Documentation for the Low Elevation Coastal Zone (LECZ) Urban-Rural Population and Land Area Estimates, Version 3

December 2021

Center for International Earth Science Information Network (CIESIN),
Columbia University
CUNY Institute for Demographic Research (CIDR),
City University of New York

Abstract

The Low Elevation Coastal Zone (LECZ) Urban-Rural Population and Land Area Estimates, Version 3 data set contains land areas with urban, quasi-urban, rural, and total populations (counts) within the LECZ for 234 countries and other recognized territories for the years 1990, 2000, and 2015. This data set updates initial estimates for the LECZ population by drawing on a newer collection of input data, and provides a range of estimates for at-risk population and land area. Constructing accurate estimates requires high-quality and methodologically consistent input data, and the LECZv3 evaluates multiple data sources for population totals, digital elevation model, and spatially-delimited urban classifications. Users can find the paper "*Estimating population and urban areas at risk of coastal hazards, 1990–2015: how data choices matter*" (MacManus, et al. 2021) in order to evaluate selected inputs for modeling Low Elevation Coastal Zones. According to the paper, the following are considered core data sets for the purposes of LECZv3 estimates: Multi-Error-Removed Improved-Terrain Digital Elevation Model (MERIT-DEM), Global Human Settlement (GHSL) Population Grid R2019 and Degree of Urbanization Settlement Model Grid R2019a v2, and the Gridded Population of the World, Version 4 (GPWv4), Revision 11. This data set is produced by the Columbia University Center for International Earth Science Information Network (CIESIN) and the City University of New York (CUNY) Institute for Demographic Research (CIDR).

Data set citation: Center for International Earth Science Information Network (CIESIN), Columbia University, and CUNY Institute for Demographic Research (CIDR), City University of New York. 2021. Low Elevation Coastal Zone (LECZ) Urban-Rural Population and Land Area Estimates, Version 3. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/d1x1-d702. Accessed DAY MONTH YEAR.

Suggested citation for this document: Center for International Earth Science Information Network (CIESIN), Columbia University and CUNY Institute for Demographic Research (CIDR), City University of New York. 2021. Documentation for the Low Elevation Coastal Zone (LECZ) Urban-Rural Population and Land Area

Estimates, Version 3. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/z0r0-gc08. Accessed DAY MONTH YEAR.

We appreciate feedback regarding this data set, such as suggestions, discovery of errors, difficulties in using the data, and format preferences. Please contact:

NASA Socioeconomic Data and Applications Center (SEDAC)
Center for International Earth Science Information Network (CIESIN)
Columbia University
Phone: 1 (845) 365-8920
Email: ciesin.info@ciesin.columbia.edu

Contents

I. Introduction

The Low Elevation Coastal Zone (LECZ) Urban-Rural Population and Land Area Estimates, Version 3 data set provides country-level estimates of urban, quasi-urban, rural, and total population (count) and land area (square kilometers) for 234 statistical areas (countries and other UN recognized territories) for the years 1990, 2000, and 2015. Estimates of populations and land area in 0-5m and 5-10m LECZs, and non-coastal areas at any elevation (outside of LECZs) can inform planning and policy-making decisions around exposure to sea level rise and storm surge. The data set also includes one spatial layer: LECZ delineations generated from the Multi-Error-Removed Improved-Terrain Digital Elevation Model (MERIT-DEM) which can be used to identify areas below 10-meter elevation along global coastlines.

This updates the Low Elevation Coastal Zone Urban-Rural Population and Land Area Estimates, Version 2 data set. Because exposure estimates vary depending on the characteristics of the input data, such as spatial resolution, root mean square error (RMSE) of elevation sources, and population allocation methods, a cross-tabulation of estimates based on five sources of elevation data, four sources of population data, and four urban proxy data sets which categorize areas as urban, quasi-urban, or rural respectively is provided. This range of estimates enables users to better understand uncertainties resulting from the choice of input data.

II. Data and Methodology

In addition to this documentation, a detailed description of the methodology along with a sensitivity analysis of data sources is available in the paper:

MacManus, K., Balk, D., Engin, H., McGranahan, G., and Inman, R.: *Estimating population and urban areas at risk of coastal hazards, 1990–2015: how data choices matter*, Earth Syst. Sci. Data, 13, 5747–5801, https://doi.org/10.5194/essd-13-5747-2021, 2021.

