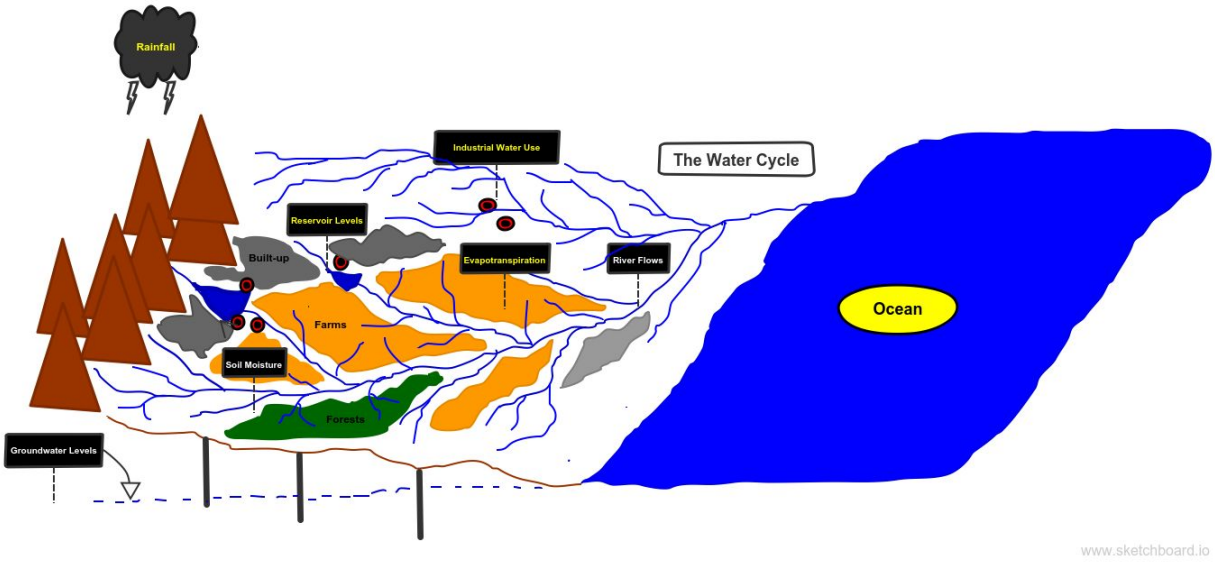
Open Water Data in India

December, 2017

Datameet & CIS



Contents

Glossary **4**

1.Open data in India **5**

2.Why the need for 'Open water data' **6**

3.Open datasets for water **7**

Water Availability 7

Water Use/Demand 7

3.1 Supporting Data 8

3.1.1 River network 8

Government - India (CWC) 8

Government - U.S. (USGS) 9

Community - Open Street Maps 9

QuickOSM - QGIS Plugin 9

Trimble Data 10

Geofabrik 12

Overpass Turbo 12

3.1.2 Watershed/River basin boundaries 13

USGS Hydrosheds 13

CWC's Watershed Atlas 13

Soil and Land Use Survey of India 15

Central Groundwater Board 16

3.1.3 Water reservoirs (dams, lakes -natural and man-made) 16

GRanD Database - Global Reservoir and Dam Database (2008) 16

Global Surface Water Explorer - European Commission - Copernicus Programme 16

Community - Open Street Maps 17

Trimble Data - Water Bodies 17

Open Street Map Data for Water 18

3.1.4 Elevation data 18

NASA - Shuttle Radar Topographic Mission (SRTM) 19

NASA - Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) 19

JAXA - Advanced Land Observing Satellite (ALOS) 20

NASA - GTOPO30 20

NRSC - CARTO-SAT 21

3.2 Water Availability 23

3.2.1 Rainfall data 23

Indian Meteorological Department 23

Customised Rainfall Information System of India (CRIS) 23

Automatic Weather Stations (AWS) 24
Raw IMD Station Data 24
Modeled IMD Data 24
Global Datasets 25
Tropical Rainfall Monitoring Mission (TRMM) 26
Global Precipitation Mission (GPM) 27
Climate Hazards Group Infrared Precipitation with Station Data (CHIRPS) 28
Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN) 29
Global Historical Climate Network: 29
Precipitation Reconstruction - Land (PREC/L) 31
Global Land Precipitation and Temperature (University of Delaware) 31
Global Precipitation Climatology Centre (GPCC) Datasets 32
Climatologies at High Resolution for Earth's Surface Area (CHELSA) 32
3.2.2 Soil moisture 32
ESA - CCI Soil Moisture 32
European Space Agency - Soil Moisture Ocean Salinity (SMOS) 33
NASA - Soil Moisture Active Passive (SMAP) 33
3.2.3 Groundwater Levels 34
Central Groundwater Board (CGWB) 34
NASA - Gravity Recovery and Climate Experiments (GRACE) 35
3.2.4 Surface Water Reservoirs 36
Central Water Commission 36
3.2.5 Surface Water Flows 37
Central Water Commission 37
Centre for Sustainability and the Global Environment - Global River Discharge Database 38
Dartmouth Flood Observatory 38
National River Linking Plan Modeled Database 40
World Meteorological Organization (WMO) - Global Runoff Data Centre 40
3.3 Water Demand 41
3.3.1 Evapotranspiration 41
Determining Evapotranspiration rates through Remote Sensing 41
MODIS Global Evapotranspiration (MOD16) 41
Global Land Evaporation Amsterdam Model (GLEAM) 42
Global Potential Evapo-Transpiration (Global-PET) and Global Aridity Index (Global-Aridity) dataset 43
3.3.2 Domestic Water Demand (Urban and Rural) 43
Census Town and Village Amenities 43
3.3.3 Industrial Water Demand 44

Openstreetmap 44
Centre for Science and Environment - Green Clearance Watch 45
Ministry of Environment and Forests - Industrial Water Allocations 46

Annexure 1: Dataset Summary Table 47

Annexure 2: Note on Datasets not included 51

Annexure 3: Resources for Introductory GIS and Remote Sensing 51

Glossary

Digital elevation model (DEM)

a digital elevation model is a gridded data format where each grid square represents the average elevation of a 30m*30m area of bare-earth's surface above mean sea level.

Raster

raster is a format wherein data is represented as an image with each pixel being a different data point. The most commonly known raster data formats are JPEG, PNG and BMP. In remote sensing and GIS however another raster format known as TIF is more common

Vector

vector is a format of data storage that is commonly used to represent points, lines or areas of interest on the ground.

1.Open data in India

Data has always been a crucial element to answer questions of how, where, when, and what nature. The level of access to data and tools to derive insights from data have always been limited. This is one truth that has been changing in the last few decades with the evolution of ever more *affordable computing power*. This phrase leads most of us think of personal computers and mobile devices. However computing power has now permeated into numerous spaces previously unimaginable. Industry has led progress in this direction. The ubiquity of computing has given us affordable tools to analyse data and offer services built on it. It has also enabled data generation itself to multiply exponentially. These technological advances hold tremendous potential for development. Larger hurdles remain in other spaces however. While data generation has multiplied, it has multiplied unevenly so in spaces which market forces favour. Hence much of the data generated is proprietary. Its commercial value helps keep businesses competitive and ensures much of it remains locked away. With this reality it falls upon government and citizen initiatives to understand these technological advances and deploy them effectively with goals of sustainability and equity in mind.

In several countries in fact it is government that leads industries in generating and disseminating data that is foundational. Their emphasis is on making data *'free and open source'* i.e. available for no cost under an open license for use, modification and dissemination by anyone. Examples of this are many, including economic data, budgets, national statistics, environmental and climate data, election data, spatial datasets such as administrative boundaries etc. Researchers, academicians and industry have benefited greatly from this in unearthing problems and innovating new solutions. Such data also benefits citizen and journalism initiatives in their efforts to hold government accountable. In addition to government, communities around the world have themselves generated open datasets, examples being, Openstreetmaps, Global Forest Watch etc.

India, as compared to other countries lags behind in the open data movement. Some of the major problems with data in India is the availability of up to date and accurate data. The need for open data is not completely unacknowledged, the last decade has seen the government formulate a 'National Data Sharing Policy' in 2012, and open its 'Open Government Data' platform for making access to data easier. The government is also formulating an 'Open Licensing Policy' under which the data can be released. However there remains a long way to go. Several critical datasets are still unavailable openly and instead sold at a hefty price and under closed licenses.

2.Why the need for ‘Open water data’

Of the many critical datasets that are still unavailable those related to water are among the most important. The central, state and lower governments hold an effective monopoly over the collection, storage and dissemination of water related data in India. This is true because the government holds primary responsibility for provisioning of water for domestic use and irrigation across the country. While they collect a lot of data their efforts at dissemination fall short. The Ministry of Water Resources recently began publishing data on its platform Water Resources Information System of India (WRIS) a web based spatial data platform for dissemination of water related data. In many ways this platform represents an advance over data availability in years preceding it. However in several key ways the platform falls short. It makes available information for perusal and examination but accessing this data comprehensively and running any sort of analysis on it is impossible. Any research efforts require complete datasets to be downloadable with adequate metadata to facilitate their use. The data must be in open formats that are inter-operable, i.e. cross platform compatible and easy to link with other datasets. This allows users to derive insights from data after analysis with other tools. Complete datasets also gives local planning efforts a greater chance of success allowing users to take the data offline and use it within software that works offline, in remote locations. The focus in the water sector has been on enabling participatory grassroots level management strategies of domestic water, irrigation projects, groundwater etc. Many of these efforts would be greatly aided with easier access to relevant data. The creation of social enterprises seeking to provide services in the rural sector, linked with agriculture and water are handicapped by the lack of water data. In many parts of the country suffering from incessant flooding accurate historical and real time flood maps would aid strategies for disaster management.

The paper henceforth presents a summary of the level of accessibility of several datasets that are relevant for the water sector, both government sources and non-government. Datasets made available by other governments in Europe, the U.S., Japan etc are freely available and could be immensely useful. However due to lack of information and some access hurdles, these are utilized only in some spaces, primarily in a few select research institutes.

3.0pen datasets for water

This section covers several datasets which directly or indirectly help to map the water cycle, but also include supporting datasets needed in analysis (such as Elevation data, Soils, River networks and Watershed boundaries). Datasets on water that aren't included, for various reasons are mentioned in Annexure 2

Remote sensing techniques are used to measure several environmental factors wherein sensors installed on devices such as satellites, drones etc capture the information from a fixed height above the surface of the earth. These sensors collect the information in several bands of wavelength and store as bands to the imagery. These band information are very useful to identify the factors such as temperature, precipitation, vegetation, water etc. Several datasets below are satellite derived. In addition to this other datasets are either digitized from older land records, or current satellite imagery or crowdsourced. There are also datasets that are compiled from ground based weather stations.

Water Availability

- Precipitation in the form of rainfall and snowfall
- Live storage capacity in reservoirs (and other man-made water bodies)
- Soil moisture
- Groundwater levels (and fluctuation)
- Surface water flows in rivers

Water Use/Demand

- Domestic water use: Human Population * estimated per capita consumption (or prescribed norm for domestic water consumption)
- Livestock water use: Livestock population * estimated per capita requirement
- Agriculture & Forests: Evapotranspiration data (derived from temperatures (daily/monthly), wind speeds, humidity (daily/monthly), soil moisture & type, type of Agricultural land use, stage of plant growth.
- Industry: Nature of industry and annual production * water required per unit of production

3.1 Supporting Data

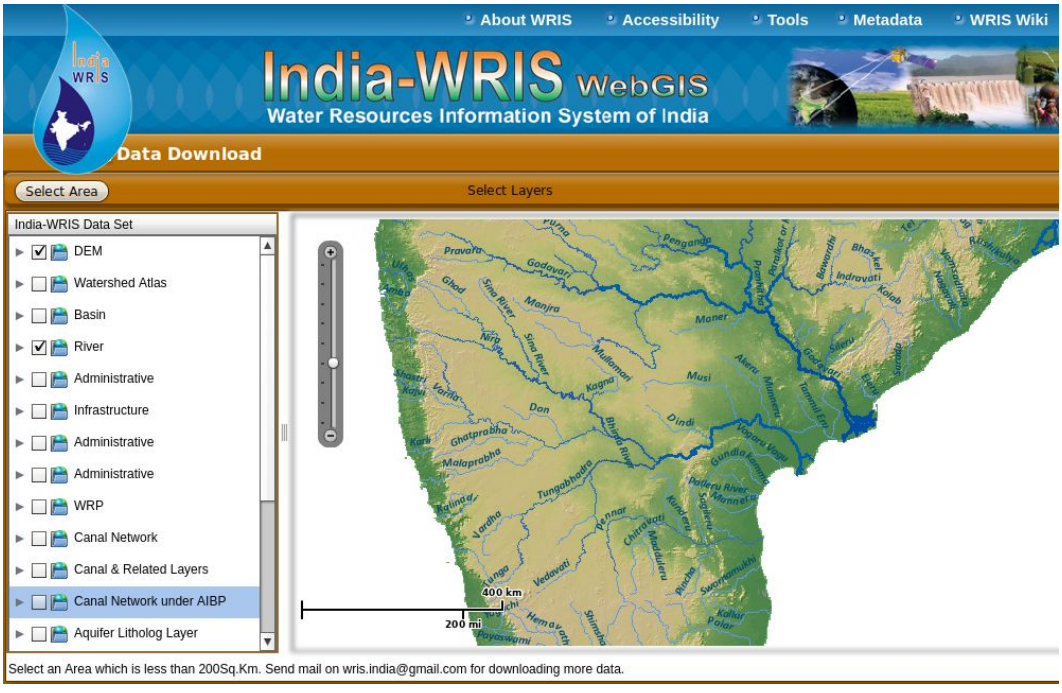
GIS data of 'vector' type allows users to create objects and areas of interest necessary for the analysis of water data and define them precisely in terms of their spatial coordinates. Objects of interest would for instance be rivers, reservoirs and areas of interest in water studies are primarily watersheds or river basins. This 'vector' data isn't the data itself but supports the analysis of water data.
Similarly there exists 'raster' data one instance of which is Digital Elevation Models (DEMs) which allow users to model the flow of water in watersheds.

3.1.1 River network

Government - India (CWC)

The Central Government (Central Water Commission) and various State Governments have been digitizing watershed and river maps since much more than a decade. This process of digitization converts paper hard copy maps into vector data depicting river networks. This process is done by remote sensing centres of the central government (National Remote Sensing Centre) and various state and regional remote sensing centres and is in different stages of completion. The regional remote sensing centre at Nagpur for instance has completed the digitization of all streams (down to first order perennial) and rivers in the state of Maharashtra. The streams and rivers all have their names associated with the vector data. This data is available for sale and the price for the entire state is several lakhs. The data itself is used to create web maps that are used in the WRIS website. While the WRIS website itself allows Data Downloads for areas smaller than 200 sq.km it notes that you can email WRIS for larger downloads. Note: This approach hasn't yielded replies in the past.

Access Link: http://india-wris.nrsc.gov.in/RiverApp.html?UType=R2VuZXJhbA==?UName=

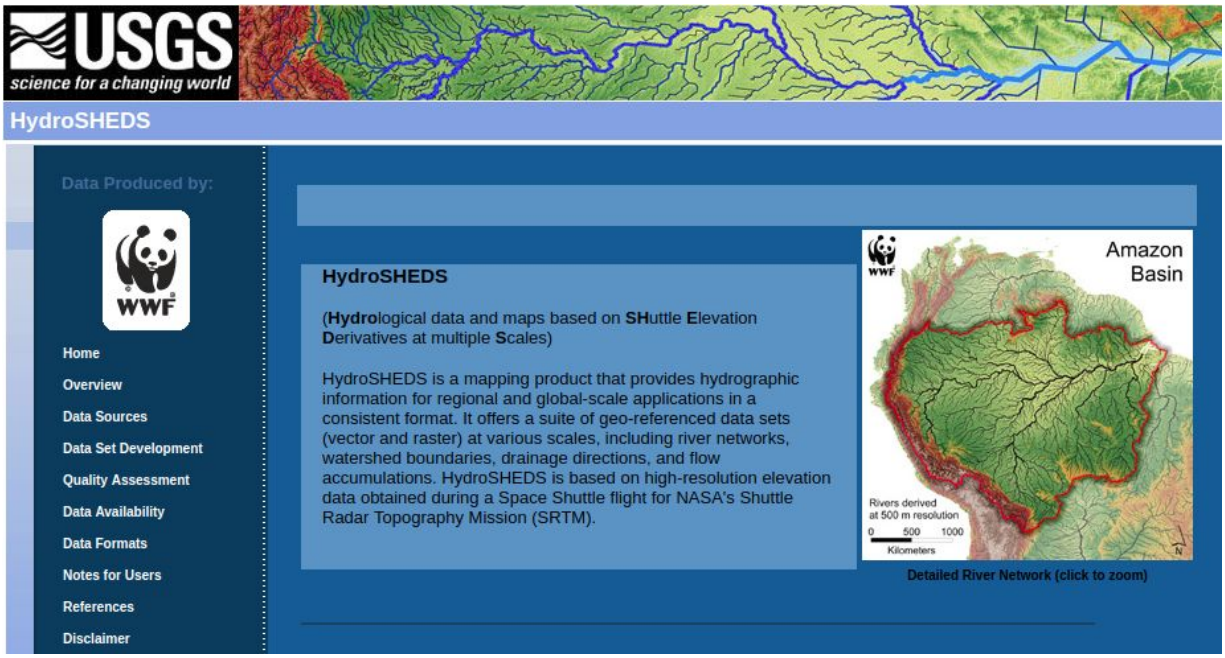


Central Water Commission - India WRIS Portal - River Network

Government - U.S. (USGS)

The United States Geological Service (U.S.G.S) has created Digital Elevation Model derived vector maps of streams of the entire world. These resemble actual stream networks for most of the world but for some areas which have significant man made interventions they may be inaccurate.

