

Gridded Population of the World Version 4 (GPWv4)

Greg Yetman
Chris Small
Erin Doxsey-Whitfield
Kytt MacManus

May 28, 2015
Global Urbanization Workshop

<http://sedac.ciesin.columbia.edu>

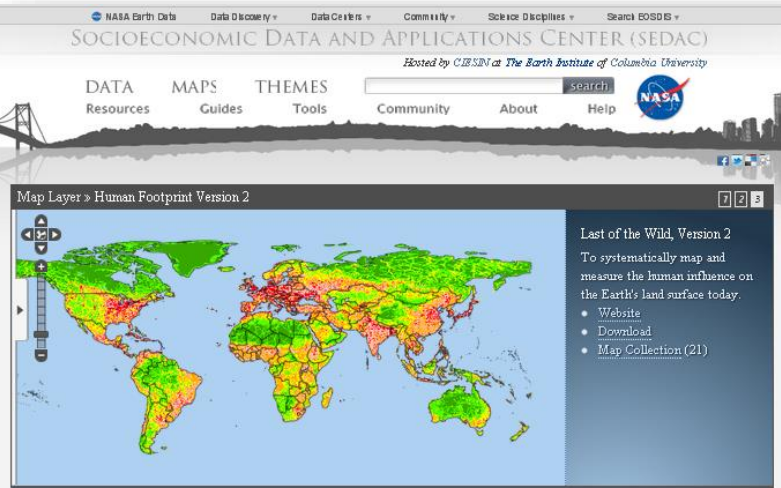
Center for International Earth Science Information Network

EARTH INSTITUTE | COLUMBIA UNIVERSITY

Focus on applying *information technology*
to pressing *interdisciplinary data, information, and research problems*
related to *human interactions in the environment*



Socioeconomic Data & Applications Center (SEDAC)



Featured Uses

- Another Way to Look at an Air Quality Problem
- World Resources Report: Decision Making in a Changing Climate
- We are Seven Billion
- Populations in Proximity to Nuclear Power Plants

News

- Earthquake Experts from China
- Treaty Multiple Status Matrix - Netscape 6
- ICA Ads and
- Man Peak Cob
- Que Sp

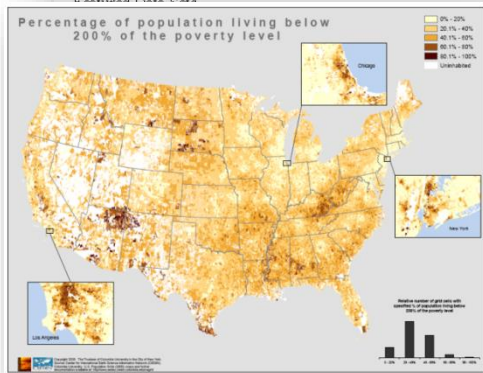
Environmental Treaties and Resource Indicators

Status Matrix

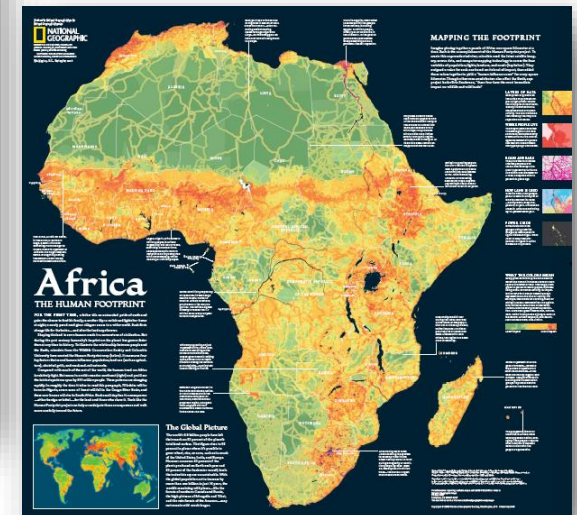
download this matrix | treaty results set | country results set

View in matrix: All Relevant Countries | Display this variable: Status

	Convention concerning the Protection of Alps	Protocol for the implementation of the Alpine Convention in the field of mountain agriculture	Protocol for the implementation of the Alpine Convention in the field of nature protection and landscape conservation	Protocol for the implementation of the Alpine Convention in the field of town and country planning and sustainable development
<input type="checkbox"/> Austria	Party
<input type="checkbox"/> EC European Communities	Party	Signatory	Signatory	Signatory
<input type="checkbox"/> France	Party	Signatory	Signatory	Signatory
<input type="checkbox"/> Germany	Party	Signatory	Signatory	Signatory
<input type="checkbox"/> Italy	Signatory	Signatory	Signatory	Signatory
<input type="checkbox"/> Liechtenstein	Party	Signatory	Signatory	Signatory
<input type="checkbox"/> Monaco	...	Signatory	Signatory	Signatory