Input data

The Low Elevation Coastal Zone (LECZ) Urban-Rural Population and Land Area Estimates, Version 3 data set updates and improves population and land area estimates for the 0-5m and 5-10m elevation zones contiguous to the coast for countries and territories globally. To produce this data set, researchers evaluated and compared input data from five elevation data sources, four population count data sources, and four data sources of urban-rural disaggregation. The thirteen data sources assessed are described in Table 1.

Table 1: Input Data

Theme	Data Set	Abbreviation	Input Spatial Resolution	Paper Reference
Elevation	Shuttle Radar Topography Mission Elevation Low Elevation Coastal Zones	SRTM	3 arc-second	(ISciences, 2003)
	Multi-Error-Removed Improved-Terrain Digital Elevation Model	MERIT-DEM	3 arc-second	(Yamazaki et al., 2017)
	TerraSAR-X add-on for Digital Elevation Measurement	TanDEM-X	3 arc-second	(Wessel et al., 2018)

	CoastalDEM90	CoastalDEM	3 arc-second	(Kulp & Strauss, 2018)
	ALOS World 3D - 30m Digital Surface Model	AW3D30	1 arc-second	(Tadono et al., 2014)
Urban Concept	Global Human Settlement - Settlement “Degree of Urbanization” Model Grid R2019a v2	GHS-SMOD	30 arc-second	(Florczyk et al., 2019)
	Global Human Settlement - Built-up Grid R2019	GHS-BUILT	9 arc-second	(Pesaresi et al., 2019), (Florczyk et al., 2019)
	Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Urban Extents Grid	GRUMPv1	30 arc-second	(Balk et al., 2005), (CIESIN et al., 2011)
	VIIRS Plus DMSP Change in Lights (VIIRS+DMSP dLIGHT)	dLIGHT	15 arc-second	(Small et al., 2020)
Population	Global Human Settlement - Population Grid R2019	GHS-POP	9 arc-second	(Freire et al., 2016), (Florczyk et al., 2019)
	Gridded Population of the World, Version 4, Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals, Revision 11	GPWv4.11 UN WPP	30 arc-second	(CIESIN et al., 2018d), (Doxsey-Whitfield et al., 2015)
	LandScan 2015 High Resolution Population Data Set	LandScan	30 arc-second	(Bright and Coleman, 2001; Bright et al., 2016)
	WorldPop Global High Resolution Population Denominators	WorldPop	3 arc-second	(Lloyd et al., 2019)

In addition to the twelve primary sources, several ancillary data inputs were also utilized to produce LECZv3:

- Gridded Population of the World, Version 4 (GPWv4): National Identifier Grid, Revision 11 (CIESIN, 2018a) (NID) is used to construct the extent of countries and territories, as well as summary statistics for those units. The horizontal resolution of this data set is 30 arc-seconds or approximately 1 km at the equator.

- Gridded Population of the World, Version 4 (GPWv4): Land and Water Area, Revision 11 (CIESIN, 2018b) forms the basis of land area estimates, as the land area grid accounts for the reduction in the underlying area of regular rectangular grid cells as they approach the poles. This allows for the accurate area measurements without requiring the use of an Equal Area projection. The resolution of this data set is 30 arc-seconds or approximately 1 km at the equator.
- Gridded Population of the World, Version 4 (GPWv4): Data Quality Indicators, Revision 11 (CIESIN, 2018c) Mean Administrative Unit Area raster, which represents the nominal resolution of the input vector geographies which were then matched to census population estimates prior to gridding. Smaller values indicate areas with high resolution input geographies, which therefore are less impacted by modelling. This raster better enables an understanding of the precision and accuracy of pixel level population estimates. The resolution of this data set is 30 arc-seconds or approximately 1 km at the equator.
- Global Human Settlement - Built-up Grid R2019 (GHS-BUILT) provides an information layer on the intensity or percentage of the built environment as derived from Landsat image collections. This data set could be used to independently evaluate how much land in the LECZ is built-up. The native resolution of GHS-Built is 250 meters and is in the Mollweide coordinate system.