Access Link: https://hydrosheds.cr.usgs.gov/datadownload.php (See River Networks)



USGS Hydrosheds Website

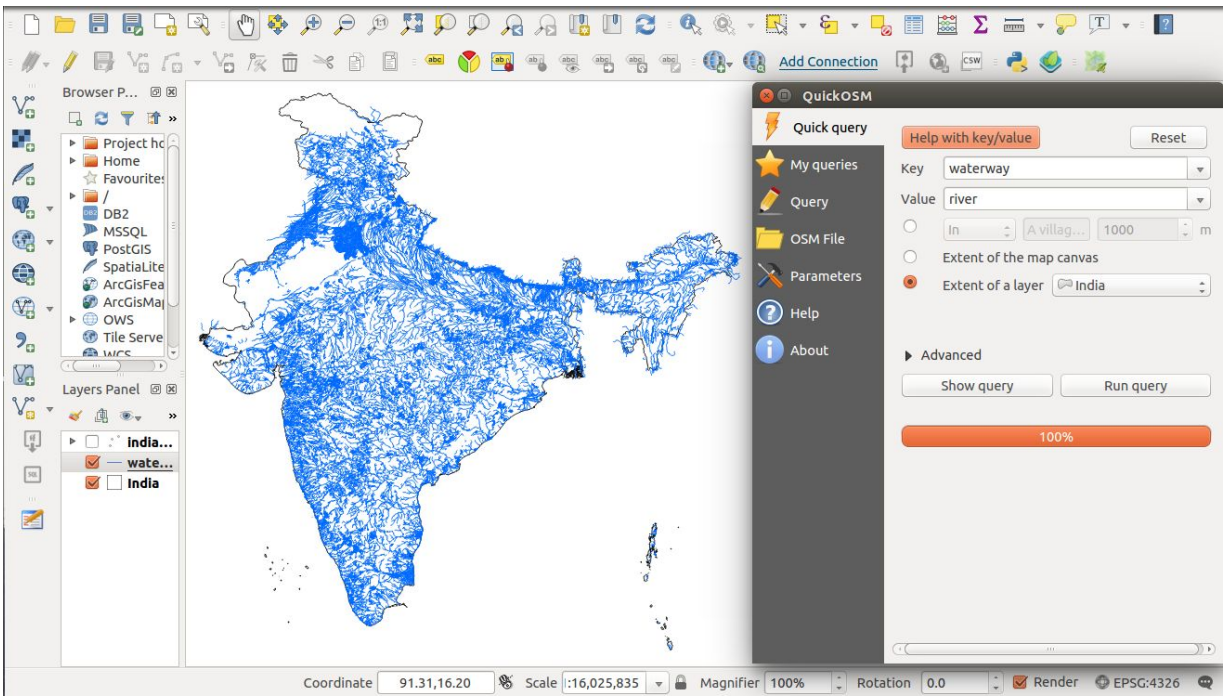
Community - Open Street Maps

openstreetmap.org has detailed waterways (river, streams and even canals) coverage for most of the country. The extent of coverage is variable and not all waterways have been named. This collection is however the most 'open' in terms of access for its level of detail. The maps are available as vector boundaries and can be accessed from one of the websites allied to OpenStreetMap. Point to be noted with OSM data is that downloading waterways data for large areas, for example the whole country of India can take time and cause certain tools to crash. Hence when facing issues here it may be better to do downloads for smaller areas. If one must download the entire country then GeoFabrik or Trimble, mentioned below are the way to go.

QuickOSM - QGIS Plugin

QuickOSM is a QGIS plugin which allows for easy download of Open Street Map data for customizable areas of interest.

Access: Through the plugins manager for QGIS

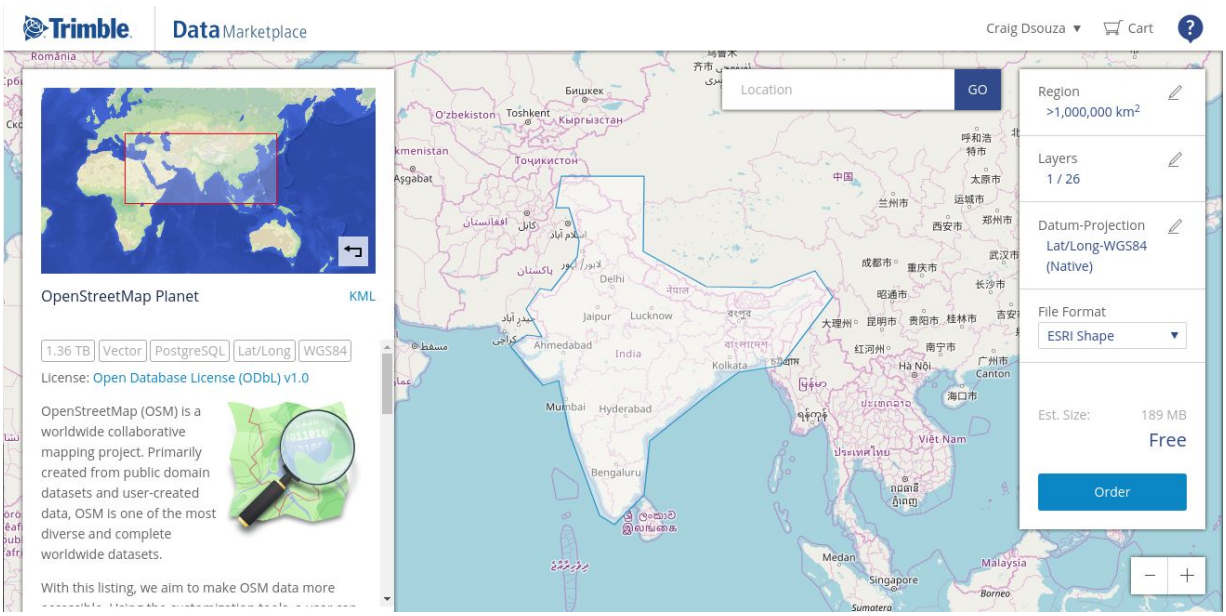


Snapshot of QuickOSM plugin being used within QGIS and the resulting waterway layers

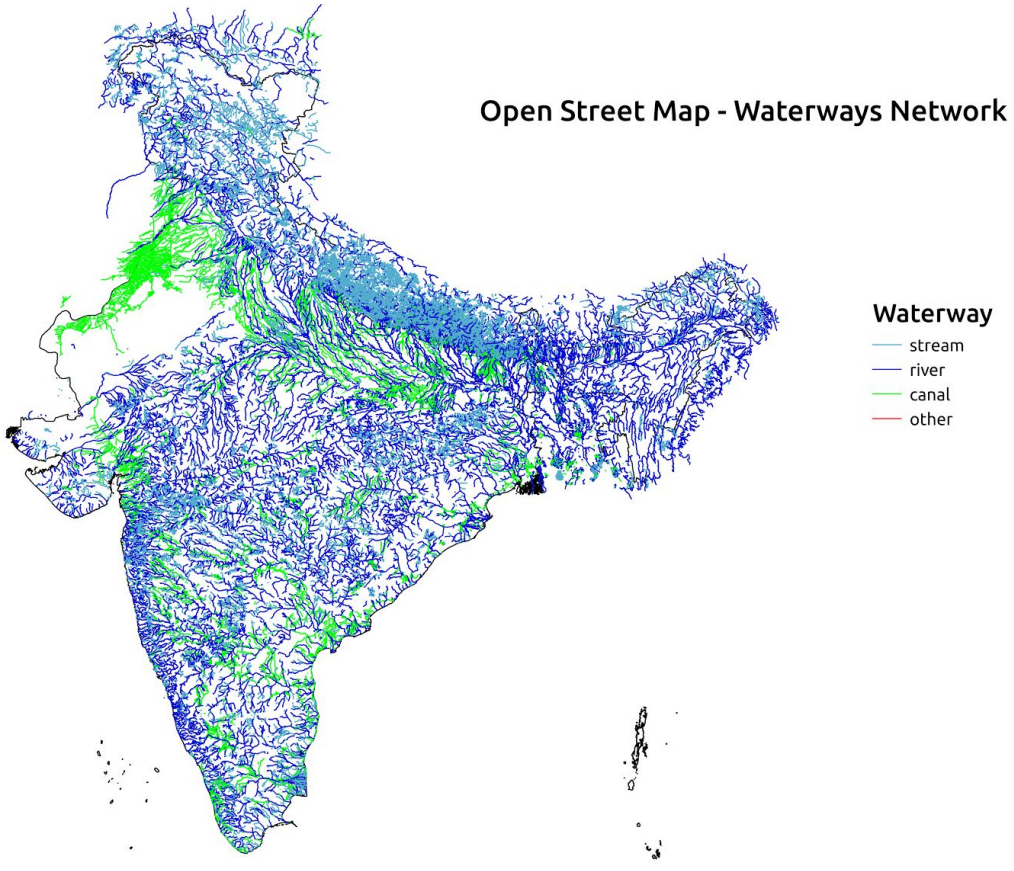
Trimble Data

Trimble is a data provider for multiple geospatial datasets. While some of their datasets are priced, they do offer free custom downloads of waterways data extracted from OpenStreetMap. This is one of the easier ways to extract data for the entire country in one go.

Access Link: https://data.trimble.com/market/provider/OpenStreetMap.html



Snapshot of order process for OSM Waterways data for India in Trimble

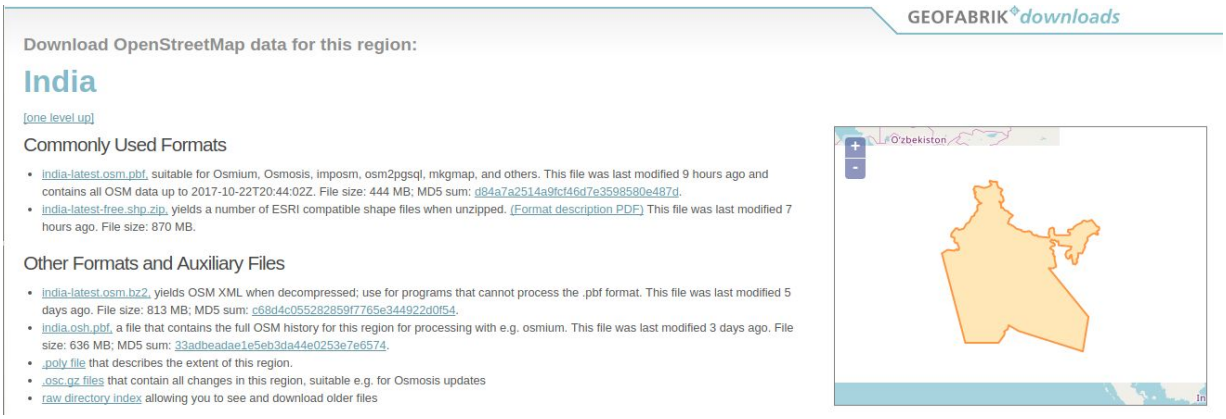


Snapshot of Open Street Map - Waterways Network

Geofabrik

Alternatively one can download daily updated extracts of OpenStreetmap data from 'Geofabrik' for India. This data comes as one large zipped file within which multiple shapefiles can be found. The data consists of all OSM data for India, not only waterways data, which will have to be filtered out in QGIS

Access Link: http://download.geofabrik.de/asia/india.html

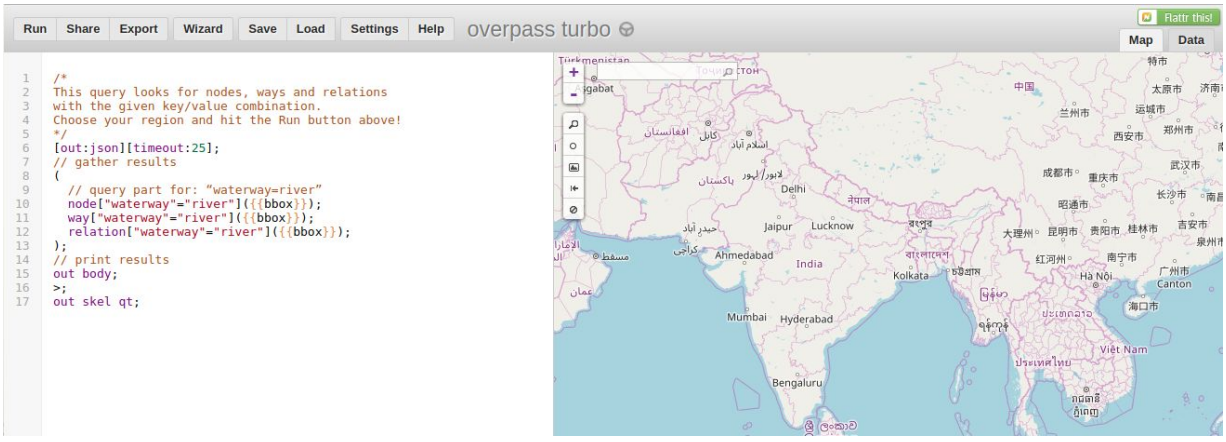


Snapshot of Geofabrik download links

Overpass Turbo

A user can also use Overpass Turbo to extract waterway data (rivers, streams, canals etc) for a limited area. Attempting to download large areas at once may cause the browser to crash. Overpass Turbo is useful to do a test run of queries for small areas before downloading OSM data

Access Link: https://overpass-turbo.eu/



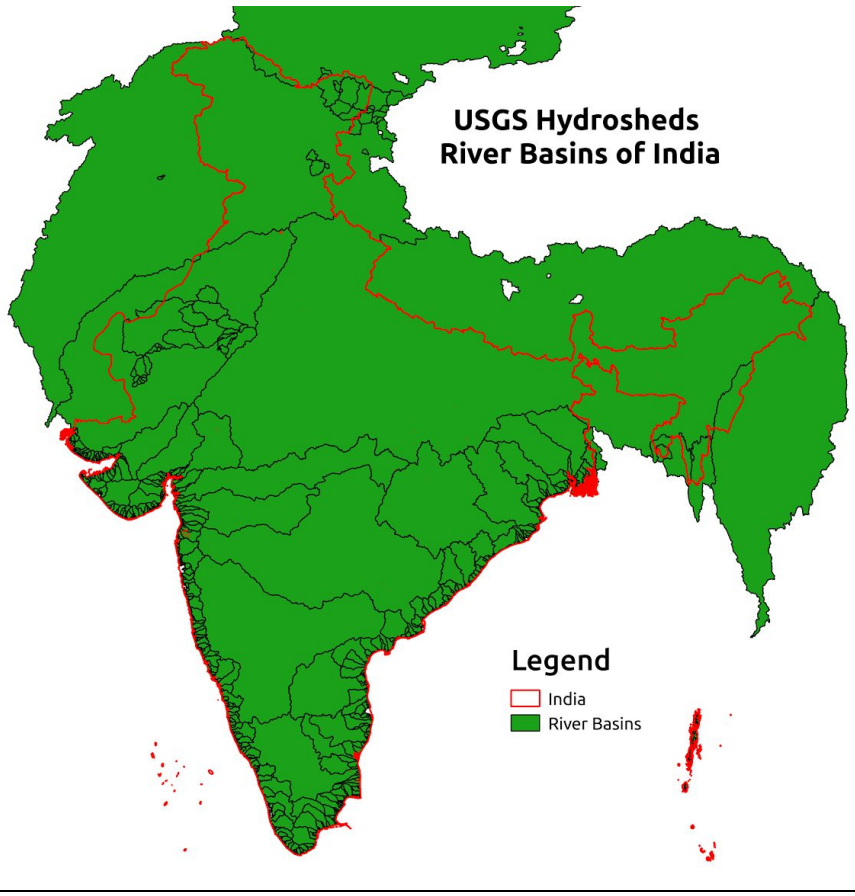
Snapshot of Overpass Turbo example for waterways data

3.1.2 Watershed/River basin boundaries

Watershed boundaries are important since they are the unit of study for research in water. Watershed vector boundaries as with river networks can be derived from Digital Elevation Models (DEMs). In fact every source of watershed boundaries follows the same technique to derive basin boundaries, however based on the definition of the size of a watershed the resulting boundaries can be different, as is evident in the compilation of sources below there are multiple sources of watershed boundaries for India.

USGS Hydrosheds

The United States Geological Survey provides such boundaries, not of watersheds but of larger river basin networks across the world.
Access Link: https://hydrosheds.cr.usgs.gov/datadownload.php (See River Networks)



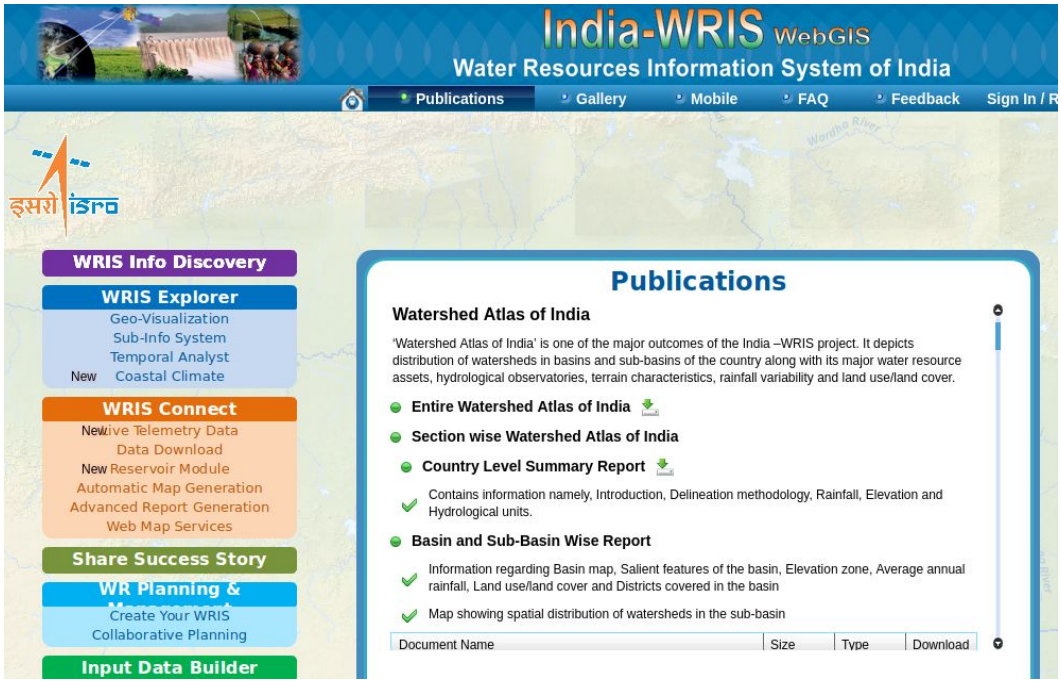
Snapshot of USGS Hydrosheds - River Basin boundaries of India

CWC's Watershed Atlas

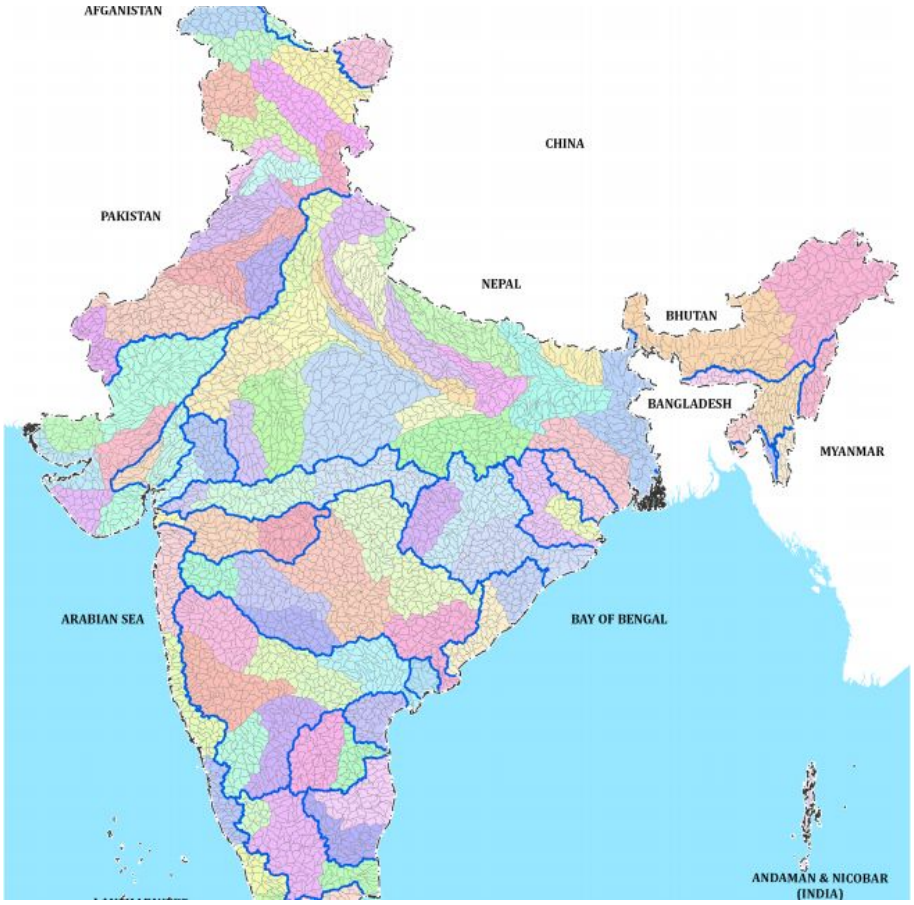
in its Watershed Atlas, the Central Water Commission provides boundaries of river basins, more detailed than USGS Hydrosheds, showing basin, sub-basin and watershed boundaries. Each approx 70,000 Ha in area and numbering a total of 4566. The project has developed a codification scheme for each watershed. These maps are available in PDF format, not as vector boundaries.