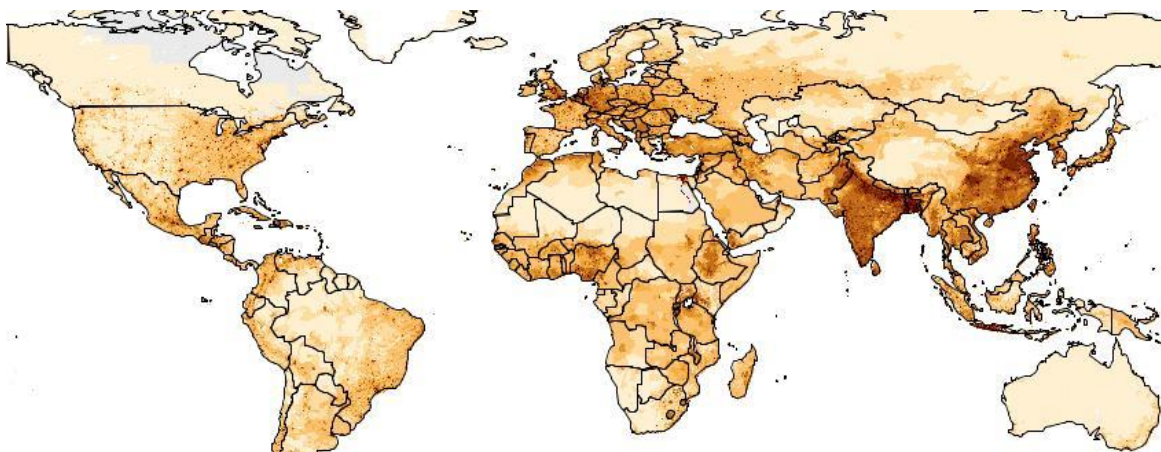


- Focus on human dimensions of environmental change
- Integration of social and earth science data, especially with remote sensing
- Direct support to scientists, applied and operational users, decision makers, and policy communities

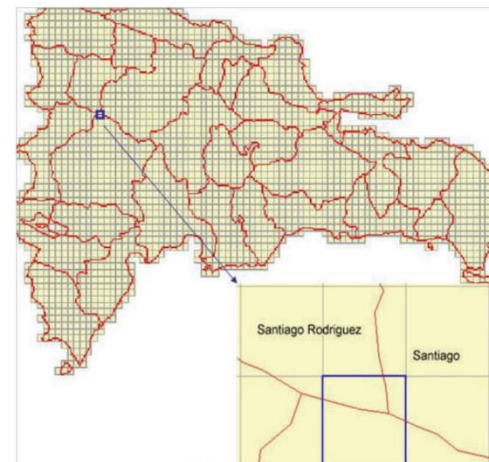


Gridded Population of the World

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



Transforming census units to a grid

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)
- Produced by Waldo Tobler, Uwe Deichmann, Jon Gottsegen, and Kelly Maloy at the UC Santa Barbara

World Population in a Grid of Spherical Quadrilaterals

Waldo Tobler^{1*}, Uwe Deichmann², Jon Gottsegen¹, and Kelly Maloy³

¹ National Center for Geographic Information and Analysis, Geography Department, University of California, Santa Barbara, CA 93106-4060, USA

² Statistical Division of the United Nations, DC2-1764, New York, NY 10017, USA

³ ESRI, Redlands, CA 92373, USA

ABSTRACT

We report on a project that converted subnational population data to a raster of cells on the earth. We note that studies using satellites as collection devices yield results indexed by latitude and longitude. Thus it makes sense to assemble the terrestrial arrangement of people in a compatible manner. This alternative is explored here, using latitude/longitude quadrilaterals as bins for population information. This format also has considerable advantages for analytical studies. Ways of achieving the objective include, among others, simple centroid sorts, interpolation, or gridding of polygons. The results to date of putting world boundary coordinates together with estimates of the number of people are described. The estimated 1994 population of 219 countries, subdivided into 19,032 polygons, has been assigned to over six million five minute by five minute quadrilaterals covering the world. These results are available over the Internet. The grid extends from latitude 57°S to 72°N, and covers 360° of longitude. Just under 31% of the (1548 by 4320) grid cells are populated. The number of people in these countries is estimated to be 5.6 billion, spread over 132 million km² of land. Extensions needed include continuous updating, additional social variables, improved interpolation methods,

correlation with global change studies, and more detailed information for some parts of the world. © 1997 John Wiley & Sons, Ltd.

