Documentation for the Global Population Density Grid Time Series Estimates

October 2017

Center for International Earth Science Information Network (CIESIN),
Columbia University

Abstract

This document outlines the basic methodology and data sets used to construct the Global Population Density Grid Time Series Estimates, v1 (1970-2000), along with use cases, limitations, and use constraints.

Data set citation: Center for International Earth Science Information Network (CIESIN), Columbia University. 2017. Global Population Density Grid Time Series Estimates. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H47M05W2. Accessed DAY MONTH YEAR.

Suggested citation for this document: Center for International Earth Science Information Network (CIESIN), Columbia University. 2017. Documentation for the Global Population Density Grid Time Series Estimates. Palisades NY: NASA Socioeconomic Data and Applications Center. https://doi.org/10.7927/H4W093VW. 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

Documentation last revised October 6, 2017.

Contents

I. Introduction

The Global Population Density Grid Time Series Estimates provide estimates of population over three decades from 1970 to 2000. Raster data sets at a resolution of 30 arc-seconds representing population densities for four time slices (1970, 1980, 1990, and 2000) were developed based on subnational growth rates from the History Database of the Global Environment (HYDE) (Klein Goldewijk *et al.* 2010) and population raster data from the Global Urban-Rural Mapping Project, Version 1 (GRUMPv1) (CIESIN *et al*. 2011). These data were developed in part to estimate net migration for the Foresight Project on Environmental Migration (CIESIN 2011), which resulted in the Global Estimated Net Migration Grids By Decade data set (de Sherbinin *et al.* 2015), also distributed by SEDAC.

II. Data and Methodology

The GRUMPv1 data are based on a fixed (circa 1995) estimate of urban extents and two census dates (circa 1990 and 2000) and are therefore not appropriate for developing a longer time series via simple extrapolation of trends. The lack of subnational migration data for the forty-year time span considered by the Foresight project meant that the authors needed to use indirect estimation methods to derive spatially explicit estimates of migration. Their basic methods can be summarized as follows, with details presented in the remainder of the section.

1. The authors utilized the History Database of the Global Environment, Version 3.1 (HYDEv3.1) population grids for the years 1970, 1980, 1990, and 2000 to create

one degree grids representing the rates of change in population for each decade. This makes optimal use of the HYDEv3.1 data set, which provides a consistent decadal time series of population distribution over several centuries.

2. The rate grids were applied to the GRUMPv1 population grids for 1995, producing "back-cast" grids to 1970, 1975, 1980, 1985, 1990, and 1995. This method has the advantage of keeping the detailed spatial distribution of the GRUMPv1 data, which are based on detailed administrative units and urban areas derived from night lights. It also enabled the analysis to be conducted at the higher resolution of the GRUMPv1 product (30 arc-second resolution for GRUMPv1 vs. 5 arc-minute resolution for HYDEv3.1).

3. The global grids were adjusted to match country totals from the UN population estimates for each given year. This was done proportionally by calculating the ratio of grids summed by country for each time slice to the UN estimate for each country for that time slice and then applying that ratio to the population count grids for each year, effectively adding or subtracting small amounts of population equally across all of the cells in each country to match the UN total.

The value of these grids is that they represent a consistent population time series at relatively high spatial resolution.

Methods

To conduct this modeling exercise, the authors chose to use the GRUMPv1 Population Count Grid, which represents an urban reallocation of the Gridded Population of the World, Version 3 (GPWv3) (CIESIN *et al*. 2005) using night-time lights plus additional urban spatial extents and an algorithm that "pulls" population from larger administrative units in rural areas into urban areas. GRUMPv1 is based on population data from GPWv3, which uses 338,863 census units outside of the U.S. and 60,884 within the U.S. It is worth noting, however, that the average population reporting unit size varies considerably by region, from 9,433 and 7,042 square km in Africa and Asia, respectively, to 5,744 square km in South America, 2,516 square km in Europe, and 1,094 square km in the rest of the Americas. This variability in the size of census unit is somewhat mitigated by the algorithm that pulls populations into urban areas, but nevertheless, in developing regions and regions with large areas of sparsely populated drylands, there is generally less certainty regarding the spatial location of populations, and this will affect estimates of net migration.

To ensure consistent rates of population change over the four decadal periods, the grids representing the rate of population change per decade based on the HYDEv3.1 population grids for the years 1970, 1980, 1990, and 2000 were applied to the GRUMPv1 data. The HYDEv3.1 grids are adjusted at the country level to match the country totals from the UN Population Division's World Population Prospects, 2008 Revision. Although HYDEv3.1 is distributed at a 5 arc-minute resolution, the rates were calculated at a one-degree resolution in order to average over a wider area and reduce the impact of decade-on-decade population variability inherent in higher resolution grid cells. A 3x3 pixel mean function was also applied to the rate grids and its values were used to fill in gaps in the HYDEv3.1-derived rates for areas that had no population in HYDEv3.1, but observed population values in GRUMPv1.

The GRUMPv1 Population Count Grid for the year 2000 was "back-cast" to 1970, 1980, and 1990, by multiplying the HYDEv3.1 rates times the population grids. For the most part, negative rates were used for back-casting, but in selected areas of declining population over the course of each decade, the sign for the rates was reversed. In each case, the gridded country totals were adjusted so that they equal the UN World Population Prospects, 2008 Revision country population totals for each time period. In this way, all population data were consistent with the UN World Population Prospects, 2008 Revision, which represents a harmonized time series of country-level demographic data.