Access Link: Entire report of watershed boundaries in the country
http://www.india-wris.nrsc.gov.in/Publications/WatershedSubbasinAtlas/Watershed%20Atlas%20of%20India.pdf (Large Download)

Access Link: Country level summary report of watershed boundaries
http://www.india-wris.nrsc.gov.in/Publications/WatershedSubbasinAtlas/1.0%20Country%20Level%20Summary%20Report.pdf (also large but not as large as the entire report)



India WRIS Portal - Access to Watershed Atlas



Snapshot of Central Water Commission - Watershed Atlas Boundaries

Soil and Land Use Survey of India

The SLUSI is another source for watershed boundaries in its Watershed Atlas of India. It has its own codification scheme for watersheds and microwatersheds which is distinct from the CWC's. The two cannot be directly compared however since both do not exist in digital form. This atlas is at a scale of 1:1 million and the watersheds number a total of 3527. The atlas is not available for free.

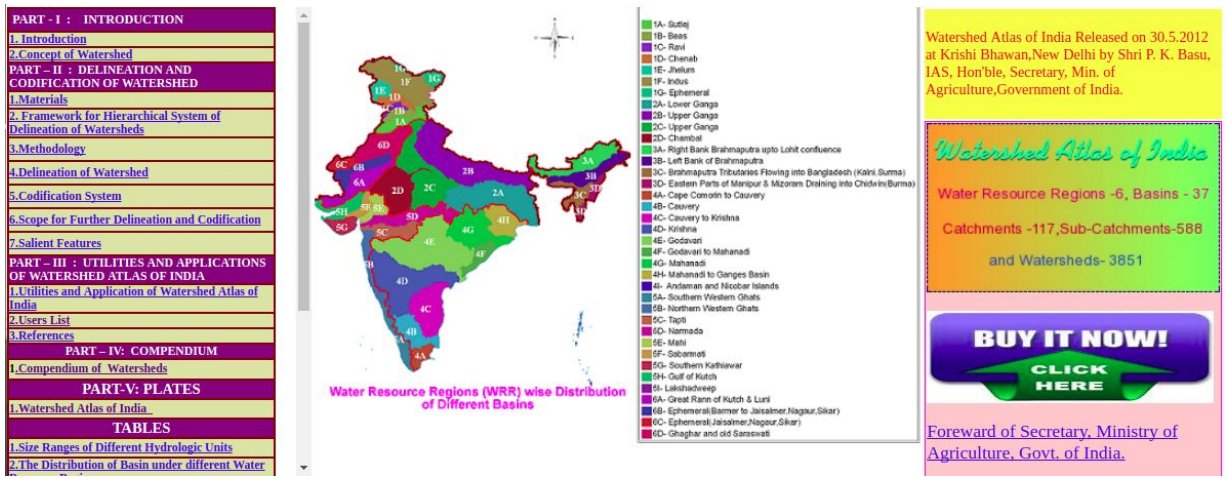
Info & Access Links

Watersheds

http://slusi.dacnet.nic.in/
http://slusi.dacnet.nic.in/watershedatlas/indexnew.html
http://slusi.dacnet.nic.in/dwainew.html

Micro Watersheds

http://slusi.dacnet.nic.in/dmwai/state.html (PNGs available here are not geo-referenced, and have no landmarks hence not very useful)



Snapshot of Watersheds Atlas webpage

Central Groundwater Board

The CGWB has maps of watershed boundaries available on their website also in a static form. Their codification scheme is again different from the SLUSI or CWC watershed maps. The maps are available for free at a scale of 1:250,000 and total watersheds number 3237.

Info Link: http://cgwb.gov.in/watershed/about-ws.html

Access Link: http://cgwb.gov.in/watershed/list-ws.html

3.1.3 Water reservoirs (dams, lakes -natural and man-made)

GRanD Database - Global Reservoir and Dam Database (2008)

This database was made under the work of the Global Water Systems Project (GWSP) In 2008 by manually drawing boundaries over satellite imagery followed by extensive error checking and cross validation. The data consists of 6,862 records globally of dams and the related reservoirs in vector format. The associated metadata consists of name of the dams and the impounded river, primary use, nearest city, area, and year of construction (or commissioning). All reservoirs with a storage capacity of more than 0.1 cubic kilometers and many smaller reservoirs have been included in the database. The data is available for free both on the project website and the FAO's website

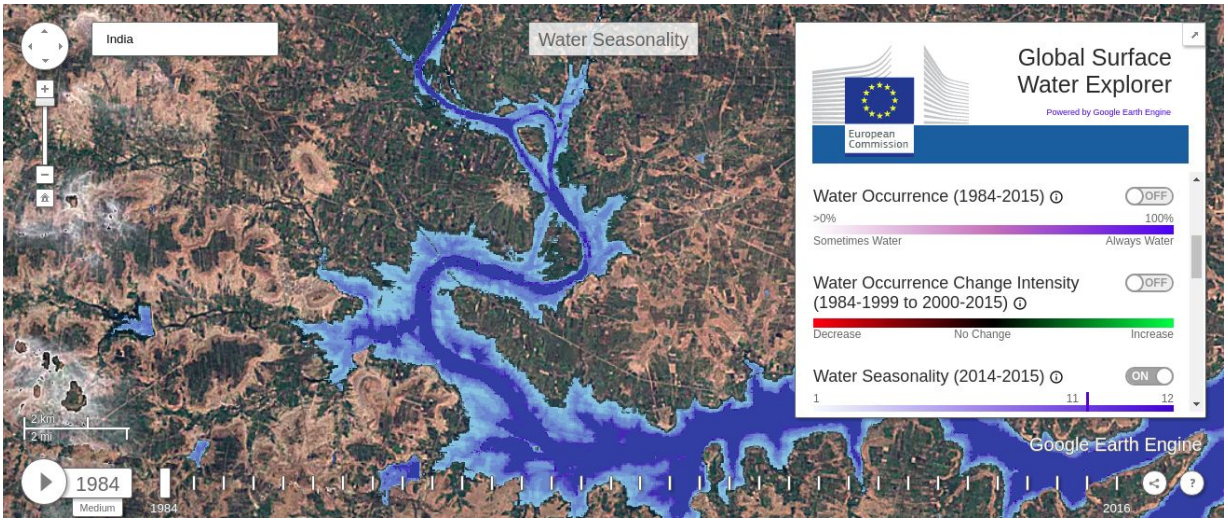
Access Link: http://www.fao.org/nr/water/aquastat/dams/index.stm

Global Surface Water Explorer - European Commission - Copernicus Programme

This dataset developed by the European Commission maps the location and temporal distribution of water surfaces at the global scale over the past 32 years and provides statistics on the extent and change of those water surfaces. The data depicts water occurrence, changes in water occurrence, water seasonality, water recurrence and water transitions. This dataset is available in the raster TIF format and can be used to identify individual reservoirs and even smaller lakes and tanks. It depicts all submerged areas and newly flooded areas as a result of man-made or natural changes within the time period from 1985-2015

Info Link: https://global-surface-water.appspot.com/

Access Link: https://global-surface-water.appspot.com/download

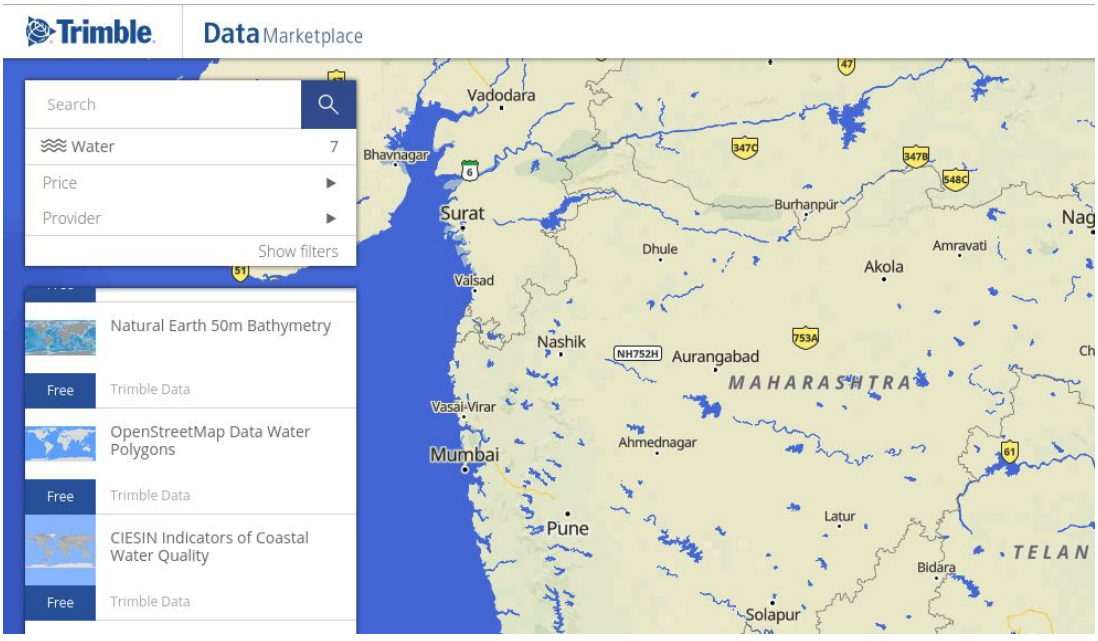


Snapshot of European Commission - Global Surface Water Explorer

Community - Open Street Maps

Trimble Data - Water Bodies

Trimble allows for easy downloads of very detailed Open Street Map water bodies data for large areas available for download for free. This is a database more comprehensive than the GRanD database mentioned above, but without useful metadata such as names of the waterbodies, capacity etc.
Access Link: https://data.trimble.com/market/provider/OpenStreetMap.html

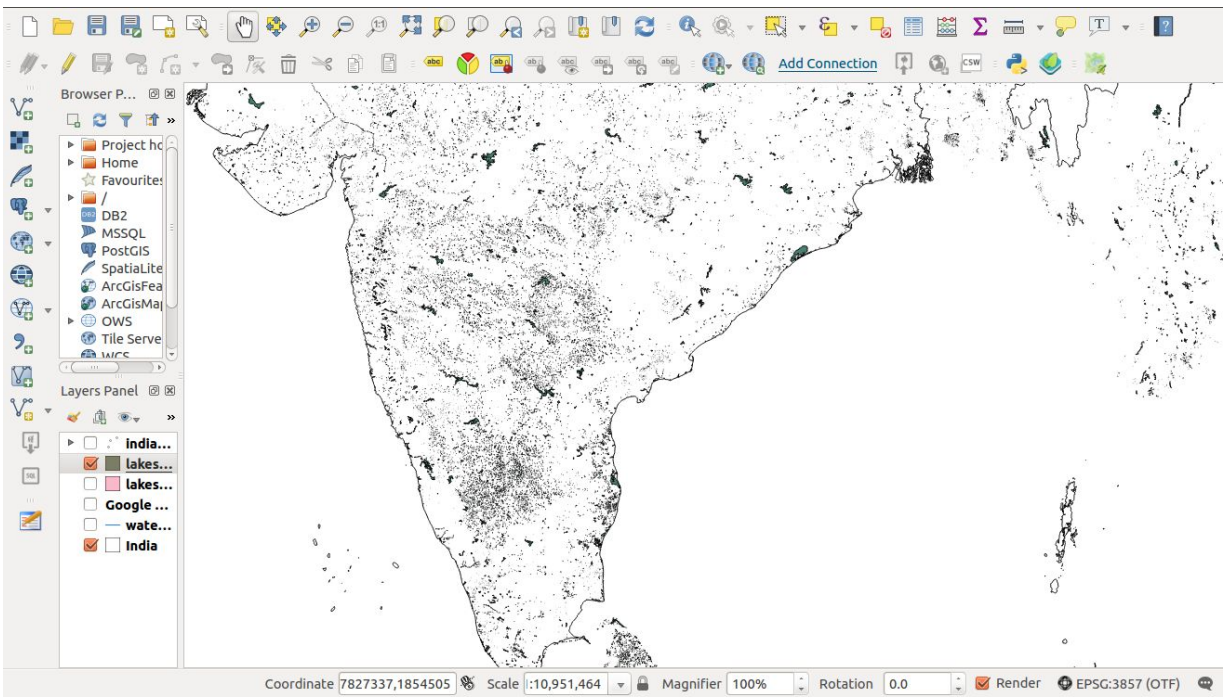


Snapshot of Trimble extract of Open Street Map Water Polygons (more are visible on zooming in)

Open Street Map Data for Water

One website has made available water bodies data from Open Street Map available for download for a global extent. This data is sampled and hence not suitable for hydrological analysis, however can be used for generating water map tiles

Access Link: http://openstreetmapdata.com/data



Snapshot of water bodies data from www.openstreetmapdata.com (for map tiles, not analysis)

3.1.4 Elevation data

In addition to river networks and boundaries, elevation is a critical dataset. Simple digital elevation models (DEMs) for the area of interest when accessed in QGIS allow a user to create watershed boundary maps and river network maps for the same area. This data is available freely through multiple sources, with a spatial resolution of 30m gridded pixels and vertical elevation resolution of 1m in raster TIF format. The accuracy of elevation models is limited, in cases the error could be as much as +- 20m. Hence this limits its usage for very accurate mapping of small streams and man made water channels. To map the course of larger streams and rivers however they are quite reliable.

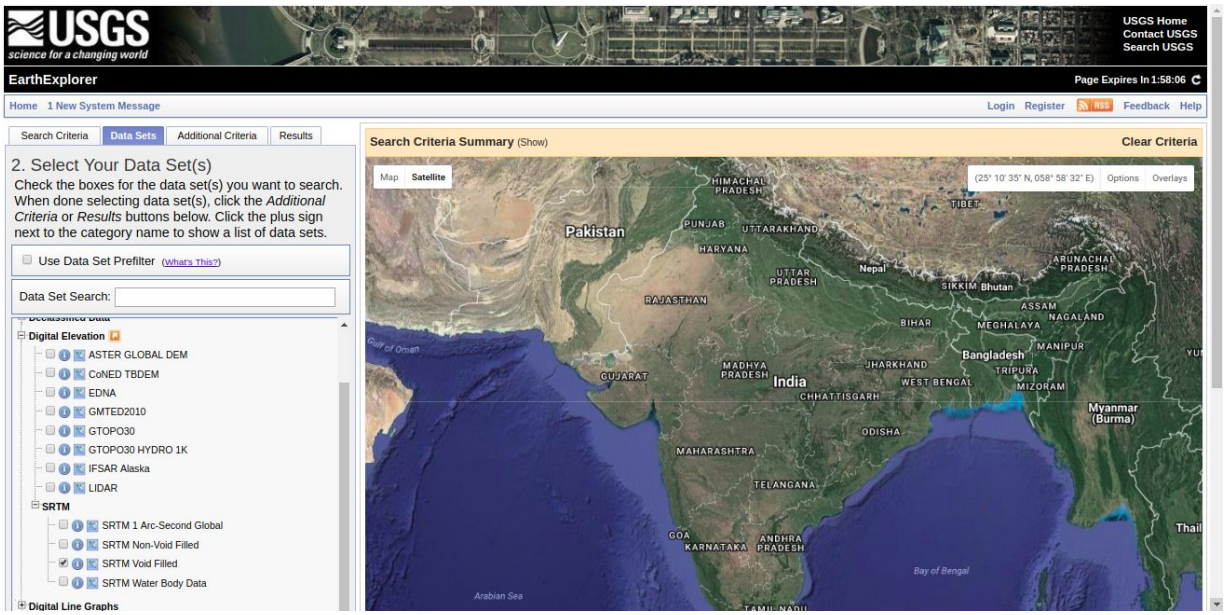
NASA has made available elevation data from two sources, the Shuttle Radar Topography Mission (SRTM) and Airborne Space Thermal Emission and Reflection Radiometer (ASTER) both on the USGS Earth Explorer Platform. In addition to this is also available GTOPO an elevation dataset of 1km resolution.

NASA - Shuttle Radar Topographic Mission (SRTM)
SRTM v 3.0 Global 1 arc second (30 m resolution) void filled is the latest version of the SRTM datasets which uses data from other datasets to fill in voids in earlier versions of the SRTM dataset. SRTM is known to have limitations in mountainous areas where elevation changes rapidly.