Received 10 August 1996; revised 30 November 1996; accepted 16 December 1996
Int. J. Popul. Geogr. 3, 203-225 (1997)
No. of Figures: 5 No. of Tables: 1 No. of Refs: 58

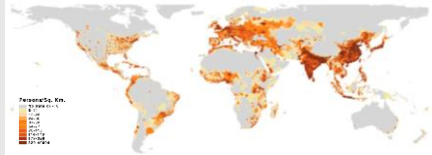
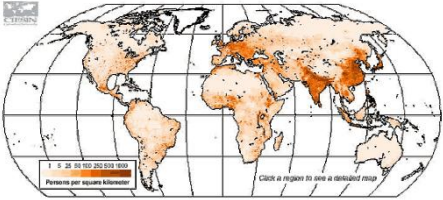
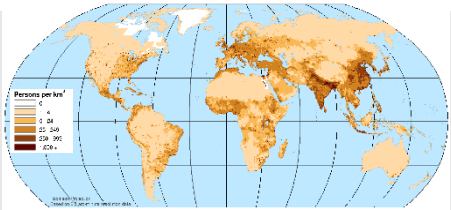

Keywords: world population; raster; five-minute quadrilaterals

INTRODUCTION

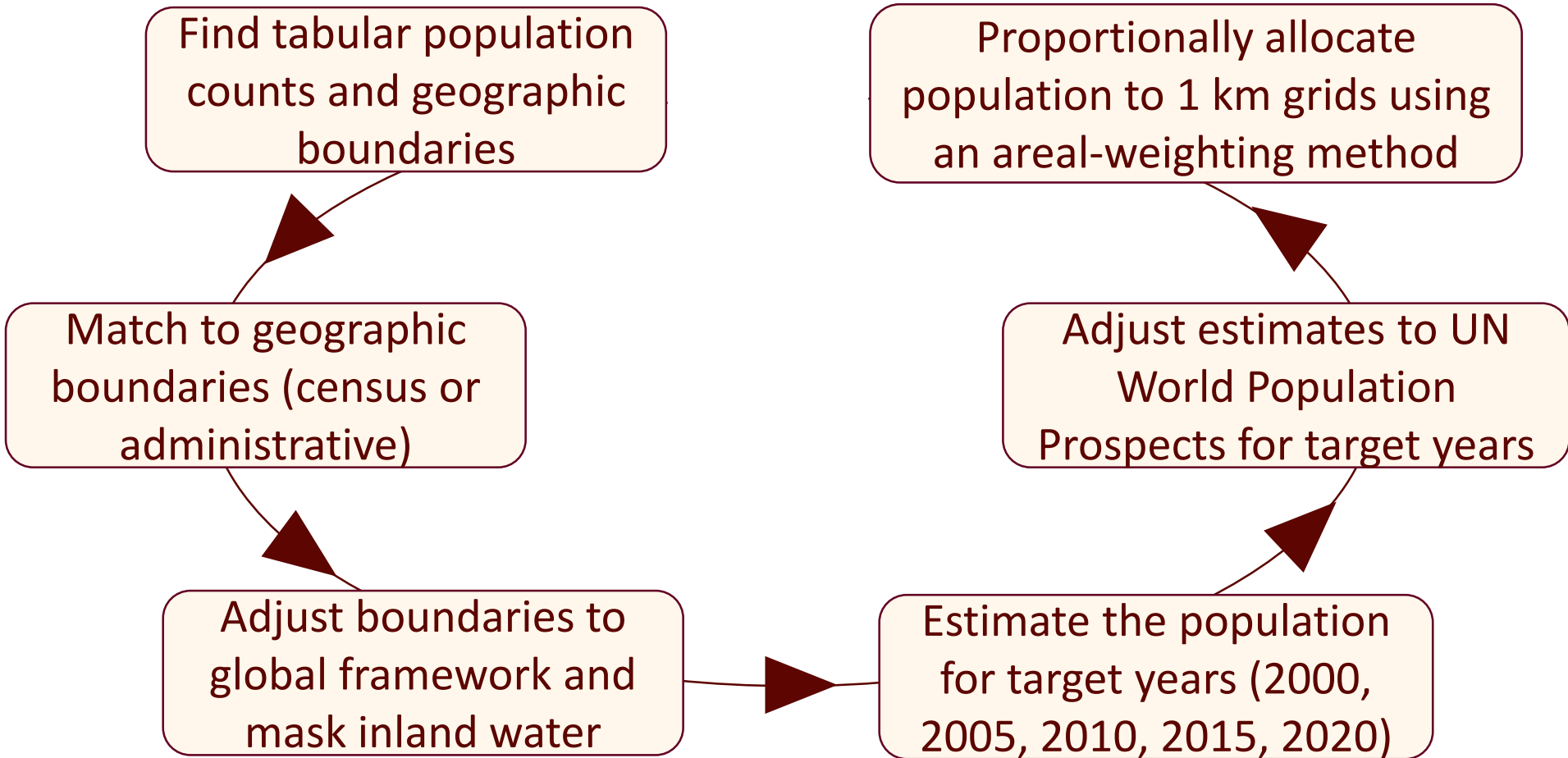
Information on the world's population is usually provided on a national basis. But we know that countries are ephemeral phenomena, and administrative partitionings of a country are irrelevant to much scientific work. As an alternative scheme one might consider ecological zones rather than nation states, yet there is no agreement as to what these zones should be. By way of contrast global environmental studies using satellites as collection devices yield results indexed by latitude and longitude. Thus it makes sense to assemble information on the terrestrial arrangement of people in a compatible manner. A recent pilot study demonstrated some practical advantages of gridded population data (Clarke and Rhind, 1992), including reporting the potential impact of sea level rise on inhabitants of the coastal region of a Scandinavian country. The project described here extends the compilation to much of the entire Earth, using latitude/longitude quadrilaterals as bins for population information. In addition to its compatibility with environmental information this data format has considerable advantages for analytical studies, and the data

