**DATASETS**

|  |  |  |  |
| --- | --- | --- | --- |
| **DATASET NAME** | **DATASET LINK** | **Resolution** | **Format** |
| HydroSHEDS upstream accumulating area | <https://www.hydrosheds.org/downloads> | 30 arcseconds | Raster |
| HydroSHEDS drainage directions | <https://www.hydrosheds.org/downloads> | 30 arcseconds | Raster |
| GPWv4 2015 (UN Adjusted)  Population Count | <https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-count-adjusted-to-2015-unwpp-country-totals-rev11/data-download> | 30 arcseconds | Raster |
| GPWv4 2015 (UN Adjusted) Population Density | <https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-adjusted-to-2015-unwpp-country-totals-rev11/data-download> | 30 arcseconds | Raster |
| GPWv4 National Identifier Grid | <https://sedac.ciesin.columbia.edu/data/set/gpw-v4-national-identifier-grid-rev11/data-download> | 30 arcseconds | Raster |

*Notes: HydroSHEDS data must be downloaded by continent. GPW data can be downloaded globally as a single file*

**GITHUB REPOSITORY:** <https://github.com/markbernhofen/riverCoastProximity>

**DATA PREPERATION**

**1. Burn coastline into the HydroSHEDS accumulation file**

1. open HydroSHEDS upstream accumulating area dataset in QGIS. Assign a value (-9999) to the nodata cells using the GRASS r.null tool. Save the dataset as (acc.tif).
2. run burnCoastline.py code (<https://github.com/markbernhofen/riverCoastProximity/blob/master/burnCoastline.py>) in directory containing acc.tif file saved in the previous step. You need to specify the no-data value from the previous step (-9999) and choose a value to assign to the coastal cells (9999).

*python burnCoastline.py -i C:\inputDirectory -d C:\outputDirectory -n -9999 -c 9999*

1. this code will output a new accumulation file (name: acc.tif) with the specified coastal value burnt in.
2. Run this for each HydroSHEDS continent

**2. Combine data into ‘mega’ continents**

1. the hydroSHEDS data prepared in the previous two steps should be separated into the following continents: Africa, Asia, Australia, Central America, Europe, North America, South America.
2. in QGIS combine the following continent datasets into ‘mega’ continents: Asia+Europe=Eurasia, North America + Central America + South America = America. So each continent (Eurasia, America, Australia, Africa) should have the following file in a continent specific directory: ‘acc.tif’ (from step 1).
3. the final step is to cut the Global GPW4 data to the individual ‘mega’ continents. Use QGIS built in tool ‘Clip Raster By Extent’ to clip the datasets by HydroSHEDS layer extent (use ‘acc.tif’ from previous step). This will ensure that the layers are the same size for the rest of the analysis. Clip the GPW4 data and save them as the following: GPWv4 2015 population count as ‘pop.tif’; GPWv4 2015 population density as ‘density.tif’, GPWv4 National Identifier grid as ‘nationID’.

**3. Run river buffer**

1. make sure that the following files are in the directory you will be working in: ‘acc.tif’ (from step 1) and ‘pop.tif’ (from step 2).
2. run the riverBuffer.py code (<https://github.com/markbernhofen/riverCoastProximity/blob/master/riverBuffer.py>) in the directory containing the relevant files. You need to specify the minimum threshold upstream accumulating area (km2) of rivers to consider (25) and also the distance from the river in km (1) that you want to consider. This code then does a neighbourhood analysis on each river (and coastal) cell in the ‘acc.tif’ file and if there is a population cell within the (specified buffer distance) vicinity of that river/coast cell the population cell is copied into the new dataset. What you end up with is essentially the ‘pop.tif’ dataset with those cells not within (1km) of a river or a coastline cropped out.

*python riverBuffer.py -i C:\inputDirectory -d C:\outputDirectory -a 25 -b 1*

1. this code will output a buffered population dataset filename: buffer.tif

**DATA ANALYSIS**

1. each ‘mega’ continent should have its own directory containing the following files: ‘buffer.tif’ (from step 3), ‘pop.tif’ (from step 2), ‘density.tif’ (from step 2), ‘nationID.tif’ (from step 2), ‘acc.tif’ (from step 1).
2. run the analysis.py code (link) in each of the ‘mega’ continent directories. You only need to specify the name of the continent when running the code (either: ‘Eurasia’, ‘Australia’, ‘Africa’, or ‘America’)

*python analysis.py -i C:\inputDirectory -n Eurasia*

1. This code will output a CSV file with the stats for each country.

*Note: some countries will span across ‘mega’ continents. For these countries you should manually sum the stats in Excel once the CSVs from each continent have been combined in a single file.*

*Note: CSVs use country ID to identify countries. A list of countries and their relevant IDs can be found here (*<https://github.com/markbernhofen/riverCoastProximity/blob/master/countryIDs.csv>)