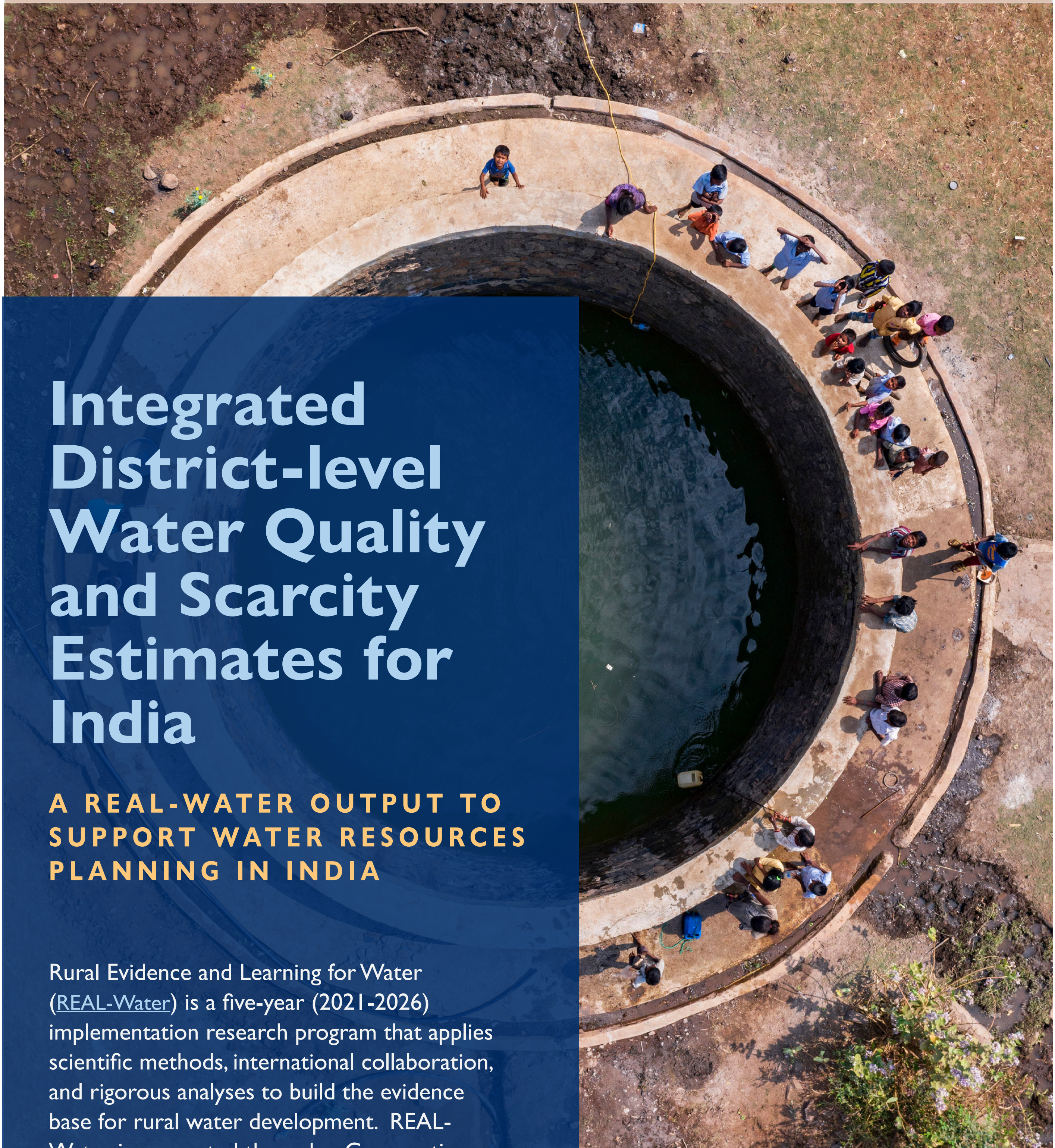


Integrated District-level Water Quality and Scarcity Estimates for India

A REAL-WATER OUTPUT TO SUPPORT WATER RESOURCES PLANNING IN INDIA

Rural Evidence and Learning for Water (REAL-Water) is a five-year (2021-2026) implementation research program that applies scientific methods, international collaboration, and rigorous analyses to build the evidence base for rural water development. REAL-Water is supported through a Cooperative Agreement between USAID and The Aquaya Institute (Aquaya).