* Correspondence to: W. Tobler.
E-mail: tobler@Geog.VCSB.edu
Contract/grant sponsor: California Space Institute.
Contract/grant sponsor: CIESIN through NASA grant; Contract/grant number: NAG1W-2901.
Contract/grant sponsor: ESRI.
Contract/grant sponsor: NCGIA.

Development of GPW

	Publication Year	Years of Estimation	Grid Resolution	Number of Input Units (subnational geographic units)	Census variables	Population Density Grid
GPWv1	1995	1994	5 arc-minutes	19,000	Total Population	
GPWv2	2000	1990, 1995	2.5 arc-minutes	127,000	Total Population	
GPWv3	2005	1990, 1995, 2000	2.5 arc-minutes	~ 400,000	Total Population	
GPWv4	2015	2000, 2005, 2010, 2015, 2020	30 arc-seconds (1 km)	~ 12,500,000	Total Population, Sex, Age, Urban/Rural status	

Basic Workflow Diagram



Census Statistics

- Although tabular census data is available in most countries, the format and structure of the data varies.

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%

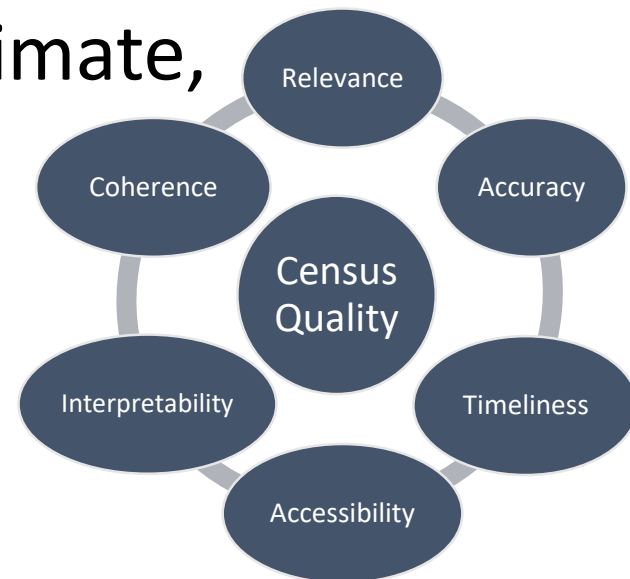
Source: 2011/2012 survey for the review of the 2010 World Programme on Population and Housing Censuses; UN Statistics Division, 2013.

