Taking Advantage of the Improved Availability of Census Data:

A First Look at the Gridded Population of the World, Version 4 (GPWv4)

Erin Doxsey-Whitfield

Kytt MacManus

Susana B. Adamo

Linda Pistolesi
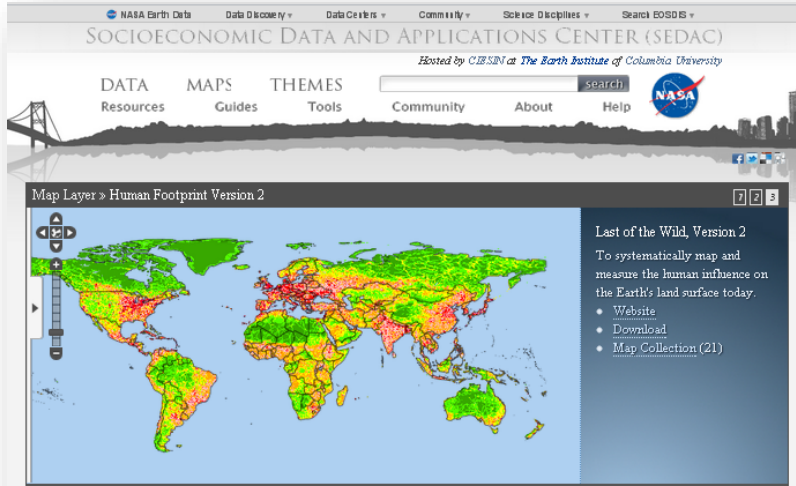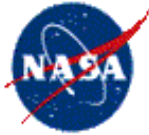
John Squires

Olena Borkovska

Sandra Baptista

October 17, 2014
Applied Geography Conference
Atlanta, Georgia

Center for International Earth Science Information Network
EARTH INSTITUTE | COLUMBIA UNIVERSITY

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
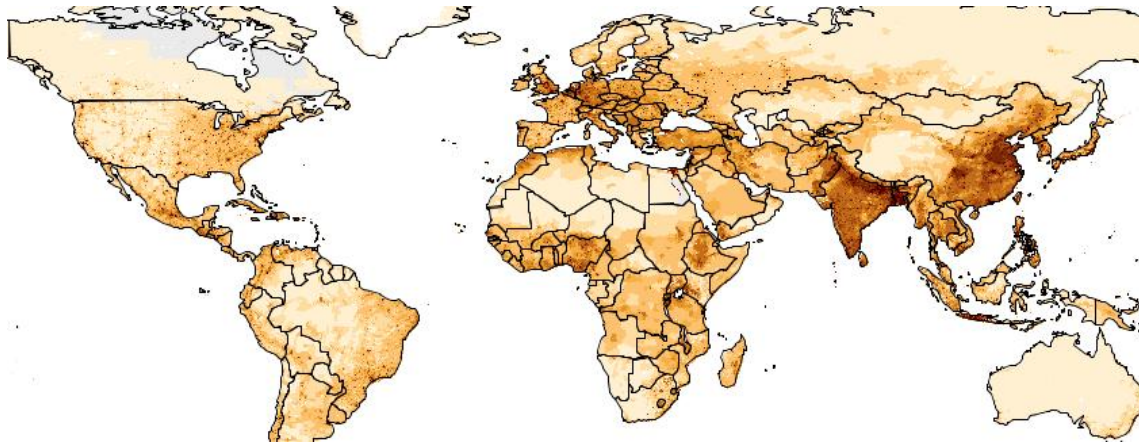
GPW version 3, 2000 population density

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)

	GPWv1	GPWv2	GPWv3	GPWv4
Publication Year	1995	2000	2005	2014/2015
Years of Estimation	1994	1990, 1995	1990, 1995, 2000	2000, 2005, 2010, 2015, 2020
Number of Input Units (subnational geographic units)	19,000	127,000	c. 400,000	~ 12,500,000
Grid Resolution	2.5 arc-minute	2.5 arc-minute	2.5 arc-minute	30 arc-second (1 km)
Census variables	Total Population	Total Population	Total Population	Total Population, Sex, Age, Urban/Rural status

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

Contents lists available at ScienceDirect

Health & Place

journal homepage: www.elsevier.com/locate/healthplace

ELSEVIER

HEALTH & PLACE

Spatial and social factors drive anemia in Congolese women

Jane P. Messina [a,*], Kashamuka Mwandagalirwa [b], Steve M. Taylor [c,d], Michael Emch [e,f], Steven R. Meshnick [c]