Methods

Elevation

Four digital elevation models were used to construct 0-5m and 5-10m low elevation coastal zones: Multi-Error-Removed Improved-Terrain Digital Elevation Model (MERIT-DEM); Shuttle Radar Topography Mission Elevation Low Elevation Coastal Zones (SRTM); TerraSAR-X add-on for Digital Elevation Measurement (TanDEM-X); and CoastalDEM90 (CoastalDEM)[1].

Multiple steps were taken in order to ensure cross-comparison between the five elevation data sets. Each of the source data sets were projected to WGS84 horizontal coordinate systems with EGM96 geoid height. Out of the 5 DEMs evaluated, 3 of them (SRTM, MERIT-DEM, and CoastalDEM) were referenced to the EGM96 Vertical Coordinate System (EPSG:5773) natively. TanDEM-X elevations were natively referenced to the WGS84 (G1150) ellipsoid (EPSG:4979) and were therefore converted to EGM96 geoid heights using the GDAL Warp tool, in order to conform with the vertical reference of the other 3 elevation sources.

---

[1] In order to obtain complete spatial coverage, CoastalDEM country files with partial spatial coverage were mosaicked with JAXA ALOS 30m global DEM (AW3D30) grids. The 1 arc-second AW3D30 data was aggregated to 3 arc-seconds in order to conform with CoastalDEM. The aggregation was accomplished by taking the mean elevation value. Countries with combined CoastalDEM and AW3D30 data were the Aland Islands, Canada, Finland, the Faroe Islands, Great Britain, Greenland, Iceland, Norway, Russia, Svalbard and Jan Mayan, Sweden, and the United States.

All elevation data were preprocessed into a common framework and subset by country boundaries as defined by the GPWv4.11 National Identifier Grid (NID). Elevation tiles were loaded into an Esri file geodatabase mosaic data set, which included corresponding vector layers (footprints) of the input raster extents and identifying metadata. A python script was used to clip the raster footprints by country boundaries from the NID. This created country level layers with attributes (i.e. file names and locations) from intersecting footprints for each of the elevation sources, which were then used to isolate a subset list of elevation tiles belonging to a given country. All elevation data were then aggregated with the MEAN method of the Esri ArcGIS Aggregate tool to a 9 arc-second horizontal resolution.

From here, steps were taken to construct the 0-5m low elevation coastal zone, 5-10m low elevation coastal zone, and non-coastal areas of any elevation:

1) The GPWv4.11 NID was buffered by 1 km on a per country basis in order to prevent the loss of population due to coastline mismatches, or the loss of the low elevation coastal zone when the elevation data source uses a coastline that is seaward of the NID.

2) The nine arc-second country elevation mosaics for each elevation source were reclassified into integers representing the following zones: 0 to 5m, 5 to 10m and greater than 10m.

3) These reclassified images were then segmented with the Esri ArcGIS region group tool with eight neighbors using the WITHIN parameter. The resultant region grouped images where groups of pixels with like values were combined such that each connected group (or region) receives its own unique identifier, along with a count of the number of pixels within the group.

4) In order to isolate coastally contiguous regions, the region grouped images were converted into polygons and selected by the location where each polygon intersects the border of a country as determined from the 1 km buffered GPWv4.11 NID. This effectively isolated all regions connected to administrative boundary coastlines.

5) Quality Assurance manual checking for spurious lowland regions contiguous with inland values.

6) The isolated, quality-assured regions were then used as extraction masks on the reclassified DEMs while non-connected inland values were coded as above 10m (the corresponding value in our resulting spatial data is coded as 31).

The resulting rasters contained coastally contiguous pixels coded into 0-5m and 5-10m LECZs, and a third category representing non-coastal areas of any elevation area (outside of LECZs). The Low Elevation Coastal Zone (LECZ) raster derived from the global MERIT-DEM data product is distributed with this data set.

