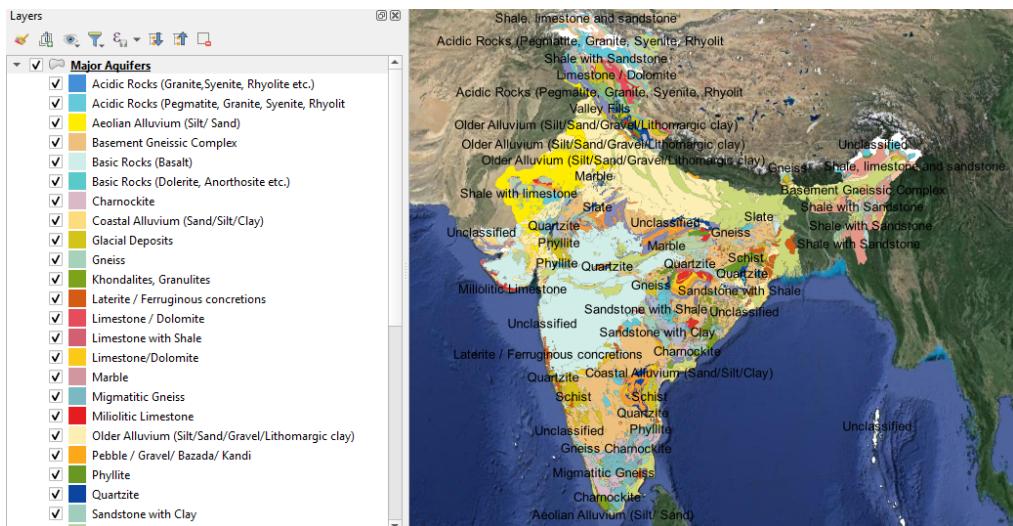


Figure 4.7: Major aquifer systems of India





---

# Climatic variables

## 5.0.1 Drought Frequency and Intensity

### Introduction

The impact of drought on Indian agriculture is significant, given the country's heavy reliance on rainfall for crop production. Drought significantly affects the crop yields and crop failure in worst case resulting in significant economic losses for farmers that can exacerbate perpetuating cycles of poverty and vulnerability. Drought-related crop failures can impact food security at both local and national levels, leading to reduced availability of food supplies and potential price spikes. This can disproportionately affect vulnerable populations, exacerbating hunger and malnutrition in affected regions. In the recent past, India has witnessed drought in 1987, 2002 and 2009 with an overall rainfall deficiency of 19%, 19% and 22% respectively which affected 59–60% of the normal cropped area alone in 1987. During 2014 and 2015, the drought encompassed major agricultural states in the country causing widespread hardships among the population.

Drought is a recurrent and sporadic feature of climate, which stems from significant rainfall deviation from normal in its spatial and temporal distribution such that it adversely impact the crops in a cropping season or successive cropping seasons. The intensity of drought depends on several factors including agro-climatic features, cropping choices and patterns, socio-economic vulnerabilities of the local population etc. The Department of Agriculture, Cooperation & Farmers Welfare under the Ministry of Agriculture & Farmers Welfare, Government of India has released the "Manual for Drought Management" in 2016 for the prevention, mitigation and management of droughts in India. The manual introduces scientific indices and parameters such as Vegetation Condition Index (VCI) for more accurate determination and assessment of drought. We have tried to implement the methodology of drought manual using various remote sensing products. The choice of product was determined by its spatial coverage, its spatio-temporal resolution and its near real time availability.

### Input layers

Based on the wide consultation with domain specialists, the drought manual [14] recommends five categories of indices [14] for determining drought incidence and severity: Rainfall, Vegetation, Water, Crop and others. Rainfall related indices include rainfall deviation, dry spell and standard precipitation index (SPI). Vegetation based indices such as Vegetation Condition Index (VCI) or NDVI/NDWI deviations from normal are derived from remote sensing data. Crop situation related indices include area under sowing and soil moisture based indices such as Percent Available Soil Moisture (PASM) and Moisture Adequacy Index (MAI). Hydrological indices include Reservoir Storage Index (RSI), Groundwater Drought Index (GWDI) and Stream-Flow Drought Index (SFDI). Other factors include socio-economic indicators such as prices of essential commodities as compared to normal prices, scarcity of

drinking water supply, agricultural wages etc. The following section outlines the methodology of using the above variables in the determination of drought at microwatershed level subject to the availability of data. We use the following publicly available datasets for building drought indicators:

- **Rainfall data** - Climate Hazards Group InfraRed Precipitation with Station data ([CHIRPS](#)) rainfall dataset with a spatial resolution of 5.566km and a temporal resolution of 1 day.
- **Vegetation data**- NDVI and NDWI are derived from [Landsat](#) dataset with a spatial resolution of 30m and a temporal resolution of 14 to 26 days.
- **Crop data**- Evapotranspiration and potential evapotranspiration required for computing MAI indicator were derived from [MODIS](#) dataset with a spatial resolution of 500m and a temporal resolution of 8 days.
- **Land use land cover**- as described in section 5.0.2

## Methodology

We compute weekly and monthly rainfall deviation using a long period average of that particular week or month across 30 years as shown in equation 11 and 12. The departure of rainfall from its long period average is considered as a credible indicator of drought as shown in table in figure [5.2](#). Weekly rainfall deviations are used to determine dry spell. The occurrence of dry spell is defined across four consecutive weeks with each week of these four weeks incurring a rainfall deviation of less than 50%. Prolonged dry spells can lead to significant reduction in crop sown area. SPI is computed on monthly scale and expresses the current monthly rainfall as a standardized departure with respect to rainfall probability distribution function. Positive SPI indicate deviations greater than median precipitation and negative SPI indicate deviations less than median precipitation as shown in table in figure [5.2](#). Rainfall indicators such as rainfall deviation, dry spell and SPI mentioned above are considered to be mandatory [\[14\]](#) for the first drought trigger to set off as shown in table in figure [5.2](#).

In the event where the first drought trigger is set off, the intensity of drought is assessed through its impact on agricultural area, crop condition, soil moisture and hydrological resources. Agricultural area is linked with the extent of sowing in the monsoon season. If the sowing fails due to delayed onset of monsoon or its deviation from normal, then it provides reliable information on the availability of water for agricultural operations. Impact on agricultural area is computed as the percentage of cropping area sown in Kharif season of that year. We compute the cropping area in kharif and total cropping area using the land use land cover classes: single kharif, single non-kharif, double and triple. Impact of drought on crop condition is often determined by spectral vegetation indices such as Normalized Difference Vegetation Index (NDVI) and Normalized Difference Wetness Index (NDWI), derived from remote sensing data. NIR and Red are the reflectance in visible and near infrared channels. NDVI values from 0.2 to 0.6 are associated with greater green leaf area and biomass while high NDWI is associated with more surface wetness. Observed NDVI and NDWI range is compared with historical range (using maximum and minimum) to derive monthly VCI for NDVI and NDWI respectively as shown in the equations 13, 14 and 15. VCI provides an idea about where the observed value is placed in the historical range as shown in table in figure [5.2](#). While combining VCI of NDVI and NDWI, the minimum of the two values is taken. Availability of moisture in soil is a very relevant indicator for drought as shown in table in figure [5.2](#) and is expressed as the ratio of Actual Evapo-transpiration (AET) to the Potential/Reference Evapo-transpiration (PET or RET) on monthly scale for MAI. The intensity of

the drought will be contingent upon the values of at least three out of four impact indicators viz, agriculture, crop condition, soil Moisture and hydrology in the following manner:

1. Severe: if all the above 3 impact indicators are in 'Severe' category.
2. Moderate: if two of the above 3 impact indicators are in 'Moderate' or 'Severe' class.
3. Mild: for all other cases.

We divide the south-west monsoon season into weeks and then into months (each of 4 weeks), starting from the first week with a stride of one week. For example, if there are 'n' weeks in the monsoon season, then there are 'n-3' months in the same monsoon season. Drought trigger is examined for each month using the dry Spell, monthly rainfall deviation and monthly SPI values as shown in table 5.2. In case of drought trigger, the intensity of drought is determined by the impact indicators of that month as shown in the flow chart 5.2. Weekly drought intensity is determined by the maximum drought intensity across all the months in which the week is considered as shown in the figure 5.1.

Monthly rainfall deviation in  $i^{th}$  month

$$RFdev_{monthly} = \frac{RF_i - RF_n}{RF_n} \quad (5.1)$$

where  $RF_n$  is the long term mean rainfall in  $i^{th}$  month and  $RF_i$  is the current rainfall in month i.

Weekly rainfall deviation in  $i^{th}$  week

$$RFdev_{weekly} = \frac{RF_i - RF_n}{RF_n} \quad (5.2)$$

where  $RF_n$  is the long term mean rainfall in  $i^{th}$  week and  $RF_i$  is the current rainfall in week i.

$$VCI_{NDVI} = \frac{NDVI_{current} - NDVI_{minimum}}{NDVI_{maximum} - NDVI_{minimum}} \quad (5.3)$$

$$VCI_{NDWI} = \frac{NDWI_{current} - NDWI_{minimum}}{NDWI_{maximum} - NDWI_{minimum}} \quad (5.4)$$

$$VCI = \min(VCI_{NDVI}, VCI_{NDWI}) \quad (5.5)$$

## Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: yearly
- Codebase: [Github repository](#)

Figure 5.1: Drought intensity is determined for each monthly window (rows) in terms of mild, moderate, severe and no drought. Weekly drought intensity (columns) is determined by the maximum drought intensity that has occurred across the corresponding monthly windows in which the week is present.