CrossMark

Journal of Biogeography (J. Biogeogr.) (2009) **36**, 2264–2278

ORIGINAL ARTICLE

Establishment success of invasive ring-necked and monk parakeets in Europe

Diederik Strubbe* and Erik Matthysen

Environmental
Toxicology and Chemistry

Hazard/Risk Assessment

Worldwide estimation of river concentrations of any chemical originating from sewage-treatment plants using dilution factors

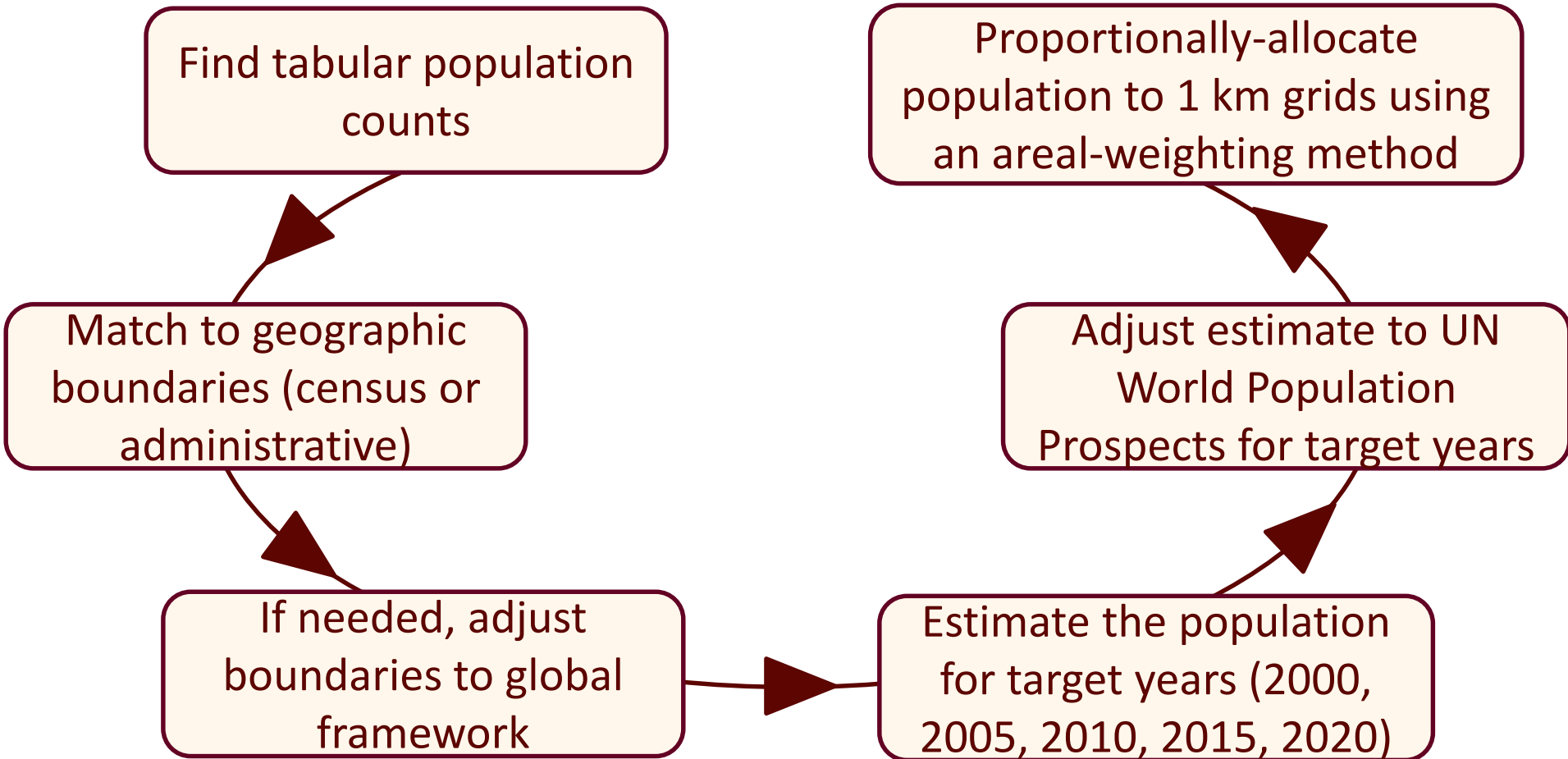Virginie D.J. Keller[1], Richard J. Williams[1], Caryn Lofthouse[2] and Andrew C. Johnson

