**Release Statement**

**Census disaggregated gridded population estimates for Sierra Leone (2015)**

**version 1.0**

29 May 2020

These data were produced by the WorldPop Research Group at the University of Southampton. This work was part of the GRID3 project with funding from the Bill and Melinda Gates Foundation and the United Kingdom's Department for International Development (OPP1182408). Project partners included the United Nations Population Fund (UNFPA), Center for International Earth Science Information Network (CIESIN) in the Earth Institute at Columbia University, and the Flowminder Foundation.

Oliver Pannell (WorldPop) supported the generation of inputs for the application of the Random Forest (RF)-based dasymetric mapping approach developed by Stevens et al. (2015). The disaggregation was done by Maksym Bondarenko (WorldPop), using the Random Forests population modelling R scripts (Bondarenko et al., 2020), with oversight from Alessandro Sorichetta (WorldPop). The sections-level administrative boundaries were provided by Jolynn Schmidt and Eniko Kelly-Voicu (CIESIN). The whole WorldPop group and GRID3 partners are acknowledged for overall project support.



**CITATION**

WorldPop and Statistics Sierra Leone. 2020. Census disaggregated gridded population estimates for Sierra Leone (2015), version 1.0. University of Southampton. doi:10.5258/SOTON/WP00668

**RELEASE CONTENT**

1. SLE_population_v1_0_gridded.tif
2. SLE_population_v1_0_agesex.zip
3. SLE_population_v1_0_README.pdf

## FILE DESCRIPTION

**SLE_population_v1_0_gridded.tif**
This geotiff raster (.tif), at a resolution of 3 arc (approximately 100m at the equator), contains estimated population counts per grid cell across Sierra Leone. The projection is Geographic Coordinate System, WGS84. 'NoData' values represent areas that were mapped as unsettled based on building footprints provided by the Digitize Africa project of Ecopia.AI and Maxar Technologies (2020). These data are stored as floating-point numbers rather than integers to avoid rounding errors in aggregated population totals for larger areas.

**SLE_population_v1_0_agesex.zip**
This zip file contains 36 geotiff rasters (.tif) at a resolution of 3 arc (approximately 100m at the equator). The projection is Geographic Coordinate System, WGS84. Each raster contains estimated counts per grid cell for an age-sex group. The file names refer to the age-sex group represented by the raster. Age-sex group labels beginning with "f" are female populations and labels beginning with "m" are male populations. The age group labels refer to the first year of the age range. For example: "f_0" is females less than one year old. "f1" is females 1 to 4 years old. "f5" is females 5 to 9 years old. "f10" is females 10 to 14 years old. This pattern continues for each 5 year interval up to 80. "f80" is females greater than 80 years old. The labelling is the same for males: "m0", "m1", "m5", "m10", … , "m80".

## RELEASE HISTORY

- This is the original release

## ASSUMPTIONS AND LIMITATIONS

This dataset was produced based on the 2015 Population and Housing Census data provided by DSTI, Sierra Leone to the GRID3 programme. Section-based population counts were disaggregated into grid cells. The observed population totals are not projected to 2020, and thus represent the 2015 population distribution.

## SOURCE DATA

- Enumeration Area (EA) level population totals based on the 2015 Population and Housing Census (Statistics Sierra Leone, 2015) were provided by DSTI (Directorate of Science, Technology & Innovation, Sierra Leone) in a shapefile format.

- Building footprints were provided by the Digitize Africa project of Ecopia.AI and Maxar Technologies (2020) and gridded building patterns derived from the building footprints by Dooley and Tatem (2020).

- WorldPop Global Gridded Age-Sex Proportions (WorldPop et al 2018)

Additional geospatial covariates (Lloyd et al., 2019; doi:10.5258/SOTON/WP00644), representing factors related to population distribution, were obtained from the "Global High Resolution Population Denominators Project" (OPP1134076). Building footprint derivatives were prepared by Dooley and Tatem (2020).