Info Links:

- https://earthdata.nasa.gov/nasa-shuttle-radar-topography-mission-srtm-version-3-0-global-1-arc-second-data-released-over-asia-and-australia
- https://lta.cr.usgs.gov/SRTMVF

Access Link:
Earth Explorer - https://earthexplorer.usgs.gov/
Land Processes DAAC - https://gdex.cr.usgs.gov/gdex/



Snapshot of USGS Earth Explorer Interface to access SRTM (30m) DEM

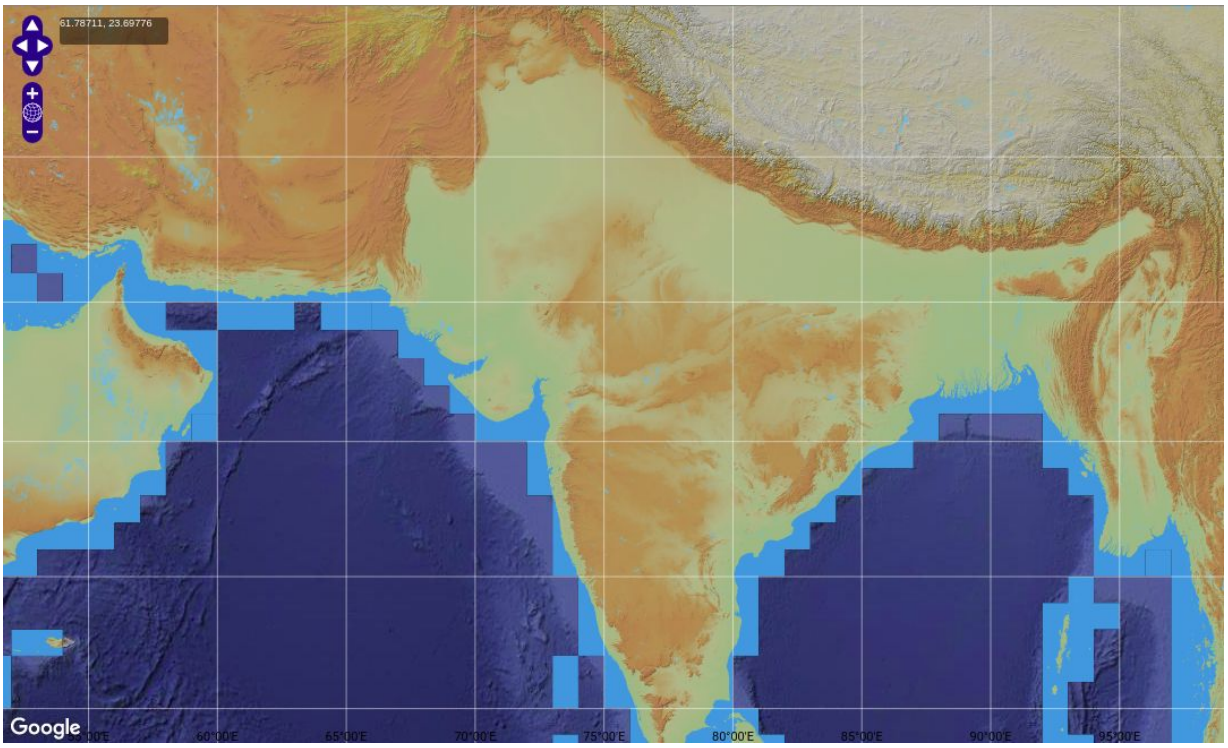
NASA - Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER)
NASA's ASTER dataset v 2.0 GDEM is derived from a second mission that utilizes different techniques to collect elevation data, it is more reliable than SRTM for mountainous regions. Its original version had many issues but its version 2.0 is a much improved product

info Link: https://lpdaac.usgs.gov/dataset_discovery/aster/aster_products_table/astgtm

Access Link:
Earth Explorer - https://earthexplorer.usgs.gov/
Land Processes DAAC - https://gdex.cr.usgs.gov/gdex/

JAXA - Advanced Land Observing Satellite (ALOS)

The AW3D30 DSM Ver.1.1 dataset released by the Japanese Space Agency in March of 2017 is the most precise global elevation dataset. Its spatial resolution is 30m, similar to the NASA - SRTM dataset.



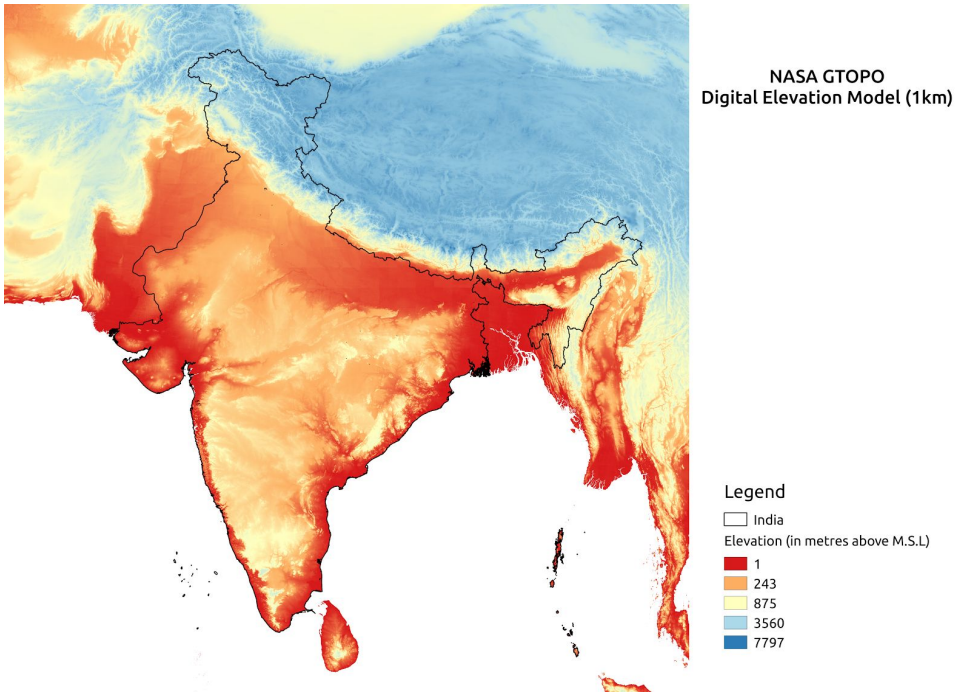
Snapshot of JAXA - ALOS dataset viewer

Info Link: http://www.eorc.jaxa.jp/ALOS/en/aw3d30/index.htm
Access Link: http://www.eorc.jaxa.jp/ALOS/en/aw3d30/l_map.htm

NASA - GTOPO30

The GTOPO30 is an early version of NASA's digital elevation products which was of lower resolution (30 arc seconds or 1 km) While it has been superseded by more accurate DEMs its smaller file size allows us to come up with quick elevation map for the entire country.

Info Link: https://lta.cr.usgs.gov/GTOPO30
Access Link: https://earthexplorer.usgs.gov/

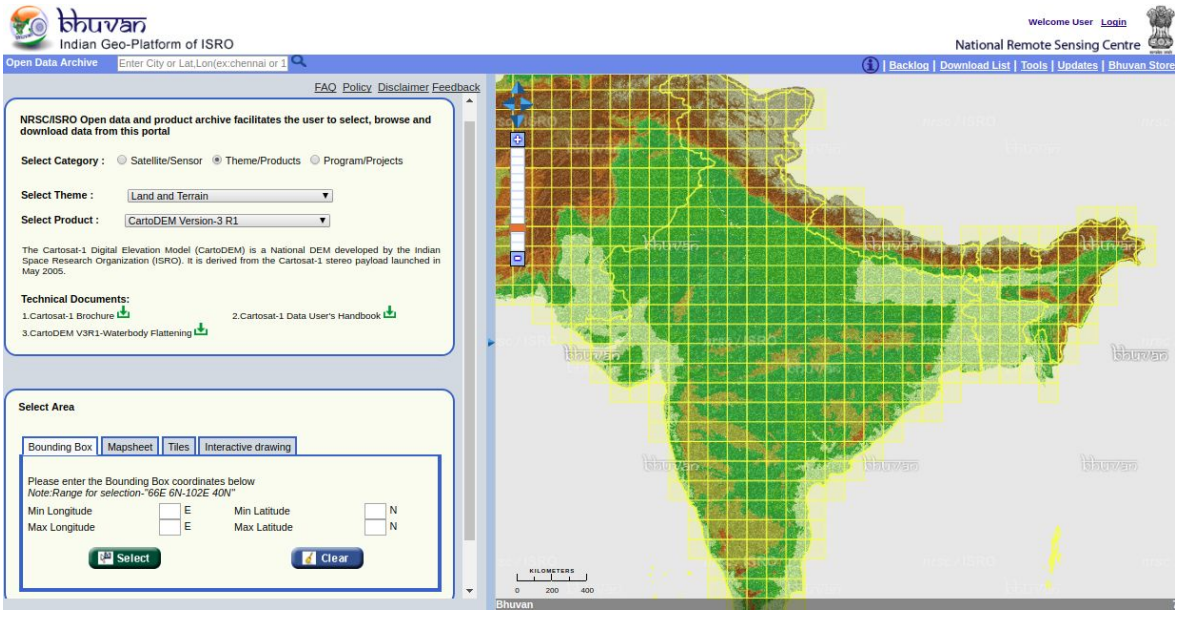


NASA GTOPO Digital Elevation Model (DEM) Low Resolution (1 km)

NRSC - CARTO-SAT

India also has its own digital elevation models, named CARTO SAT v 3 R1, which is of similar spatial resolution to NASA's datasets.

Access Link: http://bhuvan.nrsc.gov.in/data/download/index.php



Snapshot of Bhuvan Web Portal for DEM data

3.2 Water Availability

3.2.1 Rainfall data

There are several ways to measure rainfall. Traditionally, the rainfall measurements are done from various rainfall gauge stations located in several parts of the country or ground based radar (often located at airports). However, there are also satellite based sources where rainfall data can be obtained. This section describes datasets for rainfall from some of these sources that are available for analysis.

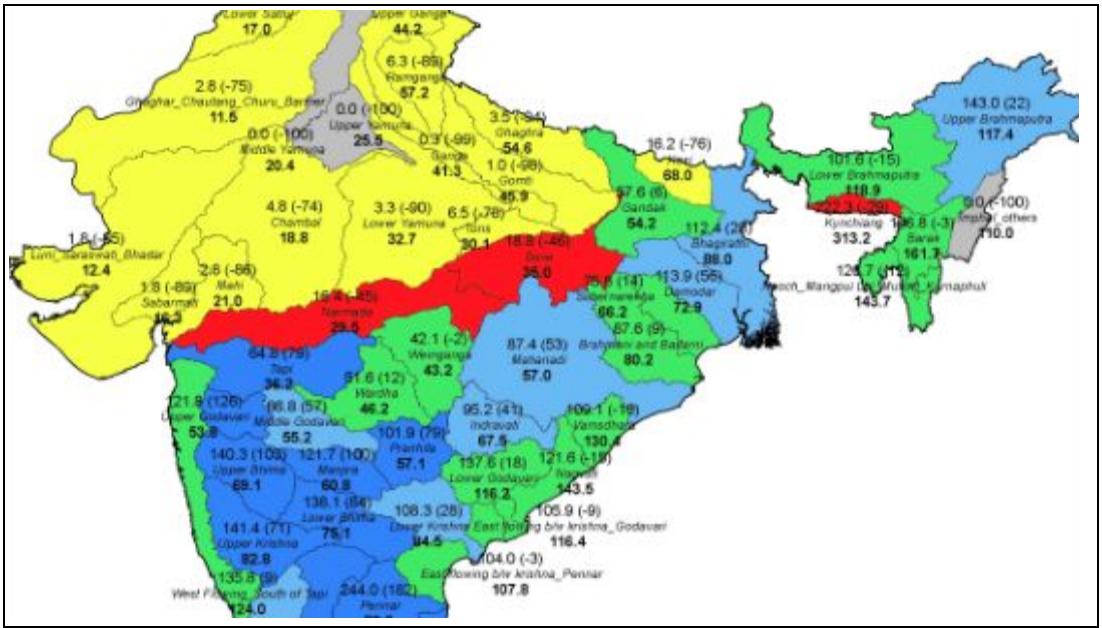
Indian Meteorological Department

The Indian Meteorological Department has a vast network of rainfall monitoring stations across the country. These are either automated weather stations (nos. 675), automatic raingauge stations (nos. 1289) or manual weather stations (nos. unknown) The total number of stations across the country is estimated to be between 2500 to 3000. The Department primarily makes available rainfall data for browsing on two platforms, discussed below, the Customised Rainfall Information System of India (CRIS) and the IMD AWS website. Neither website however provides the data in a format that is particularly useful.

Customised Rainfall Information System of India (CRIS)

This platform allows users to browse through only the present day's rainfall data for 443 district rainfall stations across the country. This data isn't available for download. The site also shows custom maps for different states, districts, hydro-meteorological sub-divisions and river basins across India based on the cumulative rainfall in the last week, month, season and year.

Access Link: http://hydro.imd.gov.in/hydrometweb/

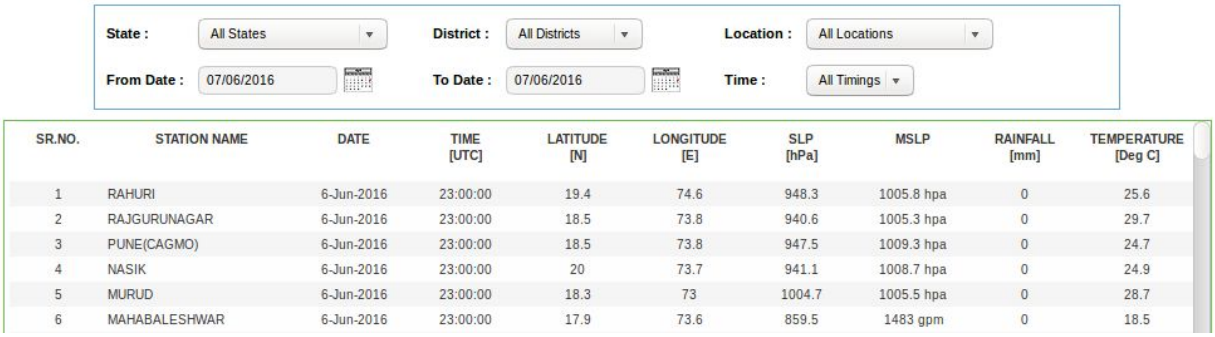


Snapshot of Customised Rainfall Information System of India (CRIS) Map

Automatic Weather Stations (AWS)

The raw data from automatic weather monitoring stations across the country have been published on the website below. This data is generally available for the past one week and for hourly time intervals. It does not allow for data download only for data viewing for one state or district at a time. The site is generally very unreliable and has been down for the entire period of August to October 2017 at least. When the site is available however the data availability is in a format better than the CRIS provides

Access Link: http://www.imdaws.com/viewawsdata.aspx



SR.NO.	STATION NAME	DATE	TIME [UTC]	LATITUDE [N]	LONGITUDE [E]	SLP [hPa]	MSLP	RAINFALL [mm]	TEMPERATURE [Deg C]
1	RAHURI	6-Jun-2016	23:00:00	19.4	74.6	948.3	1005.8 hpa	0	25.6
2	RAJGURUNAGAR	6-Jun-2016	23:00:00	18.5	73.8	940.6	1005.3 hpa	0	29.7
3	PUNE(CAGMO)	6-Jun-2016	23:00:00	18.5	73.8	947.5	1009.3 hpa	0	24.7
4	NASIK	6-Jun-2016	23:00:00	20	73.7	941.1	1008.7 hpa	0	24.9
5	MURUD	6-Jun-2016	23:00:00	18.3	73	1004.7	1005.5 hpa	0	28.7
6	MAHABALESHWAR	6-Jun-2016	23:00:00	17.9	73.6	859.5	1483 gpm	0	18.5

Snapshot of hourly rainfall data for the past week displayed on IMD AWS website

Raw IMD Station Data

If a user wishes to go beyond browsing and access all the raw data for their area of interest they must fill out a form and send it to the IMD office in Pune mentioning the number of stations, their names, and the time period of interest.
Form link: www.imdpune.gov.in/ndc_new/Request.html (note that this site is also often down)

If the data requirement is of monthly or annual rainfall for larger areas (district normals or meteorological sub divisions normals) the data can be accessed for free on the data.gov.in platform.
Access link: https://data.gov.in/catalogs/ministry_department/india-meteorological-department-imd

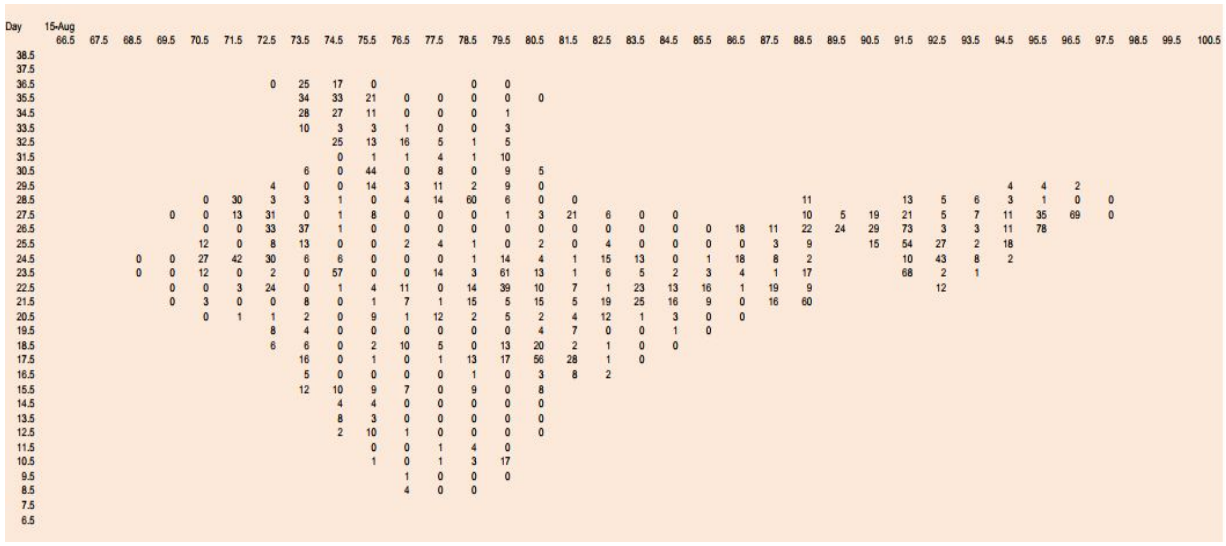
A compilation of district wise actual monthly rainfall data from 1901 to 2010 (except for 2003) of the IMD is available in easy to use spreadsheet formats on the India Water Portal platform
Access link: http://www.indiawaterportal.org/data, (http://www.indiawaterportal.org/met_data/)

Modeled IMD Data

The IMD also makes available quality controlled daily rainfall data in gridded format (binary and ASCII) - 1 degree grid and 0.25 degree grids for all of India, for the period from 1901-2015 at a price of Rs 8000 and Rs 15000 respectively.

Info Link: http://www.imd.gov.in/advertisements/20170320_advt_34.pdf

Sample data in ASCII format is visible below (1 degree grids)



Day	15-Aug																																						
	66.5	67.5	68.5	69.5	70.5	71.5	72.5	73.5	74.5	75.5	76.5	77.5	78.5	79.5	80.5	81.5	82.5	83.5	84.5	85.5	86.5	87.5	88.5	89.5	90.5	91.5	92.5	93.5	94.5	95.5	96.5	97.5	98.5	99.5	100.5				
38.5																																							
37.5																																							
36.5							0	25	17	0			0	0																										
35.5								34	33	21	0	0	0	0	0																								
34.5								28	27	11	0	0	0	1																								
33.5								10	3	3	1	0	0	3																								
32.5									25	13	16	5	1	5																								
31.5									0	1	1	4	1	10																								
30.5								6	0	44	0	8	0	9	5																								
29.5							4	0	0	14	3	11	2	9	0														4	4	2							
28.5					0	30	3	3	1	0	4	14	60	6	0	0							11			13	5	6	3	1	0	0						
27.5				0	0	13	31	0	1	8	0	0	0	1	3	21	6	0	0				10	5	19	21	5	7	11	35	69	0					
26.5					0	0	33	37	1	0	0	0	0	0	0	0	0	0	0	0	18	11	22	24	29	73	3	3	11	78							
25.5					12	0	8	13	0	0	2	4	1	0	2	0	4	0	0	0	0	3	9		15	54	27	2	18									
24.5			0	0	27	42	30	6	6	0	0	0	1	14	4	1	15	13	0	1	18	8	2			10	43	8	2								
23.5			0	0	12	0	2	0	57	0	0	14	3	61	13	1	6	5	2	3	4	1	17			68	2	1									
22.5				0	0	3	24	0	1	4	11	0	14	39	10	7	1	23	13	16	1	19	9				12										
21.5				0	3	0	0	8	0	1	7	1	15	5	15	5	19	25	16	9	0	16	60														
20.5					0	1	1	2	0	9	1	12	2	5	2	4	12	1	3	0	0																
19.5							8	4	0	0	0	0	0	0	4	7	0	0	1	0																	
18.5							6	6	0	2	10	5	0	13	20	2	1	0	0																		
17.5								16	0	1	0	1	13	17	56	28	1	0																		
16.5								5	0	0	0	0	1	0	3	8	2																				
15.5								12	10	9	7	0	9	0	8																						
14.5									4	4	0	0	0	0	0																						
13.5									8	3	0	0	0	0	0																						
12.5									2	10	1	0	0	0	0																						
11.5										0	0	1	4	0																							
10.5										1	0	1	3	17																							
9.5											1	0	0	0																							
8.5											4	0	0																								
7.5																																							
6.5																																							

Snapshot of IMD gridded modeled data (100 km resolution)

Global Datasets

Global datasets of precipitation are derived from the observation data of multiple satellites (microwave as well as Thermal IR measurements) and also actual in situ measurements of rainfall from weather monitoring stations globally. Observation data has been obtained from monitoring missions since the 1970s and in 1997 the first satellite mission specially dedicated for precipitation measurement was launched, the Tropical Rainfall Monitoring Mission (TRMM) mission. Since its decommissioning in 2015 data from the multi-satellite Global Precipitation Monitoring (GPM) mission has maintained the continuity of data.