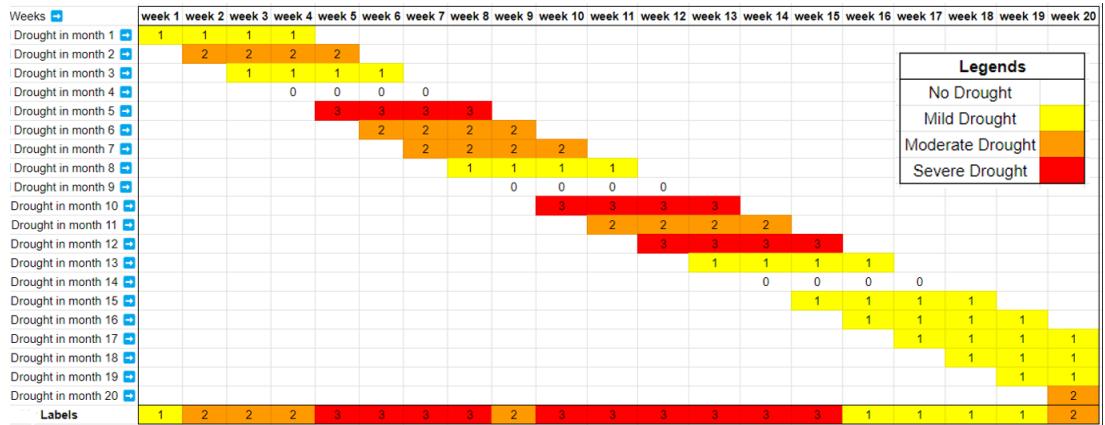
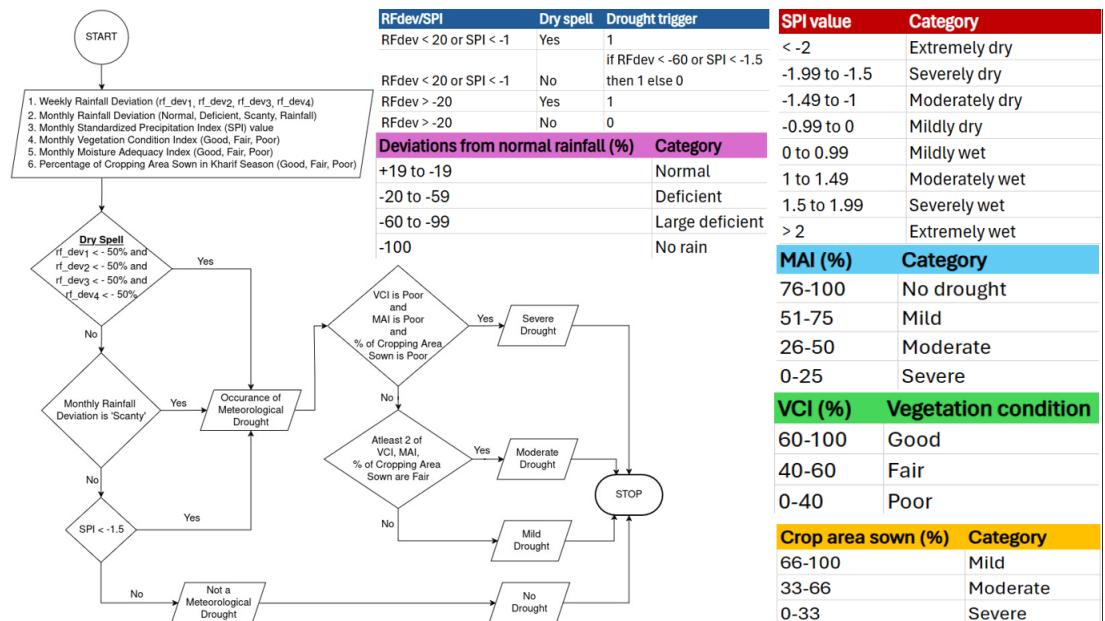


Figure 5.2: Flow chart of determining monthly drought intensity



---

# Landscape variables

## 6.0.1 Terrain classification

### Introduction

The Ministry of Rural Development, Government of India under the project of "Environmental benefits of MGNREGA", prescribes watershed development to create durable and sustainable assets [29]. The watershed development follows ridge to valley approach where the watershed is segregated into areas of different terrain for the construction of terrain specific assets to enhance the capacity of the assets for recharge and surface storage.

The CSO volunteers, following the ridge-to-valley approach identifies uplands, midlands and low lands in their landscape to intervene with specific type of assets in each of them. For example, Water Absorption Trenches (WATs) are constructed in uplands to harvest surface runoff, while farm ponds are constructed on lowlands to arrest surface and subsurface runoff for protective irrigation. The current process of identification is through visual inspection by observing the direction of flow of water during monsoon and assigning each farm plot on the cadastral map with upland, midland or lowland. Cadastral Maps are digital land records maintained by the state government. An automated classification of landscapes into broad plains, gentle slopes, steep slopes, ridges and valleys along with contour lines using digital elevation model may help in precision and efficiency of the process.

Terrain classification involves the categorization of Earth's physical features into distinct landforms such as hilltops, valley bottoms, ridges, plains, and slopes. Many natural processes such as hydrological balance and soil erosion are significantly correlated with the land forms.

### Input layers

We used NASA's SRTM Digital Elevation Model [25] at 30m resolution to generate landform classification.

### Methodology

We use landform classification method [54] by The Nature Conservancy to identify pixel level landform using Topographic Position Index (TPI) and slope at each pixel. The Topographic Position Index (TPI) compares the elevation of each pixel in a DEM to the mean elevation of a specified neighborhood around that pixel. Positive TPI values indicate positions that are elevated compared to the average of their neighboring terrain, such as ridges. Conversely, negative TPI values signify positions that are lower in elevation relative to their surroundings, like valleys. TPI values close to zero typically denote either flat areas or regions with consistent slopes.

The distance/scale at which the neighbourhood of TPI is defined, plays an important role in identifying the landform. For example, as shown in the figure 6.1, the TPI is near zero in A, while it is positive in B and negative in C. Therefore, both small-scale TPI and large-scale TPI are necessary to identify landform at a pixel. We threshold the range of small-scale TPI and large-scale TPI in a microwatershed by  $\lambda$  standard deviation and combine it with slope to determine landform classes as shown in the figures 6.2. We define  $\lambda$  using the equation XXX after fine-tuning the TPI algorithm for Indian landscapes such as Nalanda, Mohanpur, Masalia, Pindwara, Mandalgarh and Angul. We further categorize the plains into low, medium and high plains for finer landform classification by performing clustering based on their elevation of the pixels identified as plain as shown in the figure 6.3.

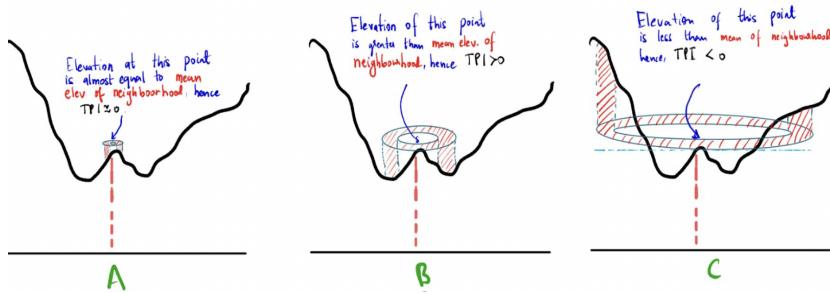
$$\lambda = \max(3 - \log(\sigma_{dc} + 1), 0.3) \quad (6.1)$$

where  $\sigma_{dc}$  is the standard deviation of elevation in a microwatershed.

### Hosting specifications

- Layer type: raster
- Spatial resolution: 30m
- Temporal resolution: static
- Codebase: [Github repository](#)

Figure 6.1: Scale dependency of TPI



## 6.0.2 Land use land cover

### Introduction

Land use land cover (LULC) classification that indicates different features on the Earth's surface, such as forests, rivers, croplands, or buildings, is used actively for monitoring and planning of land-use. For example, tracking of anthropogenic activities such as environmental degradation through deforestation and loss of farmlands are done through satellite maps annotated with LULC categories [48], and used by development practitioners to design sustainable land management policies [33].

Most global LULC products produce annual outputs and do not capture within-year dynamics. The recent Dynamic World product from Google produces a near real-time output

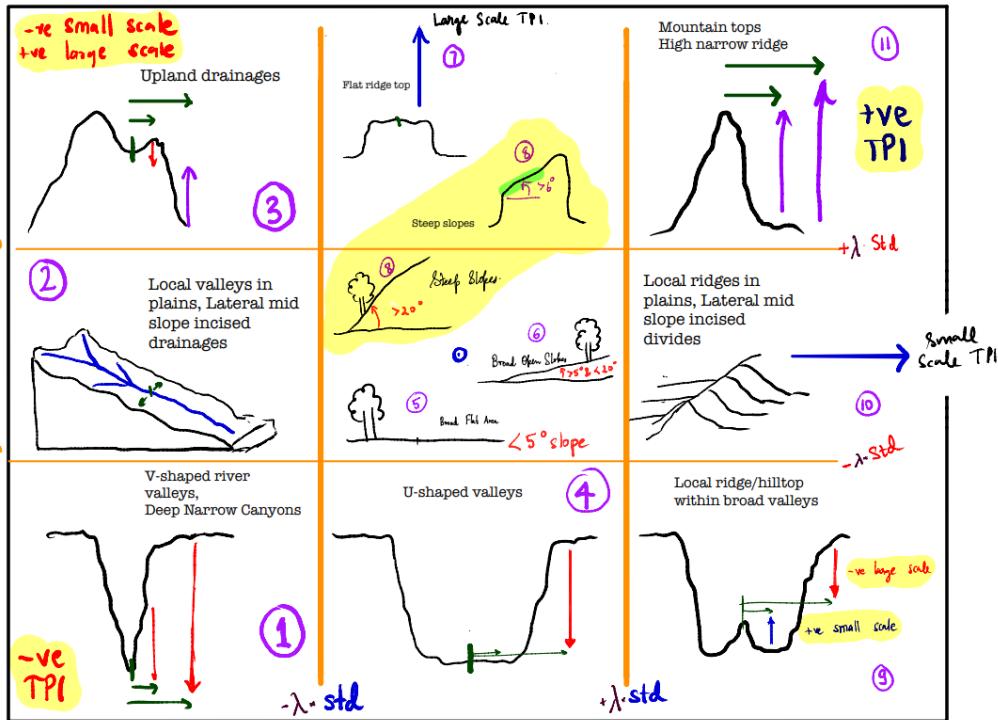


Figure 6.2: Landforms classification using slope and thresholds (denoted using  $\lambda$ ) on both small-scale and large scale TPIs

each fortnight, but the methodology has a limitation in being cross-sectional without considering past temporal values in each classification. For this reason, within-year variations are not captured well in this product.