OPEN

Issue

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

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

Higher resolution boundaries in eastern China lead to more accurate population distributions
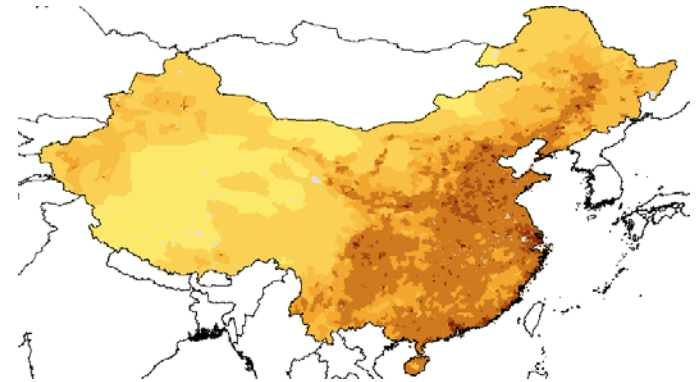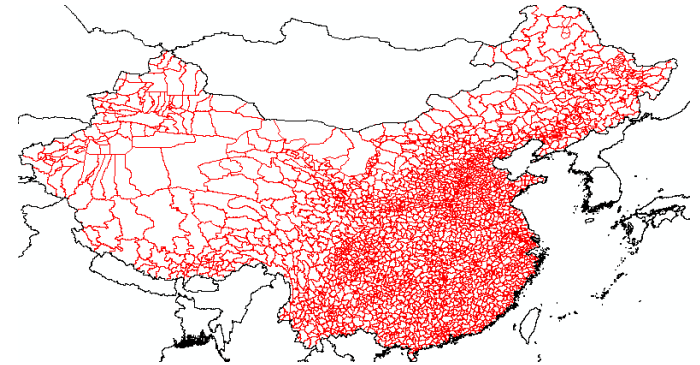
Outline

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

GPWv4 2010 Population Density Estimates
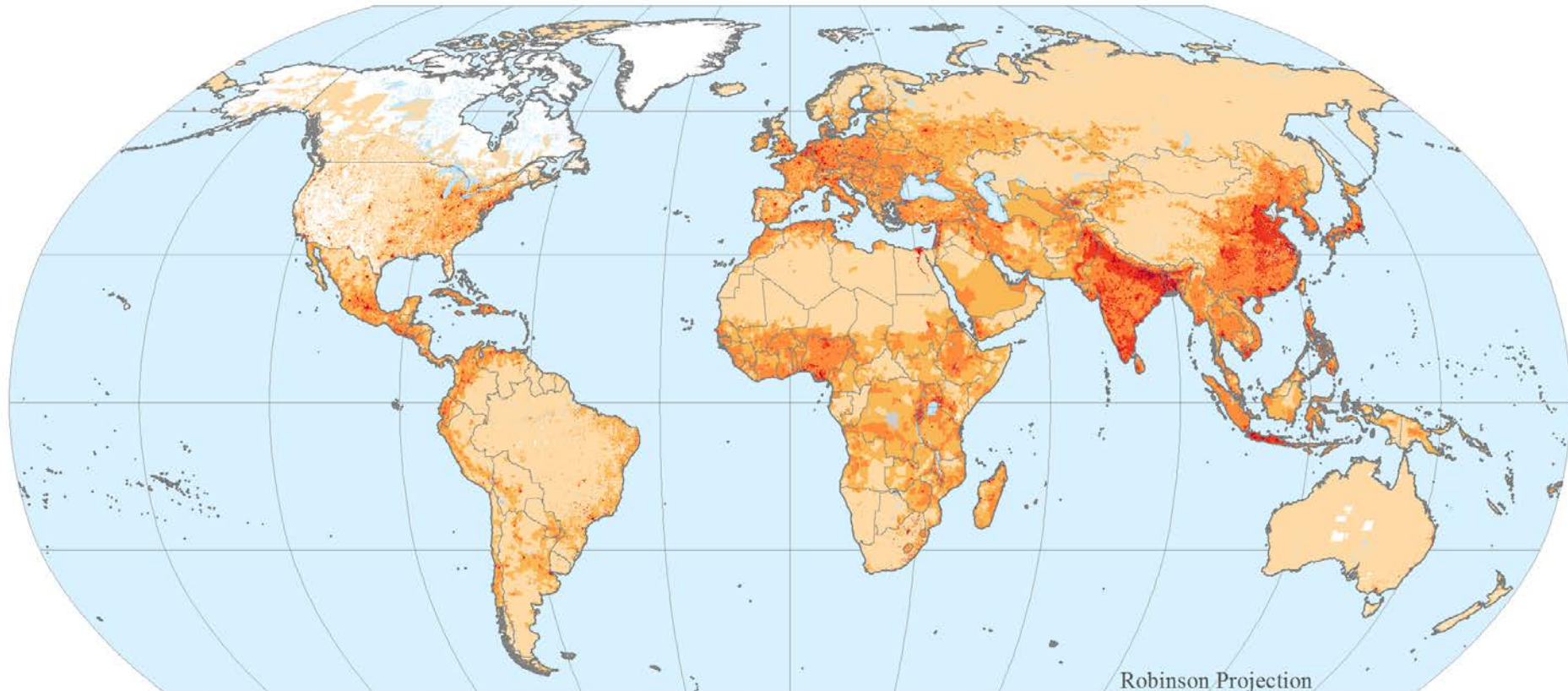
Persons per km²: 0 | 1–4 | 5–24 | 25–249 | 250–1000 | >1000 | No Data