Acquisition Methods in GPW v4

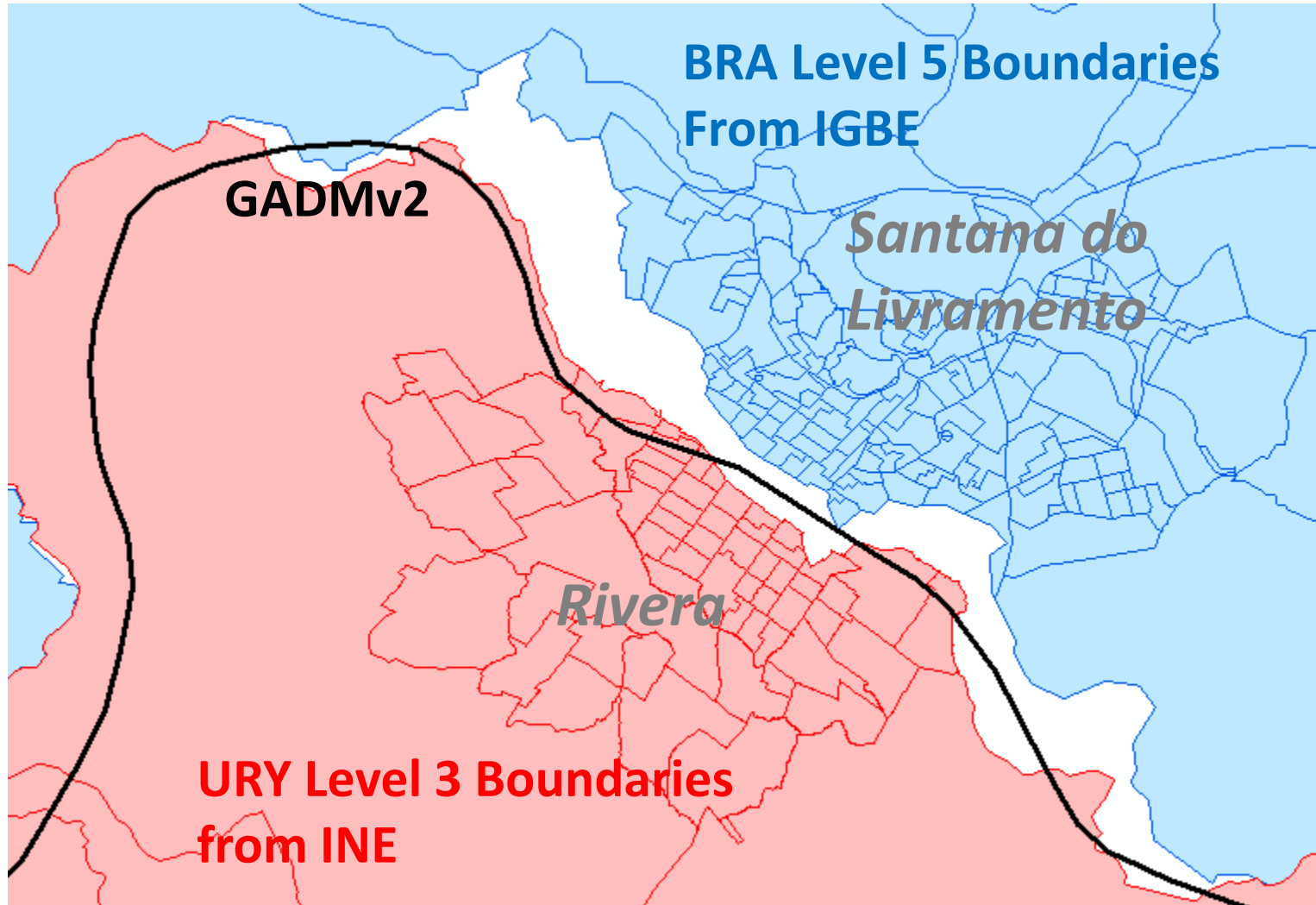
1. Many countries were acquired through basic web research
2. Other countries required more intensive personal communications (email, phone)
3. Still others could not be acquired without the assistance of in-country contacts and networks
4. In some other cases the data were licensed through purchase agreements