We have developed an intra-annual LULC classifier that uses time-series data across all seasons to classify an area ( $10m \times 10m$  pixel) in terms of its cropping intensity (accuracy 83%). Further, the model is trained on regional data sampled from across all agro-climatic zones of the country and produces a higher accuracy than other data products even on standard LULC labels such as surface water, trees, crops, etc. - accuracy 94%). We further have an ongoing groundtruth data collection effort to validate the outputs in out-of-sample geographies. We use Google Earth Engine to produce annual outputs for any area of interest, since 2003 to the current date as shown in the figure 6.4.

### Input layers

To generate land use and land cover maps, we use multiple data sources as input-

- Satellite data- We use multi-spectral data from Landsat-7 [26], Landsat-8 [27], Sentinel-2 [24], MODIS [19], and Sentinel-1 [23] satellite constellations. Different combinations of this data are used to create feature vectors that serve as input to different machine learning algorithms. We use Google Earth Engine to access these datasets.
- Dynamic World- Google's land use and land cover product is used to identify certain land cover classes like built-up, water, etc. We access the Dynamic World data repository from Google Earth Engine [9].

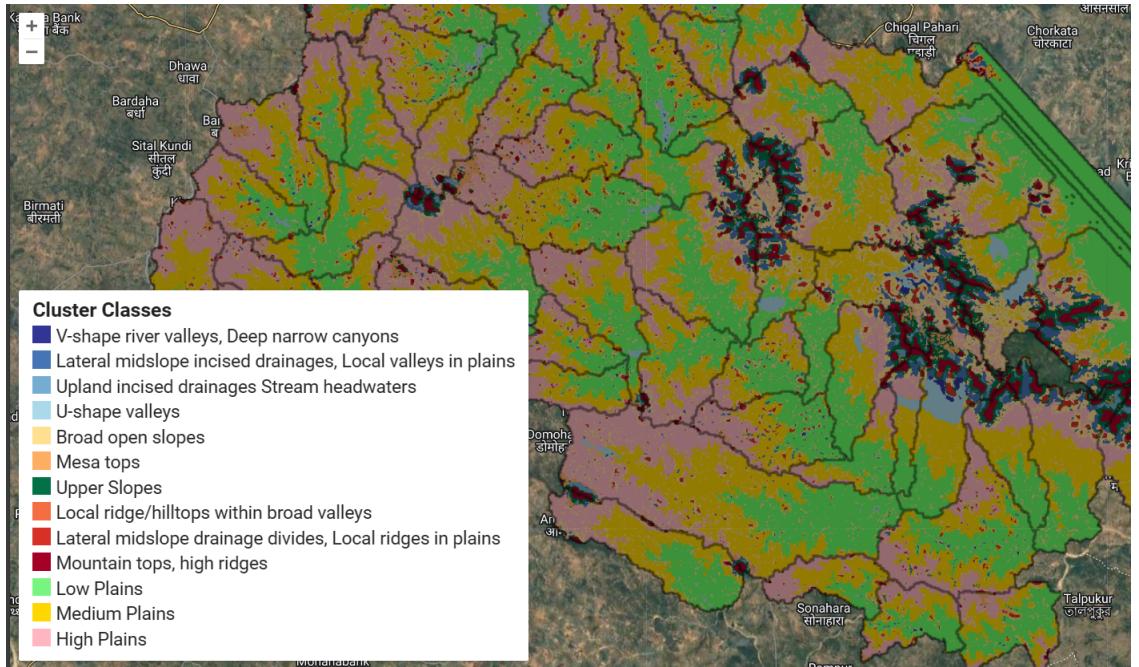


Figure 6.3: Landform classification of Masalia with micro-watershed boundaries in black.

- Open Street Maps- This data is used in the construction of the IndiaSAT groundtruth [21].
- Input Shapefiles- To run LULC for any region of interest, you require to give its shapefile (uploaded as a GEE asset) as input. For example- you can access different administrative boundaries in India at [2].

### Methodology

We perform a series of binary classifications followed by error-correction layers tailored for each of the final LULC classes under consideration. Each level of classification uses a different set of input features, datasets and methods.

- **Dynamic World based Classification:** We use Dynamic World to identify built-up regions, barren lands, and shrubs and scrubs. Segmentation-based classification methods perform well in identifying these classes and thus we use dynamic world as a proxy in our classification pipeline.
- **Surface Water Body Classification:** Our goal is to predict both seasonal and perennial water bodies and also give information on which agricultural season do these water bodies have water in. Existing LULC products like Global Surface Water and Dynamic World use optical satellite data to identify water bodies which under-performs in the Kharif season due to lower quality of optical data in monsoon season. To capture the seasonality of water bodies, we use SAR data from Sentinel-1 to identify pixels having water in kharif season and Dynamic World for Rabi and Zaid. A threshold based method on number of months predicting water in a particular agricultural season is used to identify water in all three seasons.

- **Cropland vs Tree Classification:** We classify the non-classified pixels further into 2 classes- Croplands and Forests/Trees. We use Sentinel-1 SAR data time series to perform this classification. The training data at this level is partly taken from the IndiaSAT groundtruth used in level-1 and is partly marked manually through visualization on Google Earth Pro. It is ensured that the groundtruth for both categories is geographically well distributed across different agro-climatic zones in India. We use a random forest classifier with 100 trees as the classification model that takes as input an annual 16-day time series of VV and VH bands from Sentinel-1. We further use Slope information from SRTM DEM [25] and use a threshold of 30 degrees to correct misclassifications in croplands, if any.
- **Cropping Frequency Classification:** To focus on agricultural sustainability applications, we classify cropland pixels into 4 categories based on their cropping frequency- Single Kharif, Single Non-Kharif, Double, and Triple cropping. To perform this classification, we did not have access to any training data. So, we used unsupervised classification using K-nearest neighbour algorithm. We randomly sampled cropland pixels from all agro-climatic regions in India (marked from the cropland groundtruth at level-2). The feature vector for this classification is a 16-day NDVI time series that is derived from a combination of Landsat-7, Landsat-8, Sentinel-2, and MODIS data. This time series is generated on the lines of GCI-30 paper [55]. We do not perform Whittaker smoothing to avoid missing out important peaks in the time series. Initially we created 16 clusters using KNN method and then hierarchically split them into 2 whenever the distortion exceeded the threshold of 0.23. We manually labeled each cluster by interpreting the spread of time series that belonged to that cluster. These clusters are then used to assign the classification label based on euclidean distances to their centroids.

### Hosting specifications

- Layer type: raster
- Spatial resolution: 10m
- Temporal resolution: yearly
- Codebase: [Github repository](#)

### 6.0.3 Cropping intensity

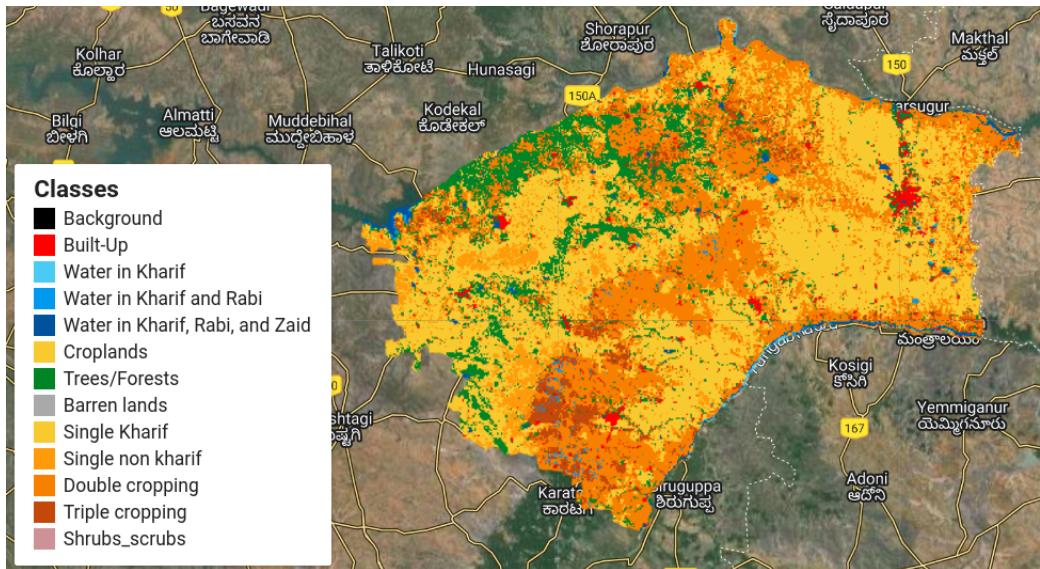
#### Introduction

Being able to classify regions in terms of cropping intensity is relevant for water security to understand changes in crop water usage over time. Most of India has three agricultural seasons in a year - during the monsoons, post-monsoon, and summer. The monsoon season is rain-fed in a large part of the country, whereas post-monsoon and summer seasons require irrigation from groundwater and other sources such as canals.

#### Input layers

We use annual land use land cover (LULC), to identify areas under single cropping, double cropping and triple cropping using pixels which are classified as single kharif, single non-kharif, double and triple classes of LULC classifier to determine cropping intensity.

Figure 6.4: Land use land cover classification of Raichur district in Karnataka for 2022



## Methodology

Using the LULC of a micro-watershed, we compute cropping intensity for a microwatershed in a year  $y$  using the following equation:

$$\begin{aligned} \text{Cropping intensity} = & 1 \times \% \text{ of area under single cropping} \\ & + 2 \times \% \text{ of area under double cropping} \\ & + 3 \times \% \text{ of area under triple cropping} \end{aligned} \quad (6.2)$$

where % of area under single cropping is computed as the summation of area under single kharif cropping and area under single non-kharif cropping divided by the total cropped area. Total cropped area is the union of cropping areas across the years. Area under a particular class is computed by the no. of pixels of that class multiplied by the pixel area. Similarly, for % of area under double cropping and % of area under triple cropping. 6.5

## Hosting specifications

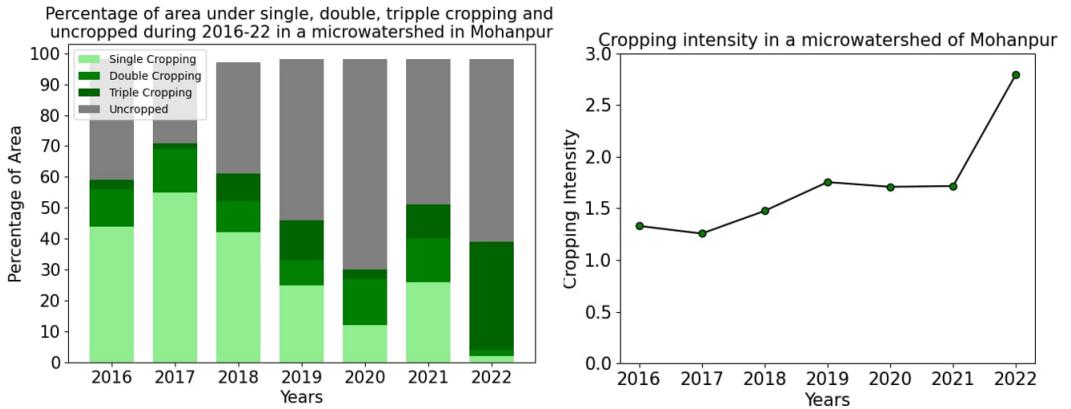
- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: annual
- Codebase: [Github repository](#)

### 6.0.4 Water bodies

#### Introduction

Knowledge about existing water storage structures, their capacity, water availability in different seasons, siltation and repair requirements, can give important indications about the

Figure 6.5: For a micro-watershed in Mohanpur, we analyze the percentage of area under single, double and triple cropping (left), and corresponding cropping intensity over the years.



resilience of a landscape to droughts and floods, especially with climate change.