Algorithms to provide us with usable precipitation datasets from satellite measurements are developed by different research groups, primarily by NASA, the NOAA (National Oceanic and Atmospheric Administration), JAXA (the Japanese Space Agency), the German Government and other governments and universities in partnership.

The International Precipitation Working Group prepared a list of datasets listed below. Here we describe some of the most commonly used satellite datasets - the TRMM, GPM-IMERG, CHIRPS and PERSIANN datasets. In addition to satellite based datasets there are also datasets based on ground precipitation monitoring stations, including GHCN and GPCC which are described here. Remote precipitation measurements are challenging as described here and although the number of datasets are plenty, however the use of precipitation data must be with caution, acknowledging the limitations of each dataset.

In addition to the datasets described, other sources not discussed here are the

- **GPCP** (Global Precipitation Climatology Project) - 1DD Dataset by NASA of global precipitation data (Similar to TRMM, but lower spatial resolution of 1 degree and temporal resolution of 1 day and global extent)
- **GOES** (Geostationary Operational Environmental Satellite) - NOAA - Precipitation Index (low spatial resolution of estimate and calculated only based on IR measurements)
- **CMAP** NOAA Climate Prediction Centre (low spatial resolution - 2.5 degree estimate based on multiple satellites and ground measurements) for the period from 1979 to present at monthly or pentad timesteps
- **CMORPH** NOAA Climate Prediction Centre (high resolution - 0.25 degree estimate which takes other satellite microwave precipitation estimates and merges them) for the period from 2002 to present at sub-daily or daily timesteps

Information for both the NASA missions (TRMM and GPM) can be found below

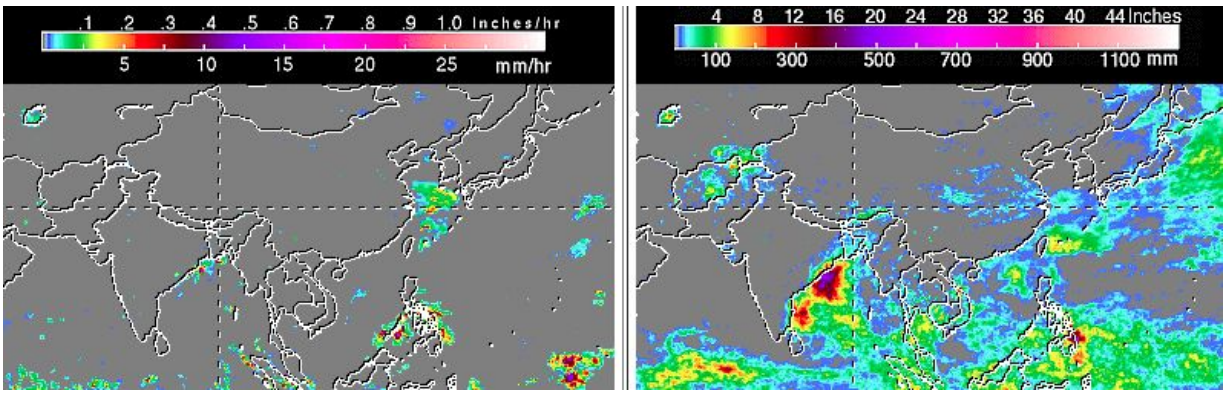
Info Link: https://pmm.nasa.gov/precipitation-measurement-missions
Trainings Link: https://pmm.nasa.gov/training
Data Access Tutorials: https://pmm.nasa.gov/data-access/tutorials

Tropical Rainfall Monitoring Mission (TRMM)

TRMM was a joint mission by NASA and the Japanese Space Agency to monitor rainfall that began in 1997 and ended in 2015. The mission was the first to provide global measurements of rainfall, as well as 3D models of storm structure and intensity. Though the TRMM satellite had rainfall measurements of its own, the final dataset released by the research team was a multi-satellite product (3B42), combining the calibrated measurements of several different American and Japanese satellites. This dataset provides estimates of rainfall with an interval of 3 hrs, both in real time (3B42RT) and more accurate research versions (3B42). The spatial resolution of this data is 0.25 deg (approx 25 km) Details of the algorithms and the dataset can be found in the documentation here

TRMM data can be downloaded here
Access Link:
https://pmm.nasa.gov/data-access/downloads/trmm (See '3B42' Research version) OR
https://mirador.gsfc.nasa.gov/cgi-bin/mirador/presentNavigation.pl?project=TRMM&tree=project

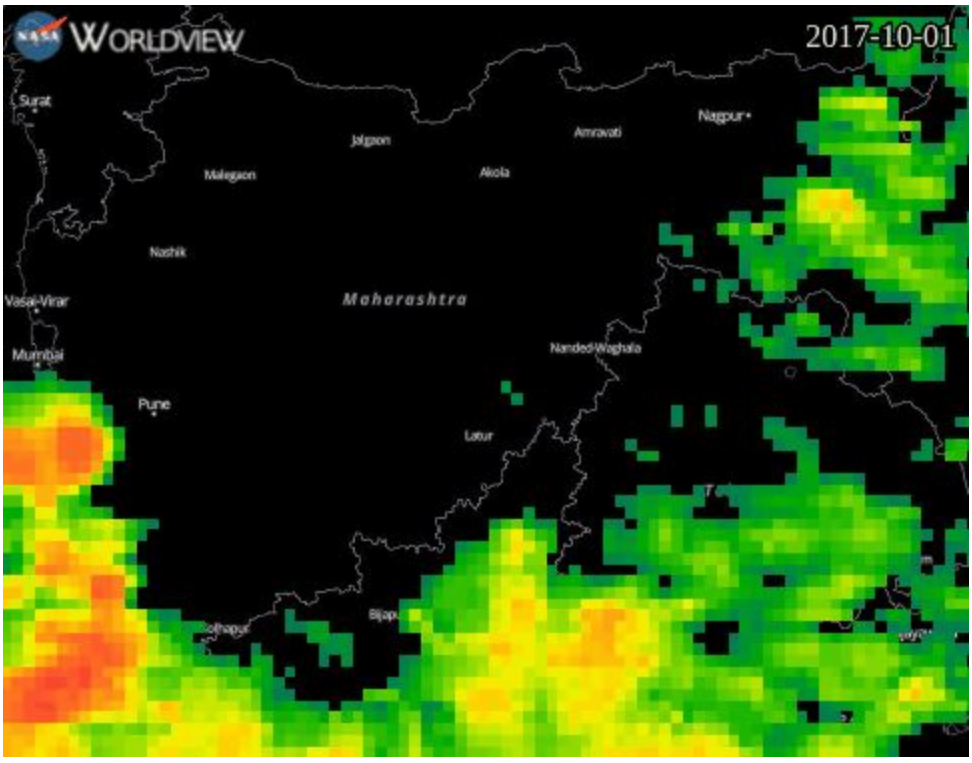
Web View Link:
https://pmm.nasa.gov/TRMM/realtime-3hr-7day-rainfall



Snapshot of 3 hourly and 7 day accumulated rainfall as available on TRMM website

Global Precipitation Mission (GPM)

The Global Precipitation Mission is a constellation of satellites that continues the measurement of rainfall after the decommissioning of the TRMM satellite in 2014. The GPM core satellite combines the measurements of several satellites (U.S., Japan, France, India, EU) to provide a merged 'GPM-IMERG' dataset, wherein the time interval is 30 minutes and the spatial resolution is 0.1 deg (approx 10km)



GIF of gridded daily rainfall as measured by GPM satellites over Maharashtra, India

Webview Link
GPM Global Viewer - https://pmm.nasa.gov/data-access/global-viewer
GPM Precipitation App - https://pmm.nasa.gov/precip-apps

Access Link
https://pmm.nasa.gov/data-access/worldview

Climate Hazards Group Infrared Precipitation with Station Data (CHIRPS)

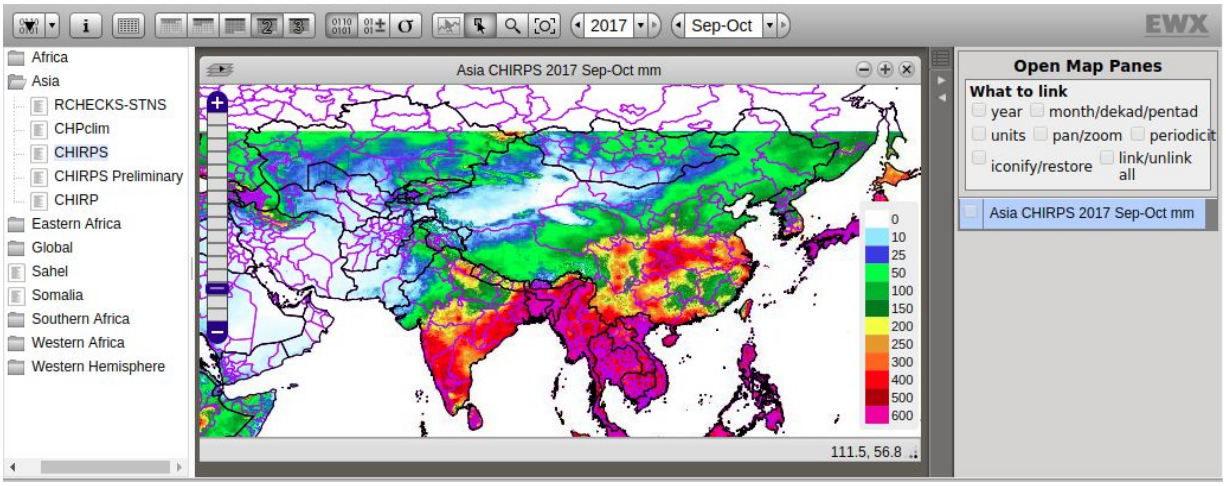
CHIRPS is a 30+ year global dataset of daily rainfall of high resolution (0.05 degree - 5 km) specially developed for drought monitoring. Its time period is from 1981 to present.

Info Link:
http://chg.geog.ucsb.edu/data/chirps/

Webview Link:
http://chg.geog.ucsb.edu/tools/ewx/
http://ewx.geog.ucsb.edu:8080/EWX/index.html

Desktop tool for Data Analysis:
http://chg.geog.ucsb.edu/tools/geoclim/

Data Link:
http://chg.geog.ucsb.edu/data/index.html OR http://clim-engine-development.appspot.com/fewsNet

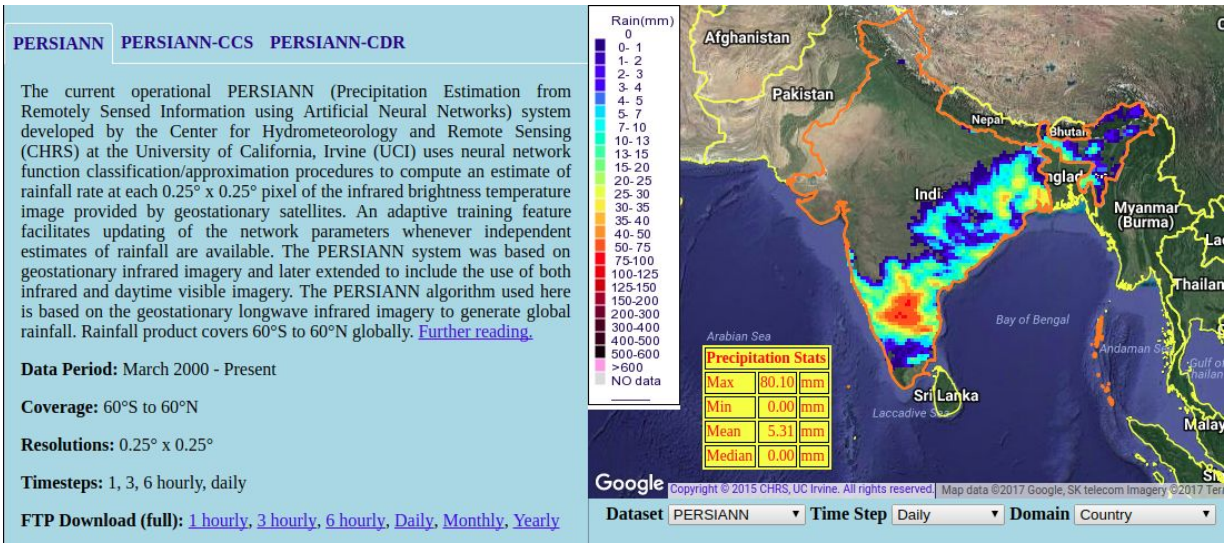


CHIRPS Web Link (Famine Early warning system) for Precipitation Data Visualization and Download

Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN)

The PERSIANN system developed at the University of California, Irvine uses infrared brightness measurements from satellites to derive hourly rainfall estimates for 0.25 * 0.25 km grids globally. The time period for this data is from the year 2000 to present

Info and Access Link:
http://chrsdata.eng.uci.edu/
https://www.ncdc.noaa.gov/cdr/atmospheric/precipitation-persiann-cdr



Daily rainfall data visualized on the PERSIANN web platform

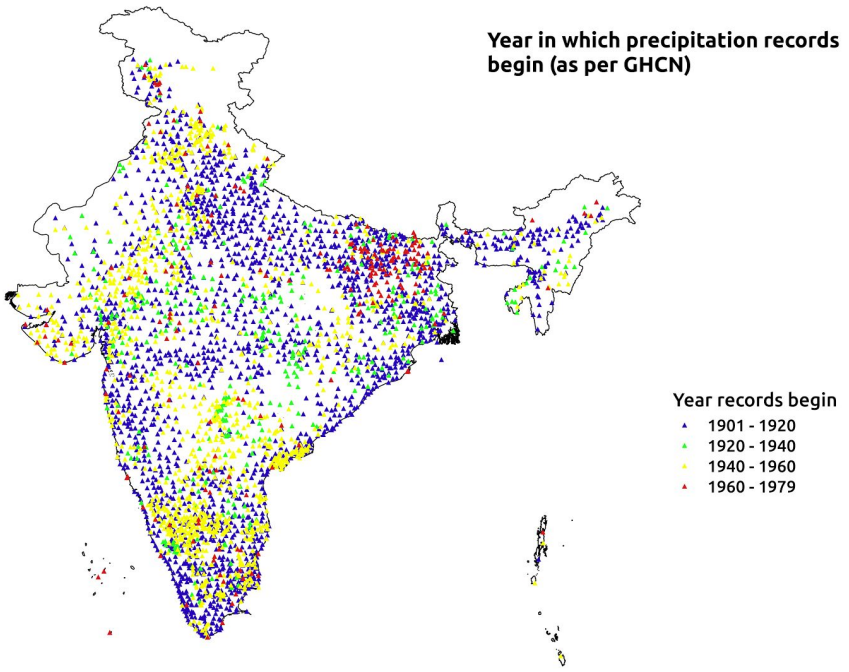
Global Historical Climate Network:

GHCN-Daily is comprised of daily climate records from land based stations across the world. The data has been integrated and subjected to a common suite of quality assurance reviews by the National Oceanic and Atmospheric Administration (NOAA), USA for international climate research purposes. The database contains records from over 100,000 stations in 180 countries and territories. Besides daily precipitation, it also provides numerous daily variables, including maximum and minimum temperature, snowfall, and snow depth; however, about one half of the stations report precipitation only. Both the record length and period of record vary by station and cover intervals ranging from less than a year to more than 175 years.

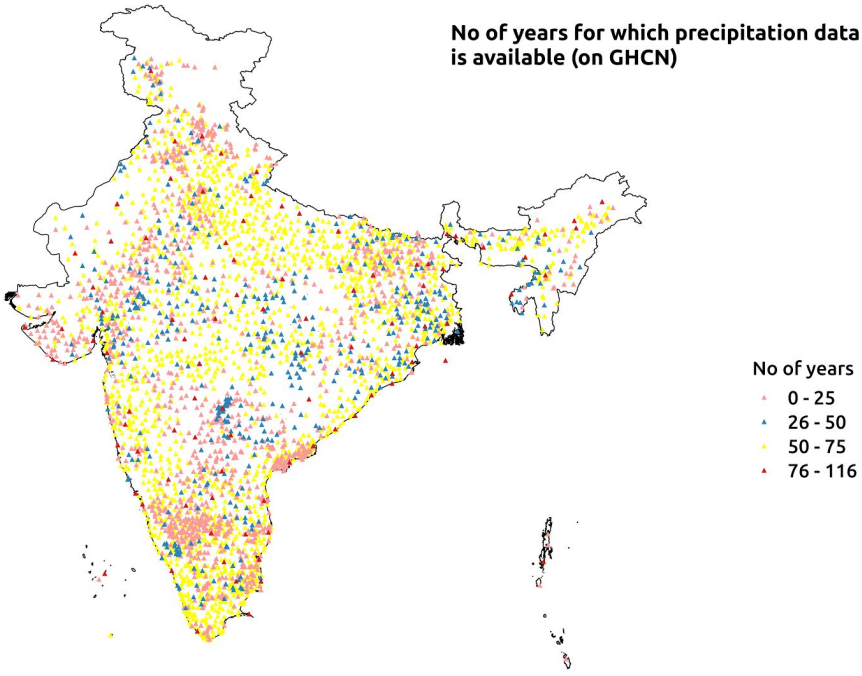
Contributions to the database are largely from the U.S.A but also include dense station networks in Brazil, South Africa, and India, however data from these countries is purely historical and not updated on a regular basis. Precipitation records end generally in the late 1990s for Brazil and South Africa, and in 1970 for India. The reason why these records end is not clear, but it could reflect a shift towards a closed data policy in India at the time.

** Of the 3807 stations in India with precipitation data the majority of them were installed in the early 20th century. The Cauvery Basin and Krishna Basins had many more stations installed in the period immediately post independence and Bihar in the period from 1960 - 1980.