Robinson Projection

Source: CIESIN, 2014

Improvements to GPWv4
1. Increased spatial resolution of input data

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

Number of Input Census Units

	GPWv3	GPWv4
Global	399,747	12,497,563
USA	60,884	10,608,747
Rest of World	338,863	1,888,816

Improvements to GPWv4
1. Increased spatial resolution of input data

Frequency of Administrative Level Used

# of countries at Level 3 or higher:

GPWv3: 47
GPWv4: 87

Country — Census block
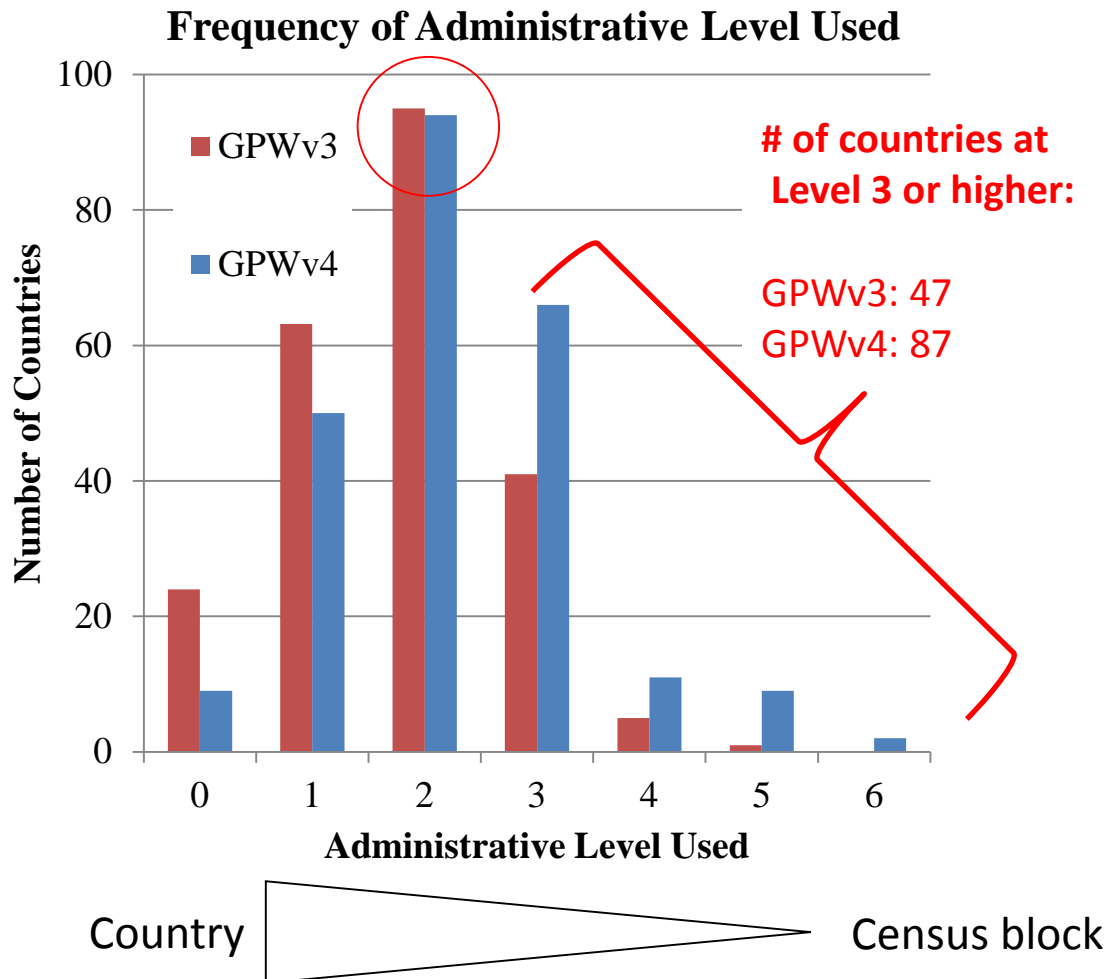
Improvements in Input Data Resolution
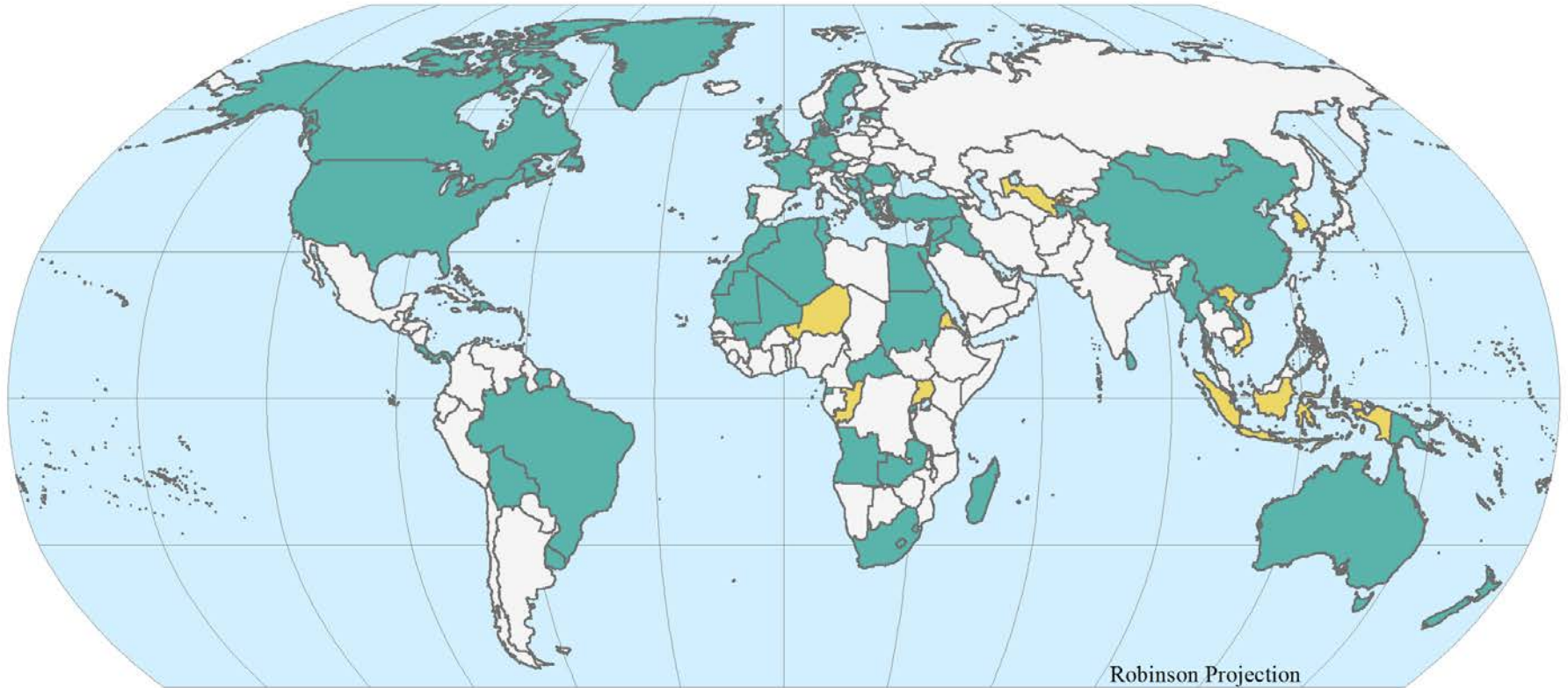
	Lower resolution than GPWv3		Same as GPWv3		Higher resolution than GPWv3
	11 countries		132 countries		98 countries