Global Surface Water (GSW) [46] maps provide information about the surface area under water in different months, but we have found that GSW maps are effective only for large water bodies and they miss out on sub-5000m<sup>2</sup> water bodies. Some years back a waterbody census was conducted in India and a dataset has been recently released, but this is a static dataset and does not provide any real-time indications of the water availability in different seasons.

The LULC method we have developed is able to accurately capture even 1000-5000m<sup>2</sup> water bodies, and many sub-1000m<sup>2</sup> water bodies as well, with seasonal water availability. We are intersecting this with the waterbody census data and will then have an accurate waterbody capacity map as well. Using this, we will develop indicators for water resilience of a landscape.

## Input layers

To generate the water bodies vector layer for a region of interest, we use the LULC raster layer generated for that region to detect pixels with water class.

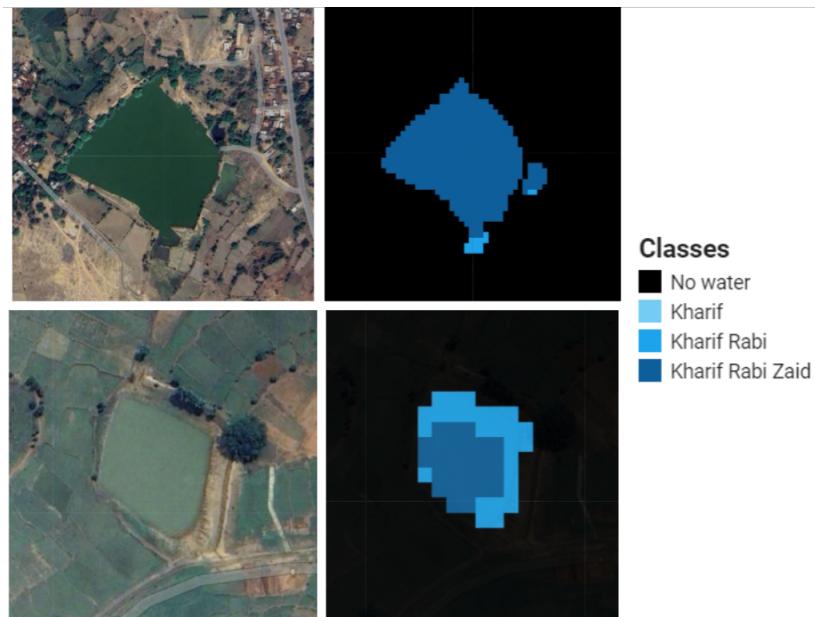
## Methodology

The hydrological year is divided into 3 agricultural seasons - Kharif (July-October), Rabi(November-February), Zaid (March-June). We use Sentinel-1 (SAR data) VV band for water pixel detection in Kharif season and Dynamic World to detect water pixels in Rabi and Zaid seasons. A threshold based method on number of months predicting water in a particular agricultural season is used to identify water in all three seasons leading to following classes per pixel as shown in the figure 6.6. Error correction was performed on top of it using Sentinel-2 bands to remove false positives of water pixels in croplands and barren lands. We perform the raster to vector conversion to generate a vector layer of water bodies with the following metadata for each water body: unique id, surface water area under Kharif, surface area under Rabi and surface area under Zaid, total area of the water body as shown in the figure 6.7.

### Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: yearly
- Codebase: [Github repository](#)

Figure 6.6: Water bodies as seen from satellite (first column) in Masalia block of Dumka district in Jharkhand and their detection using the classifier (second column)



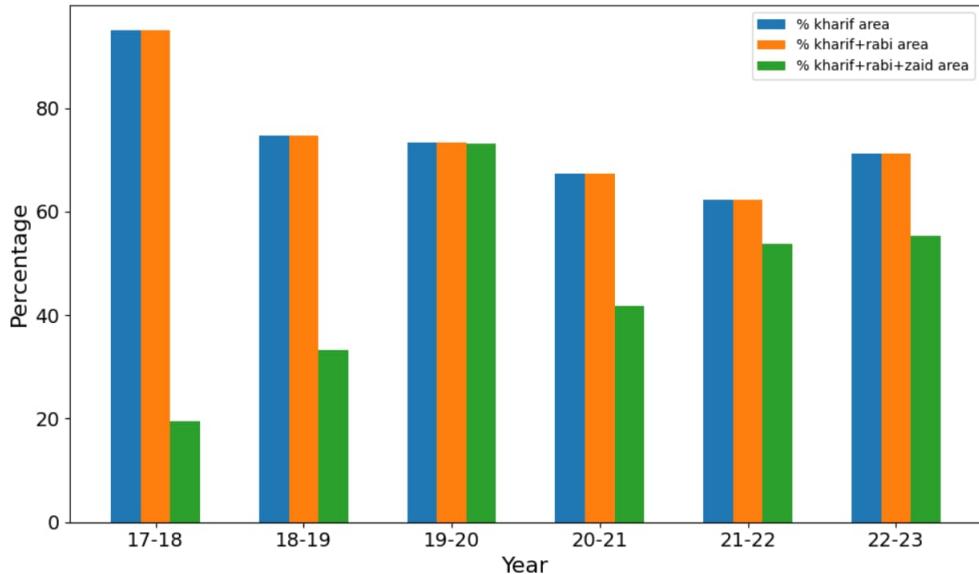
### 6.0.5 First census of water bodies

#### Introduction

The Ministry of Jal Shakti under Government of India performed the first census of water bodies [28] in 2017-18 with an objective develop a national database of all water bodies and their important properties such as storage capacity, status on whether they are functional, water spread area and type of use. However, the dataset is static and does not provide any real-time indications of the water availability in Kharif, Rabi and Zaid.

The LULC classification method mentioned in section 5.02 is able to accurately capture even  $1000\text{-}5000m^2$  water bodies, and many sub- $1000m^2$  water bodies as well, with seasonal water availability in Kharif, Rabi and Zaid. We are intersecting this with the waterbody census data to accurately generate waterbody capacity map as well. Using this, we will develop landscape-level indicators of water resilience. The capacity column of the water body can also be used to build site-level indicators of harvested runoff and infiltration.

Figure 6.7: Bar graph depicting the surface area of a water body (in percentage) in Mohanpur block of Gaya district in Bihar



### Input layers

The waterbody census data is published state-wise on Open Government Data (OGD) platform of India [11] in the form of CSV file.

### Methodology

We converted the state-wise CSV files of waterbody census into shapefiles using the geo-coordinate (latitude, longitude) of each waterbody as shown the figure 6.8.

### Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: static as of 2017-18
- Dataset: [Google drive folder](#)

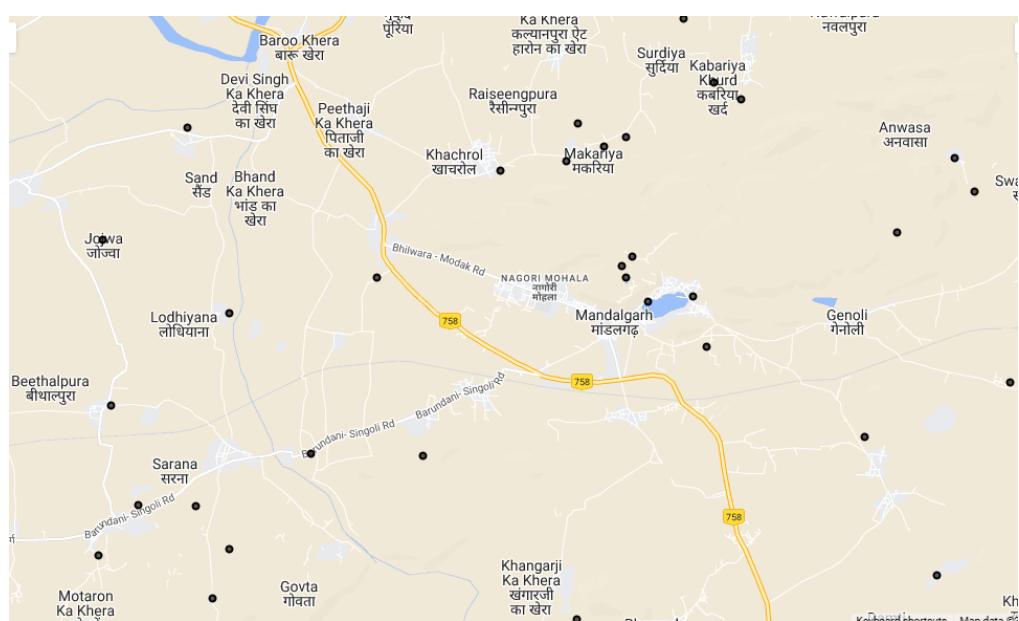


Figure 6.8: The map shows geo-coordinate of waterbodies from the waterbody census data of Rajasthan, specifically zoomed into Mandalgarh block of Bhilwara district for visualization.

---

# Welfare variables

## 7.0.1 NREGA assets categorization

### Introduction

The Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA) in India provides guaranteed wage employment to rural households for various types of public works. MGNREGS Master Circular for the financial year 2018–19 by Ministry of Rural Development, specifies 260 permissible works [15], including the construction and maintenance of water conservation assets. We categorized (as shown in table 7.1) 260 works from the lens of Natural Resource Management (NRM) where the works can provide site level and landscape level impact on agriculture, and their ability to generate livelihood. The categorization of existing works will be helpful in examining the current distribution of assets and can facilitate the planning of new assets to ensure water security and sustainable livelihoods in rural India.