Population

Population distribution data were sourced from Global Human Settlement - Population Grid R2019 (GHS-POP), Gridded Population of the World, Version 4, Revision 11 (GPWv4.11 UN WPP), LandScan 2015 High Resolution Population Data Set (Landscan), and WorldPop Global High Resolution Population Denominators (WorldPop). The original resolutions for the data sets were as follows: WorldPop (3 arc-second); GHS-POP (9 arc-seconds); GPWv4.11 UN WPP (30 arc-seconds); and LandScan (30 arc-seconds). If required, these data sets were processed from their native resolutions to comparable 9 arc-second (nominally 300m) resolution data sets. Accordingly, WorldPop was aggregated from 3 arc-seconds to 9 arc-seconds using the Esri ArcGIS aggregate tool. GPWv4.11 UN WPP and LandScan were uniformly disaggregated by a factor of 100 and quality-assured to have the same total population before and after the sampling. Once processed in a standardized 9 arc-second resolution, population data sets were subset by country.

Urban-Rural Concept

The four data sets used in this estimation of urban and rural populations were the Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Urban Extents Grid, VIIRS Plus DMSP Change in Lights (VIIRS+DMSP dLIGHT), Global Human Settlement - Built-Up Grid R2019 (GHS-BUILT), and the Global Human Settlement - Settlement "Degree of Urbanization" model Grid R2019a v2 (GHS-SMOD). These data sets were preprocessed in order to conform to a comparable 9 arc-second horizontal resolution. dLIGHT is natively 15-arc seconds, while GHS-SMOD and GRUMPv1 are natively 30 arc-seconds. These data sets were uniformly up-sampled to a 9 arc-second resolution, using a Nearest Neighbor approach since the underlying data is categorical. Additionally, GHS-BUILT was projected into the WGS84 coordinate system with a Nearest Neighbor cell assignment. Afterwards, the data sets were subset by country using the GPWv4.11 NID.

In working with heterogeneous urban proxy data sets, it was necessary to establish an urban-rural classification scheme to facilitate inter-comparison. For this reason, our urban-rural population estimates rely on three simplified thematic categories; Urban, Quasi-Urban, and Rural. The GHS-SMOD data set already contains classification and aggregation schemas that align with Urban, Quasi-Urban, and Rural categorical designations, the so called Level 1 schema. GHS-BUILT was reclassified into Urban (> 50% built-up), Quasi-Urban (> 3% and ≤ 50% built-up), and Rural (≤ 3% built-up) using the Esri ArcGIS Reclassify tool. dLIGHT was visually cross-referenced with GHS-SMOD and GHS-BUILT in order to determine the following thresholds; Urban (> 100 dn), Quasi-Urban (> 3 and < 100 dn), and Rural (< 3 dn) using the Esri ArcGIS Reclassify tool, where digital numbers (dn), 0 to 255, represent the relative luminosity of pixels across the time periods represented in the data set (1992, 2002, 2013). GRUMPv1 Urban Extents Grid was constructed as a dichotomous urban-rural grid and cannot be applied to classifications outside of this binary. It is included here for comparison with LECZv1 and LECZv2 data.

Processing auxiliary data sets
GPWv4.11 Land Area grid had a native horizontal resolution of 30 arc-seconds. It was uniformly up-sampled to 9 arc-seconds resolution by a factor of 100 and quality assured to have the same total land area per pixel both before and after the sampling. After up-sampling, the land area data was subset by country

The GPWv4.11 Mean Administrative Unit Area grid also had a native horizontal resolution of 30 arc-seconds, but because the values in this grid represent the average size of input population units, there was no need to up-sample. These data were simply resampled at 9 arc-second resolution. After resampling, the data was subset by country.

GHS-BUILT was used here not only to discriminate between Urban, Quasi Urban, and Rural as a categorical data set, but also to summarize built-up intensities as a measure in its own right. It was projected from the Mollweide coordinate system into WGS84 coordinates using Nearest Neighbor at 9 arc-seconds. After reprojection, the data was subset by country.

Results: Population and Land Area Estimates Workbook
Country-level LECZ population and land area estimates were produced for each permutation of the 12 data sources using the Esri ArcGIS Zonal Statistics as Table tool at the country level. Continent-level and UN geo-regional population and land area estimates were also derived from the zonal statistics corresponding to each permutation. A python script was then used to compile these data into a single master table. In the workbook, pivot tables are used to summarize data from the intersection of Digital Elevation Models (DEMs), population data sets, and urban-rural disaggregation. This data is further summarized by country, UN region, and continent. Summary tables for data sources and global-level results are provided in a separate worksheet.