GPWv4 Administrative Level

Robinson Projection
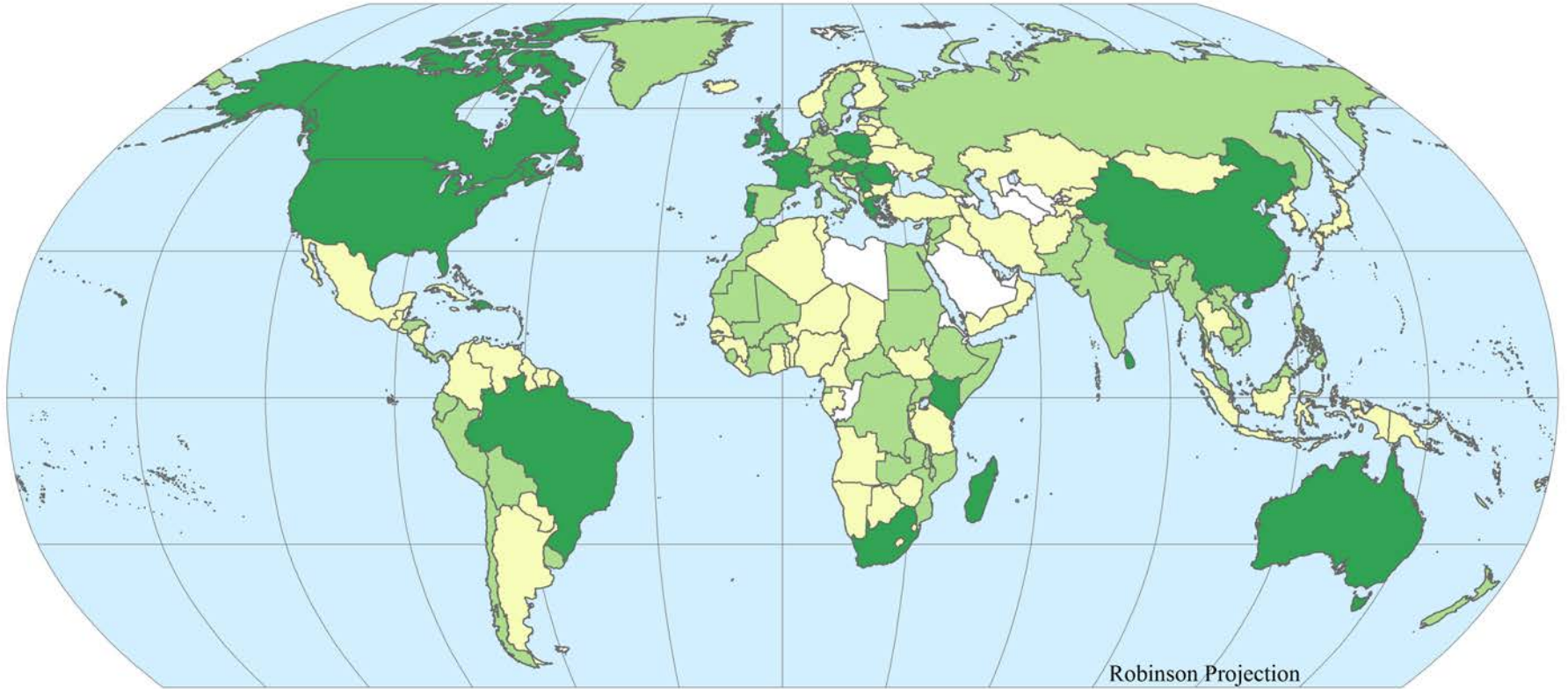
Level 0–1 ☐ Level 2 ☐ Level 3 ☐ Level 4–6 ☐