Table 7.1: MGNREGA category and their corresponding keywords

Category	Keywords
Soil and water conservation	Bund, canal, channel, check dam, drainage structures, dykes, embankments, gully plugs, recharge pits, soak pit, soakage channels, spurs, terrace, trench, water courses, water drain
Irrigation on farms	Bunds, farm ponds, harvesting ponds, open well, percolation tank, stabilization pond
Other on-farm works	Sand filter
Plantations	Farm forestry, forest, horticulture, nursery, sericulture, shelter belt
Land restoration	Azola cultivation, biomanure, chaur land, compost, fallow land, levelling, storage building, waterlogged land
Off-farm livelihood assets	Cattle shelters, drying yards, fisheries pond, goat shelter, pigery shelter, poultry shelter, worksched
Community assets	Building, compound wall, crematorium, haat, play fields, roads, shelter, toilets

### Input layers

- MGNREGS permissible work list [18, 17].

- MGNREGA assets: We obtained geotagged MGNREGA assets from Bhuvan [6]. The geotagging of existing NREGA works began in September 2016. The data contains the assets built from January 2005 to November 2022. The metadata of assets such as work name, work type, expenditure and total person days employment provided for work was obtained from NREGA MIS[16]. Figure 7.2 shows metadata for available for an asset. Details of downloading and processing the asset data can be found here [22].

### Methodology

We employ keyword matching algorithm to categorize geotagged MGNREGA assets. The MGNREGS permissible work list uniquely identifies each work with its one line description [18]. For example, “Construction of Farm Ponds for Individuals” is a uniquely identified work in the permissible work list. For each permissible work, we extract keywords from its description and assign them to a category. The assignment results in a bag of words for each category as shown in table 7.1. In the above example, we mapped the “pond” keyword to “Irrigation on farms” category. The assigned keywords of each category are used to categorize a geotagged asset. The keywords present in the work type, work name and asset name columns (in order) of a geotagged asset were used to match the assigned keywords of each category. We further refined the bag of words for each category by adding stem words of the keywords and some region specific keywords in Hindi such as “pokhar” in Bihar which is used for pond while geotagging the asset. We performed random sampling of categorized geotagged assets from each category to manually examine the performance and the performance was found to be satisfactory. Figure ?? shows geotagged assets in Mohanpur block of Gaya district in Bihar, color coded based on their category.

### Hosting specifications

- Layer type: vector
- Spatial resolution: NA
- Temporal resolution: yearly
- Dataset: [Google drive folder](#)
- Codebase: [Github repository](#)

Figure 7.1: MGNREGA assets in Pindwara block of Sirohi district in Rajasthan

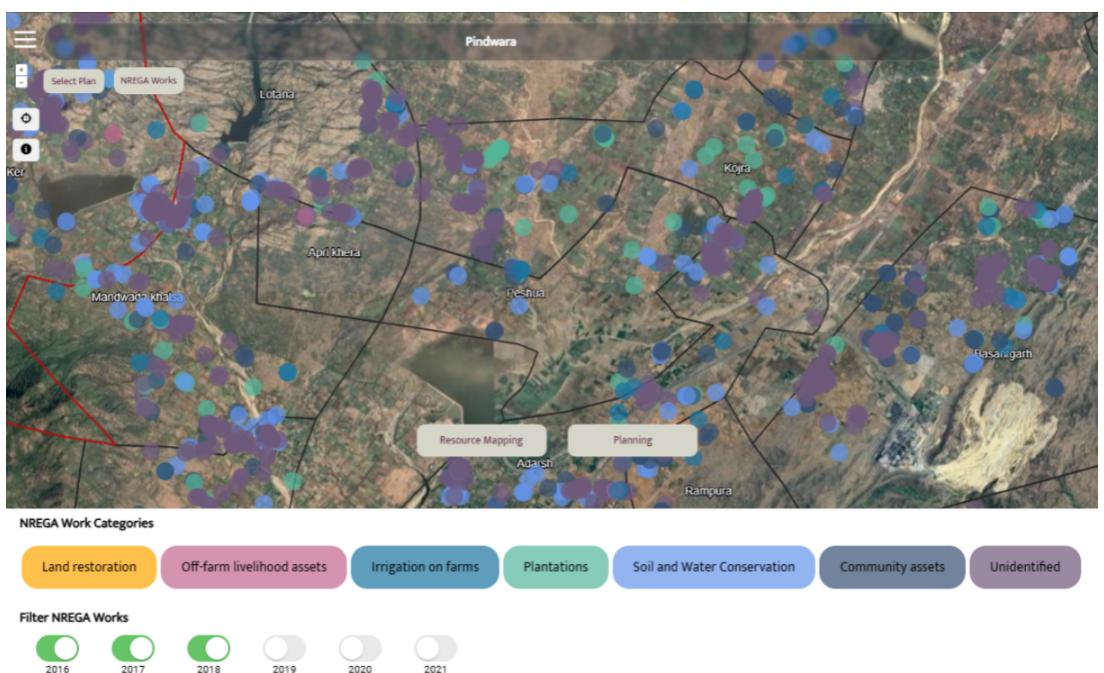


Figure 7.2: Metadata of a MGNREGA asset in Masalia block of Dumka district in Jharkhand

The screenshot shows a mobile application interface for viewing MGNREGA work information. At the top, there is a black header bar with the time '11:53' and various system icons (signal strength, battery level at 92%). Below this is a navigation bar with a back arrow labeled '< Home' and a 'Maps' button. The main content area has a title 'NREGA Work Information' and a close button ('X'). A table lists the following details:

<b>State</b>	JHARKHAND
<b>Panchayat</b>	ranga
<b>Panchayat ID</b>	40727
<b>Asset Name</b>	wc
<b>Work Name</b>	DOVA NIRMANA AT LATABAR OF POLAK CHANDRA MANDAL
<b>Work Type</b>	Renovation of Fisheries Ponds for Community
<b>Work Category</b>	Household Livelihood
<b>Total Expenditure</b>	45325
<b>Latitude</b>	24.11516017
<b>Longitude</b>	87.14008435
<b>End Location</b>	-1
<b>Unskilled Labor</b>	0
<b>Semi-Skilled Labor</b>	2221
<b>Estimated Cost</b>	0.46108
<b>Start Location</b>	-1
<b>Work Start Date</b>	02/02/2017

At the bottom of the table, there is a grey footer bar with the text 'Latitude : 28.55 Longitude : 77.19' and a small location icon. The footer also includes standard Android navigation icons: three horizontal dots, a square, and a triangle.

---

# Site assessment

## 8.0.1 Lineament

### Introduction

In geology, lineaments refer to linear features on the Earth's surface such as faults, fractures, and joints. Lineament is also an important hydrological variable because of its potential to provide groundwater recharge [34]. The groundwater recharge occurs due to the infiltration of rainfall and runoff through the linear features [35].

### Input Layers

The state-wise lineaments data has been taken from Bhuvan [7] by sending out a WMS request and generating raster files from the packets.

### Methodology

The raster files of lineament are processed to generate a proximity mask. The Proximity (Raster Distance) algorithm on QGIS computes the distance from the center of each pixel to the center of the nearest pixel on a target pixel (a lineament in this case). The generated raster proximity raster is now scanned for pixels with a distance of less than a buffer value (currently 2m) and selected pixels are exported as raster to form a lineaments buffer mask. This mask is now georeferenced to the respective state for ease of visualization. Lineaments of Mandalgarh block is shown in the figure 8.4a.

### Hosting specifications

- Layer type: raster
- Spatial resolution: 194m.
- Temporal resolution: static
- Dataset: [Google drive folder](#)

## 8.0.2 Lithology

### Introduction

Lithology is a branch of geology that encompasses the study of various types of rocks, based on their composition, texture, structure, and other properties. The porosity and permeability of rocks influences the groundwater recharge [52]. For example, sandstone has much higher porosity and hence have higher groundwater prospect as compared to granulite rock which is hard rock and has compact structure [45].

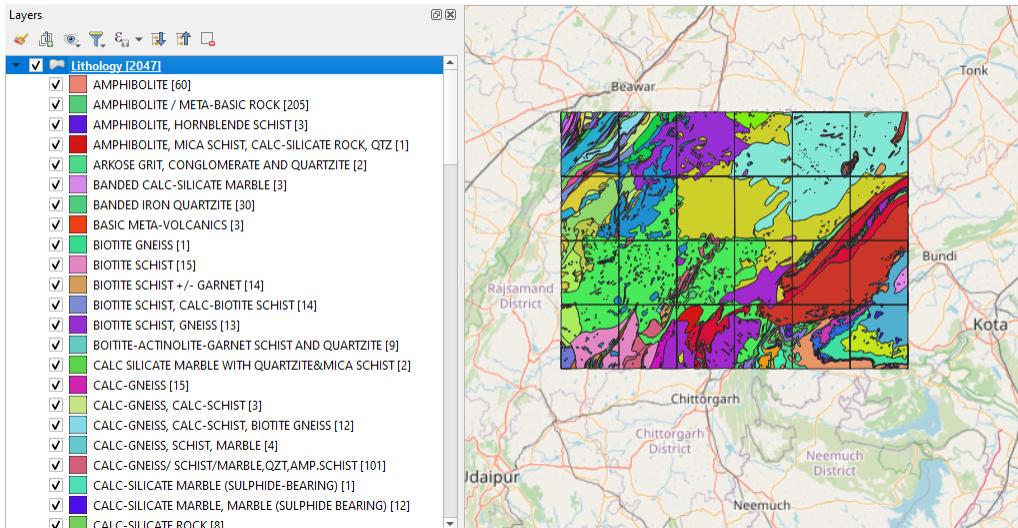
## Input Layers

The state-wise lithology data has been downloaded manually from Bhukosh [5] hosted by Geological Survey of India in the form of shapefile. The lithology shapefile from Bhukosh for Bhilwara district of Rajasthan is shown in the figure ??.

