Indian Village Satellite Imagery and Energy Access Dataset

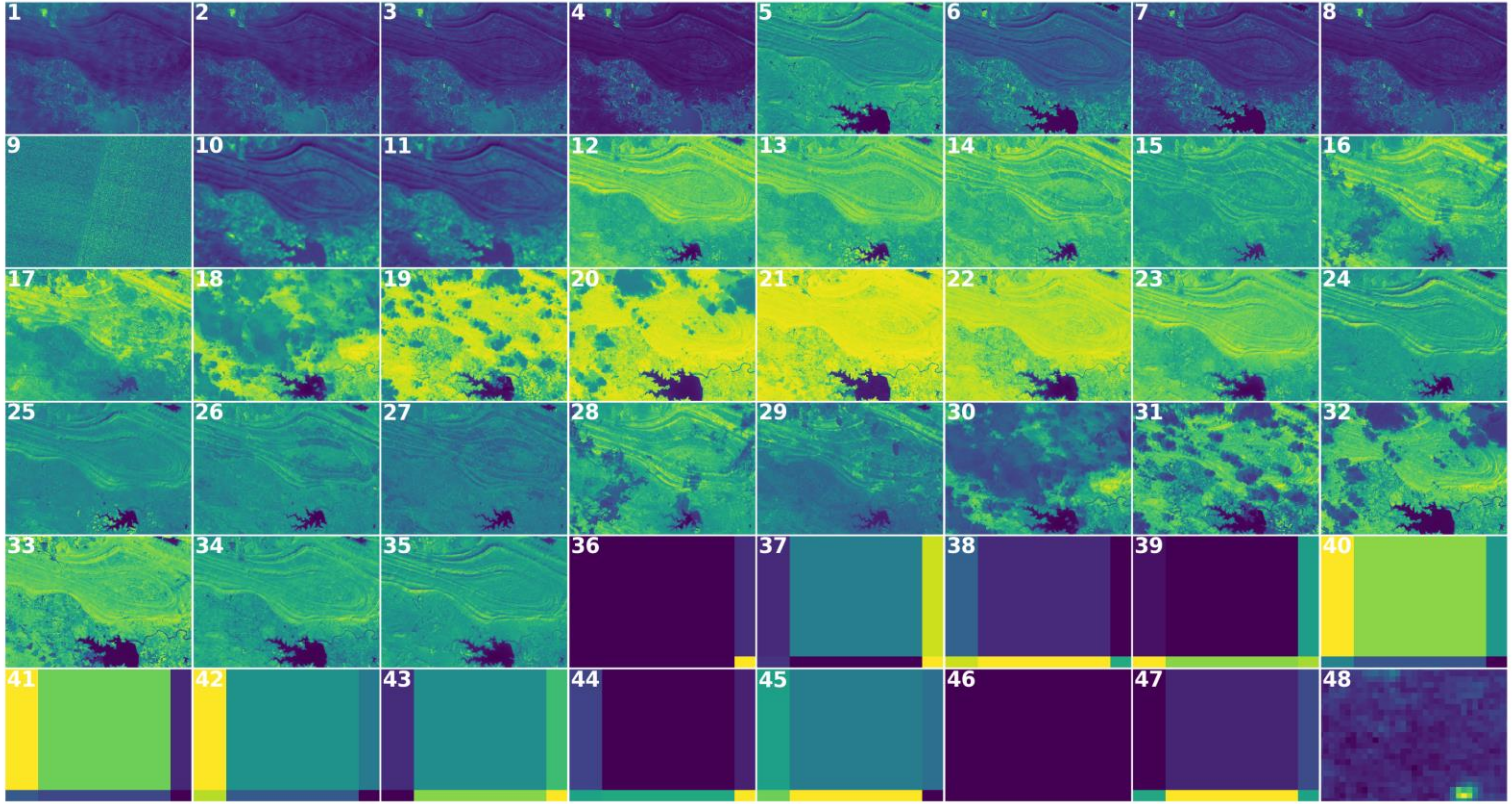
This dataset contains remote sensing data for every village in the state of Bihar, India, and the corresponding electrification rate for most of those villages, as reported by the Garv data platform from the Indian government as of July 2017. This dataset contains imagery data for 45,220 villages and the electrification rate data for 32,817 of those villages. This dataset may be of particular interest to those investigating the use of remote sensing data in the process of electricity access expansion.

