Gridded Population of the World, Version 4: A First Look

Erin Doxsey-Whitfield and Linda Pistolesi

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

Introduction

The Gridded Population of the World (GPW) data set displays the global distribution of the human population on a continuous surface. The GPW series, which began in 1995 and is now in its fourth version (GPWv4), is a valued asset for the research community. The third version of GPW has been used extensively in integrated human-environment research, including vulnerability mapping, disaster impacts, and health dimensions of environmental change, often with satellite remote sensing or other biophysical data. GPWv4 will be publicly available at http://sedac.ciesin.columbia.edu/. This poster describes the GPW methodology and the key highlights of GPWv4.

Methods

The development of GPWv4 builds upon previous versions of the data set (Figure 2). The two basic inputs of GPW are non-spatial population data (i.e., tabular counts of population listed by administrative area) and spatially-explicit administrative boundary data. GPWv4 has been updated using census data and cartography from the 2010 round of censuses. Data are gridded at 30 arc-seconds (~1 km).

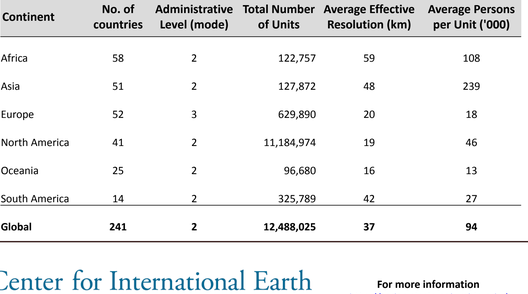


Fig. 2. Steps used in GPW to convert tabular census data into a uniform grid.

GPWv4 uses the areal-weighting method (also known as uniform distribution or proportional allocation) to disaggregate population from census units into 1km (30 arc-second) grid cells through the simple assumption that the population of a grid cell is an exclusive function of the land area within that pixel. Unlike more highly-modeled methods that incorporate ancillary geographic data (e.g. land cover, urban extent, etc.), the areal-weighting method does not use any additional data to allocate the population within a grid cell (Bhaduri et al., 2002; Balk et al., 2006; Tatem et al., 2007).

Census information modeled with this approach may be freely and easily incorporated into global analyses that make use of ancillary data sets that might be endogenous to more highly-modeled surfaces. However, the precision and accuracy of a given pixel is a direct function of the size of the input areal unit, which is why a large effort was made to increase the spatial resolution of input units in GPWv4.