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
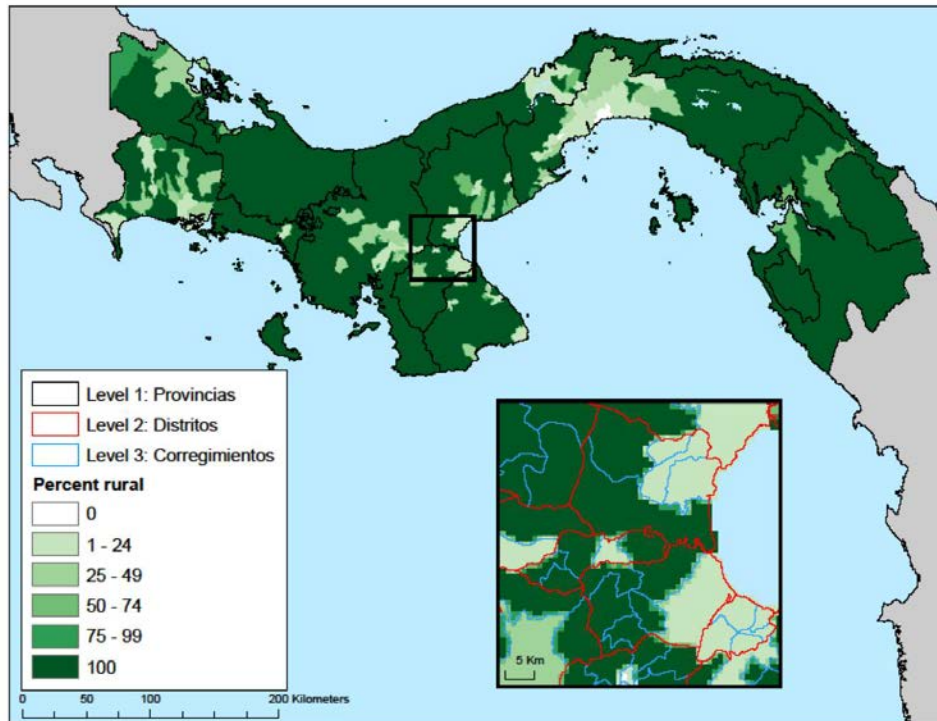
2010 % urban population, Panama