File/Folder Name	Description
IndianVillagesDataset_30m.zip	**30-meter resolution data**
└ imagery_res30_48bands	Folder containing a 30-meter resolution .tif files with remote sensing data for each village
└ masks_res30	Folder containing a binary mask representing the location of each village within the imagery
IndianVillagesDataset_15m.zip	**15-meter resolution data**
└ imagery_res15_b8	Landsat's panchromatic Band 8 which is 15-meter resolution
└ imagery_res15_rgb	15-meter Pan-sharpened versions of Landsat's RGB bands
└ masks_res15	Folder containing a binary mask representing the location of each village within the imagery
ElectrificationMap_Bihar.geojson	Geospatial data containing the political boundary of each of the 45,220 villages as well as the census ID, number of households, and number of electrified households
garv_data_bihar.csv	CSV file with the number of households and electrified households for each village

30-meter resolution data

Each village file in the 'imagery_res30_48bands' folder is named with the village name and the village ID (example: *Bihar(vil)-Harkhar-259167.tif*). This file contains 48 bands, the first 11 from Landsat 8, the 12-35 derived from Landsat 8 data over a 12 month period in 2016 computing normalized difference vegetation index (NDVI) and the green index. Bands 36-47 NASA's Global Precipitation Measurement rainfall data, and band 48 is VIIRS lights at night data from the 2016 annual composite. Note that rainfall data and nighttime lights data are lower resolution and have been upsampled for inclusion in this dataset.

Example: Bihar(vil)-Harkhar-259167.tif



Bands

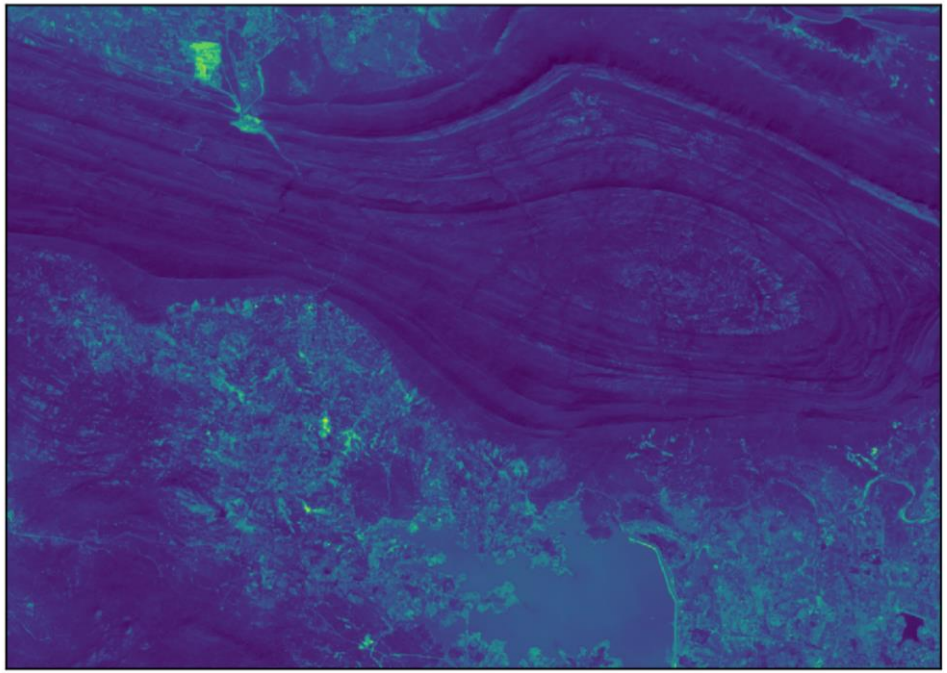
1: Coastal aerosol
(0.43 - 0.45 µm)
2: Blue
(0.45 - 0.51 µm)
3: Green
(0.53 - 0.59 µm)
4: Red
(0.64 - 0.67 µm)
5: Near Infrared
(0.85 - 0.88 µm)
6: Short-wave Infrared 1
(1.57 - 1.65 µm)
7: Short-wave infrared 2
(2.11 - 2.29 µm)
8: Panchromatic (30-meter)
(0.50 - 0.68 µm)
9: Cirrus (1.36 - 1.38 µm)
10: Thermal Infrared 1
(10.60 - 11.19 µm)
11: Thermal Infrared 2
(11.50 - 12.51 µm)
12-23: NDVI
Jan-Dec 2016
24-35: green index
Jan-Dec 2016
36-47: rainfall data
Jan-Dec 2016
48: Nighttime lights