Results: Raster data set
LECZs derived from Multi-Error-Removed Improved Terrain Digital Elevation Model (MERIT-DEM) are provided as a spatial layer in a 9 arc-second resolution raster (GeoTIFF). This data categorically demarcates the two low elevation coastal zones (0-5m and 5-10m LECZs) and a third category representing the non-coastal areas at any elevation (outside of LECZs) coded as 31. This raster data is also provided as a web map and image service. It is in the WGS84 coordinate system.

III. Data Set Description(s)

1. Excel Workbook

Data set description:
The Excel-based data sets include the full LECZv3 population and land are estimates database in a macro-enabled workbook, as well as a worksheet of summary tables. The Excel workbook for LECZv3 contains a database of the resulting population estimates from each permutation of the 12 data sources, as well as land area estimates, estimates of the average built-up percentages, and data quality information about the average size of input census geographies. This data is available for the years 1990, 2000, and 2015. Pivot tables are used to summarize the database, and summaries are available at the county, continent, and UN region level. These pivot tables facilitate data exploration by allowing a user to expose the underlying data behind each cell. By double-clicking on a cell, the pivot table creates a new spreadsheet that contains all of the data that generated the value of that cell.

Pivot tables allow a user to filter data based on chosen criteria, excluding information outside of the specified criteria. Multiple filters can be used at once, to further refine the data. The filters for LECZv3 are:

- Continent filter
- UN Region Filter
- Country Filter
- Elevation Source Filter
- Urban Concept Filter
- LECZ Description filter
- Settlement Description Filter
- Year Filter

Furthermore, tabular data is provided to users in summary tables. These tables summarize data sources, as well as global-level results in overall aggregate terms and as a percent change over time (1990-2015). A list of tables is included in the front sheet of the summary table excel workbook.

2. Raster Data Set

Data set description:
The Low Elevation Coastal Zone (LECZ) Urban-Rural Population and Land Area, Version 3 MERIT-LECZs raster represents the spatial extent of the Low Elevation Coastal Zone, by classifying land areas as the 5 meter LECZ (0-5m above sea level), 5-10 meter LECZ (5-10m above sea level), and all elevations >10 meters and/or non-contiguous to the coastline. The Multi-Error-Removed Improved-Terrain Digital Elevation Model (MERIT-DEM) is the core input data set for these estimations. This data is coded categorically, according to the codebook included below. The data are provided

as a spatial layer in a 9 arc-second resolution raster in the WGS84 coordinate system (GeoTIFF). This raster data is also provided as a web map and image service.

2.1 Python Jupyter Notebook

Data set description:
Included in the LECZv3 GeoTIFF data set zip is a python jupyter notebook with original code for producing Low Elevation Coastal Zone spatial data. This code demonstrates the analytic process by which coastline data and Digital Elevation Models generate LECZ outputs. The example data analysis applies MERIT-DEM and GPWv4.11 coastline data sources and uses Sri Lanka as the area of interest.

The python jupyter notebook, titled *create-leczs.ipynb*, is stored in the subfolder *code*. This folder contains example data, including an open sample of MERIT-DEM raster data, GPWv4.11 land mask and coastline data for Sri Lanka. This code requires the open source python modules *os* and *sys*, as well as the proprietary Esri *arcpy* module and spatial analyst extension. While the Esri modules are proprietary, analogous tools in open source python modules and R packages exist. The python jupyter notebook is presented in executed form, so that users without access to these modules can examine intermediate and final outputs, which can guide their own implementations in other languages or using other packages.

The code underlying this analysis can be rerun, substituting areas of interest, digital elevation models, land mask, and coastline data sources so long as users have access to *arcpy*. To produce new output, users should update data paths or delete the sample outputs provided. This demonstration serves to facilitate reproducibility and the wider application of these novel methods.