October 2023
www.globalwaters.org/real-water

BACKGROUND

REAL-Water has compiled a rich spatially-explicit dataset of water quality and scarcity parameters for India to assist decisionmakers and researchers in understanding the distribution of water resources risks across the country. Pulling together publicly available data collected and distributed by the government of India as well as a satellite-derived dataset from the European Commission's Joint Research Centre, our team has generated a single district-level Geographic Information System (GIS) shapefile which allows for visualization and queries of 226 water quality and scarcity-related variables processed at the district level in India. The file type allows visualization and analysis on most GIS software platforms.

The district shapefile for India integrates information from multiple datasets in different spatial formats. Its groundwater quality parameters are drawn from a network of monitoring **points** (typically water supply wells). Its groundwater utilization parameters correspond to district-level polygon summaries of values collected at local Indian administrative units (Blocks/Mandals/Taluks). Its aquifer properties are drawn from aquifer formation boundaries which are also represented as **polygons**. And its precipitation and surface water extent information are collected as **gridded** data. Source data were compiled at the district level for production of the shapefile, and are publicly available for download along with detailed metadata descriptions.

SOURCE DATA

The integrated shapefile for India draws on multiple Indian government datasets as well as a dataset of the European Commission's Joint Research Centre.

The Central Groundwater Board (CGWB)

1. Groundwater quality
Water Quality Reports on the Central Ground Water Board's website.[1]
Parameters: As, F, Cl^-, NO_3^-, TDS
Spatial unit: individual wells, point lat/longs
Years: 2010, 2012, 2014, 2016, 2018

2. Aquifer Properties
Aquifer Systems of India.
Parameters: "principal aquifer systems" or formations
Spatial unit: contiguous formations (as polygons) originally mapped at roughly 1:250,000 scale.
Year: 2012

3. Groundwater exploitation
National Compilation on Dynamic Ground Water Resources of India, 2020.
Parameters: Extraction and estimated recharge
Spatial unit: "Assessment units" – either watershed boundaries (with high (>20%) slope areas removed) or administrative blocks in alluvial areas
Years: 2020

1 These data were available for view and download during the assembly of our integrated dataset, but as of November 2023 they no longer appear on the CGWB website. We are unaware if or when these data may be restored.

The European Commission's Joint Research Centre's Copernicus Programme's Global Surface Water Dataset

See Pekel et al (2016). High-resolution mapping of global surface water and its long-term changes. *Nature* 540, 418-422. (doi:10.1038/nature20584)
Parameters: Surface water occurrence change
Spatial unit: gridded, 30 x 30m resolution
Years: 1999-2020

The Indian Meteorological Department (IMD) New High Spatial Resolution Long Period Daily Gridded Rainfall Data Set Over India

See Pai et al (2014). Development of a new high spatial resolution long period grided rainfall data set over India and its comparison with existing data sets over the region. *MAUSAM* 65 (1) 1-18 (doi: 10.54302/mausam.v65i1.851)
Parameters: Daily rainfall.
Spatial unit: gridded 15 x 15arc-sec (between roughly 390 x 390m and 450 x 450m, depending on latitude)
Years: 1901-2022

DATA PROCESSING

Our data integration involves different operations for different input data formats. While some input information was already computed a district level (such as estimates of groundwater extraction and recharge, which India's Central Groundwater Board summarizes in "Dynamic Groundwater Resources of India 2020" (specifically Annexure III(B), "District-wise Categorization of Blocks/Mandals/Taluks in India (as in 2020)"), the point and grid information (as well as the aquifer formation polygon data that do not match district boundaries) required processing to compute district-level values.