### Hosting specifications

- Layer type: vector
- Spatial resolution: NA.
- Temporal resolution: static
- Dataset: [Google drive folder](#)

Figure 8.1: Lithology shapefile from Bhukosh for Bhilwara district of Rajasthan



### 8.0.3 CLART

#### Introduction

The Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA), promises unskilled manual work to every adult for at least 100 days in rural parts of the country. To ensure water security in the rural landscape assets such as check dams, percolation tanks, and irrigation channels are constructed and maintained. The overall focus of the scheme is on ensuring higher incomes for farmers by increasing the water availability and productivity of land. The existing method of suitable site identification for these structures has been based on local people's opinion, and taking into consideration factors such as geomorphology, climate, annual rainfall, vegetation cover, distances from farms and so on. However, misidentification of suitable locations for building water assets has resulted in non-operation, non-utilization and inefficient investment. Therefore, there is a need for site suitability assessment in terms of whether the site provides groundwater recharge or surface water storage or has a high runoff.

The India Observatory (IO), the technological branch of the Foundation for Ecological Security (FES), has created Composite Land Assessment & Restoration Tool (CLART) with the goal of empowering rural communities for planning the sites for soil and water conservation measures. The primary objective is to facilitate the restoration of degraded ecosystems and enhance economic opportunities in rural areas. It achieves this by providing access to geospatial datasets such as rock type, slope and landcover in a geographic information system (GIS) based android application. We intend to reproduce CLART in a reprogrammable/reconfigurable manner to allow dynamic scoring based on community requirements for their landscape.

## Input Layers

- **Lineaments:** as mentioned in section 7.0.1.
- **Lithology:** as mentioned in section 7.0.2.
- **Drainage density:** Drainage network of the watershed helps in visualizing which areas have a high groundwater recharge. Drainage density is directly related to slope and inversely proportional to permeability. The steeper the slope with low permeability, the higher the drainage density, thus less infiltration and more surface runoff. We used QGIS Line Density tool, which calculates drainage density for each pixel as the sum of lengths of the drainage lines in the search radius weighted by their stream orders. We used elevation raster [25] as input to Line Density tool with 1km and 10m as the search radius and pixel size respectively.
- **Slope percentage:** Slope represents the rate of change of elevation for each DEM pixel. Steep slopes generally reduce recharge as runoff flows very rapidly and would not permit infiltration. Plains, however, enhance groundwater recharge because higher retention time is provided for rainwater to infiltrate the soils. We used SRTM DEM [25] to obtain pixel-wise slope on Earth Engine.

## Methodology

We processed the input layers above to generate pixel level score for each layer as shown below:

- The Central Ground Water Board (CGWB) document defines rainfall infiltration factor (RIF) for each major aquifer [20]. Each lithologic group in the shapefile was mapped to one of the major aquifers using the string matching algorithm to assign scores of recharge. The range of RIF was equally binned into three bins to assign scores of 1 (low recharge), 2 (moderate recharge), 3 (high recharge). The higher the RIF, the more the recharge. The shapefiles were then rasterized with the recharge score as the burn in value.
- The observed range of drainage density values was binned into three equal parts to assign scores of 1 (high recharge), 2 (moderate recharge), 3 (low recharge). The lower the drainage density, the more the no. of higher stream orders that enable more groundwater recharge.
- The lineament layer was processed to generate a proximity mask by taking a 2m buffer around the lineament to assign a score of 10 within the buffer and a score of 1 outside the buffer.

Recharge potential is a measure of the ability of an area to recharge the groundwater. The scores of lithology, drainage density and lineaments are multiplied to determine the recharge potential as shown in the equation 18 and table 8.1. Recharge potential is combined with different slope percentages and land cover categories to derive treatment codes as shown in the figure 8.2. The flowchart of methodology is shown in the figure 8.3, where community based score reassignment can be performed to produce CLART output accordingly.

$$\text{Recharge potential} = \text{Lineament} * \text{Lithology} * \text{Drainage density} \quad (8.1)$$

where lineament scores can have values of 1, 10. Lithology and drainage density scores can have values of 1, 2, 3 resulting the table 8.1.

### Hosting specifications

- Layer type: raster
- Spatial resolution: 30m
- Temporal resolution: static
- Codebase: [Github repository](#)

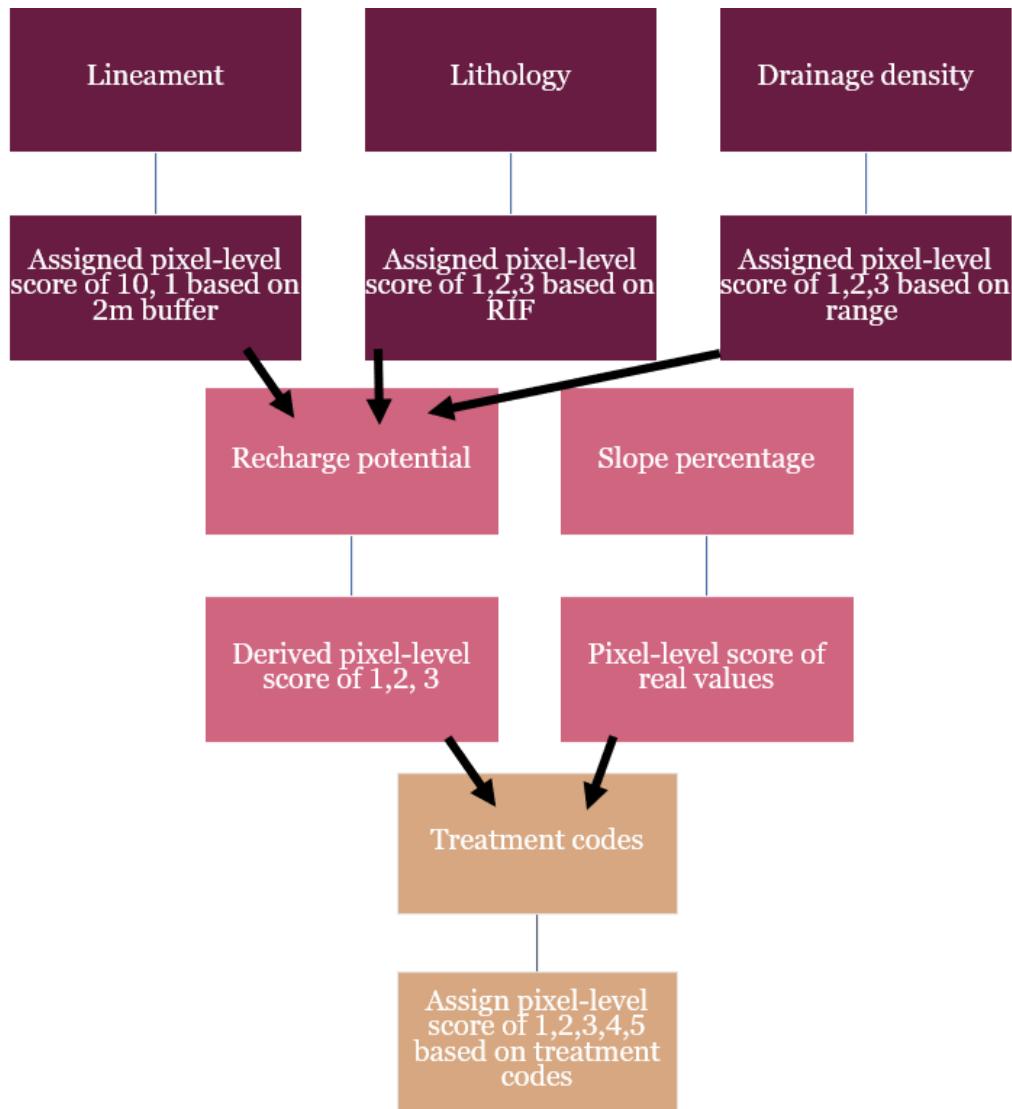
Table 8.1: Recharge potential values and their corresponding scores for recharge

Recharge potential values	Score
1,2,10,20,30,40,60,90	1 (high recharge)
3,4	2 (moderate recharge)
6,9	3 (low recharge)

Figure 8.2: The recharge potential scores are combined with different slope percentages and land cover categories to derive treatment codes in CLART.

Recommended Treatment Code	Recommended Treatment Type	Recharge Potentiality	Slope	Land Use/Land Cover
<b>1</b>	Good Recharge structure (Percolation tank, WHS, CCT etc)	Very High (5) High (4)	3-5 % (2) 5-7% (3)	Current fallow (5), Other Waste land (9), Gullied (10), Scrubland (11)
<b>2</b>	Moderate Recharge structure (WAT, GP, LBCD etc)	Moderate (3)	0-3% (1), 3-5 % (2) 5-7% (3)	Current fallow (5), Other Waste land (9), Gullied (10), Scrubland (11)
<b>3</b>	Surface water Harvesting structure (WHS, FP, FB etc)	Low (2) Very Low (1)	0-3% (1), 3-5 % (2)	Current fallow (5), Other Waste land (9), Gullied (10), Scrubland (11)
<b>9</b>	Regeneration (Plantation, Veg Int, Grass seeding, stone bunding, bench terracing etc)	1,2,3	Slope >10%	Current fallow (5), Other Waste land (9), Gullied (10), Scrubland (11) Mixed, degraded forest, Deciduous forest

Figure 8.3: Flowchart of methodology to produce reprogrammable CLART using score assignment