Data set web page:
SEDAC URL: https://sedac.ciesin.columbia.edu/data/set/lecz-urban-rural-population-land-area-estimates-v3/data-download
Permanent URL: https://doi.org/10.7927/d1x1-d702

Data set format:
The data set are available as global grids in GeoTIFF format and as statistical data in Microsoft Excel (XLSM) tabular spreadsheet format. These files are packaged in separate compressed zip files for download. The GeoTIFF zip includes source code and sample data in a python jupyter notebook titled *create-leczs.ipynb*. Each downloadable is a compressed zip file containing: 1) GeoTIFF or XLSM file, 2) Source code, if applicable, 3) Readme.TXT file, and 4) PDF Documentation. Specifically:

- The Low Elevation Coastal Zone (LECZ) Urban-Rural Population and Land Area Estimates, Version 3 for the years 1990, 2000, and 2015 are available in a macro-enabled Excel Workbook (.xlsm), along with Summary Tables in an Excel Worksheet (.xlsx).

- Low Elevation Coastal Zone (LECZ) Land Area Estimates, Version 3 MERIT LECZs Raster Layer is available in GeoTIFF format in the WGS84 coordinate system, along with the source code used to produce rasterized LECZ data.

Data set downloads:

- lecz-urban-rural-population-land-area-estimates-v3-xlsm.zip
- lecz-urban-rural-population-land-area-estimates-v3-merit-leczs-geotiff.zip

IV. How to Use the Data

This data set is global in extent. The raster data is intended for use in any software that can read and analyze geospatial data. This data can be subset spatially. The Excel data set can be used in statistical and Geographic Information System (GIS) software. Excel data can be subset by country or by any of the fields listed in the Codebook. The python jupyter notebook can be opened in the jupyter notebook environment, Esri ArcGIS Notebook software, or other suitable Integrated Development Environments (such as Microsoft Visual Studio Code) and hosted notebook servers (such as Google Colab).

V. Potential Use Cases

The Low Elevation Coastal Zone (LECZ) Land Area Estimates, Version 3 data set assesses current and past risks related to sea level rise and increased natural disasters. With this information, adaptation and mitigation strategies can be tailored to the geographic locations that exhibit risk factors. Some potential applications include:

- Evaluating population and development trajectories in “high risk” coastal zones. This information can be evaluated in comparison to non-coastal areas, in order to assess high-risk development practices in the context of broader development trends.

- Assessing current sovereign, regional, and continental sea level rise exposure based on population counts and urban-proxy measures.

- Identifying where increasing development in the built environment is intersecting with increased vulnerability to sea level rise and/or natural disasters.

VI. Limitations

The elevation data is produced and distributed in the World Geodetic System 1984 (WGS84) Geographic Coordinate System. The GHSL data product, however, is produced and distributed in the Mollweide Equal Area Projected Coordinate System (not including GHS-POP, which is also released in a WGS84 version). In order to conduct analyses on these data sources, it is necessary to harmonize their coordinate systems, but the projection of raster data is not without complications.

When a raster data set is projected from one coordinate system to another, the registration and total number of pixels represented are altered. In other words, the number of pixels may change along with the location of those pixels relative to ground truth. The projection of the elevation data source (WGS84) was maintained to avoid introducing uncertainties about the location of the LECZs. Therefore, it was needed to project GHS-BUILT and GHS-SMOD to conform with that elevation source. The thematic layers (GHS-BUILT, GHS-SMOD) were not simple to validate owing to the fact that there is no available alternative source for these data to compare with. Any error introduced by projecting these data from Mollweide to WGS84 using a Nearest Neighbor approach is quite minimal, however, it should be noted that because of the fact that the LECZs represent small swaths of land area, they are also more sensitive to any apparent shifts of pixel locations. Although the projection issue does produce some uncertainty, it would not have been possible to use these data sources together without taking this approach.

Uncertainty associated with coastlines, and how they align with other data sets, is a known problem because there is no authoritative, global data set on coastlines, and because coastlines are dynamic. This is especially an issue for island nations where such mismatches are aggravated, and where low-lying coastal land may be found in high proportions.

VII. Acknowledgments

These data were produced with funding from SEDAC and the World Resources Institute's Center for Urban Transitions. The production team was led by Kytt MacManus (CIESIN) and Deborah Balk (CIDR), with critical contributions from Hasim Engin (CIDR), Rya Inman (CIESIN), and Gordon McGranahan from the Institute of Development Studies (IDS). Mairead Milan, Alexandra Hays, Sarah Colenbrander, Catlyne Haddaoui, and Leah Lazar provided additional support and comments.