Info Link: https://www.ncdc.noaa.gov/ghcn-daily-description
Access Link: https://www.ncdc.noaa.gov/ghcnd-data-access OR http://climexp.knmi.nl



Meteorological stations in India which have contributed to the GHCN (arranged by first year)



Meteorological stations in India which have contributed to the GHCN (arranged by no. of years of data)

Precipitation Reconstruction - Land (PREC/L)

This dataset is a gridded monthly global estimate of rainfall derived from the precipitation data in GHCN v2. The record extends from 1948 till present.

Info and Access Link: https://www.esrl.noaa.gov/psd/data/gridded/data.precl.html

Global Land Precipitation and Temperature (University of Delaware)

This dataset is also a gridded dataset of monthly timesteps and of global coverage derived from the ground station data from GHCN v2, the only difference is in the length of the data record, which begins in 1901 and runs until 2014

Info Link:
https://climatedataguide.ucar.edu/climate-data/global-land-precipitation-and-temperature-willmott-matsuura-university-delaware
http://climate.geog.udel.edu/~climate/

Access Link: http://climate.geog.udel.edu/~climate/html_pages/download.html#P2014

Global Precipitation Climatology Centre (GPCC) Datasets

These are global historical datasets of precipitation built using measurements from ground stations worldwide. GPCC provides three different products, which are gridded datasets (0.5, 1 degrees and 2.5 degrees) made by interpolating ground station data.

Full Data Reanalysis Product: Gridded data product of monthly precipitation derived from quality controlled daily precipitation data from 67,200 stations worldwide from 1901 to 2013. Used for regional or global long term trend analyses.

Monitoring Product: Gridded data product of monthly precipitation derived from quality controlled precipitation data from 7,000 stations worldwide from 1982 to present. Intended to be used for calibration of satellite data and drought monitoring.

FIrst Guess Product: Gridded data product derived from daily station data from a limited number of stations worldwide since 2009.

Info Link:
https://www.dwd.de/EN/ourservices/gpcc/gpcc.html AND
https://www.esrl.noaa.gov/psd/data/gridded/data.gpcc.html#detail

Download Link: ftp://ftp-anon.dwd.de/pub/data/gpcc/html/download_gate.html

Climatologies at High Resolution for Earth's Surface Area (CHELSA)

A gridded monthly precipitation dataset of 0.5 degrees spatial resolution beginning in 1979 until 2013 The dataset is maintained by the Swiss Federal Institute for Forest, Snow and Landscape Research and is derived from a Global Circulation Model names ERA.

Info Link: http://chelsa-climate.org/about/
Access Link: http://chelsa-climate.org/downloads/

3.2.2 Soil moisture

ESA - CCI Soil Moisture

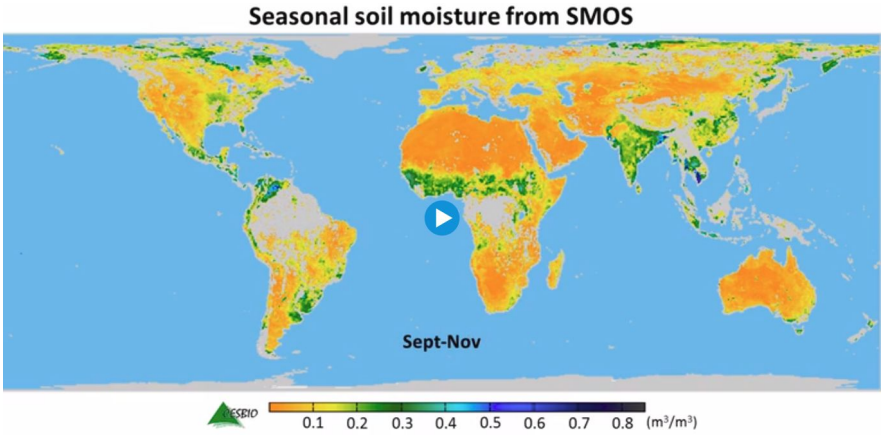
The first dedicated mission to measure soil moisture was initiated by the European Space Agency (ESA) in 2009. To address the deficiency of long term soil moisture data the ESA worked on modeled soil moisture datasets derived from past satellite missions that weren't especially dedicated for soil moisture measurements. In 2012, they released the results of this effort, a 32 year long dataset spanning from 1978 till 2010.

Access Link:
http://www.esa-soilmoisture-cci.org/node/145 (for download)
https://wci.earth2observe.eu/ (to view online)

European Space Agency - Soil Moisture Ocean Salinity (SMOS)

This mission of the ESA begun in 2009. Its radio telescope has a microwave imaging radiometer which passively picks up microwave emissions from the earth's surface that record the degree of soil moisture. The satellite is able to measure soil moisture as low as 4% and has a spatial resolution of 35-50 kms and a temporal frequency of 1-3 days

Info Link: http://www.esa.int/Our_Activities/Observing_the_Earth/SMOS/Mapping_moisture
Access Link: https://smos-ds-02.eo.esa.int/oads/access/
Resources: Special Journal Issue http://www.sciencedirect.com/science/journal/00344257/180?sdc=1

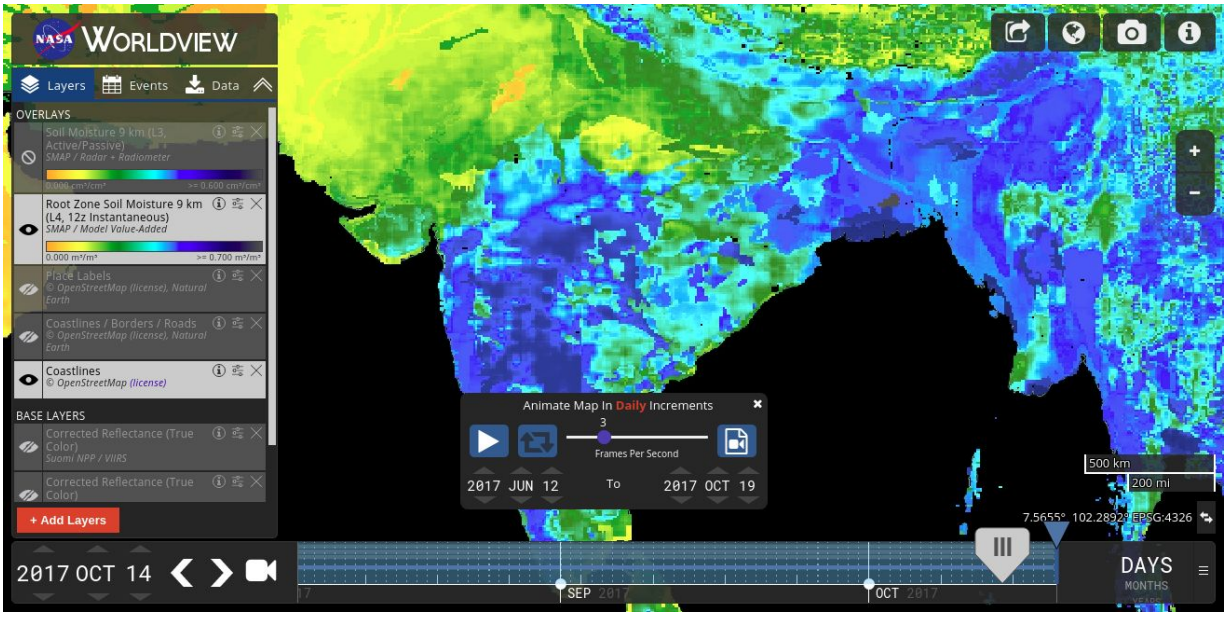


Soil Moisture data of the ESA's Soil Moisture Ocean Salinity Satellite mission

NASA - Soil Moisture Active Passive (SMAP)

In addition to the European Union NASA also has a dedicated mission for measuring Soil Moisture called SMAP. The mission which began in April 2015 suffered a setback just a few months after its launch when one of its radar instruments failed. The mission however still continues with the second radar being complemented by radar data from the Sentinel 1A satellite of the ESA. The composite data from SMAP is available at a resolution of 9 kms and with a temporal frequency of 3 days.

Info Link: https://smap.jpl.nasa.gov/
Access Link: https://worldview.earthdata.nasa.gov/ & https://nsidc.org/data/smap/smap-data.html



Soil Moisture data of the NASA's Soil Moisture Active Passive Satellite mission

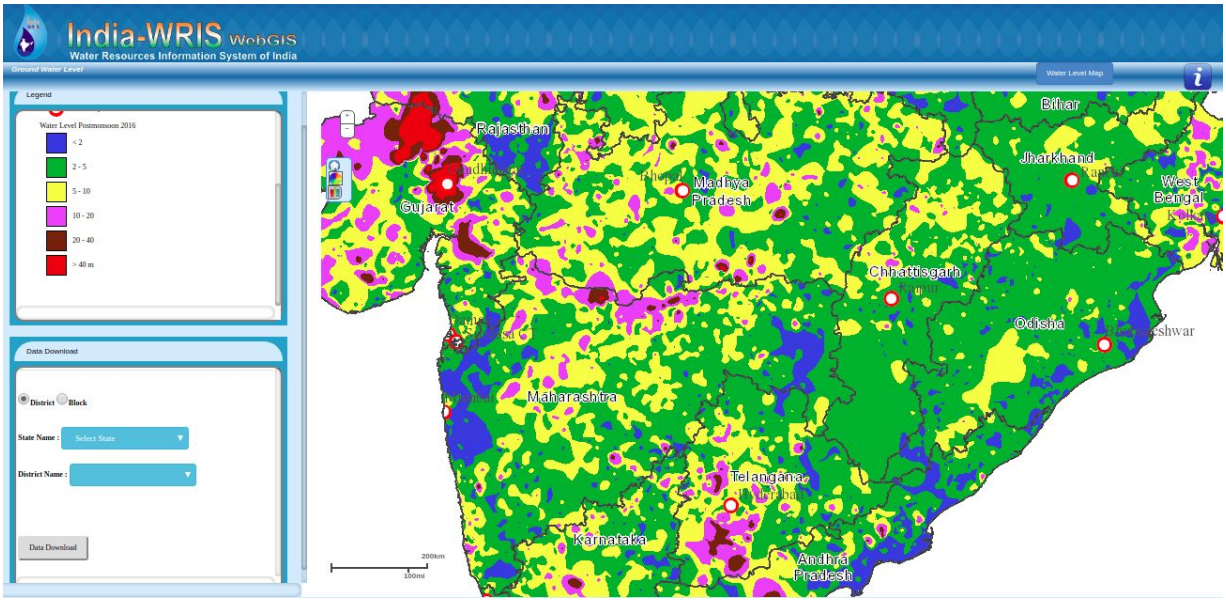
While the sources mentioned above are pre-prepared soil moisture datasets, the literature shows that soil moisture can be estimated with a higher degree of resolution based on indices composed of measurements from bands of other satellites such as LANDSAT and SENTINEL. These indices use land surface temperature as a proxy for soil moisture. These approaches should be explored further in another paper.

3.2.3 Groundwater Levels

Central Groundwater Board (CGWB)

The CGWB has monitoring wells across the country and takes measurement at each well four times a year, January, March-May, August and November. This data is made available through the India WRIS Platform. Bulk data can be downloaded from the following link after selecting the relevant district. In many cases the data available begins in the 1990s and continues till 2016

Access Link: http://www.india-wris.nrsc.gov.in/GWL/GWL.html



Snapshot of the Groundwater level map as available through the India WRIS website

The data describes the depth below groundwater level for each monitoring well and the precise latitude and longitude of the well itself.

	A	B	C	D	E	F	G	H	I	J	K	L
1	STATE	DISTRICT	TEH_NAME	BLOCK_NAME	LAT	LON	SITE_NAME	SITE_TYPE	WLCODE	YEAR_OBS	MONSOON	POSTMONSOONKHARIF
2	MS	Pune	Bhor	Baramati	18.10056	74.48056	Dhumalwadi	Dug Well	W17549	2016	8.25	8.33
3	MS	Pune	Bhor	Bhor	18.20417	73.90222	Kasurdi-1	Dug Well	W25766	2016	0.4	
4	MS	Pune	Manchar	Pune City	18.53889	73.84583	Pune-1	Dug Well	W07385	2016	3.8	3.15
5	MS	Pune	Baramati	Ambegaon	19.06667	73.925	Mahalunge P	Bore Well	W21438	2016	1.44	7.84
6	MS	Pune	Lonavala	Mulshi	18.58194	73.53194	Kolwan	Dug Well	W17572	2016		1.55
7	MS	Pune	Lonavala	Mulshi	18.49639	73.72	Bukum	Dug Well	W17571	2016		0.2
8	MS	Pune	Khadkale	Junnar	19.16667	74.21667	Ane	Dug Well	W17567	2016	3.17	2.6
9	MS	Pune	Khadkale	Junnar	19.13333	73.98333	Narayangaon	Dug Well	W07372	2016	6.45	7.37
10	MS	Pune	Khadkale	Junnar	19.13333	74.19306	Belhe	Dug Well	W17568	2016		
11	MS	Pune	Khadkale	Junnar	19.27083	73.96667	Otur	Dug Well	W07373	2016	14.65	13.4
12	MS	Pune	Indapur	Haveli	18.63333	74.06111	Parne	Dug Well	W07357	2016	6.4	5.12
13	MS	Pune	Dehu Road	Velhe	18.3	73.63333	Wehle	Dug Well	W07354	2016	0.01	0.52
14	MS	Pune	Indapur	Haveli	18.33333	73.85	Shivpur Khed	Dug Well	W07360	2016	1.4	2.07
15	MS	Pune	Khadkale	Junnar	19.20611	73.86944	Junnar	Dug Well	W17569	2016	0.87	3.7

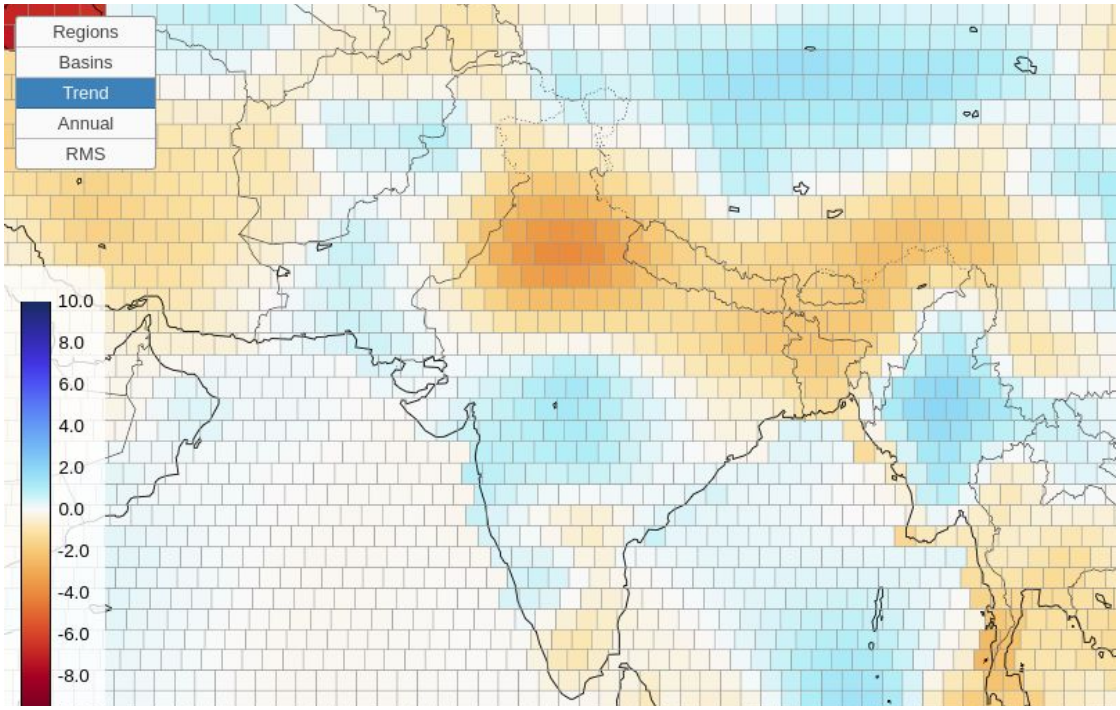
Snapshot of the Groundwater level data as available through the India WRIS website

NASA - Gravity Recovery and Climate Experiments (GRACE)

The GRACE mission is a pair of satellites which together measure minute changes in the earth's gravitational field and use these measurements to infer changes in water and ice across the world, including groundwater resources. The spatial resolution of the data is comparatively low (about 300km), which makes it unsuitable for watershed level analysis but river basin analysis is possible. Moreover the temporal resolution is high, with monthly data available.

Info link: https://grace.jpl.nasa.gov/mission/grace/ ,
https://grace.jpl.nasa.gov/applications/groundwater/

Access link: https://grace.jpl.nasa.gov/data/get-data/
Resources for working with GRACE data
Mascon Visualization Tool: http://ccar.colorado.edu/grace/gsfc.html



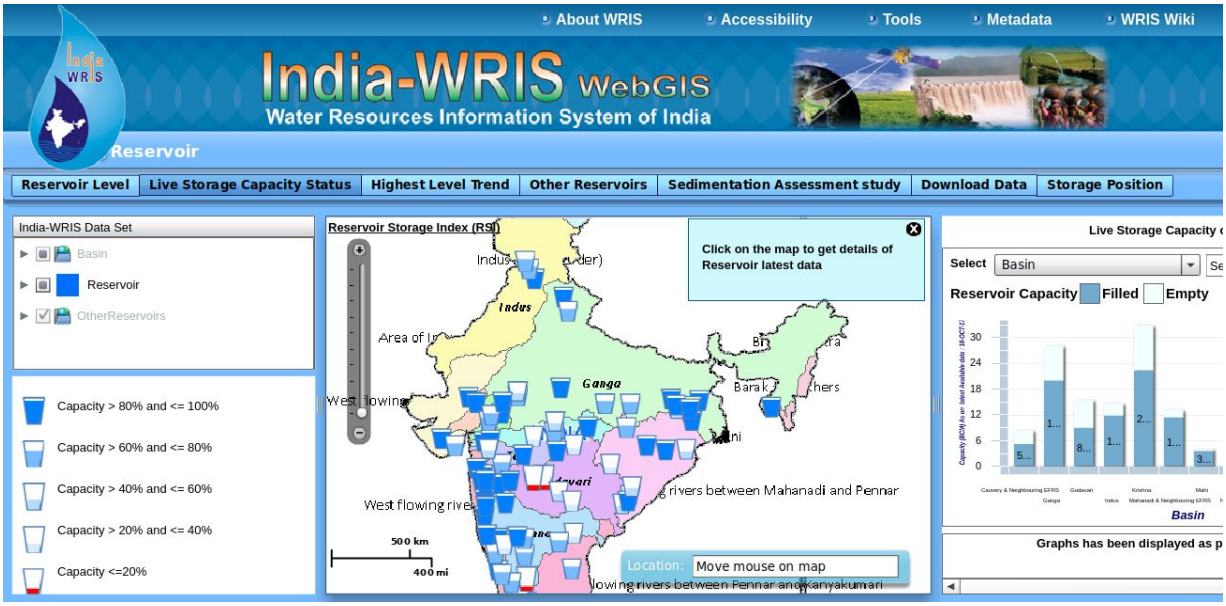
Long term changes in groundwater levels over India with data from GRACE

3.2.4 Surface Water Reservoirs

Central Water Commission

Reservoir levels of major reservoirs in the country are published on the India WRIS platform. The data is near real time and goes back to historical record. It is of irregular frequency, in some cases it is of daily frequency but in many cases it isn't. The duration is different for each reservoir and in many cases doesn't go back to the date of commissioning of the reservoir.

