Release Statement

Bottom-up gridded population estimates for Nigeria, version 2.1

19 July 2023

Original Release: 10 July 2019

ABSTRACT

This data release provides gridded population estimates (spatial resolution of 3 arc-seconds, approximately 100 m grid cells) with national coverage for Nigeria, along with estimates of the number of people belonging to various age-sex groups. Version 2.1 is an update to the previous version 2.0 gridded population estimates and is based on a correction of the settlement map. These model-based population estimates most likely represent the time period around 2019, corresponding to the period when the satellite imagery was processed to generate building footprints. Populations are mapped only into areas where residential settlements are predicted.

These data were produced by the WorldPop Research Group at the University of Southampton. This work was part of the GRID3 Bridge Funding project with funding from the Bill and Melinda Gates Foundation (INV-045694). Project partners included the GRID3 Inc and the Center for International Earth Science Information Network in the Earth Institute at Columbia University. Statistical modelling was originally led by Chris Jochem and Doug Leasure with additional support and oversight from Attila Lazar and Andy Tatem. Ortis Yankey led the population map update with additional support from Edith Darin.

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

1. NGA_population_v2_1_gridded.tif
2. NGA_population_v2_1_agesex.zip

LICENSE

SUGGESTED CITATIONS

WorldPop. 2023. Bottom-up gridded population estimates for Nigeria, version 2.1. WorldPop, University of Southampton.

FILE DESCRIPTIONS

The projection for all GIS files is the geographic coordinate system WGS84 (World Geodetic System 1984: EPSG 4326).

NGA_population_v2_1_gridded.tif

This geotiff raster, at a spatial resolution of 3 arc-seconds (approximately 100m at the equator), contains estimates of total population size per grid cell across Nigeria. NA values represent areas that were mapped as unsettled based on a gridded settlement layer derived from building footprints (Maxar Technologies, Inc. and Ecopia Tech Corporation, 2021). These data are stored as floating point numbers rather than integers to avoid rounding errors in aggregated population totals for larger areas.

NGA_population_v2_1_agesex.zip

This zip file contains 40 geotiff rasters at a spatial resolution of 3 arc-seconds (approximately 100 m). Each raster provides gridded population estimates for an age-sex group per grid cell across Nigeria. 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 by 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 using age-sex proportions from the WorldPop project (WorldPop et al., 2018). The age-sex proportions were applied to the gridded population estimates

(NGA_population_v2_1_gridded.tif) to allocate the population to the different age-sex classes. 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

RELEASE HISTORY

- **Version 2.1** (19 July 2023, https://dx.doi.org/10.5258/SOTON/WP00765)
    - The population map was updated using amended settlement data due to a change in the feature extraction algorithm
    - Totals at ward level are the same as previous version therefore they are representative of the year 2019.

- **Version 2.0** (17 November 2021)[https://dx.doi.org/10.5258/SOTON/WP00729]
    - Refinement of gridded population estimates using more recent settlement data based on building footprints.
    - Predictions of residential and non-residential buildings incorporated in the settlement map.
    - A different regional boundary definition was used in the model, corresponding with Nigerian statistical regions.
    - Representative of the year 2019.

- **Version 1.2** (15 September 2020)
    - A peer-reviewed article (Leasure et al., 2020b) was added to describe the statistical methods that were developed to produce the population estimates (https://doi.org/10.1073/pnas.1913050117).

- **Version 1.2** (20 May 2020)
    - Gridded population estimates were added to NGA_population_v1_2_agesex.zip for the following demographic groups: children under 1, children under 5, children under 15, and women 15 to 49 years of age.

- **Version 1.2** (26 March 2020) [https://dx.doi.org/10.5258/SOTON/WP00661]
    - Gridded population estimates were added for individual age-sex groups (NGA_population_v1_2_agesex.zip).
    - The SQL database “NGA_population_v1_2_sql.sql” that is used in WOPR applications was updated to remove unnecessary data (e.g. covariate values, names of administrative units).
    - Population tiles were updated with a revised color palette. This file was renamed from “NGA_population_v1_2_tiles_population.zip” to “NGA_population_v1_2_tiles.zip”.

- Uncertainty tiles “NGA_population_v1_2_tiles_uncertainty.zip” were removed because they were discontinued for use in WorldPop web applications (e.g. https://apps.worldpop.org/woprVision).

- **Version 1.2** (10 July 2019) [https://dx.doi.org/10.5258/SOTON/WP00655]
  - The previous release contained a few grid cells with erroneously high population estimates that resulted from the way the statistical model was summarised (based on 1,000 samples from posterior predictions as opposed to 10,000 samples used here).
  - This update changes the population estimates slightly in every grid cell. State and LGA totals have changed marginally but remain within 1% of previous estimates.
  - Representative of the time period from 2016 to 2017.