For point data, this entailed grouping all observations (e.g. nitrate concentrations) within a district for a given year and specifying the number of observations, minimum value, maximum value, mean and median values, and standard deviations, with each value corresponding to a field in the shapefile's tabular data.

For gridded data (precipitation from the Indian Meteorological Department and surface water extents from the Copernicus Programme's Global Surface Water dataset), we assigned to the district a value equivalent to the average of all grid cells falling within its boundaries. For boundary grid cells (which do not sit entirely within a district polygon), only those cells whose center is located within the district polygon are included.

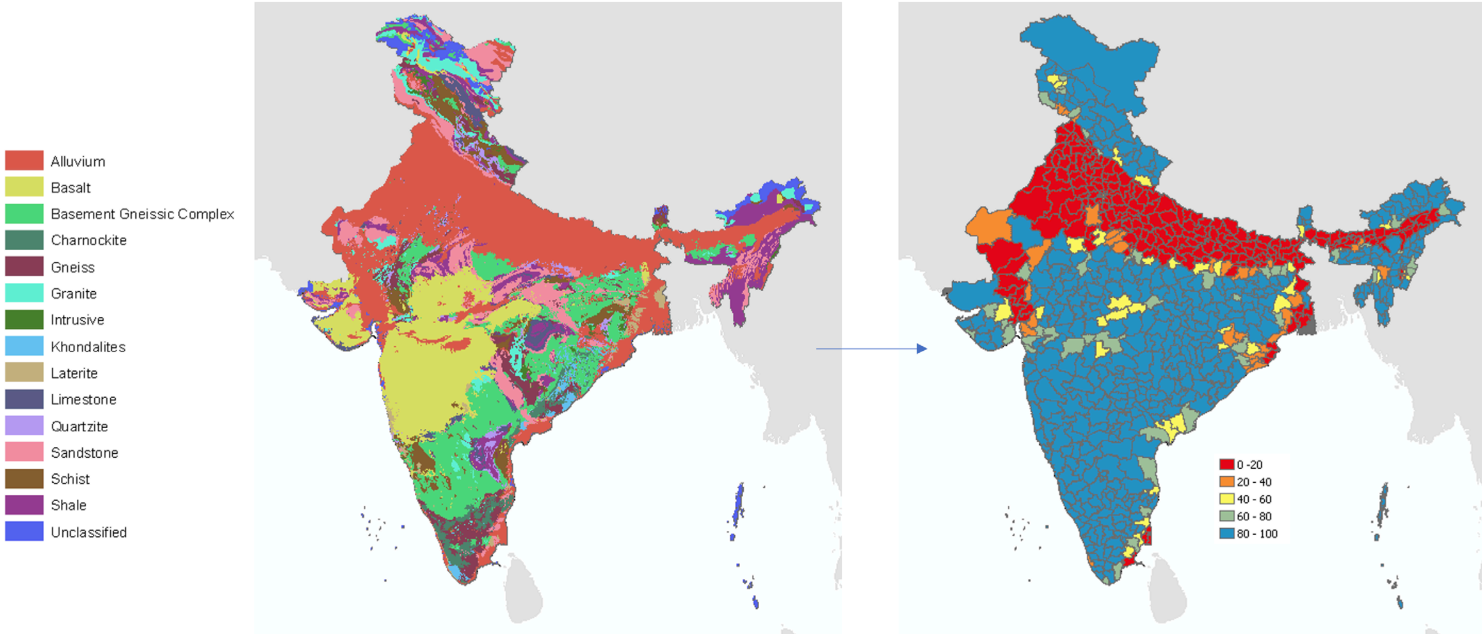


Figure 1. Conversion of aquifer formation polygons (left) to district-level aquifer statistics (right) in the REAL-Water integrated shapefile for India. The maroon category at left corresponds to alluvium, dominated by the Indo-Gangetic plain as well as some coastal districts. At right, values are proportions of each district comprised of consolidated (hard rock) medium. The distinction between consolidated and unconsolidated aquifer is important because of their very different recharge properties and depletion patterns, with consolidated systems often more difficult to predict because of the abundance of fractured (vs uniform) flow as well as more localized dynamics