To develop the population density grid, the population in each grid cell was divided by the area (in square km) of each grid cell to yield a population density per square kilometer.

III. Data Set Description(s)

Data set description:
The population density grid time series estimates consist of the number of persons per 30 arc-second (~1 km) grid cell divided by the area of the grid cell in kilometers for each of the four target years: 1970, 1980, 1990, 2000.

Data set web page:
http://sedac.ciesin.columbia.edu/data/set/popdynamics-global-pop-density-time-series-estimates

Data set format:
The data are available in GeoTiff format as downloadable zip files. Each downloadable is a compressed zip file, containing: 1) global GeoTiff for the year of estimate, 2) PDF documentation.

Data set downloads:
popdynamics-global-pop-density-time-series-estimates-1970-geotiff.zip
popdynamics-global-pop-density-time-series-estimates-1980-geotiff.zip
popdynamics-global-pop-density-time-series-estimates-1990-geotiff.zip
popdynamics-global-pop-density-time-series-estimates-2000-geotiff.zip

IV. How to Use the Data

The data can be used in any standard geographic information system (GIS) software package.

V. Potential Use Cases

Consistent time series gridded population data are difficult to come by owing to the challenge imposed by changing underlying census geographies. If census reporting units are spatially divided or combined with each succeeding census, the resulting grids will reflect these changes rather than changes in the actual populations on the ground. This data set is appropriate for research requiring high resolution (30 arc-second, or approximately 1km) population data for the decades 1970, 1980, 1990 and 2000. However, users are advised to review the limitations in the next section before applying the data to any particular spatial analysis.

VI. Limitations

The population grids are consistent internally within the time series, but are not recommended for use in creating longer time series with any other population grids, including GRUMPv1, Gridded Population of the World, Version 4 (GPWv4), or non-SEDAC developed population grids.

One drawback of HYDEv3.1 is that many small island states are not included in the data set, so population estimates are not included for these countries.

Furthermore, owing to data limitations, these population time series estimates contain uncertainties. Uncertainties are highest at the individual grid cell level and for small areas, such as second administrative level units (e.g., counties, municipals). Therefore, these data are not recommended for use in analyses of small areas. They are best used, as they were in the original Foresight study, for analyses over broader areas such as large ecosystems or climate zones. Users are encouraged to read the full documentation in order to understand the limitations.

VII. Acknowledgments

This work reflects the contributions of a large number of individuals. Alex de Sherbinin served as PI and lead author of the report (de Sherbinin *et al*. 2012), and Marc Levy served as co-PI and also led the development of the methodology. Greg Yetman led the geospatial processing with assistance from Kytt MacManus, Liana Razafindrazay, Cody Aichele and John Squires.

Funding for development of this data set was provided by the United Kingdom Government Office for Science as part of the Foresight Project on Global Environmental Migration. Funding for documentation and dissemination was provided under the U.S. National Aeronautics and Space Administration (NASA) contract NNG13HQ04C 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

X. Recommended Citation(s)

Data set:

Center for International Earth Science Information Network (CIESIN), Columbia University. 2017. Global Population Density Grid Time Series Estimates. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H47M05W2. Accessed DAY MONTH YEAR.

Scientific publication:

de Sherbinin, A., M. Levy, S. Adamo, K. MacManus, G. Yetman, V. Mara, L. Razafindrazay, B. Goodrich, T. Srebotnjak, C. Aichele, and L. Pistolesi. 2012. Migration and Risk: Net Migration in Marginal Ecosystems and Hazardous Areas. *Environmental Research Letters* 7(4): 045602. http://iopscience.iop.org/1748-9326/7/4/045602.

XI. Source Code

Please contact SEDAC User Services at ciesin.info@ciesin.columbia.edu for information about the code used to develop the data set.

XII. References

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): Population Count Grid. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4VT1Q1H. Accessed DAY MONTH YEAR.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2011. *Foresight Project on Migration and Global Environmental Change, Report MR4: Estimating Net Migration by Ecosystem and by Decade, 1970-2010*. London: UK Government Foresight.

Center for International Earth Science Information Network (CIESIN) Columbia University, United Nations Food and Agriculture Programme (FAO), and Centro Internacional de Agricultura Tropical (CIAT). 2005. Gridded Population of the World, Version 3 (GPWv3): Population Count Grid. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4639MPP. Accessed DAY MONTH YEAR.

de Sherbinin, A., M. Levy, S. Adamo, K. MacManus, G. Yetman, V. Mara, L. Razafindrazay, B. Goodrich, T. Srebotnjak, C. Aichele, and L. Pistolesi. 2015. Global Estimated Net Migration Grids by Decade: 1970-2000. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4319SVC. Accessed DAY MONTH YEAR.

Klein Goldewijk, K., A. Beusen, and P. Janssen. 2010. Long term dynamic modeling of global population and built-up area in a spatially explicit way, HYDE 3.1. *The Holocene* 20(4):565-573. http://dx.doi.org/10.1177/0959683609356587.

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	Contributors	Revisions
October 6, 2017	John Squires, Gregory Yetman, Alex de Sherbinin	This document is the 1st instance of documentation