Climate Gender in East Africa - Land Cover Analysis

Juliana Gómez Consuegra

2024-04-29

The following files shows the pipeline used to analyze land cover data from the Copernicus Land Monitoring Service, and Afrobarometer data for East African countries. The goal of this project is to analyze the relationship between land cover and gender issues in East Africa. Given the time constraints for this project, the full analysis was not carried out. However, the csv files for the years 2020-2022, so the pipeline is available for future use.

Methodology

1. Download desired rasters from Copernicus Land Monitoring Service.
2. Georeferenced coordinates from cities in the Afrobarometer data.
3. Clip rasters to desired area, using georeferenced coordinates.
4. Convert rasters to dataframes.
5. Merge dataframes.
6. Output as csv and netcd (in case we want to use clipped maps in QGIS).
7. Load into R and analyze.

Associated notebooks

Steps	Notebook	Folder
1	`Climate_workbook.ipynb`	/Python
2	`georeferencing.ipynb`	/Python
3-6	`clip_based_on_georef.ipynb`	/Python
7	`landcover.RMD (this notebook)`	/r
8	`afrobarometer_analysis.Rmd`	/r
9	`afrobarometer_2022_analysis.Rmd`	/r

Associated files

File	Description	Folder
`land_cover_20.csv`	Land cover csv file for 2020, for countries of interest	`/csv`
`land_cover_21.csv`	Land cover csv file for 2021, for countries of interest	`/csv`
`land_cover_22.csv`	Land cover csv file for 2022, for countries of interest	`/csv`
`satellite_land_cover_2016.nc`	Worldwide land cover data for 2016	`/Maps`

File	Description	Folder
satellite_land_cover_2017.nc	Worldwide land cover data for 2017	/Maps
satellite_land_cover_2018.nc	Worldwide land cover data for 2018	/Maps
satellite_land_cover_2019.nc	Worldwide land cover data for 2019	/Maps
satellite_land_cover_2020.nc	Worldwide land cover data for 2020	/Maps
satellite_land_cover_2021.nc	Worldwide land cover data for 2021	/Maps
satellite_land_cover_2022.nc	Worldwide land cover data for 2022	/Maps
r7_merged_data_34ctry.dta	Afrobarometer data-2019	/Afrobarometer/data/
afrobarometer_release-dataset_merged-34country_rb_en_2023-03-01.sav	Afrobarometer data-2019	/Afrobarometer/data/
ken_admbnda_adm2_iebc_20191031	Administrative boundaries for Kenya	/Shape_files
moz_admbnda_adm2_ine_20190607	Administrative boundaries for Mozambique	/Shape_files
tza_admbnda_adm2_20181019	Administrative boundaries for Tanzania	/Shape_files
uga_admbnda_adm2_ubos_20200824	Administrative boundaries for Uganda	/Shape_files
landcover19_countries.nc	Land cover raster for 2019, for countries of interest	/Clipped_files
landcover20_countries.nc	Land cover raster for 2020, for countries of interest	/Clipped_files
landcover21_countries.nc	Land cover raster for 2021, for countries of interest	/Clipped_files
landcover22_countries.nc	Land cover raster for 2022, for countries of interest	/Clipped_files
Land cover classification gridded maps.pdf	Metadata for land cover data	/Metadata
kenya_georef.csv	Georeferenced coordinates for Kenya	/georeferenced
mozambique_georef.csv	Georeferenced coordinates for Mozambique	/georeferenced
tanzania_georef.csv	Georeferenced coordinates for Tanzania	/georeferenced
uganda_georef.csv	Georeferenced coordinates for Uganda	/georeferenced

Land Cover data

Land cover classification gridded maps derived from satellite observations were downloaded from the Copernicus Land Monitoring Service, for the years 2016-2022. However, because of time constraints, only the last 3 years were sliced to the countries of interest and converted to csv. However, all files (2016-2022) will be included as deliverables of this project as zip files, for future use.

This dataset characterizes the land cover of a particular year. The land cover was derived from the analysis of satellite data time series of the full period. The land cover classification is based on the LCCS (Land Cover Classification System) and consists of 22 classes.

The dataset is in csv format, under the folder `csv`. Files can be read using `readr` in R, with the read_csv function.

Variables of interest

Here are the following variables of interest in the dataset:

Name	Units	Description
Change count	Dimensionless	Number of years where land cover class changes have occurred, since 1992. 0 for stable, greater than 0 for changes.
Current pixel state	Dimensionless	Pixel identification from satellite surface reflectance observations, mainly distinguishing between land, water, and snow/ice. Six values are used: 1, 2, 3, 4, 5, 6; respectively meaning: clear land, clear water, clear snow ice, cloud, cloud shadow, filled.
Land cover class	Dimensionless	Land cover class per pixel, according to a legend of 22 classes, defined using the Land Cover Classification System developed by the United Nations Food and Agriculture Organization. Distinct values are encoded as unsigned byte (0..255). The complete legend is available in the NetCDF files metadata and in the Product User Guide documentation.
Observation count	Dimensionless	Number of valid satellite observations that have contributed to each pixel's classification
Processed flag	Dimensionless.	Flag to mark areas that could not be classified. Two values are used: 0, 1; respectively meaning: not_processed, processed.

More information on the dataset can be found at: https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-land-cover?tab=overview

Additional data from previous attemps at a pipeline can be found in the `Older` folder of this project.

Afrobarometer data

The 2019 Afrobarometer data was analyzed in the `afrobarometer_analysis.Rmd` notebook, in the `r` folder of this project. See said notebook for summary statistics.

The 2022 Afrobarometer data was downloaded, and variables of interest (related to gender) were selected. See `afrobarometer_2022_analysis.Rmd` for variables of interest.

Ideally, each of the land cover files should be joined to the georef file, based on latitude and longitude, so they can then be joined by city name to the Afrobarometer data. Because of time constraints of the project, this step was not carried out. See the `clip_based_on_georef.ipynb` notebook for more information.

For additional questions, you can reach out to the author at julianagc@berkeley.edu.