15-meter resolution data

Each village file in the ‘imagery_res15_b8’ and ‘imagery_res15_rgb’ folders are named with the village name and the village ID (example: *Bihar(vil)-Harkhar-259167.tif*). While much of the Landsat 8 data are 30-meter resolution, Band 8 is 15-meter resolution and can be used to pan-sharpen the RGB bands

Example: Bihar(vil)-Harkhar-259167

Band 8 – Landsat 8 Panchromatic layer at full 15-meter resolution



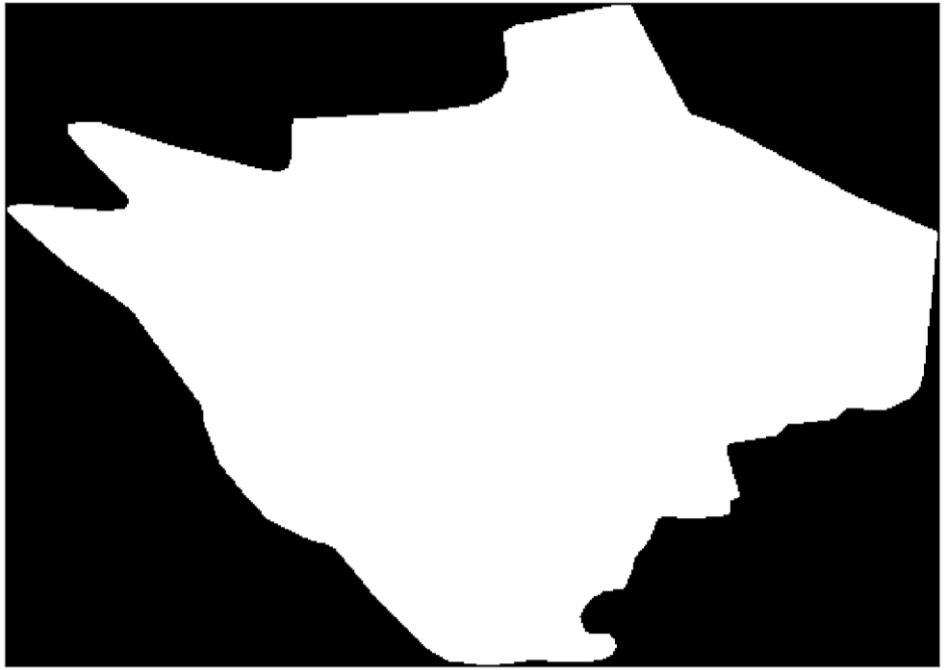
RGB – Landsat RGB layers pan-sharpened to 15-meters
(Shown here contrast-stretched for optimal display)

Village boundary binary masks (15- and 30-meter)

Example: Bihar(vil)-Harkhar-259167

The folders ‘masks_res15’ and ‘masks_res30’ contain a file for each village named with the village name and the village ID (example: *Bihar(vil)-Harkhar-259167.tif*), as all of the data files are named. These files contain the political boundaries of the village that corresponds to the image. This can be used to isolate only those pixels that belong to the village within an image of that resolution.