Evaluation of Census Data Quality

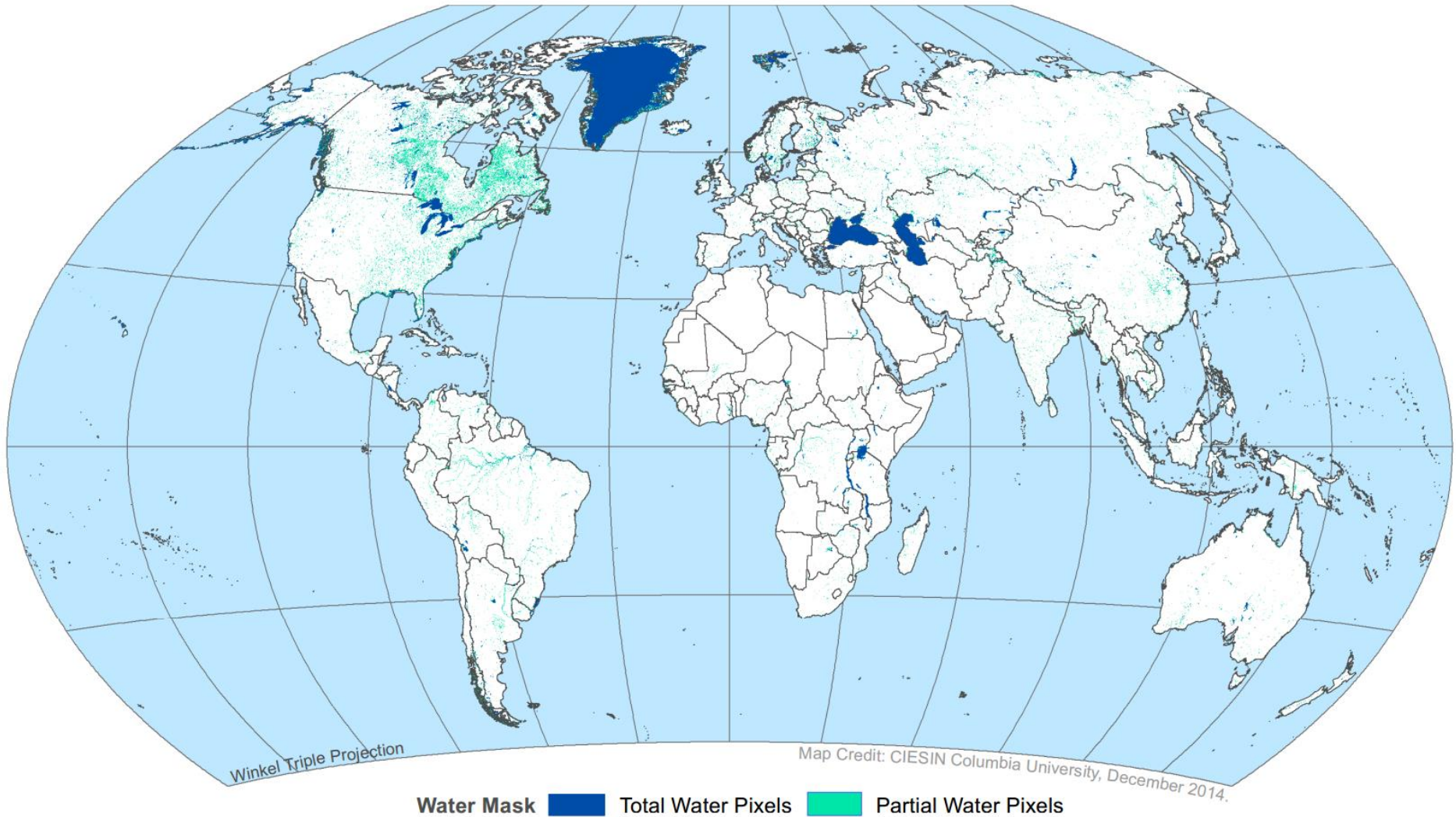
- GPWv4 will include national level metadata and other qualitative indicators which will inform users of:
 1. The currency of the population data
 2. Whether the data is from estimate, preliminary, or final census
 3. Coverage errors evaluated in post-enumeration surveys