Funding for development and dissemination of this data set was provided under the U.S. National Aeronautics and Space Administration (NASA) contract 80GSFC18C0111 for the continued operation of the Socioeconomic Data and Applications Center (SEDAC), which is operated by the Center for International Earth Science Information Network (CIESIN) of Columbia University.

VIII. Disclaimer

CIESIN follows procedures designed to ensure that data disseminated by CIESIN are of reasonable quality. If, despite these procedures, users encounter apparent errors or misstatements in the data, they should contact SEDAC User Services at ciesin.info@ciesin.columbia.edu. Neither CIESIN nor NASA verifies or guarantees the accuracy, reliability, or completeness of any data provided. CIESIN provides this data without warranty of any kind whatsoever, either expressed or implied. CIESIN shall not be liable for incidental, consequential, or special damages arising out of the use of any data provided by CIESIN.

IX. Use Constraints

This work is licensed under the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0).

Users are free to use, copy, distribute, transmit, and adapt the work for commercial and non-commercial purposes, without restriction, as long as clear attribution of the source is provided.

X. Recommended Citation(s)

Data set(s):

Center for International Earth Science Information Network (CIESIN), Columbia University, and CUNY Institute for Demographic Research (CIDR), City University of New York. 2021. Low Elevation Coastal Zone (LECZ) Urban-Rural Population and Land Area Estimates, Version 3. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/d1x1-d702. Accessed DAY MONTH YEAR.

Scientific publication:

MacManus, K., Balk, D., Engin, H., McGranahan, G., and Inman, R.: *Estimating population and urban areas at risk of coastal hazards, 1990–2015: how data choices matter*, Earth Syst. Sci. Data, 13, 5747–5801, https://doi.org/10.5194/essd-13-5747-2021, 2021.

XI. Source Code

Many of the techniques utilized to generate estimates of populations by elevation, population source, and along the urban continuum leverage well known workflows and geo-processing tools. The code provided here focuses on the novel aspect of the work, namely how to produce a LECZ from some DEM and Coastline data for a country or other area of interest. License restrictions on some of the data utilized in this work also prevent their redistribution, therefore the sample code utilizes sample data from the open MERIT-DEM product and coastline and land area of interest data files from GPWv4.11, which is also open. The sample code is provided as a Python Jupyter Notebook, which utilizes the Esri *arcpy* module. Sample input and output data are also included. To run the code to produce new outputs, users should update data paths or delete the sample outputs provided.

XII. References

Balk, D., F. Pozzi, G. Yetman, U. Deichmann, and A. Nelson. 2005. The distribution of people and the dimension of place: methodologies to improve the global estimation of urban extents. In *Proceedings of the Urban Remote Sensing Conference, International Society for Photogrammetry and Remote Sensing*, Tempe AZ, 14-16 March 2005.

Bright, E. A., and P. R. Coleman. 2001. LandScan 2000. Oak Ridge National Laboratory.

Bright, E. A., A. N. Rose and M. L. Urban. 2016. LandScan 2015. Oak Ridge National Laboratory.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2018a. Gridded Population of the World, Version 4 (GPWv4): National Identifier Grid, Revision 11. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4TD9VDP.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2018b. Gridded Population of the World, Version 4 (GPWv4): Land and Water Area, Revision 11. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4Z60M4Z.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2018c. Gridded Population of the World, Version 4 (GPWv4): Data Quality Indicators, Revision 11. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H42Z13KG.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2018d. Gridded Population of the World, Version 4 (GPWv4): Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals, Revision 11. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4PN93PB.

Center for International Earth Science Information Network (CIESIN), Columbia University, International Food Policy Research Institute (IFPRI), The World Bank, and

Centro Internacional de Agricultura Tropical (CIAT). 2011. Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Urban Extents Grid. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4GH9FVG.