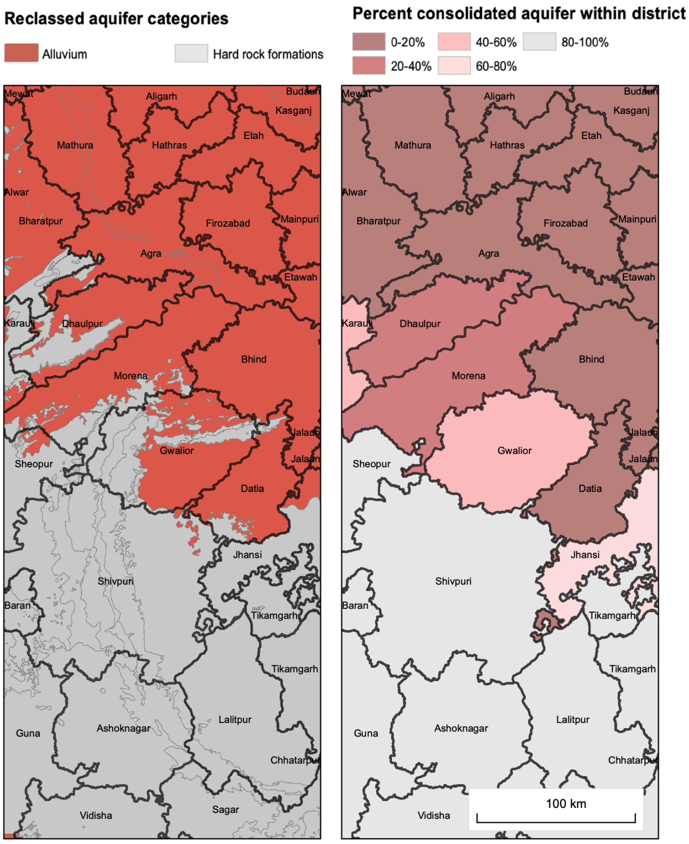
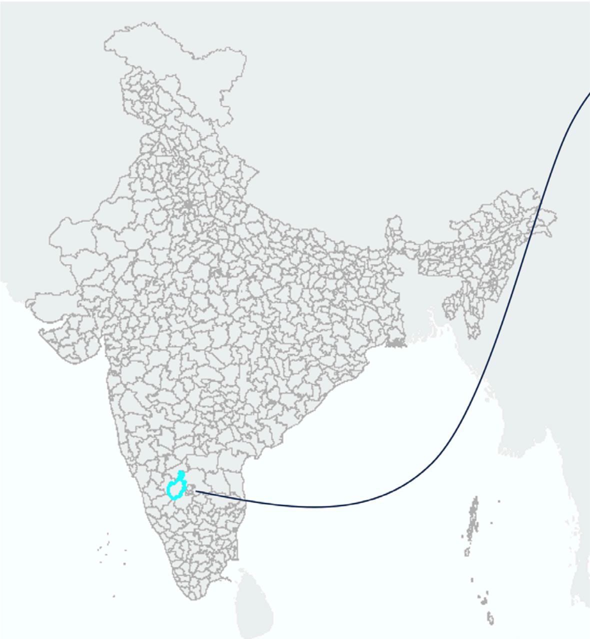


Figure 2. Reclassed aquifer map (left), with overlain district boundaries in bold stroke, and computed consolidated aquifer percentages (right).