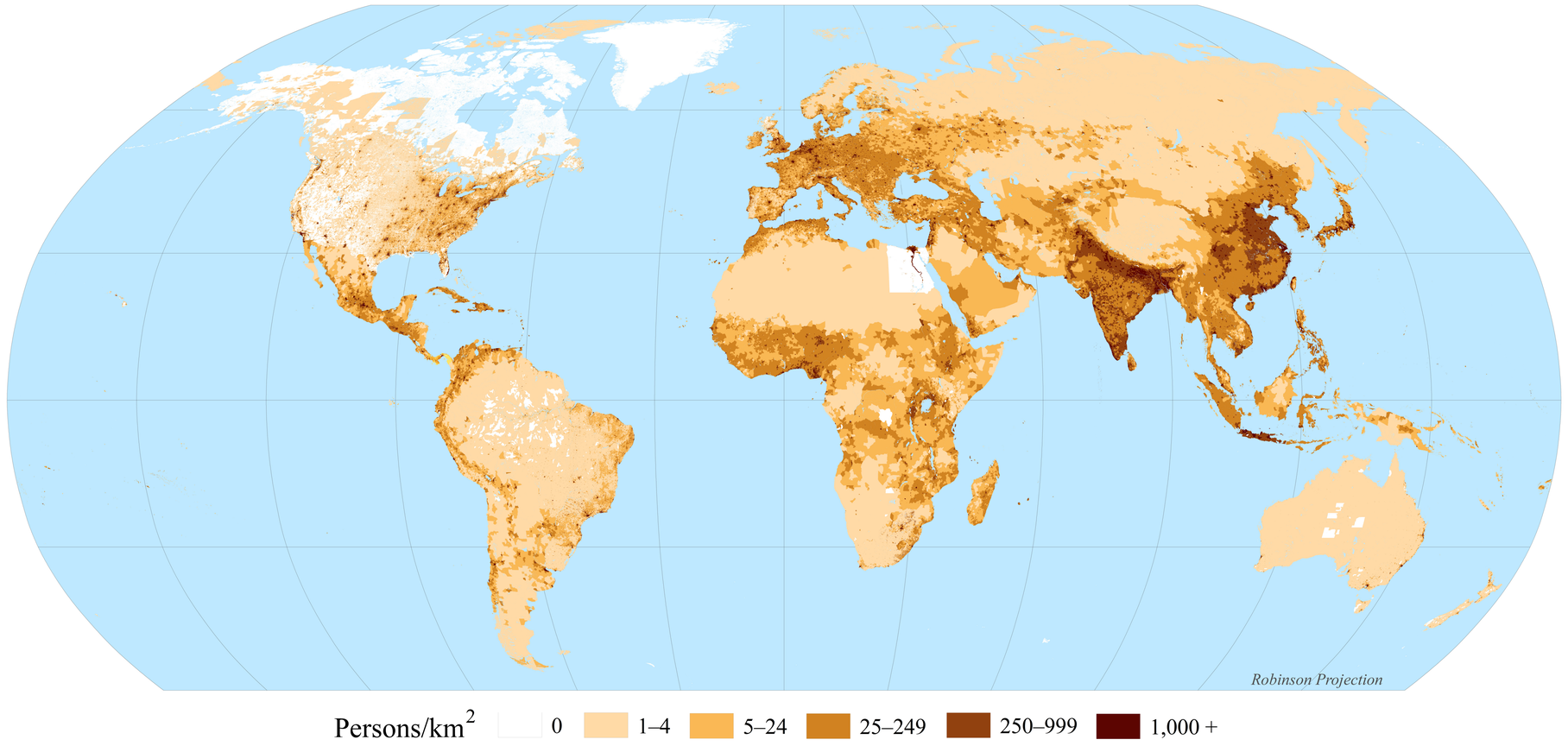


Fig. 1. Population Density, 2010.

Increased Spatial Resolution of Input Units

In the last decade, there have been considerable advances in the availability of online and open source census data. Improvements in technology have allowed census bureaus to more easily distribute their results to the public using electronic and online formats. The increase in freely accessible tabular and boundary data has allowed us to greatly improve the accuracy of GPWv4 by incorporating census data at a higher spatial resolution than was previously possible.

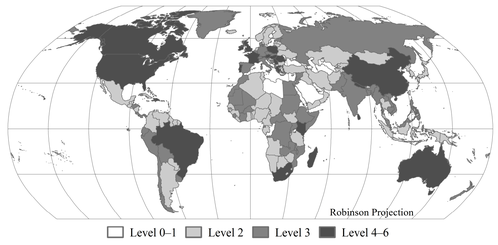


Fig 3. Administrative level used in GPWv4 by country.

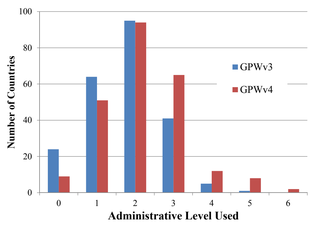


Fig 4. Frequency of administrative level used. Higher numbers denote more units of a smaller size.

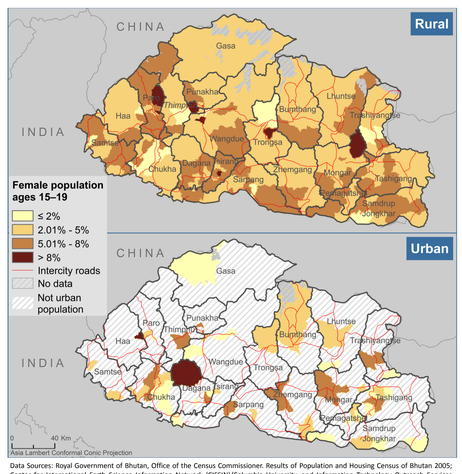
Table 1. Summary information of input units by continent.