Doxsey-Whitfield, E., K. MacManus, S. B. Adamo, L. Pistolesi, J. Squires, O. Borkovska, and S. R. Baptista. 2015. Taking advantage of the improved availability of census data: A first look at the Gridded Population of the World, Version 4. *Papers in Applied Geography*, 1(3): 226-234. https://doi.org/10.1080/23754931.2015.1014272.

Florczyk A. J., C. Corbane, D. Ehrlich, S. Freire, T. Kemper, L. Maffenini. M. Melchiorri, M. Pesaresi, P. Politis, M. Schiavina, F. Sabo, and L. Zanchetta. GHSL Data Package 2019, EUR 29788 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-13186-1, JRC 117104. https://doi.org/10.2760/290498. https://publications.jrc.ec.europa.eu/repository/handle/JRC117104.

Freire S., K. MacManus, M. Pesaresi, E. Doxsey-Whitfield, and J. Mills. 2016. Development of new open and free multi-temporal global population grids at 250 m resolution. Proceedings of the *19th AGILE Conference on Geographic Information Science*. Helsinki, Finland, June 14-17, 2016. https://agileonline.org/conference_paper/cds/agile_2016/shortpapers/152_Paper_in_PDF.pdf.

ISciences. 2003. SRTM30 Enhanced Global Map -Elevation/Slope/Aspect (release 1.0). Ann Arbor: ISciences, LLC (based on the raw SRTM data from Jet Propulsion Laboratory).

Kulp, S. A., and B. H. Strauss. 2018. CoastalDEM: A Global Coastal Digital Elevation Model Improved from SRTM Using a Neural Network. *Remote Sensing of Environment* 206, 231–39. https://doi.org/10.1016/j.rse.2017.12.026.

Lloyd, C. T., H. Chamberlain, D. Kerr, G. Yetman, L. Pistolesi, F. R. Stevens, and A. Tatem. 2019. Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. *Big Earth Data*, 108-139. https://doi.org/10.1080/20964471.2019.1625151.

Pesaresi, M., A. Florczyk, M. Schiavina, M. Melchiorri, and L. Maffenini. 2019. GHS Settlement Grid, Updated and Refined REGIO Model 2014 in Application to GHS-BUILT R2018A and GHS-POP R2019A, Multitemporal (1975-1990-2000-2015), R2019A. European Commission, Joint Research Centre (JRC) [Dataset]. https://doi.org/10.2905/42E8BE89-54FF-464E-BE7B-BF9E64DA5218. http://data.europa.eu/89h/42e8be89-54ff-464e-be7b-bf9e64da5218.

Small, C., and Center for International Earth Science Information (CIESIN), Columbia University. 2020. VIIRS Plus DMSP Change in Lights (VIIRS+DMSP DLIGHT).

Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/9ryj-6467.

Tadono, T., H. Ishida, F. Oda, S. Naito, and K. I. Minakawa. 2014. Precise Global DEM Generation by ALOS PRISM. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II–4, 71-76. https://doi.org/10.5194/isprsannals-II-4-71-2014.

Wessel, B., M. Huber, C. Wohlfart, U. Marschalk, D. Kosmann, and A. Roth. 2018. Accuracy Assessment of the Global TanDEM-X Digital Elevation Model with GPS Data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 171–82. https://doi.org/10.1016/j.isprsjprs.2018.02.017.

Yamano, H., H. Kayanne, T. Yamaguchi, Y. Kuwahara, H. Yokoki, H. Shimazaki, and M. Chikamori. 2007. Atoll island vulnerability to flooding and inundation revealed by historical reconstruction: Fongafale Islet, Funafuti Atoll, Tuvalu. *Global and Planetary Change*, 407–416. https://doi.org/10.1016/j.gloplacha.2007.02.007.

Yamazaki, D., D. Ikeshima, R. Tawatari, T. Yamaguchi, F. O'Loughlin, J. C. Neal, and P. D. Bates. 2017. A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, 5844– 5853. https://doi.org/10.1002/2017GL072874.

XIII. Documentation Copyright and License

Appendix 1. Data Revision History

No revisions have been made to this data set.

Appendix 2. Contributing Authors & Documentation Revision History

Revision Date	ORCID	Contributors	Revisions
December 14, 2021		Alexandra Hays, Kytt MacManus, Mairead Milán	This document is the 1st instance of documentation.