For the aquifer data, we executed a two-step process. First, we reclassified the input layer from the original 14 types to two values ("unconsolidated" corresponding to Alluvium, and "consolidated" corresponding to all 13 hard rock categories), as depicted in Figure 1. We then assigned to each district a value corresponding to the proportion of the district area made up of consolidated aquifer and the proportion made up of unconsolidated aquifer (Figure 2). We computed the values via the ***Summarize Within*** operation in ArcGIS Pro 2.8 and the results are captured as one of the fields of the 226 field shapefile ("Consol_p" – percentage of the area of the district comprised of consolidated aquifer).

DATA STRUCTURE AND ILLUSTRATIVE DATA VISUALIZATIONS AND QUERIES

The final shapefile is a map of district polygons with a corresponding tabular data file with 226 water resource-related parameters (Figure 3).



State Name	Karnataka
District	Chitradurga
Total groundwater quality observations for yr 2018	14
Total chloride observations for yr 2018	14
# of chlorine observations above limit for yr 2018	8
Minimum chloride value for yr 2018 (mg/L)	21.3
Maximum chloride value for yr 2018 (mg/L)	791.65
Average chloride value for yr 2018 (mg/L)	327.1
Median chloride value for yr 2018 (mg/L)	291.1
Standard deviation of chloride values for yr 2018 (mg/L)	239.1
Total nitrate observations for yr 2018	14
# of nitrate observations above limit for yr 2018	9
Minimum nitrate value for yr 2018 (mg/L)	10.9
Maximum nitrate value for yr 2018 (mg/L)	306.64
Average nitrate value for yr 2018 (mg/L)	85.8
Median nitrate value for yr 2018 (mg/L)	68
Standard deviation of nitrate values for yr 2018 (mg/L)	77
Groundwater recharge observations in the “Safe” category	1
Groundwater recharge observations in the “Semi Critical” category	0
Groundwater recharge observations in the “Critical” category	0
Groundwater recharge observations in the “Overexploited” category	5
Groundwater recharge observations in the “Saline” category	0
Average surface water extents for the period 1999-2003 (km2)	110.576953
Average surface water extents for the period 2016-2020 (km2)	82.8583147
Percentage consisting of consolidated aquifer (%)	100
Percentage of assessed groundwater units in a district classified as either “Critical” or “Overexploited” (%)	83

Figure 3. Relational database structure of the integrated district-level shapefile for India. The black arrow indicates a selection of the tabular data associated with the district of Chitradurga in Karnataka state. Only 22 of the 226 parameters per district are included in the figure, for display clarity purposes.

Image 1: Woman filling drinking water in rural Rajasthan, India.

In Figure 4 we offer examples of single parameter displays that are represented in the integrated shapefile.