Boundary Matching

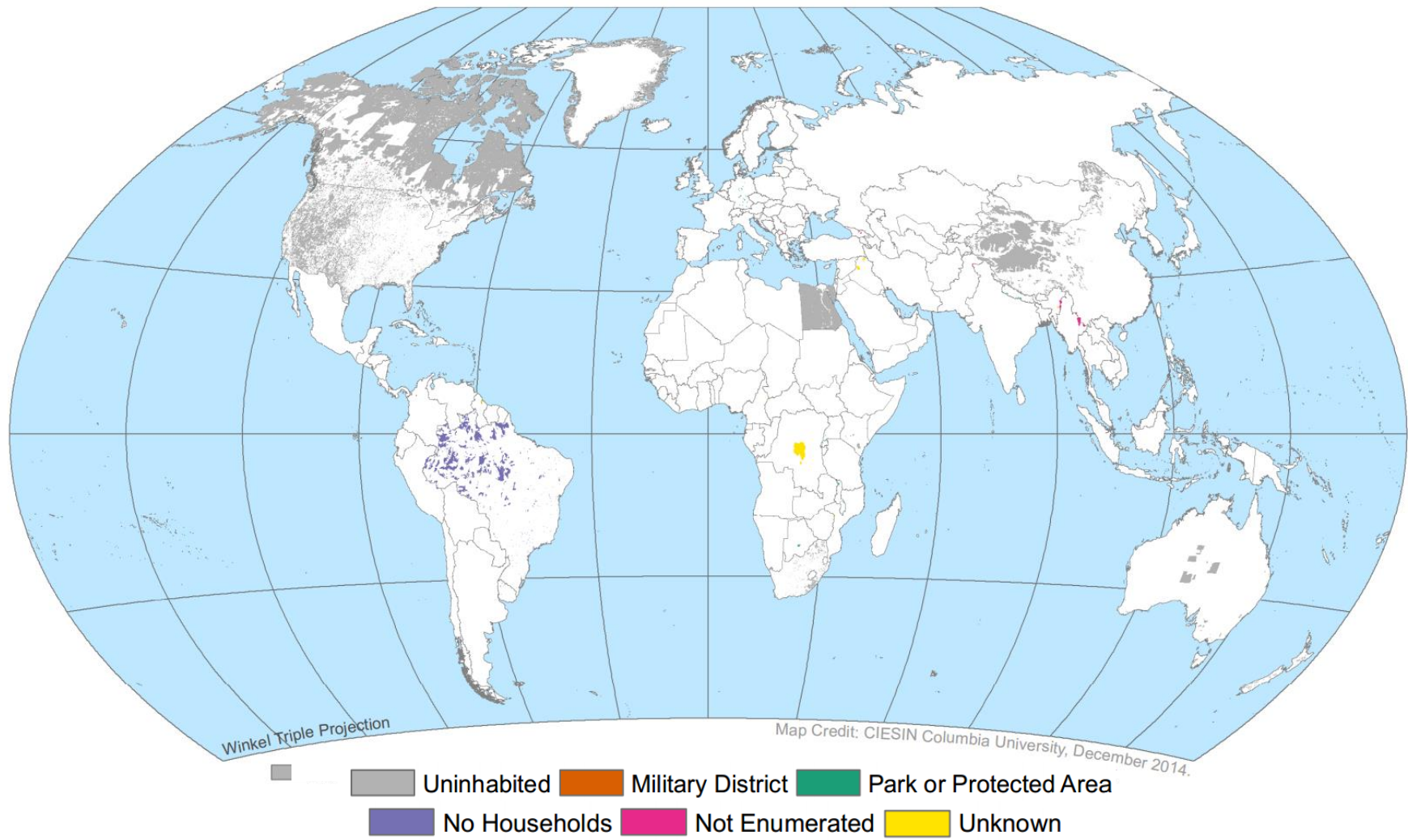


Data Quality Indicators: Water Mask



Distinguishes between pixels that are completely water and/or ice (Total Water Pixels) and pixels that also contain land (Partial Water Pixels).

Data Quality Indicators: Data Context

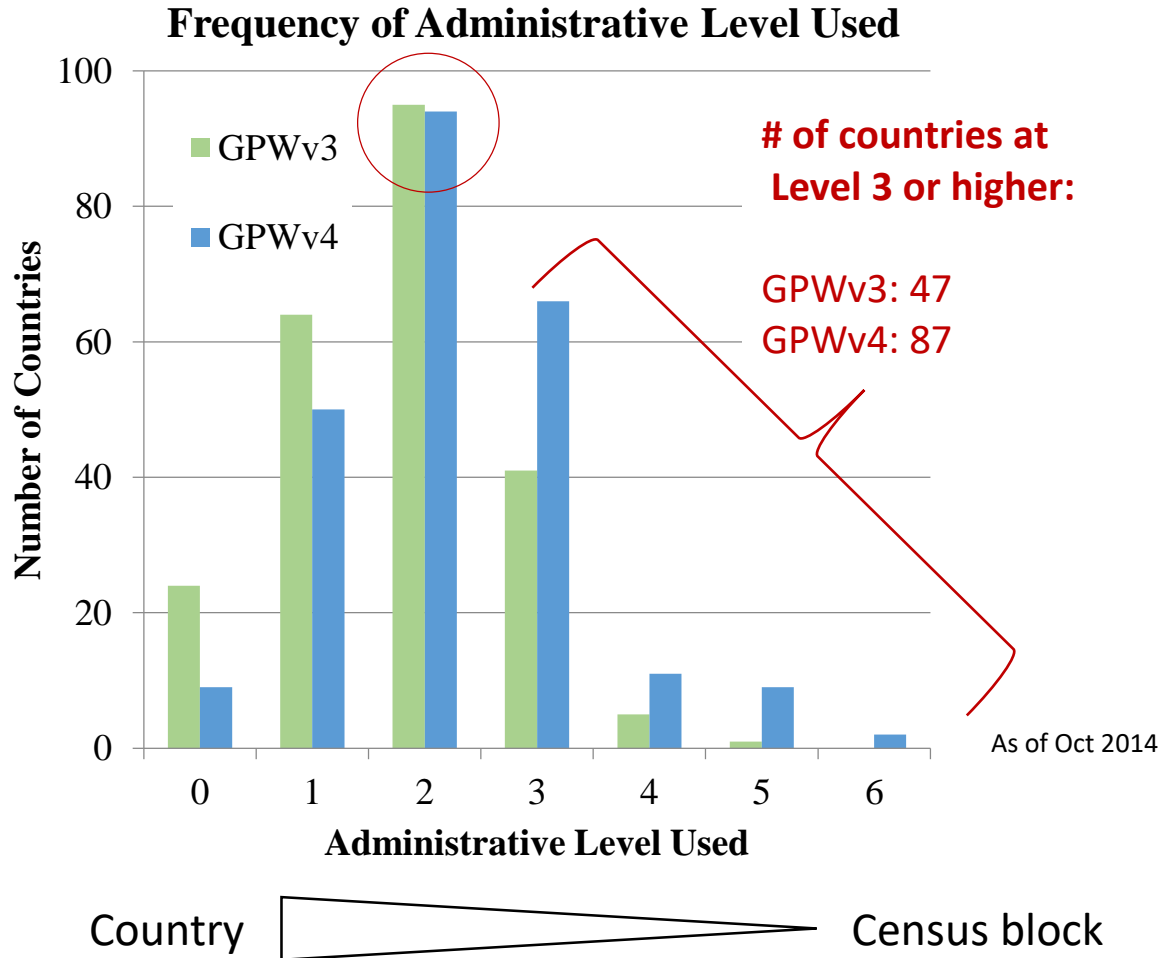


Categorizes pixels with a 0 population estimate in the population count and density grids, based on information included in the census documents.