## METHODS OVERVIEW

The EA units in the DSTI shapefile were aggregated to the 'Section' level because some EAs were overlapping. Following the Random Forest (RF)-based dasymetric mapping approach (Stevens et al., 2015; Bondarenko et al., 2020), the RF algorithm (Breiman, 2001) was used to produce a gridded population density dataset, with a spatial resolution of 3 arc seconds (approximately 100 m at the equator), that was subsequently used to dasymetrically disaggregate population counts from Sections into grid cells. Observed male/female totals were combined with the WorldPop gridded age proportions for Sierra Leone (WorldPop et al. 2018) to produce gridded population estimates for each age-sex group. The WorldPop gridded age-sex proportions were produced using the methods of Pezzulo et al. (2017) and Carioli et al. (in prep). Zonal sums were calculated for all admin units to check that the disaggregated values add up to the observed totals.

## REFERENCES


Bondarenko M., Nieves J. J., Stevens F. R., Gaughan A. E., Tatem A. and Sorichetta A. 2020. wpgpRFPMS: Random Forests population modelling R scripts, version 0.1.0. University of Southampton: Southampton, UK. https://dx.doi.org/10.5258/SOTON/WP00665

Breiman, L. Random forests. Mach. Learn. 45, 5–32 (2001).

Carioli A, Pezzulo C, Hanspal S, Hilber T, Hornby G, Kerr D, Tejedor-Garavito N, Nielsen K, Pistolesi L, Adamo S, Mills J, Nieves JJ, Chamberlain H, Bondarenko M, Ves N, Wigley A, Koper P, Lloyd C, Yetman G, Gaughan A, Stevens F, Linard C, James W, Sorichetta A, Tatem AJ. In prep. Population structure by age and sex: a multi-temporal subnational perspective.

Dooley, C. A. and Tatem, A.J. 2020. Gridded maps of building patterns throughout sub-Saharan Africa, version 1.0. University of Southampton: Southampton, UK. Source of building Footprints "Ecopia Vector Maps Powered by Maxar Satellite Imagery"© 2020. https://dx.doi.org/10.5258/SOTON/WP00666

Ecopia.AI and Maxar Technologies. 2020. Digitize Africa data. http://digitizeafrica.ai

Lloyd, C.T., Chamberlain, H., Kerr, D., Yetman, G., Pistolesi, L., Stevens, F.R., Gaughan, A.E., Nieves, J.J., Hornby, G., MacManus, K., Sinha, P., Bondarenko, M., Sorichetta, A., and Tatem A.J., 2019.Global spatio-temporally harmonised datasets for producing highresolution gridded population distribution datasets. Big Earth Data,

3(2), 108-139. https://dx.doi.org/10.1080/20964471.2019.1625151

Pezzulo C, Hornby GM, Sorichetta A, Gaughan AE, Linard C, Bird TJ, Kerr D, Lloyd CT, Tatem AJ. 2017. Sub-national mapping of population pyramids and dependency ratios in Africa and Asia. Sci. Data 4:170089. https://dx.doi.org/10.1038/sdata.2017.89

Statistics Sierra Leone (2015) Sierra Leone 2015 Population and Housing Census. https://www.statistics.sl/index.php/census/census-2015.html

Stevens, F. R., Gaughan, A. E., Linard, C. & Tatem, A. J. Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. PLoS ONE 10, e0107042 (2015). https://doi.org/10.1371/journal.pone.0107042

WorldPop (www.worldpop.org - School of Geography and Environmental Science, University of Southampton; Department of Geography and Geosciences, University of Louisville; Departement de Geographie, Universite de Namur) and Center for International Earth Science Information Network (CIESIN), Columbia University (2018). Global High Resolution Population Denominators Project - Funded by the Bill and Melinda Gates Foundation (OPP1134076). https://dx.doi.org/10.5258/SOTON/WP00646