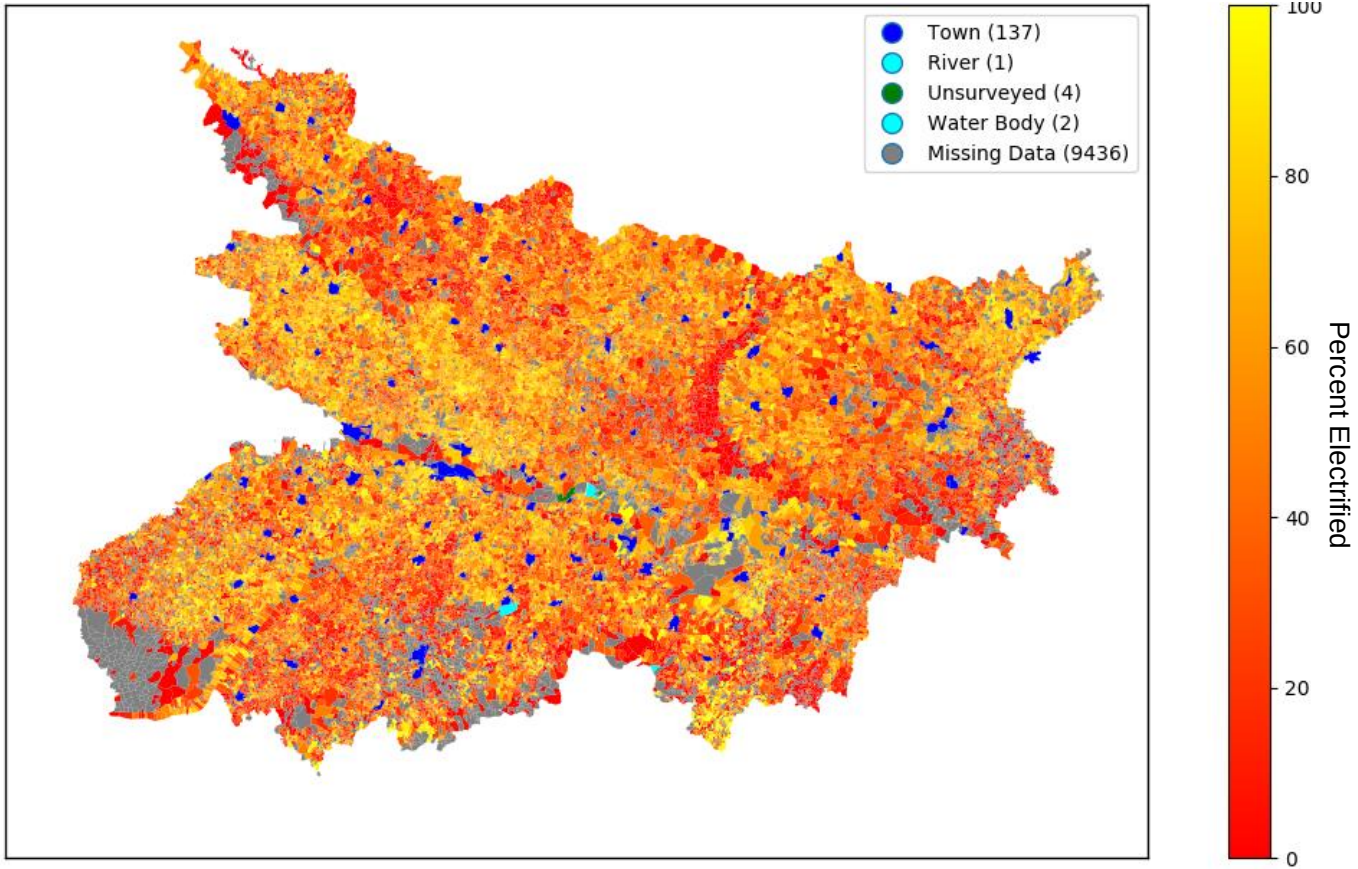
Village boundary binary mask (example shown is 30-meter resolution)



Corresponding RGB Image for reference

Geospatial data in 'ElectrificationMap_Bihar.geojson'

This geospatial data file containing the political boundary of each of the 45,220 villages and the electrification rate data for 32,817 of those villages. These data re plotted in the map shown below



