
Erin Doxsey-Whitfield
Kytt MacManus
Susana B. Adamo
Linda Pistolesi
John Squires
Olena Borkovska
Sandra Baptista

October 17, 2014
Applied Geography Conference
Atlanta, Georgia
Socioeconomic Data & Applications Center (SEDAC)

- Integration of social and earth science data, especially with remote sensing
- Focus on human dimensions of environmental change
- Direct support to scientists, applied and operational users, decision makers, and policy communities
- Strong links to geospatial data community
Outline

1. What is Gridded Population of the World (GPW)?
2. Methods for GPWv4
3. Improvements made in GPWv4
4. Challenges with acquiring/integrating census data
5. Recommendations for National Statistics Offices
Gridded Population of the World (GPW)

- Global population grid (raster) developed to provide a **spatially-disaggregated population layer** that is compatible with data sets from social, economic, and Earth science fields.

- Input data: census population counts and census geography

- Census population data are transformed from their native spatial units to a global grid of quadrilateral latitude-longitude cells (Balk et al. 2010)

- Free and openly available

GPW version 3, 2000 population density
History of GPW

- GPWv1 was an outgrowth of a Global Demography Workshop held at CIESIN in 1994 (produced by Waldo Tobler, Uwe Deichmann, and others).
- Consensus that a consistent global database of population totals in raster format would be invaluable for interdisciplinary study (Deichmann et al., 2001)

<table>
<thead>
<tr>
<th></th>
<th>GPWv1</th>
<th>GPWv2</th>
<th>GPWv3</th>
<th>GPWv4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Input Units</td>
<td>19,000</td>
<td>127,000</td>
<td>c. 400,000</td>
<td>~ 12,500,000</td>
</tr>
<tr>
<td>(subnational geographic units)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grid Resolution</td>
<td>2.5 arc-minute</td>
<td>2.5 arc-minute</td>
<td>2.5 arc-minute</td>
<td>30 arc-second (1 km)</td>
</tr>
<tr>
<td>Census variables</td>
<td>Total Population</td>
<td>Total Population</td>
<td>Total Population</td>
<td>Total Population, Sex, Age, Urban/Rural status</td>
</tr>
</tbody>
</table>
Applications of GPW

Broad use in research, policy making, communications, and human and environmental problem-solving

Often combined with satellite remote sensing or other biophysical data

Areas of application:
- Health dimensions of environmental change
- Vulnerability mapping
- Disaster impacts
Outline

1. What is Gridded Population of the World (GPW)?

2. Methods for GPWv4

3. Improvements made in GPWv4

4. Challenges with acquiring/integrating census data

5. Recommendations for National Statistics Offices
Methods for GPWv4

2. Adjust estimate to UN World Population Prospects for target years
3. Proportionally-allocate population to 1 km grids using an areal-weighting method
4. Match to geographic boundaries (census or administrative)
5. If needed, adjust boundaries to global framework
6. Find tabular population counts
7. If needed, adjust boundaries to global framework
GPW is minimally-modeled

- GPW uses the areal-weighting method
  - Does not incorporate ancillary data (e.g. land use/land cover, transportation networks, elevation, etc.)
  - Uniformly distributes population based on land area

- Maintains fidelity to input data

- The accuracy of GPW pixel estimates is directly related to the size of the input census units
Outline

1. What is Gridded Population of the World (GPW)?

2. Methods for GPWv4

3. Improvements made in GPWv4

4. Challenges with acquiring/integrating census data

5. Recommendations for National Statistics Offices
Improvements to GPWv4

1. Increased spatial resolution of input data

Accuracy of GPW depends upon the size of the input census units

<table>
<thead>
<tr>
<th></th>
<th>GPWv3</th>
<th>GPWv4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Global</strong></td>
<td>399,747</td>
<td>12,497,563</td>
</tr>
<tr>
<td><strong>USA</strong></td>
<td>60,884</td>
<td>10,608,747</td>
</tr>
<tr>
<td><strong>Rest of World</strong></td>
<td>338,863</td>
<td>1,888,816</td>
</tr>
</tbody>
</table>
Improvements to GPWv4