Improvements to GPWv4
2. Additional census variables

Adolescent Literacy in Nigeria, 2006
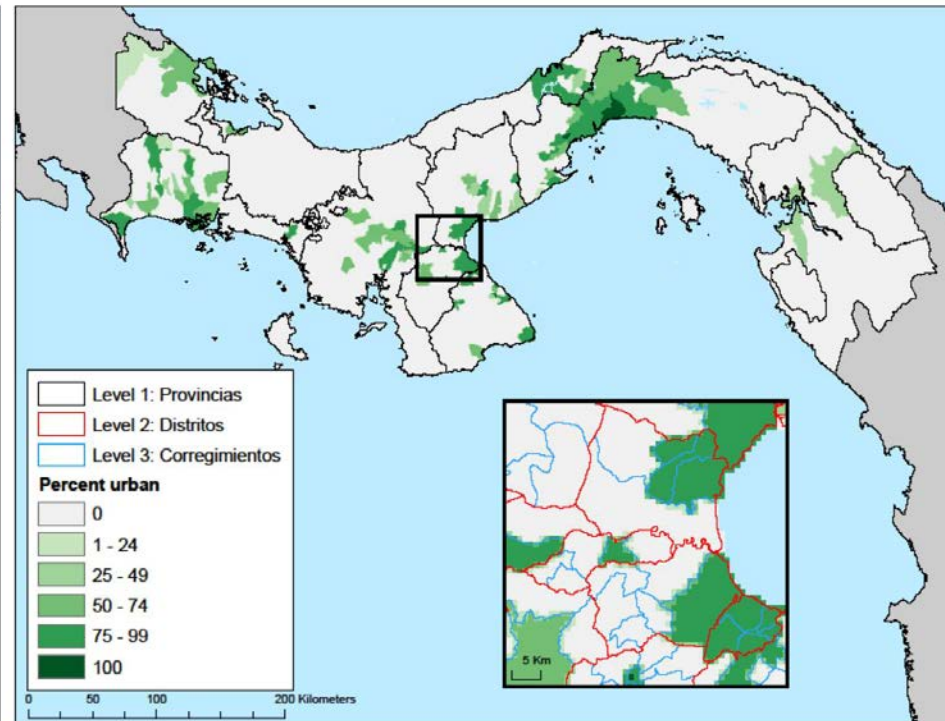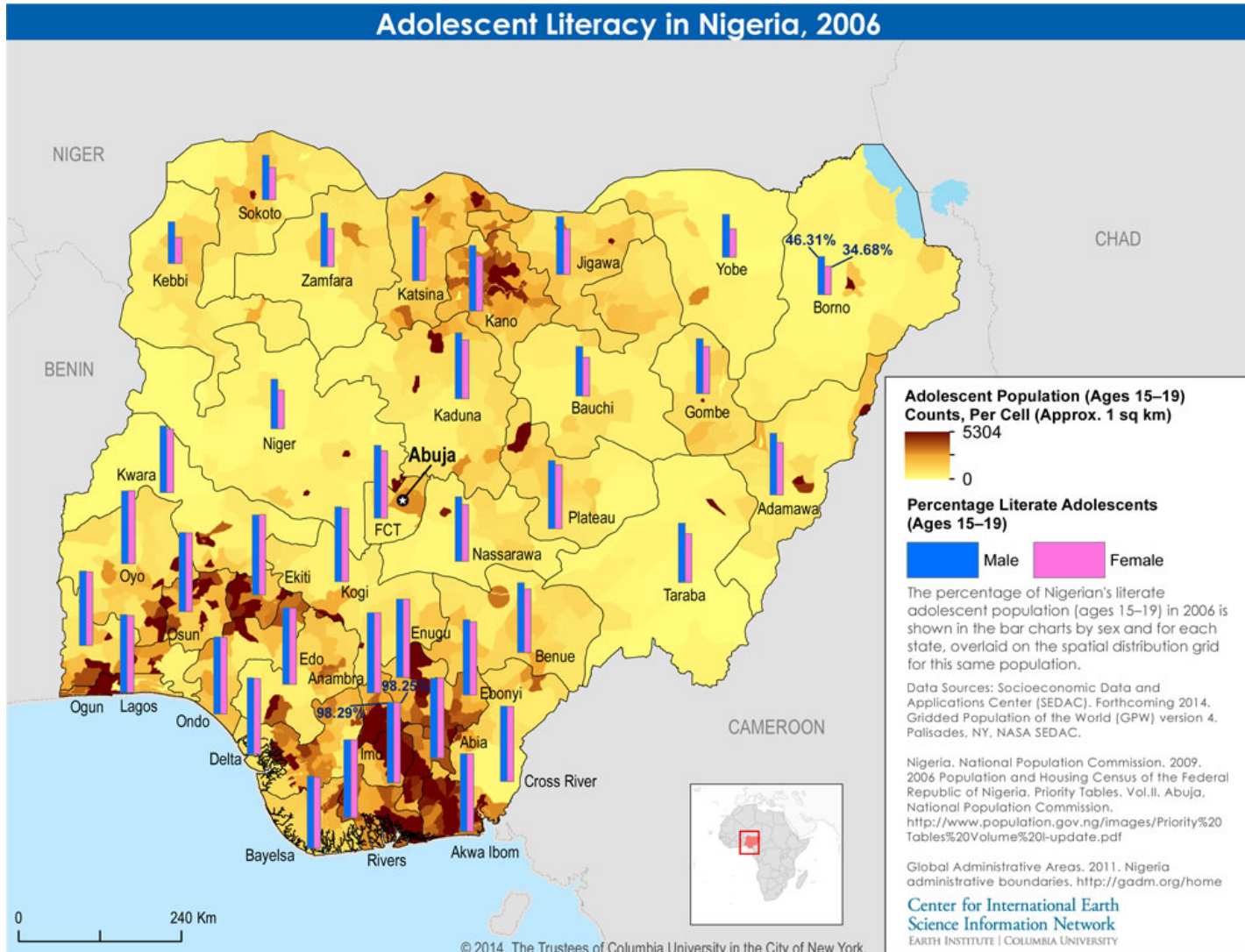
Reasons for Improvements

- 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

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

Source: 2011/2012 survey for the review of the 2010 World Programme on Population and Housing Censuses; UN Statistics Division, 2013.

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

Robinson Projection

Persons per km² □ 0   1–4   5–24   25–249   250–1000   >1000   □ No Data

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