Data Fields

Polygon coordinates. Latitude and longitude coordinates of the village political boundary
SUB_DIST. Sub-district name
CEN_2001. Census 2001 ID
TYPE. Type of region (Village, Town, River, Unsurveyed, Water Body, Missing Data)
STATE. Indian state name
NAME. Name of unit
DISTRICT. Name of district
CEN_2011. Census 2011 ID
HH. Total number of households.
eH. Number of electrified households.
Perc. Percent of electrified households

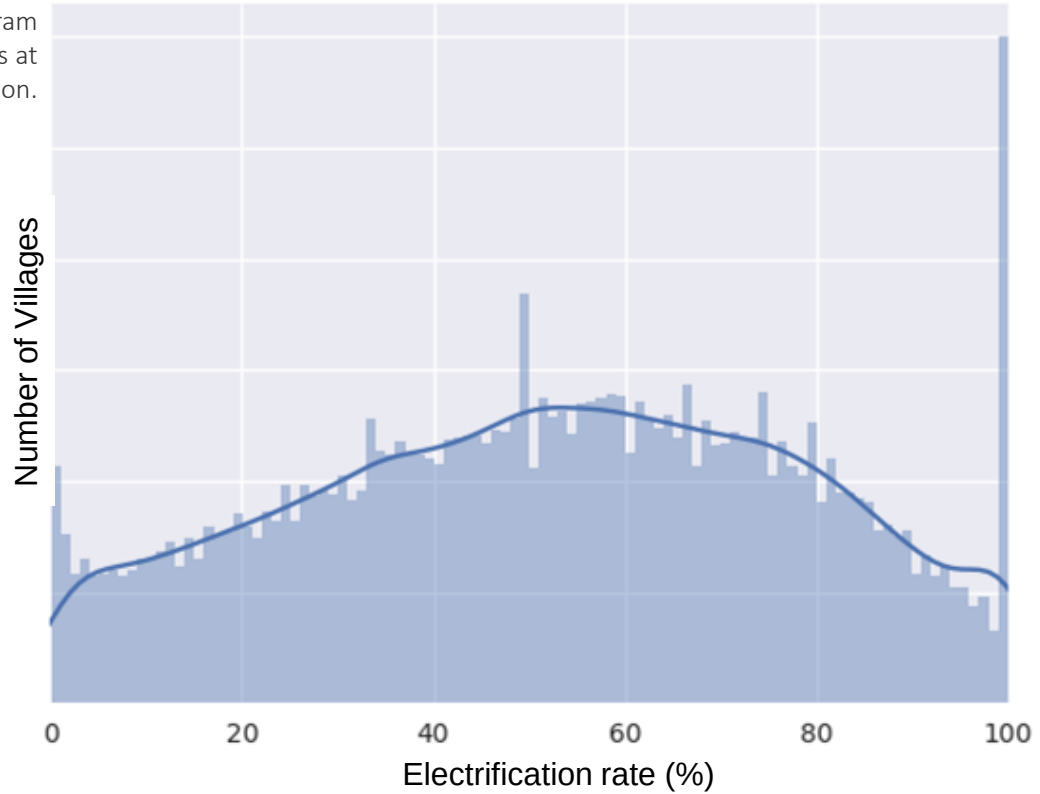
Garv_data_bihar.csv

This flat file contains the Bihar village names, Census 2011 ID (which is used throughout this dataset for uniquely labeling villages) as well as the total number of households and the number of electrified households.

Data Fields

Census 2011 ID
Village Name
District Name
State Name
Number of Households
Number of Electrified Households

This figure shows a histogram of the number of villages at each rate of electrification.



Dataset Sources

USGS Landsat 8 Data
Link: https://landsat.usgs.gov/landsat-8
Download link via USGS EarthExplorer: https://earthexplorer.usgs.gov/

NASA Global Precipitation Measurement (GPM) v4
Link: https://www.nasa.gov/mission_pages/GPM/main/index.html

NOAA VIIRS Stray Light Corrected Nighttime Day/Night Band Version 1
Link: https://ngdc.noaa.gov/eog/viirs/download_dnb_composites.html

Indian Government Garv Electrification Data
Link: https://garv.gov.in/

Datameet Indian Village Boundary Data
Link: http://projects.datameet.org/indian_village_boundaries/