- **Version 1.1** (22 February 2019) [https://dx.doi.org/10.5258/SOTON/WP00657]
  - Updated to include floating-point rasters rather than integer rasters to resolve rounding errors when calculating population totals for larger areas (e.g. zonal sums)

- **Version 1.0** (11 November 2018) [https://dx.doi.org/10.5258/SOTON/WP00656]
  - Original release of Nigeria population dataset

ASSUMPTIONS AND LIMITATIONS

These population estimates mostly likely represent the time period of 2018 to 2019, reflecting when the majority of the satellite imagery scenes were processed to produce the bottom-up population estimates version 2.0. The updated settlement layer (Ecopia version 2) used in this gridded population estimates version 2.1 means that the spatial allocation of population has changed compared with previous data release. The population totals at ward level remain however the same as for the bottom-up population estimates version 2.0.

The method adopted to reallocate the population (see Methods section) is based on the mean estimate of previous version and do not incorporate any of the previously modelled population distribution and thus any quantification of uncertainty. The variation in population density is based on the bottom-up population estimates version 2.0. This model assumes that population densities observed during the earlier time period are still representative of the more recent period. Similarly, the age and sex structures data come from an earlier time point and were assumed to be stable.

Furthermore, the population estimates assume that no people live in areas that are predicted to be either unsettled or non-residential. Compared with version 1.2, the residential/non-residential classification used in the version 2.0 and v2.1 estimates is, in general, less restrictive. It considers more areas to be potential residential in order to avoid excluding and underestimating the population in urban centres. This could lead to over- or under-estimates in misidentified areas.

Due to differences in the spatial data of administrative boundaries, population estimates may be missing from some areas near the Nigerian border.

SOURCE DATA

Bottom-up gridded population estimates for Nigeria, version 2.0 (WorldPop and National Population Commission of Nigeria, 2021)

We used previous gridded population estimates generated using a bottom-up model-based approach that linked population, observed at 1142 microcensus survey locations collected in 2016 – 2017, and high-resolution geospatial datasets with national coverage.

Settlement Data

Settled grid cells and the number of buildings per grid cell were defined as a raster. These gridded layers were derived from building footprint polygons extracted from high-resolution satellite imagery. A first version was released in 2020 and used in the bottom-up gridded population estimates for Nigeria, version 2.0. The current gridded population leverages a second version incorporating corrections in the feature extraction algorithm (Maxar Technologies, Inc. and Ecopia Tech Corporation, 2021).

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

We used WorldPop gridded age-sex proportions for Nigeria 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). We multiplied our gridded population estimates (NGA_population_v2_1_gridded.tif) by the gridded age-sex proportions to produce NGA_population_v2_1_agesex.zip.

METHODS OVERVIEW

Building on the gridded population model for Nigeria version 2.0 (WorldPop and National Population Commission of Nigeria, 2021), we adapted the population map to match the settlement map version 2 provided by Maxar Technologies, Inc. and Ecopia Tech Corporation AI (2021).

We implemented a two-step approach:

1. The population size at grid-cell level was updated by multiplying the mean population estimate from version 2.0 with the ratio of the number of buildings observed in the settlement map version 2 over the number of buildings observed in the settlement map version 1 in the grid cell. If no building was present in version 1, we posited that one building was observed in the building footprint version 1 and one person was estimated in the gridded population version 2.0 and applied the same ration method. In summary we computed at grid cell level an intermediary population estimate at grid cell level, pop:

$$pop = \begin{cases} \overline{pop}_{v2.0} * \dfrac{building_{v2}}{building_{v1}}, & pop_{v2.0} > 0 \\ building_{v2}, & pop_{v2.0} = 0 \end{cases}$$

2. Because we consider that the settlement map consisted in a correction, we don't want the change in spatial allocation to impact the total population sizes previously estimated. We therefore calibrate the new population estimates at grid-cell level such that they sum up to ward population estimates of version 2.0 following this formula:

$$pop_{v2.1} = \sum_{ward} \overline{pop}_{v2.0} * \frac{pop}{\sum_{ward} pop}$$

In other words, we use the gridded population estimated by the first step to disaggregate ward population estimated by the bottom-up model used in the gridded population version 2.0.

We implemented the model using R (R Core Team, 2023).

WORKS CITED

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.

Maxar Technologies, Inc. and Ecopia Tech Corporation. 2021. Ecopia Landbase Africa powered by Maxar. Available at https://www.ecopiatech.com Version 2

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.

R Core Team. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.

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).

WorldPop and National Population Commission of Nigeria. 2021. Bottom-up gridded population estimates for Nigeria, version 2.0. WorldPop, University of Southampton. doi: 10.5258/SOTON/WP00729