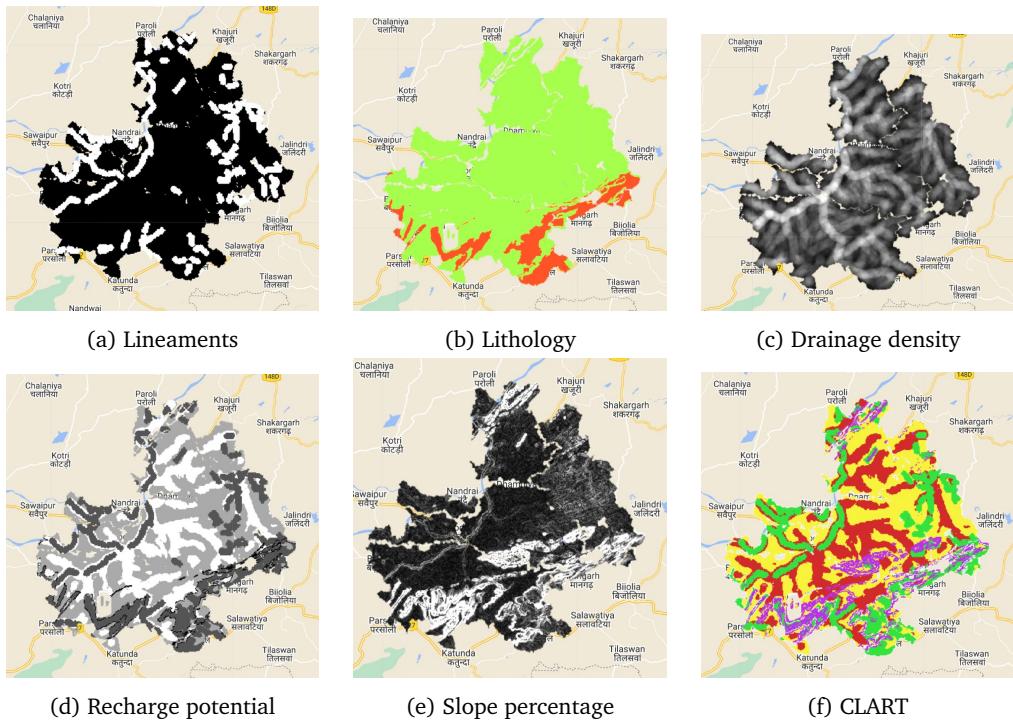


Figure 8.4: The figure shows the component layers for generating CLART output for Mandalgarh block in Bhilawara district of Rajasthan



---

# Administrative boundaries

Administrative boundaries in India refer to the demarcated lines that separate different administrative regions within the country, established to facilitate the governance and administration of various administrative units such as states and districts. The administrative units of India is hierarchical where India is divided into 28 states and 8 union territories. States are subdivided into districts, districts are sub-divided into blocks, blocks are sub-divided into panchayats and panchayats being further sub-divded into villages.

## 9.0.1 State

The shapefile represents the administrative boundary of Indian states. The Survey of India provides data upto district level. The Survey of India is the country's National Survey and Mapping agency under the Department of Science and Technology, Govt. of India. The dataset can be downloaded from [here](#). State boundaries of India are shown in the figure [9.1](#)

## 9.0.2 District

The shapefile represents the administrative boundary of Indian districts. The Survey of India provides data upto district level. The Survey of India is the country's National Survey and Mapping agency under the Department of Science and Technology, Govt. of India. The dataset can be downloaded from [here](#). District boundaries of India are shown in the figure [9.1](#)

## 9.0.3 Block

The block shape files for 15 Indian states were generated by performing a union over the constituent village shape files for each block and can be found [here](#). The village shapefiles for the year 2001 were obtained from the [NASA SEDAC website](#). The village to block mapping was obtained from the [Local Government Directory data](#). The data was processed to store the mapping of villages to a block as shown [here](#). Block boundaries of India are shown in the figure [9.1](#)

## 9.0.4 Panchayat

The panchayat shape files for 15 Indian states were generated by performing a union over the constituent village shape files for each panchayat and can be found [here](#). The village shapefiles for the year 2001 were obtained from the [NASA SEDAC website](#). The village to panchayat mapping was obtained from the [Local Government Directory data](#). The data was processed to store the mapping of villages to a panchayat as shown [here](#). Panchayat boundaries of Mohanpur block, in Gaya district of Bihar are shown in the figure [9.2](#)

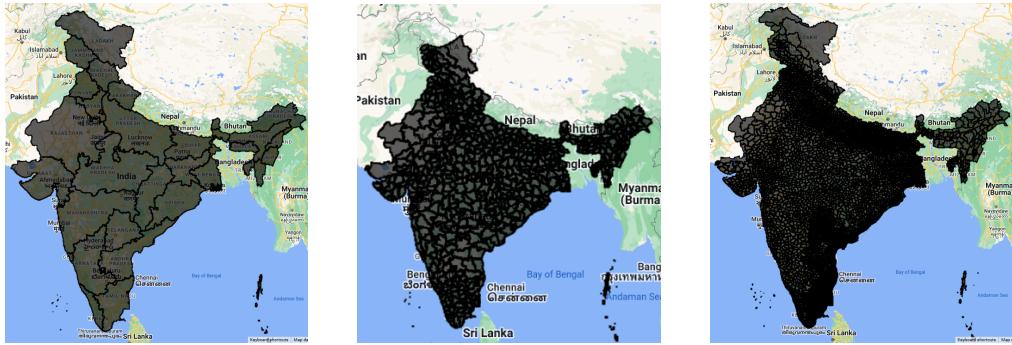


Figure 9.1: State, district and block boundaries (from left to right) of India

### 9.0.5 Village

Shapefiles at the village level for the year 2001 were obtained from the [NASA SEDAC website](#). Village shapefiles hosted on NASA SEDAC website were generated by digitizing the village boundaries from the official analog maps published by the Survey of India [42] for 2001 to compile the socio-economic variables [41] collected during the Indian Census in 1991 and 2001 at village level.

The [folder](#) contains state wise village shapefiles compiled with village level household variables of census 2011 (as shown [here](#)) and Aggregate Development Index (ADI) [37] of 2011 and 2019. A socio-economic development index called the Aggregate Development Index (ADI) [30] was developed, which is composed of indicators related to asset ownership, primary source of drinking water, primary source of lights, access to bathroom facilities, and literacy rate [37] at a village level. ADI ranges from 5 to 15, with a higher ADI indicating better socio-economic development. Village boundaries of Mohanpur block, in Gaya district of Bihar are shown in the figure [9.2](#)

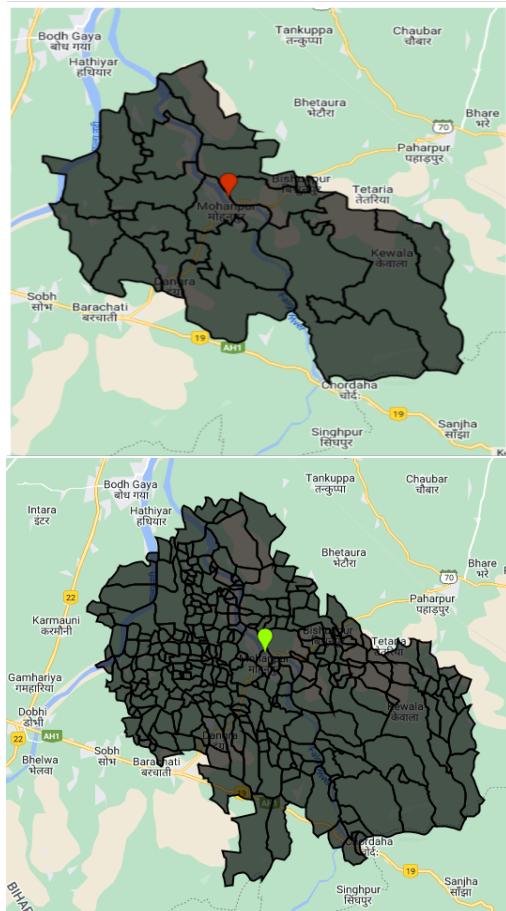


Figure 9.2: Panchayat (top) and village (bottom) boundaries of Mohanpur block, in Gaya district of Bihar



---

## Bibliography

- [1] NSIDC DAAC's Data Access and Service API. <https://github.com/nsidc/NSIDC-Data-Access-Notebook>.
- [2] Administrative Boundaries of India. [https://drive.google.com/drive/folders/1HJef1T8wJg8eoZ4kUU8u1HHNoH3510Zv?usp=drive\\_link](https://drive.google.com/drive/folders/1HJef1T8wJg8eoZ4kUU8u1HHNoH3510Zv?usp=drive_link).
- [3] Aquifer properties and decription. [https://csciitd-my.sharepoint.com/:x/g/personal/anz208490\\_iitd\\_ac\\_in/EYzjXM0tf1BCjPH5x5IwtDgBEPYuGu-VTNi655cHIpWZsg?e=Sk643b](https://csciitd-my.sharepoint.com/:x/g/personal/anz208490_iitd_ac_in/EYzjXM0tf1BCjPH5x5IwtDgBEPYuGu-VTNi655cHIpWZsg?e=Sk643b).
- [4] Aquifer Systems of India, Central Ground Water Board, Ministry of Water Resources, Government of India. <https://www.cgwb.gov.in/cgwpnm/public/uploads/documents/1687419512680023437file.pdf>.
- [5] Bhukosh, Geological Survey of India. <https://bhukosh.gsi.gov.in/Bhukosh/Public>.
- [6] Bhuvan GEO-MGNREGA. [https://bhuvan-app2.nrsc.gov.in/mgnrega/nrega\\_dashboard\\_phase2/](https://bhuvan-app2.nrsc.gov.in/mgnrega/nrega_dashboard_phase2/).
- [7] Bhuvan, Indian Geo-Platform of ISRO, NRSC. <https://bhuvan-app1.nrsc.gov.in/thematic/thematic/index.php>.
- [8] Chapter 2 - FAO Penman-Monteith equation. <https://www.fao.org/3/x0490e/x0490e06.htm>.
- [9] Dynamic World V1. [https://developers.google.com/earth-engine/datasets/catalog/GOOGLER\\_DYNAMICWORLD\\_V1](https://developers.google.com/earth-engine/datasets/catalog/GOOGLER_DYNAMICWORLD_V1).
- [10] Famine Early Warning Systems Network (FEWS NET) Land Data Assimilation System FLDAS Noah Land Surface Model L4 Central Asia Daily 0.01 x 0.01 degree (FLDAS\_NOAH001\_G\_CA\_D). [https://disc.gsfc.nasa.gov/datasets/FLDAS\\_NOAH001\\_G\\_CA\\_D\\_001/summary](https://disc.gsfc.nasa.gov/datasets/FLDAS_NOAH001_G_CA_D_001/summary).
- [11] First Census of Water Bodies Data, Open Government Data (OGD) platform of India. <https://data.gov.in/catalog/first-census-water-bodies-data>.
- [12] Global Hydrologic Soil Groups (HYSGOs250m) for Curve Number-Based Runoff Modeling, ORNL DAAC. [https://daac.ornl.gov/SOILS/guides/Global\\_Hydrologic\\_Soil\\_Group.html](https://daac.ornl.gov/SOILS/guides/Global_Hydrologic_Soil_Group.html).
- [13] Global Satellite Mapping of Precipitation (GSMap), JAXA Earth Observation Research Center. [https://developers.google.com/earth-engine/datasets/catalog/JAXA\\_GPM\\_L3\\_GSMaP\\_v6\\_operational](https://developers.google.com/earth-engine/datasets/catalog/JAXA_GPM_L3_GSMaP_v6_operational).