1. Increased spatial resolution of input data

<table>
<thead>
<tr>
<th>Administrative Level Used</th>
<th>Frequency of Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPWv3</td>
<td></td>
</tr>
<tr>
<td>GPWv4</td>
<td></td>
</tr>
</tbody>
</table>

- # of countries at Level 3 or higher:
  - GPWv3: 47
  - GPWv4: 87
Improvements in Input Data Resolution

<table>
<thead>
<tr>
<th>Resolution Level</th>
<th>Number of Countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower resolution than GPWv3</td>
<td>11 countries</td>
</tr>
<tr>
<td>Same as GPWv3</td>
<td>132 countries</td>
</tr>
<tr>
<td>Higher resolution than GPWv3</td>
<td>98 countries</td>
</tr>
</tbody>
</table>
Improvements to GPWv4

2. Additional census variables

Demand to include demographic information in global population grids

GPWv4 will include detailed global grids for:

- Sex
- Age (single year or 5-year age groups)
- Urban/Rural status
Improvements to GPWv4

2. Additional census variables

2010 % rural population, Panama

2010 % urban population, Panama
Improvements to GPWv4

2. Additional census variables

Adolescent Literacy in Nigeria, 2006

The percentage of Nigeria’s literate adolescent population (ages 15–19) in 2006 is shown in the bar charts by sex and for each state, overlaid on the spatial distribution grid for this same population.


Center for International Earth Science Information Network

EARTH INSTITUTE | COLUMBIA UNIVERSITY

http://blogs.ei.columbia.edu/
Reasons for Improvements

• GPWv4 has finer grid resolution
  – 30 arc-second (~1km) vs 2.5 arc-minute (~4km)

• Technological advancements by NSOs
  – More census data is available online
  – More is open source
  – More interactive, online databases
Outline

1. What is Gridded Population of the World (GPW)?

2. Methods for GPWv4

3. Improvements made in GPWv4

4. Challenges with acquiring/integrating census data

5. Recommendations for National Statistics Offices
Challenges with acquiring census data

- Room for ongoing improvements with online dissemination

- Asymmetry between availability of tabular census data and census geography

- Challenges can lead to compromises in data processing

<table>
<thead>
<tr>
<th>Primary method of census data dissemination</th>
<th>Percent of respondents (121 countries)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper publications</td>
<td>52%</td>
</tr>
<tr>
<td>Static web pages (html, excel, PDF)</td>
<td>28%</td>
</tr>
<tr>
<td>Interactive online databases</td>
<td>14%</td>
</tr>
<tr>
<td>CD/DVD</td>
<td>4%</td>
</tr>
<tr>
<td>Other</td>
<td>2%</td>
</tr>
</tbody>
</table>

Challenges with integrating census data

Ideally, the NSO releases census population data and digital census geography, *with a common identifying code*.

- **GIS data from non-Census Source**
  - Refer to the same unit by different names
  - GIS data might not capture changes in boundaries over time and therefore requires editing
  - Significant effort is needed to reconcile census areas with those present in the GIS data and to assign common identifiers
Outline

1. What is Gridded Population of the World (GPW)?
2. Methods for GPWv4
3. Improvements made in GPWv4
4. Challenges with acquiring/integrating census data
5. Recommendations for National Statistics Offices
Recommendations for NSOs and National Mapping Agencies

1. Developing a digital data dissemination strategy at the highest resolution possible
2. Providing census data in a database format (e.g. Excel)
3. Including common identifying codes that are shared by tabular and boundary data
4. Documenting changes in administrative areas

Obtaining high-resolution census data and matching census geography remains a priority for improving global gridded population data sets
Acknowledgements

The authors would like to thank 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. This work was funded by NASA under contract NNG08HZ11C for the continued operation of the Socioeconomic Data and Applications Center (SEDAC) at CIESIN, a center of Columbia University.

Erin Doxsey-Whitfield  edwhitfi@ciesin.columbia.edu  http://sedac.ciesin.columbia.edu/