Access Link: http://www.india-wris.nrsc.gov.in/ReservoirApp.html



Data of reservoir levels available through the India WRIS website

	A	B	C
1	Reservoir Data: Hirakud Reservoir		
2	Period: 2005-2017		
3	No of Records: 4240		
4	Date	LEVEL (m)	STORAGE (BCM)
4233	27-JUL-17	185.38	1.47
4234	28-JUL-17	185.7	1.58
4235	02-AUG-17	185.03	1.35
4236	09-AUG-17	186.09	1.72
4237	16-AUG-17	187.02	2.1
4238	23-AUG-17	188.08	2.57
4239	07-SEP-17	189.45	3.271
4240	14-SEP-17	189.31	3.19
4241	20-SEP-17	189.44	3.264
4242	27-SEP-17	192.02	4.823
4243	04-OCT-17	192.02	5.378
4244	11-OCT-17	192.02	5.378

Data of reservoir levels available through India WRIS, of almost daily frequency

3.2.5 Surface Water Flows

Central Water Commission

The India WRIS platform also makes available historical data of the Central Water Commission (CWC) for some of its monitoring stations. The Ganga and Brahmaputra River Basins have the bulk of the CWC monitoring stations however the data from them isn't available in the public domain. For other stations, where available, data is of daily frequency, in many cases going back 20-40 years and contains information on gage level, discharge, and sediment flow.

Access Link: http://www.india-wris.nrsc.gov.in/HydroObservationStationApp.html

Map showing monitoring stations of the CWC for river flows (data from all stations is not open)

Centre for Sustainability and the Global Environment - Global River Discharge Database

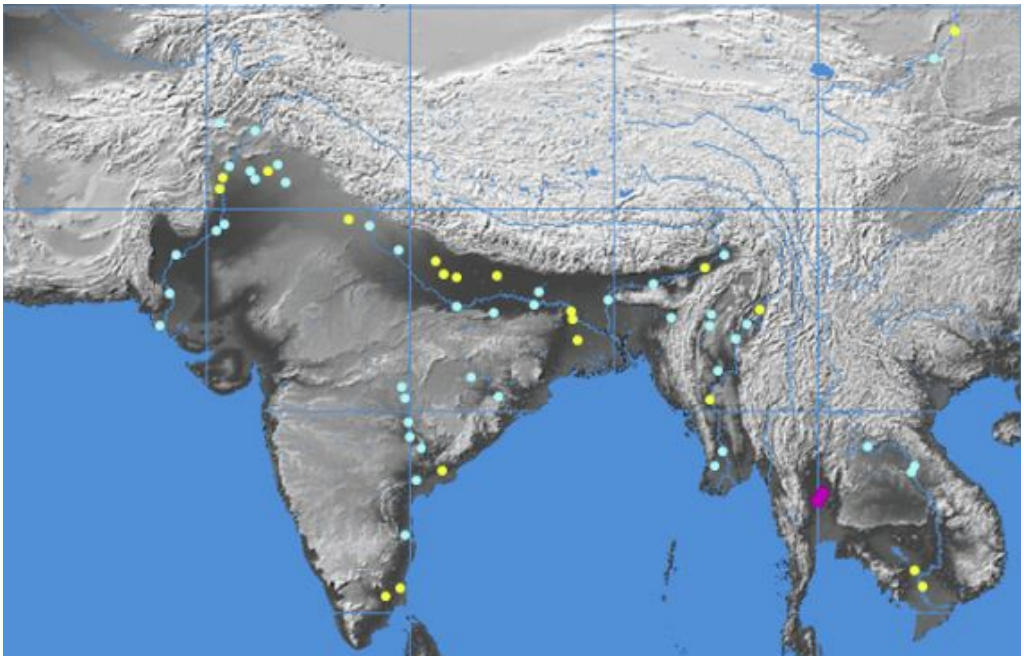
The University of Wisconsin-Madison has put together a database of global river discharge at about 3500 stations around the world. The data is actual monthly discharge available for 45 stations in India. These stations provide data for anywhere between 10 to 80 years and the latest year for any station is 1979.

Access Link: https://nelson.wisc.edu/sage/data-and-models/riverdata/index.php
Original data: https://daac.ornl.gov/RIVDIS/guides/rivdis_guide.html

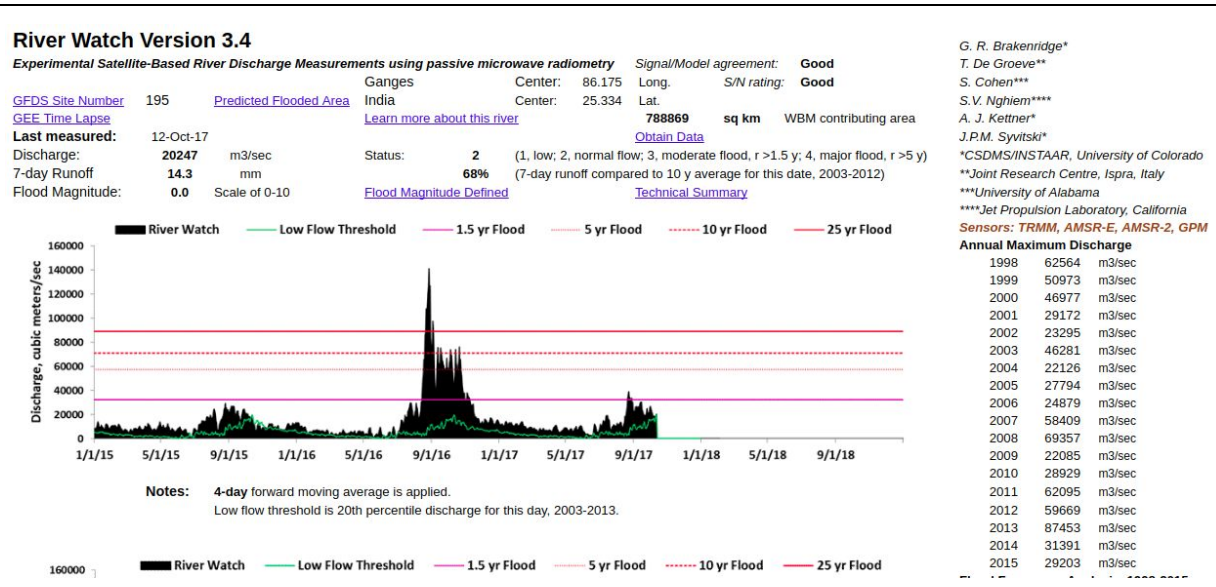
Dartmouth Flood Observatory

The Dartmouth Flood Observatory (Colorado University) does not have any directly measured discharge data but it instead employs remote sensing techniques. Its River Watch v3.4 obtains real time remotely sensed microwave data that is used to determine the quantum of water discharge at 2500 carefully selected sites globally of which several are in India, many in the Ganga and Brahmaputra basins. This is a record of daily estimated river discharge with the records starting in the year 1998.

Project Website: http://floodobservatory.colorado.edu
University of Colorado page: http://csdms.colorado.edu/wiki/Data:Dartmouth_Flood_Observatory
Select a monitoring site: http://floodobservatory.colorado.edu/DischargeAccess.html (full page takes time to load)



Map of river gauge stations for which data is available on River Watch

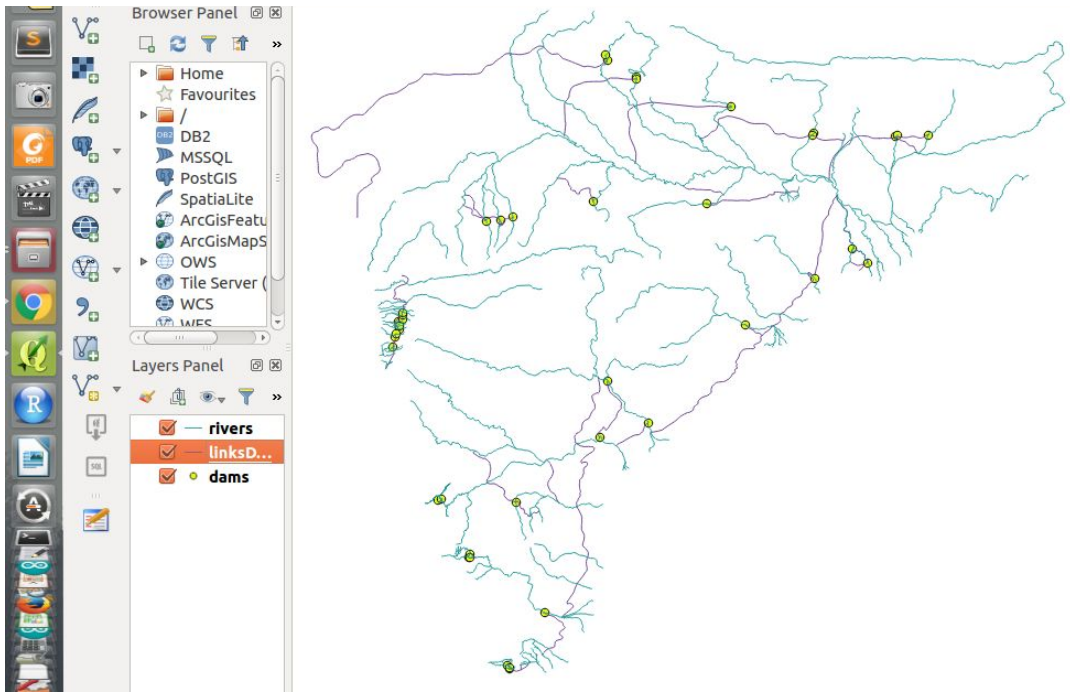


Charts of historical river gauge data produced from data available on River Watch

National River Linking Plan Modeled Database

This database contains modeled information by the Community Surface Dynamics Modeling System (CSDMS) at the University of Colorado. The database contains data on each of the links and dams and relevant literature on river discharges.

Access Link: http://csdms.colorado.edu/wiki/Data:NRLP_India



Modeled river discharge data available on proposed National RIver Linking Plan

World Meteorological Organization (WMO) - Global Runoff Data Centre

The GRDC maintains on behalf of the WMO data from approximately 9000 stations globally river discharge data, which is available for free for research purposes on signing a User Declaration. This data hasn't been thoroughly explored.

Access Link: http://www.bafg.de/GRDC/EN/02_srvcs/21_tmsrs/riverdischarge_node.html

3.3 Water Demand

3.3.1 Evapotranspiration

Evapotranspiration is essentially the largest component of 'water demand'.
It can be thought of as

1) Evaporation: the amount of water directly taken up by heat of the atmosphere from the bare surface of the earth.
2) Transpiration: the amount of water taken up by vegetation of different types.

Hence water taken away by evapotranspiration becomes unavailable for other uses. There are multiple factors which affect the rate of evapotranspiration, including solar radiation, atmospheric temperature, wind speed, humidity, pre-existing soil moisture and importantly vegetation type (crop or forest)

Evapotranspiration rates can be determined from ground measurements of these parameters and also from remote sensing measurements. Rates of evapotranspiration for a river basin or watershed can be estimated once one knows the Land Cover pattern in the area. Multiplying the area under each Land use type (urban, agriculture, forest, grassland, water body etc) by the estimated rates of ET gives us an approximation of 'water demand' by ET (as a volume)
These approximations of 'water demand' can be made more accurate if we knew not only whether a piece of land was agricultural, but also if we knew the specific crop growing there, since different crops have different rates of ET. Similarly for forests, different forest types have different rates of ET.

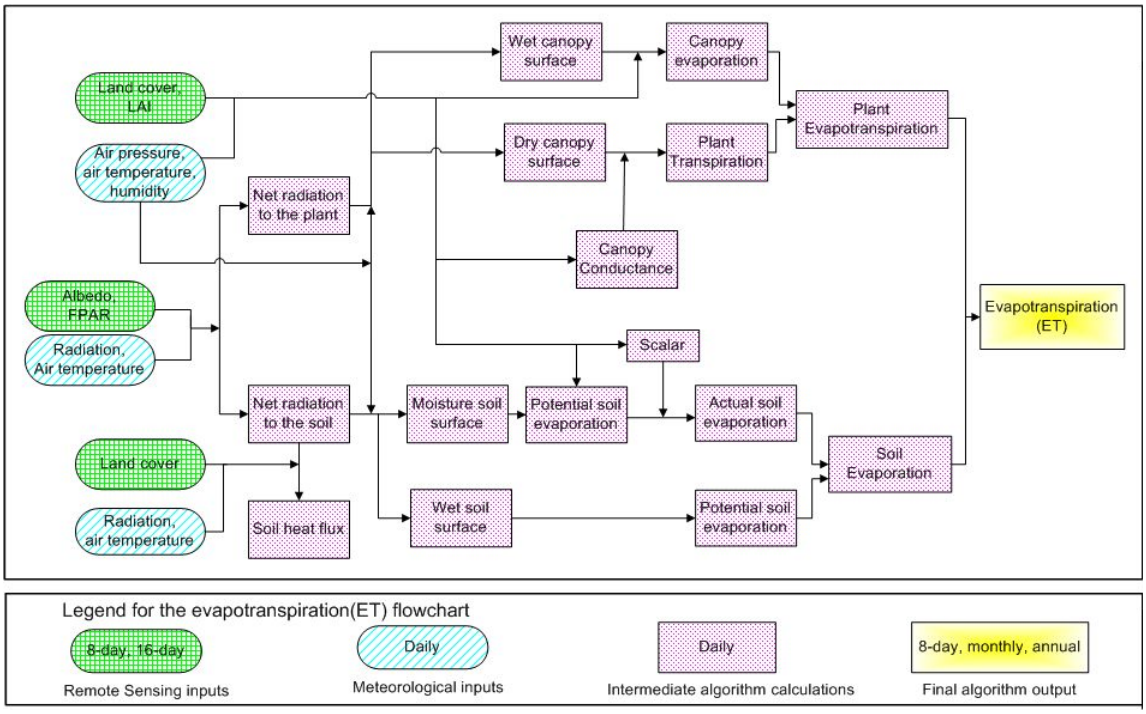
Determining Evapotranspiration rates through Remote Sensing

ET rates can be determined directly using RS, by observing all relevant parameters (solar radiation, air temperature, wind speed, humidity, moisture and land use/cropping patterns) from satellite measurements or by a mixture of remote and ground observations. Some pre-prepared ET datasets are listed below. These estimates can be improved on by incorporating more local data measurements that validate ET.

MODIS Global Evapotranspiration (MOD16)

Prepared by the '*Numerical Terradynamic Simulation Group*' at the University of Montana this dataset uses a selection of remote sensing and meteorological datasets to derive estimates for Evapotranspiration (ET) for the entire vegetated area of the earth at a spatial resolution of 0.5km^2 for 8 day timesteps for the period from 2000 to present.

Info Link: http://www.ntsg.umt.edu/project/modis/mod16.php
Access Link: https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod16a2_v006

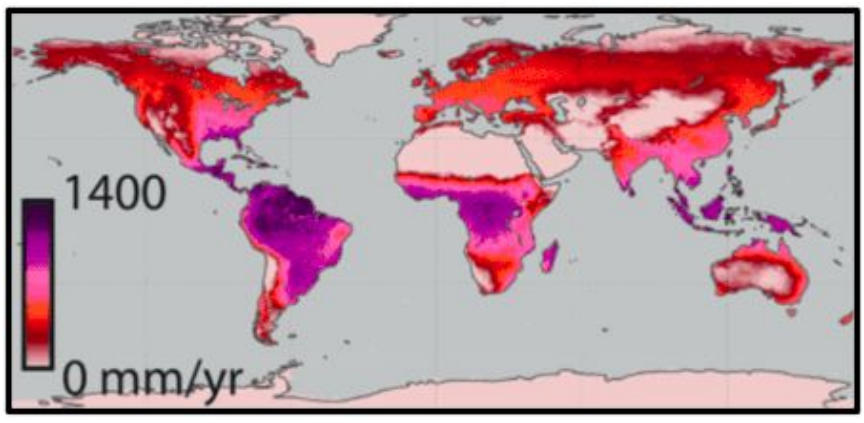


Mu et al. (2011) model used to estimate ET by NTSG group, University of Montana

Global Land Evaporation Amsterdam Model (GLEAM)

GLEAM is a set of models that separately estimate land evaporation, transpiration, open water evaporation, surface and root zone soil moisture, etc with the goal to maximize evapotranspiration data availability using. The models have been used to create three versions of the GLEAM dataset, v3.1a, 3.1b, 3.3c of which v3.1a is derived from both satellite and meteorological data and spans the period from 1980 -2016. This data has been shown to have a correlation of about 0.64 with ground measurements of ET which makes it suitable for regional studies but not watershed level analysis.