Continent	No. of countries	Administrative Level (mode)	Total Number of Units	Average Effective Resolution (km)	Average Persons per Unit ('000)
Africa	58	2	122,757	59	108
Asia	51	2	127,872	48	239
Europe	52	3	629,890	20	18
North America	41	2	11,184,974	19	46
Oceania	25	2	96,680	16	13
South America	14	2	325,789	42	27
Global	**241**	**2**	**12,488,025**	**37**	**94**

Overall improvements between GPWv3 and GPWv4:

- Substantial increase in total input units
  - **Globe**: 12,500,000 total input units (vs. 375,000 in GPWv3)
  - **U.S.**: 10,500,000 census block groups (compared to 60,000 census tracts in GPWv3)
  - **Outside of U.S.**: 1,900,000 units (more than five-fold increase over GPWv3)
- 86 countries gridded at higher level
- 35 more countries gridded at level 3 or higher (Figures 3 and 4)
- 29 countries saw an increase of 100–999 units
- 25 countries increased by more than 1,000 units
- Average resolution of all countries improved from 46 km to 37 km
- Higher overall resolution and increased accuracy

Addition of New Census Variables



Data Sources: Royal Government of Bhutan, Office of the Census Commissioner. Results of Population and Housing Census of Bhutan 2005; Center for International Earth Science Information Network (CIESIN)/Columbia University, and Information Technology Outreach Services (ITOS)/University of Georgia. 2013. Global Roads Open Access Data Set, Version 1 (gROADSv1). Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1; GADM database (www.gadm.org), version 2.0, December 2011.

In order to broaden the applicability of GPW, version 4 is expanding to include three census variables, circa 2010:

- age, as single- or five-year age groups
- sex
- urban/rural status

Where possible, the variables will be cross-tabulated, resulting in a consistent global gridded population data set with detailed estimates of age, sex, and urban/rural distribution within each country. These detailed grids will be a vital tool for investigating a range of issues, including sustainable development, urbanization, migration, hazard vulnerability, disaster preparedness, and health. Examples of these grids are given for Bhutan and Costa Rica (Figures 5 and 6).

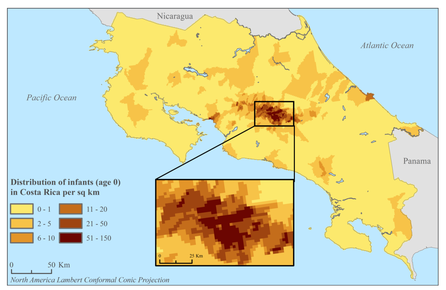


Fig. 6. Distribution of infants (age 0) in Costa Rica, 2010. This grid can be used as an input to calculate infant mortality rates, an indicator used to measure the health and well-being of a population.

Fig. 5. 2010 estimate of urban and rural female adolescent population in Bhutan, as percentage of total population. Used in conjunction with intercity roads and adolescent fertility rates, these maps provide increased insight into adolescent pregnancy (CIESIN, 2013). Reducing adolescent birth rates worldwide is part of the Millennium Development Goals, the internationally-agreed development targets for 2015.

Center for International Earth Science Information Network
Earth Institute | Columbia University

For more information
http://sedac.ciesin.columbia.edu/
Contact: Erin Doxsey-Whitfield
(edwhitfi@ciesin.columbia.edu)

This work was carried out under NASA Contract NNG08HZ11C for the Continued Operation of the Socioeconomic Data and Applications Center (SEDAC).

The authors would like to thank Kytt MacManus, Susana B. Adamo, John Squires, Olena Borkovska, Lisa Lukang, Valentina Mara, Jared Talkin, Jane Mills, Ilyun Koh, Sneha Rao, Chenyang Zhao, and Evan Drewry, who all contributed to the data set production.

References: Balk, D.L., U. Deichmann, G. Yetman, F. Pozzi, S.I. Hay, and A. Nelson. 2006. *Advances in Parasitology* 62: 119–156. Bhaduri, B., E. Bright, P. Coleman, and J. Dobson. 2002. *Geoinformatics* 5 (2): 34–37. CIESIN. 2013. *Analysis of spatial data combined with teen pregnancy rates may provide increased insights*. Palisades, New York: CIESIN, The Earth Institute, Columbia University. http://blogs.ei.columbia.edu/2013/08/06/analysis-of-spatial-data-combined-with-teen-pregnancy-rates-may-provide-increased-insights/. Tatem, A.J., A.M. Noor, C. von Hagen, A. Di Gregorio, and S.I. Hay. 2007. *PLoS One* 2 (12): e1298.