- [14] Manual for Drought Management 2016, Ministry of Agriculture and Farmers Welfare, Government of India. <https://sdma.cg.gov.in/ManualDrought2016.pdf>.
- [15] MASTER CIRCULAR A GUIDE FOR PROGRAMME IMPLEMENTATION FY 2018-2019. [https://nregs.kerala.gov.in/mgnregs/wp-content/uploads/2021/01/AMC\\_18-19.pdf](https://nregs.kerala.gov.in/mgnregs/wp-content/uploads/2021/01/AMC_18-19.pdf).
- [16] MGNREGA MIS. <https://mnregaweb4.nic.in/netnrega/error.html?aspxerrorpath=/netnrega/MISreport4.aspx>.
- [17] MGNREGA work category. [https://csciitd-my.sharepoint.com/:x/g/personal/ramneekk\\_iitd\\_ac\\_in/EfM9Wta6TVpAnxbMADWb5ZIBg8fIVp9NhHmfHOLjmgMnlA?rtime=DBETvY\\_V20g](https://csciitd-my.sharepoint.com/:x/g/personal/ramneekk_iitd_ac_in/EfM9Wta6TVpAnxbMADWb5ZIBg8fIVp9NhHmfHOLjmgMnlA?rtime=DBETvY_V20g).
- [18] MGNREGS Permissible Work List. <https://megsres.nic.in/sites/default/files/mgnrega-permissible-work-list.pdf>.
- [19] MODIS Collections in Earth Engine. <https://developers.google.com/earth-engine/datasets/catalog/modis>.
- [20] National Compilation on Dynamic Ground Water Resources of India, 2022. [https://static.pib.gov.in/WriteReadData/userfiles/file/GWRA2022\(1\)HIDO.pdf](https://static.pib.gov.in/WriteReadData/userfiles/file/GWRA2022(1)HIDO.pdf).
- [21] Open Street Maps. <https://www.openstreetmap.org/>.
- [22] Procedure to download and process NREGA geotagged assets. [https://drive.google.com/file/d/1vSi8c-orzVuZH9Zd3-\\_X8Z-vfukrIDE/view?usp=sharing](https://drive.google.com/file/d/1vSi8c-orzVuZH9Zd3-_X8Z-vfukrIDE/view?usp=sharing).
- [23] Sentinel-1 SAR GRD. [https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS\\_S1\\_GRD](https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD).
- [24] Sentinel-2 MSI. [https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS\\_S2\\_SR](https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR).
- [25] SRTM Digital Elevation Data Version 4. [https://developers.google.com/earth-engine/datasets/catalog/CGIAR\\_SRTM90\\_V4](https://developers.google.com/earth-engine/datasets/catalog/CGIAR_SRTM90_V4).
- [26] USGS Landsat 7 Level 2. [https://developers.google.com/earth-engine/datasets/catalog/LANDSAT\\_LE07\\_C02\\_T1\\_L2](https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LE07_C02_T1_L2).
- [27] USGS Landsat 8 Level 2. [https://developers.google.com/earth-engine/datasets/catalog/LANDSAT\\_LC08\\_C02\\_T1\\_L2](https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C02_T1_L2).
- [28] Water Bodies, First Census Report, Ministry of Jal Shakti, Government of India. <https://cdnbbsr.s3waas.gov.in/s3a70dc40477bc2adceef4d2c90f47eb82/uploads/2023/04/2023040672.pdf>.
- [29] Watershed Development for Rainfed Areas under MGNREGS MGNREG, Designed and Developed by WASSAN as part of GIZ Project "Environmental benefits of the Mahatma Gandhi National Rural Employment Guarantee Act". <https://e-saksham.nic.in/wp-content/uploads/2018/07/Watershed-development-for-rainfed-areas-under-MGNREGS-A-resource-booklet-for-pdf>.

- [30] Chahat Bansal, Arpit Jain, Phaneesh Barwaria, Anuj Choudhary, Anupam Singh, Ayush Gupta, and Aaditeshwar Seth. Temporal prediction of socio-economic indicators using satellite imagery. In *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*, pages 73–81. 2020.
- [31] Ritu Bharadwaj, Simon Addison, and Mohan Reddy. Climate resilience information system and planning tool for mahatma gandhi national rural employment guarantee scheme: the crisp-m tool, 2021.
- [32] Kantha Rao Bhimala, GK Patra, and Sheshakumar Goroshi. Annual and seasonal trends in actual evapotranspiration over different meteorological sub-divisions in india using satellite-based data. *Theoretical and Applied Climatology*, 152(3-4):999–1017, 2023.
- [33] Chao Chen, Xinyue He, Zhisong Liu, Weiwei Sun, Heng Dong, and Yanli Chu. Analysis of regional economic development based on land use and land cover change information derived from landsat imagery. *Scientific Reports*, 10(1):12721, 2020.
- [34] Sumit Das and Sudhakar D Pardeshi. Comparative analysis of lineaments extracted from cartosat, srtm and aster dem: a study based on four watersheds in konkan region, india. *Spatial Information Research*, 26:47–57, 2018.
- [35] Arjun B Doke, Rajendra B Zolekar, Hemlata Patel, and Sumit Das. Geospatial mapping of groundwater potential zones using multi-criteria decision-making ahp approach in a hardrock basaltic terrain in india. *Ecological Indicators*, 127:107685, 2021.
- [36] Sheshakumar Goroshi, Rohit Pradhan, Raghavendra P Singh, KK Singh, and Jai Singh Parihar. Trend analysis of evapotranspiration over india: Observed from long-term satellite measurements. *Journal of Earth System Science*, 126:1–21, 2017.
- [37] Anant Gulgulia, Aman Gupta, Akshay P Sarashetti, Aaditya Sinha, and Aaditeshwar Seth. Tracking socio-economic development in rural india over two decades using satellite imagery. *ACM J. Comput. Sustain. Soc.*, 1(2), dec 2023.
- [38] P K Gupta, Sam Chauhan, and M P Oza. Modelling surface run-off and trends analysis over india. *Journal of Earth System Science*, 125:1089–1102, 2016.
- [39] Stephanie K Kampf, Stephen J Burges, John C Hammond, Aditi Bhaskar, Tim P Covino, Abby Eurich, Hannah Harrison, Michael Lefsky, Caroline Martin, Daniel McGrath, et al. The case for an open water balance: Re-envisioning network design and data analysis for a complex, uncertain world. *Water Resources Research*, 56(6):e2019WR026699, 2020.
- [40] Amy McNally, Kristi Arsenault, Sujay Kumar, Shraddhanand Shukla, Pete Peterson, Shugong Wang, Chris Funk, Christa D Peters-Lidard, and James P Verdin. A land data assimilation system for sub-saharan africa food and water security applications. *Scientific data*, 4(1):1–19, 2017.
- [41] Prasanth Meiyappan, Parth S Roy, Yeshu Sharma, Reshma M Ramachandran, Pawan K Joshi, Ruth S DeFries, and Atul K Jain. Dynamics and determinants of land change in india: integrating satellite data with village socioeconomics. *Regional Environmental Change*, 17:753–766, 2017.
- [42] Prasanth Meiyappan, Parth S Roy, Aiman Soliman, Ting Li, Pinki Mondal, Shaowen Wang, and Atul K Jain. India village-level geospatial socio-economic data set: 1991, 2001. *Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC)*, 10:H4CN71ZJ, 2018.

- [43] Surendra Kumar Mishra, Manoj Kumar Jain, RP Pandey, and VP Singh. Catchment area-based evaluation of the amc-dependent scs-cn-based rainfall-runoff models. *Hydrological Processes: An International Journal*, 19(14):2701–2718, 2005.
- [44] J.L. Monteith. Evaporation and environment. *Symposia of the Society for Experimental Biology*, 19:205 – 234, 1965. Cited by: 4759.
- [45] Preetilata Murmu, Mukesh Kumar, Deepak Lal, Irjesh Sonker, and Sudhir Kumar Singh. Delineation of groundwater potential zones using geospatial techniques and analytical hierarchy process in dumka district, jharkhand, india. *Groundwater for Sustainable Development*, 9:100239, 2019.
- [46] Jean-François Pekel, Andrew Cottam, Noel Gorelick, and Alan S Belward. High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633):418–422, 2016.
- [47] Charles Henry Brian Priestley and Robert Joseph Taylor. On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly weather review*, 100(2):81–92, 1972.
- [48] ISRO SAC. Desertification and land degradation atlas of india (based on irs awifs data of 2011-13 and 2003-05). *Ahmedabad: Space Applications Centre, ISRO, Ahmedabad, India*, 219, 2016.
- [49] Diego Salazar-Martínez, Friso Holwerda, Thomas RH Holmes, Enrico A Yépez, Christopher R Hain, Susana Alvarado-Barrientos, Gregorio Ángeles-Pérez, Tulio Arredondo-Moreno, Josué Delgado-Balbuena, Bernardo Figueroa-Espinoza, et al. Evaluation of remote sensing-based evapotranspiration products at low-latitude eddy covariance sites. *Journal of Hydrology*, 610:127786, 2022.
- [50] U Scs. Urban hydrology for small watersheds, technical release no. 55 (tr-55). *US Department of Agriculture, US Government Printing Office, Washington, DC*, 1986.
- [51] Andrew N Sharpley. Epic-erosion/productivity impact calculator: 1, model documentation. *USDA Techn. Bull. 1759*, 235, 1990.
- [52] Prafull Singh, Jay Krishna Thakur, and UC Singh. Morphometric analysis of morar river basin, madhya pradesh, india, using remote sensing and gis techniques. *Environmental earth sciences*, 68(7):1967–1977, 2013.
- [53] Kaicun Wang and Robert E Dickinson. A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability. *Reviews of Geophysics*, 50(2), 2012.
- [54] Andrew D. Weiss. Topographic position and landforms analysis. 2001.
- [55] Miao Zhang, Bingfang Wu, Hongwei Zeng, Guojin He, Chong Liu, Shiqi Tao, Qi Zhang, Mohsen Nabil, Fuyou Tian, José Bofana, et al. Gci30: a global dataset of 30 m cropping intensity using multisource remote sensing imagery. *Earth System Science Data*, 13(10):4799–4817, 2021.