Release Statement

Census disaggregated gridded population estimates for Niger (2021), version 1.0.

7 February 2022

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 Foreign, Commonwealth & Development Office (INV 009579, formerly OPP 1182425). Project partners included the United Nations Population Fund, Center for International Earth Science Information Network in the Columbia Climate School at Columbia University, and the Flowminder Foundation. Thomas Abbott (WorldPop) led the input processing and the modelling work following the Random Forest (RF)-based dasymetric mapping approach developed by Stevens et al. (2015). Heather Chamberlain, Sarchil Qader, and Attila N Lazar advised on the modelling procedure. The Institut National de la Statistique du Niger (INS) released the census-based total population projection using the results of the 2012 census of population and digital Commune boundaries. Engagement with INS was lead by Mathias Kuepie (UNFPA). The work was overseen by Attila N. Lazar and Andy J Tatem.

The authors followed rigorous procedures designed to ensure that the used data, the applied method and thus the results are appropriate and of reasonable quality. If users encounter apparent errors or misstatements, they should contact WorldPop at release@worldpop.org.

RELEASE CONTENT

NER_population_v1_0_gridded.tif
NER_population_v1_0_agesex.zip
NER_population_v1_0_mastergrid.tif

LICENSE

SUGGESTED CITATIONS

Abbott T. J., Chamberlain H., Qader S. H., Lazar A. N., Kuepie, M., Tatem A. J. 2022. Census disaggregated gridded population estimates for Niger (2021), version 1.0. University of Southampton. doi:10.5258/SOTON/WP00733

FILE DESCRIPTIONS

The projection for all GIS files is the geographic coordinate system WGS84 (World Geodetic System 1984).

NER_population_v1_0_gridded.tif

This geotiff raster, at a spatial resolution of 3 arc-seconds (approximately 100m at the equator), contains estimates of the total population size per grid cell across Niger. NA values represent areas that were mapped as unsettled based on gridded building patterns derived from building footprints (Dooley and Tatem, 2020). These data are stored as floating-point numbers rather than integers to avoid rounding errors in aggregated population totals for larger areas.

NER_population_v1_0_agesex.zip

This zip file contains 40 GeoTIFF rasters representing estimated population counts for specific age and sex groups within grid cells of approximately 100m. We provide 36 rasters for the commonly reported age-sex groupings of sequential age classes for males and females separately. These are labelled with either an “m” (male) or an “f” (female) followed by the number of the first year of the age class represented by the data. “f0” and “m0” are population counts of under 1-year olds for females and males, respectively. “f1” and “m1” are population counts of 1 to 4 year olds for females and males, respectively. Over 4 years old, the age groups are in five year bins labelled with a “5”, “10”, etc. Eighty year olds and over are represented in the groups “f80” and “m80”. We provide four additional rasters that represent demographic groups often targeted by programmes and interventions. These are “under1” (all females and males under the age of 1), “under5” (all females and males under the age of 5), “under15” (all females and males under the age of 15) and “f15_49” (all females between the ages of 15 and 49, inclusive).

These data were produced *post-hoc* by multiplying the total population counts provided in the *NER_population_v1_0_gridded.tif* raster and age and sex proportions derived from the microcensus data for each province. While this data represents population counts, values contain decimals, i.e. fractions of people. This is because both the input population data and age-sex proportions contain decimals. For this reason, it is advised to aggregate the rasters at a coarser scale. For example, if four grid cells next to each other have values of 0.25 this indicates that there is 1 person of that age group somewhere in those four grid cells.

NER_population_v1_0_mastergrid.tif

This geotiff raster contains 1s for each approximately 100m grid cell (0.0008333 decimal degrees) across the study area. Zero values indicate grid cells that did not contain buildings and were therefore assumed to be unpopulated. NAs show grid cells considered as outside the study area.

RELEASE HISTORY

Version 1.0 (7 February 2022) doi:10.5258/SOTON/WP00733-
Original release of this data set.

SOURCE DATA

- Digital Commune boundary and their projected population totals for 2021 based on the 2012 Population and Housing Census were provided by INS in a shapefile format.
- Gridded building patterns (building count, building total area, building mean area, building area variance, building density, building length, building mean length and building length variance) were derived from building footprints by Dooley and Tatem (2020).
- Additional geospatial covariates (Lloyd et al., 2019), representing factors related to population distribution (distance to land cover maps, mean precipitation and temperature, slope and elevation, motorized friction surface, walking friction surface, travel time to city, distance to coastline, protected areas, schools, health facility, market place, place of worship, local roads, main roads, railway station, road intersection, and built settlement, and night-time lights), were obtained from the "Global High-Resolution Population Denominators Project" (OPP1134076) and the Copernicus Climate Data Service (Muñoz Sabater, 2021).
- The WorldPop gridded age-sex proportions for Niger (WorldPop et al., 2018) were used 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).

METHODS OVERVIEW

Pre-processing: Overlapping Commune boundaries and minor existing gaps were fixed to ensure accurate linkage between population count and spatial unit. The building pattern and other covariates were adjusted so that extent precisely covers the INS national boundary.

Modelling: Following the Random Forest (RF)-based dasymetric mapping approach (Stevens et al., 2015), the popRF 'R' package (Bondarenko et al., 2021) based on Breiman (2001) algorithm was used to model Commune total population density as a combination of the geospatial covariates and then to estimate the total population density in each approximately 100 m grid cell (0.0008333 decimal degrees grid or 3 arc seconds). The model could explain 91.6% of the total population input variance.

The gridded population estimates were then combined with the WorldPop age/sex pyramid table for Niger (WorldPop et al., 2018) to produce gridded population estimates for each sex group (female and male) at regular age intervals.

ASSUMPTIONS AND LIMITATIONS

This dataset was produced based on the projected 2021 population totals for Communes derived from the 2012 Population and Housing Census, and provided by INS. Although the enumerated population totals have been projected to 2021, the estimate of population in each Commune may not reflect the current population, given the time elapsed since the last census and the necessary assumptions made in projecting the population estimates.

The gridded population estimates are constrained within the settled area derived from gridded building metrics (Dooley and Tatem, 2020). We assumed that the building footprint data (Ecopia.AI and Maxar Technologies, 2020), from which the gridded building metrics were derived, is accurate and that each building polygon corresponds to a building structure. In addition, the distribution of buildings might not represent the current building landscape because of the necessity to use satellite imagery from different years in extraction of the building footprints (e.g. due to cloud coverage) (Dooley and Tatem, 2020). In locations which have recently experienced rapid settlement changes, for example, establishment of new settlements, rapid urban growth or abandonment of settlements, the population estimates are likely to be less accurate. The gridded building metrics and the building footprints from which they were derived, do not provide any information on the building types or use. Therefore in areas that are predominantly non-residential, there is the potential for the estimated population to be overestimated.

WORKS CITED

Bondarenko M., Nieves J.J., Forrest R.S., Andrea E.G., Jochem C., Kerr D., and Sorichetta A. (2021): popRF: Random Forest-informed Population Disaggregation R package, _Comprehensive R Archive Network (CRAN)_, url:https://cran.rproject.org/package=popRF.

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, 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 high resolution gridded population distribution datasets. Big Earth Data, 3(2), 108-139. https://dx.doi.org/10.1080/20964471.2019.1625151

Muñoz Sabater, J., (2021): ERA5-Land monthly averaged data from 1950 to 1980. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.68d2bb30

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.

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