Lessons Learned from the production of Gridded Population of the World Version 4 (GPW4)

Columbia University, CIESIN, USA
EFGS
October 2014
Gridded Population of the World

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

- Population data are transformed from their native spatial units to a global grid of quadrilateral latitude-longitude cells (2.5 arc minutes in GPW3) (Balk et al. 2010)
History of GPW

- GPWv1 was an outgrowth of a Global Demography Workshop held at CIESIN in 1994

- 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
Methods for GPWv4

Find tabular population counts → Match to geographic boundaries (census or administrative) → Adjust boundaries to global framework → Estimate the population for target years (2000, 2005, 2010, 2015, 2020) → Adjust estimate to UN World Population Prospects for target years → Proportionally-allocate population to 1 km grids using an areal-weighting method.
Complication

• Higher resolution boundaries can sometimes make regional integration difficult.
Complication

• Option 1: Select either URY or BRA as framework and adjust other boundary
Complication

• Option 2: Select smaller scale international framework and adjust both countries
Our Solution

• We choose
  – Option 2: Select smaller scale international framework and adjust both countries

• But this is not without issue
  – Changes the density of border units
  – Increases uncertainty of estimates of border pixels
Challenges with census data and boundaries

- Room for ongoing improvements with online dissemination
- Different levels of availability of tabular census data and census geography
- Significant research required to reconcile census areas with those present in the GIS data and to assign common identifiers

<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
India Census Atlas

Documents boundary change over time

Rich historical data ranging back to the 1800’s

Makes the idea of a Census time-series imaginable
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.)
  – Distributes population based on land area

• The accuracy of GPW pixel estimate is directly related to the size of the input areal units

Higher resolution boundaries in eastern China lead to more precise pixel estimates
Characteristics of GPWv4
Increased spatial resolution of input data

Pixel level accuracy of GPW depends upon the size of the input census units

<table>
<thead>
<tr>
<th>Number of Input Census Units</th>
<th>GPWv3</th>
<th>GPWv4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>399,747</td>
<td>12,497,563</td>
</tr>
<tr>
<td>USA</td>
<td>60,884</td>
<td>10,608,747</td>
</tr>
<tr>
<td>Outside USA</td>
<td>338,863</td>
<td>1,888,816</td>
</tr>
</tbody>
</table>
Characteristics of GPWv4

Leverages greater availability of Census information

![Frequency of Administrative Level Used](image)

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

11 countries: Lower resolution than GPWv3
132 countries: Same as GPWv3
98 countries: Higher resolution than GPWv3
Characteristics of GPWv4

Additional census variables

Demographic information in population grids

GPWv4 will include grids for:

• Sex
• Age (single year or 5-year age groups)
• Urban/Rural status
Panama Example
Sub-National Urban Rural Fractions

2010 % rural population, Panama

2010 % urban population, Panama
Costa Rica Example

Concentrations of demographic variables

2010 distribution of infants (age 0) in Costa Rica per sq km

- Yellow: 0 - 1
- Orange: 2 - 5
- Brown: 6 - 10
- Dark Brown: 11 - 20
- Gray: 21 - 50
- Maroon: 51 - 150

Provinces

North America Lambert Conformal Conic Projection
Nigeria Example

Extending Demographic 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.


http://blogs.ei.columbia.edu/
Census Drama

• Mauritania attempts Census in 2011 but is blocked by protest from minority groups who claim they are not being enumerated

Standoff in Mauritania – West Africa.

The population census engaged since June 2011 in Mauritania has poured a lot of ink and created a lot of talk throughout the months in our Country. Today, it is pouring the blood of Mauritanian youth in the different cities of the country.

Since the beginning, the process has been, by all observers, qualified as unjust and unrepresentative of the interest of all the components of the population. The Black populations, in the light of the discrimination they have been submitted to all the time, have seen it coming. The first citizens who went to the census centers to enroll faced a barrage of questions that are supposed to determine their “mauritanity”. The issue: only Black people are asked to prove their citizenship while the Moors (Arabs) have just to show their “skin”. I am speaking literally showing their skin, speaking hassanya and the doubt is cleared about their citizenship.
Census Drama

• Chile releases Census in 2013 which is later withdrawn due to serious errors

Chilean President Apologizes for Census

By PASCALE BONNEFOY
Published: August 9, 2013

SANTIAGO, Chile — President Sebastián Piñera’s administration had promised the “best census ever.”

Previous censuses were held in one day, with volunteers. But last year, his government allocated $32 million to the task and introduced vast changes, like paying census takers to go door to door over a three-month span. New questions asked about gay couples who lived together, knowledge of other languages, the use of the Internet and recycling.

Yet this week, the president was forced to apologize to the country for serious mistakes in the nationwide census, which, despite the intense effort, failed to survey over 9 percent of the population.
GPW4 Population Density 2010