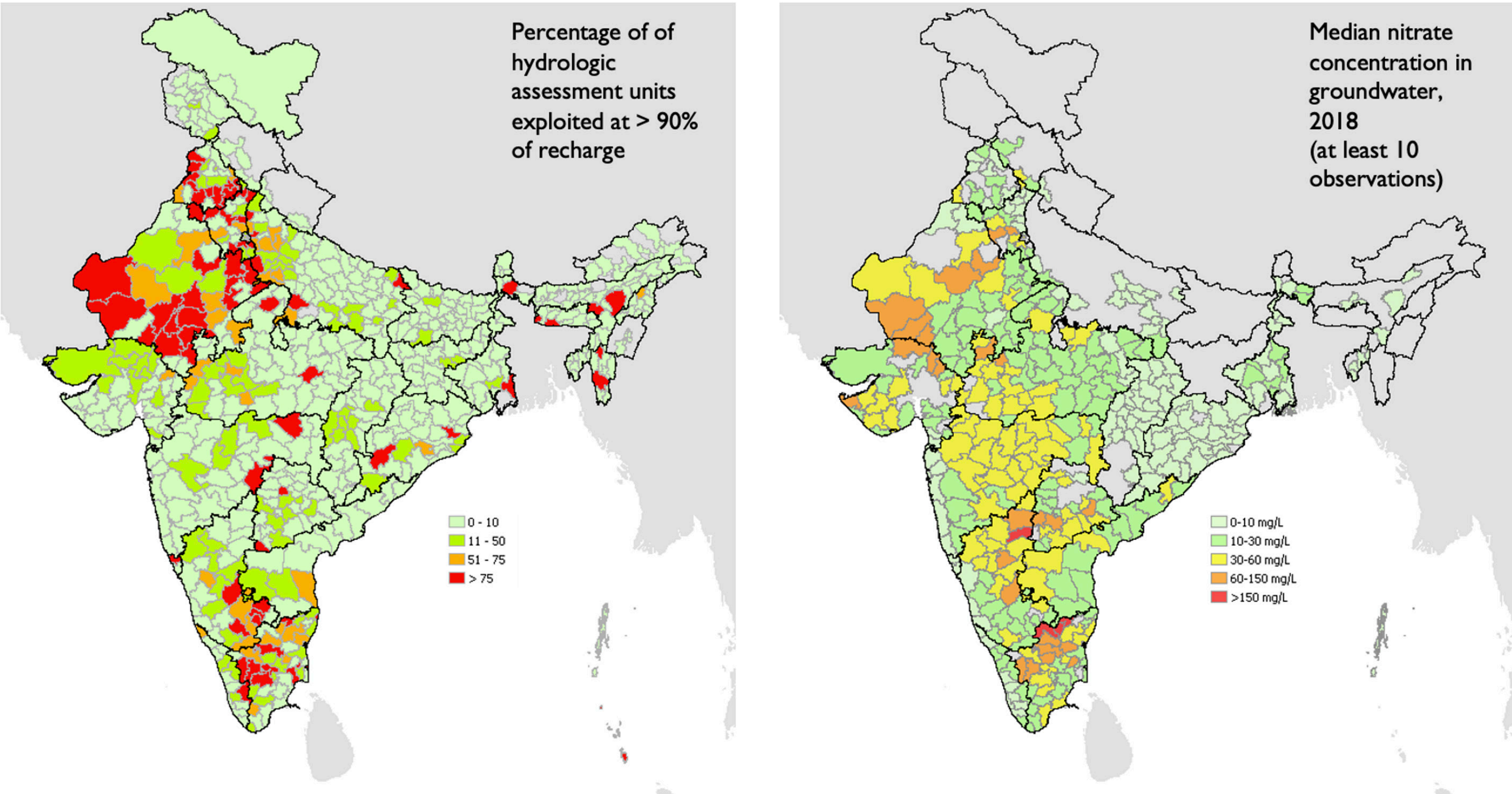


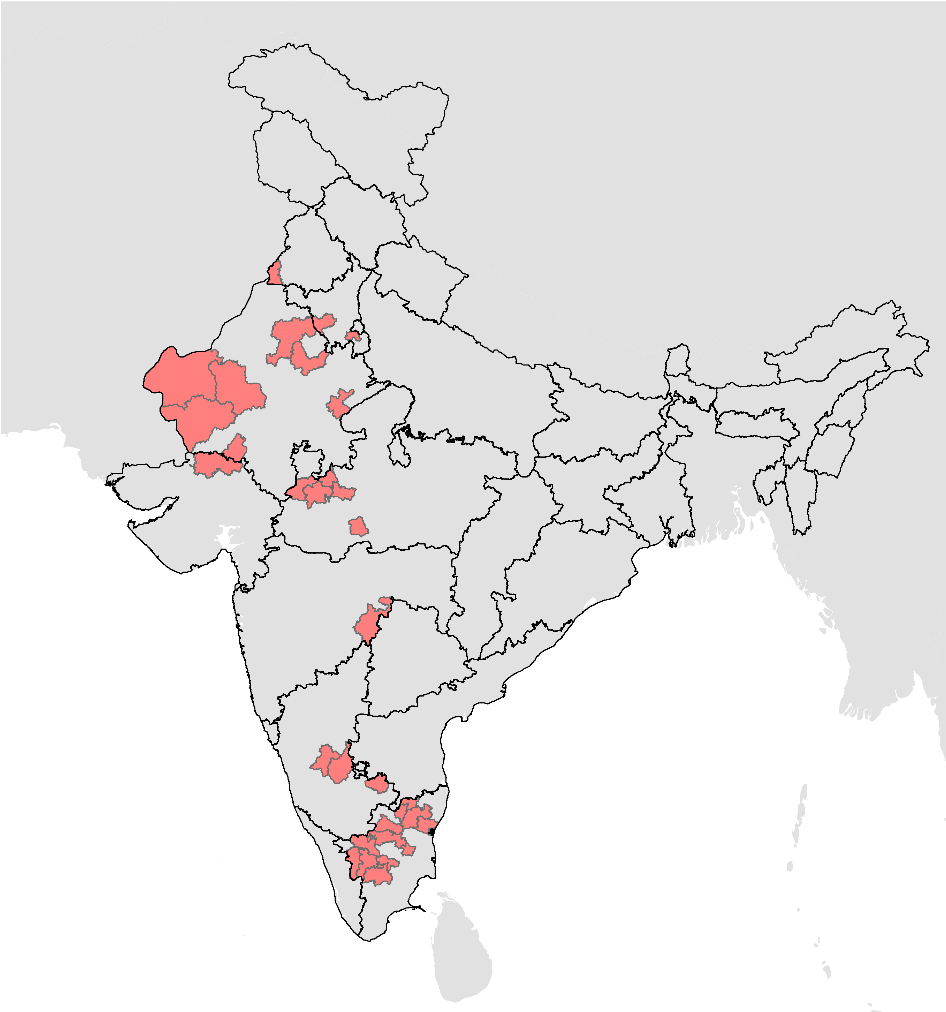
Figure 4. Single parameter displays for groundwater exploitation and nitrate concentration. Left: districts exhibiting groundwater exploitation risks, with red polygons having more than 75% of assessed hydrologic units either "critical" (withdrawals at >90% of recharge) or "over-exploited" (>100% of recharge). Right: the distribution of nitrate contamination of groundwater (with yellow, orange, and red polygons indicating median values above 30 mg/L in 2018).

The figure's left panel illustrates district-level risks for groundwater depletion as measured by comparisons of withdrawal and recharge rates in sub-district administrative units (block, mandal, or taluk). The numerical values represent the percentage of sub-district administrative units where groundwater withdrawals exceed 90% of recharge. The panel on the right displays median nitrate concentrations for each district.

With 226 parameters, the shapefile enables users to identify and map districts according to specific water quality and scarcity risks that can be assessed individually or in combination. For example, in Figure 5, we highlight those districts exhibiting *both* elevated nitrate (above an arbitrary threshold of 30 mg/L for the median concentration among observations in a district) and high groundwater exploitation (specifically, those with more than half of all assessed administrative units exhibiting withdrawal rates that exceed 90% of recharge rates).

The shapefile also includes the number of observations for specific parameters by year, enabling users to identify periods during which the utility of district-level averages are limited by low observation numbers.

Figure 5. Multi-parameter query result indicating the co-occurrence of anthropogenic groundwater quality risks and groundwater withdrawal stress. Highlighted districts are those with nitrate concentrations above 30 mg/L and groundwater extraction above 90% of recharge in greater than half of all assessed hydrologic units.



In sum, this integrated shapefile allows visualization and querying that will assist decisionmakers, researchers, and other interested stakeholders in spatial analyses of water resources across India. These analyses will, in turn, support the implementation of policies and interventions designed to mitigate water quality and scarcity risks.

PREFERRED CITATION

REAL-Water. (2023). Integrated District-level Water Quality and Scarcity Estimates for India. United States Agency for International Development (USAID) Rural Evidence and Learning for Water.

CONTACTS

Ranjiv Khush, REAL-Water Project Director, ranjiv@aquaya.org
Jeff Albert, REAL-Water Deputy Project Director, jeff@aquaya.org