Info and Access Link: https://www.gleam.eu/

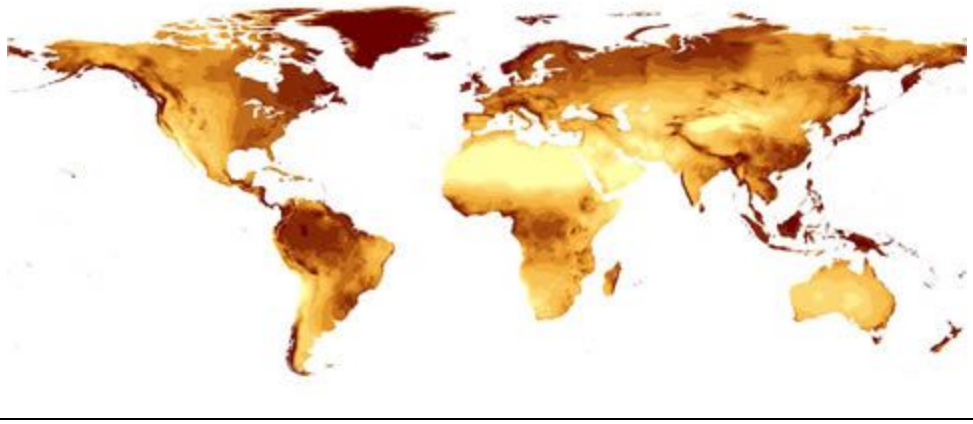


Snapshot of Global transpiration data using the GLEAM model

Global Potential Evapo-Transpiration (Global-PET) and Global Aridity Index (Global-Aridity) dataset

The Global PET database is a dataset that provides high resolution PET data and a measure of rainfall deficit for vegetation growth. The data is available at 1 km^2 resolution and at monthly timesteps. The model used to derive the dataset is Zomer et al (2007 and 2008)

Info and Access Link:
http://www.cgiar-csi.org/data/global-aridity-and-pet-database
https://www.dropbox.com/sh/e5is592zafvovwf/AAAijCvHNiE4mYvYqWDpeJ3Ga/Global%20PET%20and%20Aridity%20Index?dl=0



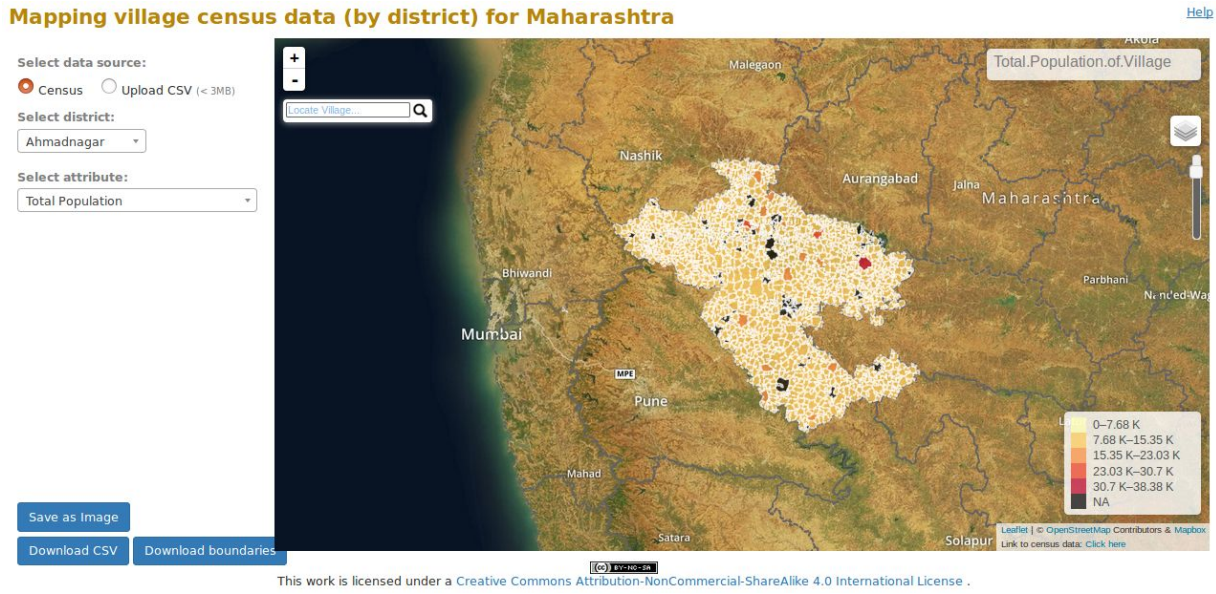
Snapshot of the Global PET and Aridity database

3.3.2 Domestic Water Demand (Urban and Rural)

Census Town and Village Amenities

Domestic water demand in each watershed across the country is primarily a factor of the human population located in the watershed. The data for population in India is usually available for larger administrative areas (districts, states etc) not for hydrological units (watersheds, river basins etc) However data for population at the smallest administrative unit (village boundaries) available with the Census of India can be used to infer the population within each watershed when juxtaposed with watershed boundary maps.

Access Link: http://projects.datameet.org/indian_village_boundaries/



Village level population data can help estimate domestic water demand in watersheds

3.3.3 Industrial Water Demand

Data on Industrial water demand requires two individual datasets to come together.

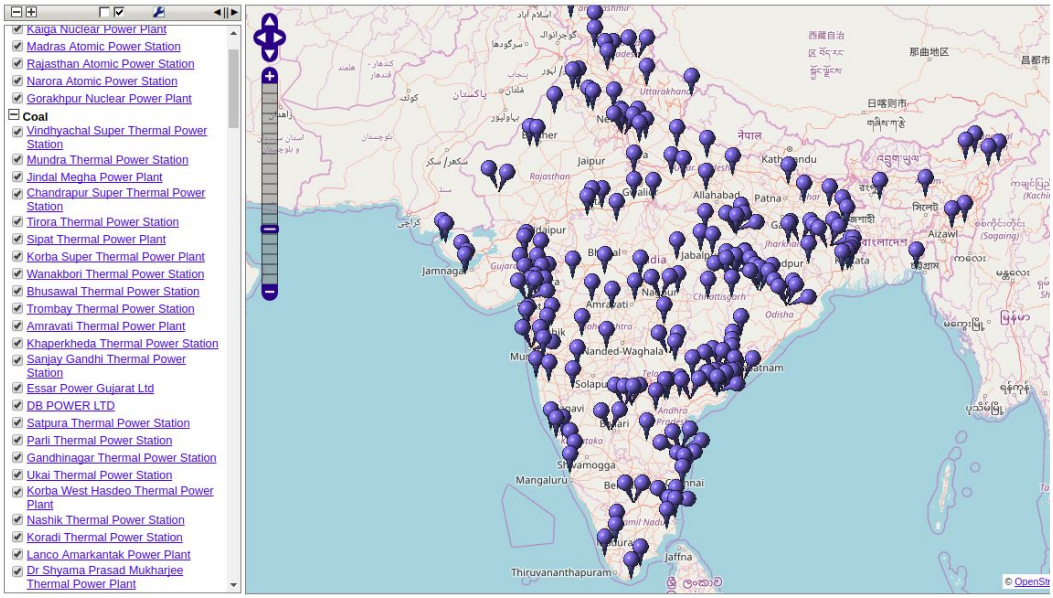
1. Spatial data on locations of large industries in the country
2. Actual water use/Allocations of water to each industry

Studies show that from a water use perspective, thermal power is the largest industrial user of water.

Openstreetmap

Openstreetmap is one such database which provides location information for thermal power plants across the country. Since this is a community generated database it is unlikely to be up to date, but it does contain what appears to be an exhaustive list of power plants in the country. This database contains no information on the water use of or allocation to each industry

Access Link:
https://tools.wmflabs.org/osm4wiki/cgi-bin/wiki/wiki-osm.pl?project=en&article=List_of_power_stations_in_India

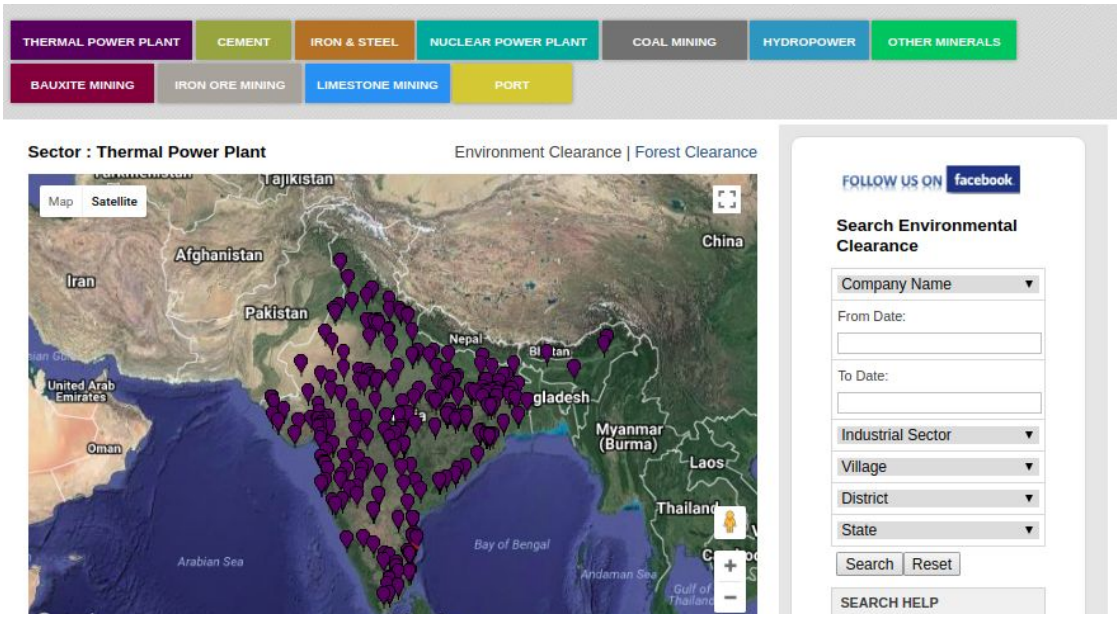


Locations of thermal power stations as available on Open Street Map

Centre for Science and Environment - Green Clearance Watch

Another source for location information for industries is the database compiled by the Centre for Science and Environment. This database has been compiled as part of the Green Clearance Watch project and contains information on water allocation to each industry

Access Link: http://www.greenclearancewatch.org/node/3222/44



Locations of thermal power stations as seen on Centre for Science and Environment

Ministry of Environment and Forests - Industrial Water Allocations

This database contains information on the water allocation (i.e. clearance given by the MoEF) to each industry in the form of its Environmental Clearance data

Access Link: http://environmentclearance.nic.in/

S.No	Proposal Details	Location	Important Dates	Company/Proponent	Type of project	* Attached Files
1	**Proposal No** : IA/JH/THE/32025/2015 **File No** : J-13012/21/2015-IA.I **Proposal Name** : Patratu Super Thermal Power Project, Phase-I (3 x 800 MW) **Category** : Thermal Projects	**State** : Jharkhand **District** : Ramgarh **Tehsil** : Patratu	**Date of Submission for TOR** : 23 Oct 2015 **Date of TOR Granted** : 07 Jun 2016 **Date of Submission for EC** : 09 Jun 2017 **Proposal Accepted by MS** : 19 Jun 2017 **Date of EC Granted** : 07 Nov 2017	PATRATU VIDYUT UTPADAN NIGAM LIMITED	New	
2	**Proposal No** : IA/JH/THE/54853/2016 **File No** : J-13012/01/2016 - IA. I (T) **Proposal Name** : 1600 (2x800) MW Godda Thermal Power Plant of Adani Power (Jharkhand) Limited **Category** : Thermal Projects	**State** : Jharkhand **District** : Godda **Tehsil** : Godda	**Date of Submission for TOR** : 02 Jun 2016 **Date of TOR Granted** : 26 Jul 2016 **Date of Submission for EC** : 19 May 2017 **Proposal Accepted by MS** : 20 May 2017 **Date of EC Granted** : 31 Aug 2017	Santosh Kumar Singh	New	

Snapshot of data available on industrial environmental clearances on MoEF

Annexure 1: Dataset Summary Table

Each dataset in this assessment has been given a simple 'open data' score based on the indicators 5 to 11. The maximum possible score is 12

Name	Time Period	Area	Spatial resolution	Temporal resolution	License to use	Data format	Web map	Web Map Service	API	Free	Metadata & Documentation	Open data score
No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	
Points for 'open data' score					0-Closed/ Unknown 2 - NC 3 - C	0-Physical copy 1-PDF 2-Digital	0-No 1-Yes	0-No 1-Yes	0-No 1-Yes	0-No 2-Yes	0-None 1-Minimal 2-Good	
River Network												
GoI - Central Water Commission	NA	India	All 1st order perennial streams	NA	Unknown	.pdf	Yes	Yes	No	Yes	None	5
U.S.G.S. Hydrosheds	NA	Global	25 km^2 watershed	NA	NC	.shp	No	No	No	Yes	Good	8
Open Street Map	NA	Global	Variable	NA	NC (ODbL)	.osm, .shp	Yes	Yes	Yes	Yes	Good	11
Watershed Boundaries												
USGS Hydrosheds	NA	Global	Major river basins	NA	NC	.shp	No	No	No	Yes	Good	8
CWC Watershed Atlas	NA	India	700 km^2 watersheds	NA	Unknown	.pdf	Yes	Yes	No	Yes	Minimal	6
SLUSI Watershed Atlas	NA	India	Larger than CWC	NA	Unknown	.pdf	No	No	No	No	Good	3
CGWB Watershed Maps	NA	India	Similar to SLUSI	NA	Unknown	.jpg	No	No	No	Yes	Minimal	4
Water Reservoirs												
GRanD Database	upto 2008	Global	NA	NA	NC (ask for C)	.kml, .csv	Yes	No	No	Yes	Good	9
Global Surface Water Explorer - European Commission	1985 to 2015	Global	NA	Annual	C	.TIFF	Yes	Yes	Yes	Yes	Good	12
Open Street Map	NA	Global	Variable	NA	NC (ODbL)	.osm, .shp	Yes	Yes	Yes	Yes	Good	11
Elevation												
NASA - SRTM	2000	Global	30 m	NA	C	TIFF	Yes	Yes	Yes	Yes	Good	12
NASA - ASTER	NA	Global	30 m	NA	C	TIFF	Yes	No	No	Yes	Good	10
JAXA - ALOS	2006	Global	30 m	NA	C	TIFF	Yes	No	No	Yes	Good	10
NASA - GTOPO30	NA	Global	1 km	NA	C	TIFF	Yes	No	No	Yes	Good	10

Name	Time Period	Area	Spatial resolution	Temporal resolution	License to use	Data format	Web map	Web Map Service	API	Free	Metadata & Documentation	Open data score
No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	
NRSC - CartoSat	2005	India	30 m	NA	C	TIFF	Yes	No	No	Yes	Good	10
Rainfall												
IMD - CRIS	Last 1 day	India	District wise	Daily	Unknown	.html	Yes	No	No	Yes	None	5
IMD - AWS	Last 1 week	India	Station wise	Hourly	Unknown	.csv	No	No	No	Yes	None	4
IMD - Raw Data Daily	1900 - 2017	India	Station wise	Hourly/Daily	C	.csv	No	No	No	No	None	5
IMD - Raw Data Monthly	1901 - 2010	India	District	Monthly	Unknown	.csv	No	No	No	Yes	None	4
IMD - Modeled data	1901 - 2015	India	25 km^2 grid	Daily	C	.grd	No	No	No	No	Minimal	6
TRMM	1997 to 2015	Global	25 km^2 grid	3 hours	C	netCDF TIFF	Yes	No	No	Yes	Good	10
GPM - IMERG	2015 to present	Global	10 km^2 grid	0.5 hours	C	HDF5 netCDF TIFF	Yes	No	No	Yes	Good	10
CHIRPS	1981 to present	Global	5 km^2 grid	Daily	C	BIL netCDF TIFF	Yes	No	No	Yes	Good	10
PERSIANN	2000 to present	Global	25 km^2 grid	hourly	C	netCDF	Yes	No	No	Yes	Good	10
GHCN	1900 to present	Global	Station wise	variable	NC	.dly	Yes	No	No	Yes	Good	9
PREC/L	1948 to present	Global	50 km^2 grid	monthly	C	netCDF	Yes	No	No	Yes	Good	10
GLPT - UD	1900 to 2014	Globa;	50 km^2 grid	monthly	NC	netCDF ascii	No	No	No	Yes	Good	8
GPCC	1901 to 2013	Global	50 km^2 grid	monthly	NC	netCDF ascii	Yes	No	No	Yes	Good	9
CHELSA	1979 to 2013	Global	50 km^2 grid	monthly	C	TIFF	No	No	No	Yes	Good	9
Soil Moisture												
ESA - Soil Moisture	1978 to 2010	Global	25 km^2 grid	daily	C	netCDF	Yes	No	No	Yes	Good	10
ESA - SMOS	2010 to present	Global	40 km^2 grid	2 days	C	netCDF	Yes	No	No	Yes	Good	10
NASA - SMAP	2015 to present	Global	9 km^2 grid	3 days	C	HDF	Yes	No	Yes	Yes	Good	11
Groundwater Levels												
CGWB	1990s to 2016	India	Station wise	3 months	Unknown	.csv	Yes	Yes	No	Yes	Minimal	7
NASA - GRACE	2002 to present	Global	300 km^2 grid	monthly	C	netCDF	Yes	No	No	Yes	Good	10

Name	Time Period	Area	Spatial resolution	Temporal resolution	License to use	Data format	Web map	Web Map Service	API	Free	Metadata & Documentation	Open data score
No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	
Surface Water Storage												
CWC - WRIS	1980s to present	India	reservoir wise	daily	Unknown	.csv	Yes	No	No	Yes	None	4
Surface Water Flows												
CWC - WRIS	1970s to present	India	station wise	daily	Unknown	.csv	Yes	No	No	Yes	None	5
CSGE - Global River Discharge Database	1900 to 1979	Global	station wise	monthly	Unknown	ascii	No	No	No	Yes	Good	6
Dartmouth Flood Observatory	1998 to present	Global	station wise	daily	C	Digital	No	No	No	Yes	Good	9
National River Linking Model of Flows	Unknwon	India	Gridded	NA	Unknown	.shp .csv	No	No	No	Yes	Minimal	5
WMO - Global Runoff Data Centre	1900 to present	Global	station wise	NC	NC	.xls	No	No	No	Yes	Minimal	7
Evapo-transpiration												
MODIS 16	2000 to present	Global	0.5km^2 grid	8 days	C	HDFEOS2	Yes	No	No	Yes	Good	10
GLEAM	1980 to 2016	Global	Unknown	Unknown	NC	Digital	No	No	No	Yes	Good	8
Global PET		Global	1 km^2 grid	monthly	NC	Arc/INFO grid	Yes	No	No	Yes	Good	9
Domestic Water Demand												
Census Town and Village Amenities Database	2001, 2011	India	Village	10 yrs	NC	.csv	Yes	No	No	Yes	Good	6
Industrial Water Allocations												
Open Street Map	1990s to present	India	NA	NA	NC (ODbL)	.osm, .shp	Yes	Yes	Yes	Yes	Minimal	9
CSE - Green Clearance Watch	2007 to present	India	NA	NA	Unknown	HTML	Yes	No	No	Yes	Minimal	5
MoEF - Environmental Clearances	2000 to present	India	NA	NA	Unknown	PDF	No	No	No	Yes	Minimal	4

Annexure 2: Note on Datasets not included

The compilation in this paper is intended to be an introduction to satellite (remote sensing) and other meteorological datasets that are most directly useful for the study of water resources from a physical sciences perspective. The datasets included here are those that can be immediately plugged into water balance equations to derive estimates of water availability and demand in a given area of interest.

Other datasets that may also be useful for water resource studies have been excluded if the information they provide cannot directly be used but require to be incorporated in a model to estimate the water cycle component. For instance: temperature, wind speed, humidity, etc are all datasets that can be used to infer evapotranspiration rates which gives us an idea of water demand. However such datasets have been left out for the purpose of keeping the paper concise and to the point. These could perhaps be included in an extended version of this paper.

Similarly for the purpose of keeping things concise state level datasets have not been included, the focus has been only on datasets published by the Central Government. Water quality datasets have also not been included, since this paper focuses more on the quantitative aspects, i.e. quantum of water available and water demand.

Besides this the paper has also primarily focused on listing free sources of water data (An exception has been made if the priced source is government in origin). Some examples of data providers of priced remote sensing products that can be directly or indirectly used for water studies are listed below.

- Planet Labs
- Urthecast
- Vandersat
- Digital Globe
- Climate Corporation

Annexure 3: Resources for Introductory GIS and Remote Sensing

Much of Introductory GIS becomes clear when playing with the data firsthand in QGIS (the most popular free and open source GIS platform)
Video tutorials: https://www.youtube.com/channel/UCrBM8Ka8HhDAYvQY1VX2P0w/videos
Popular Website for Intro to GIS: http://www.qgistutorials.com/en/docs/introduction.html
Gentle Intro to QGIS: http://docs.qgis.org/2.18/en/docs/gentle_gis_introduction/introducing_gis.html
QGIS Official documentation: http://www.qgis.org/en/docs/index.html#documentation-for-qgis-2-0

Remote sensing requires a familiarity with the physics of light and energy
Useful Book: https://www.itc.nl/library/papers_2009/general/PrinciplesRemoteSensing.pdf
NASA Webinars on Remote Sensing:
https://arset.gsfc.nasa.gov/webinars/fundamentals-remote-sensing