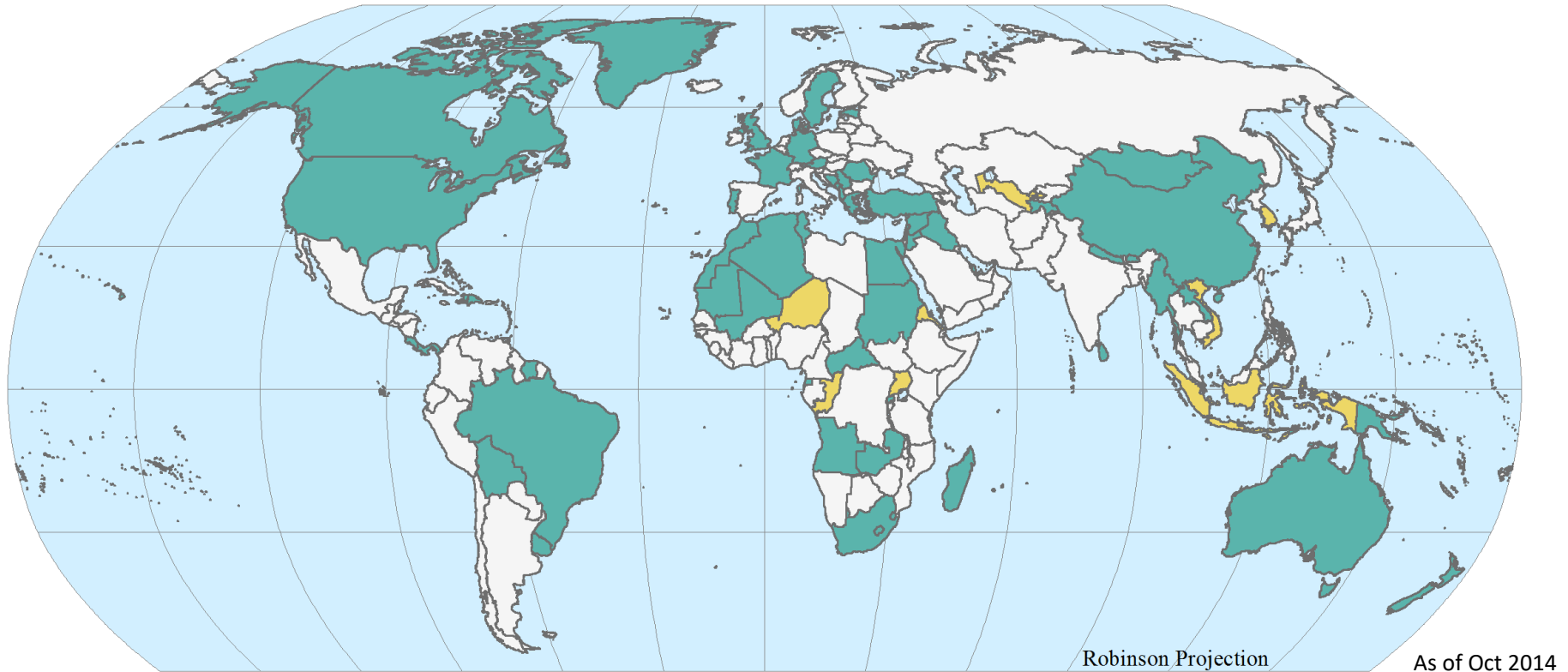
Additional data sets

- Data Quality Indicators: Mean Admin Unit Area
 - The mean unit area data provides a quantitative surface that indicates the size of the input unit(s) from which population count and density grids are derived.
- State/Territory Identifier Grid
 - The identifier boundary data are derived from the pixels as polygons and thus have rectilinear boundaries at large scale.
- Centroids
 - Estimates of human population counts and densities for the years 2000, 2005, 2010, 2015, and 2020 by administrative unit centroid location.

Increased resolution of input census units





Improvements in Input Data Resolution




Robinson Projection

As of Oct 2014

 Lower resolution than GPWv3
11 countries

 Same as GPWv3
132 countries

 Higher resolution than GPWv3
98 countries

Addition of gridded census variables

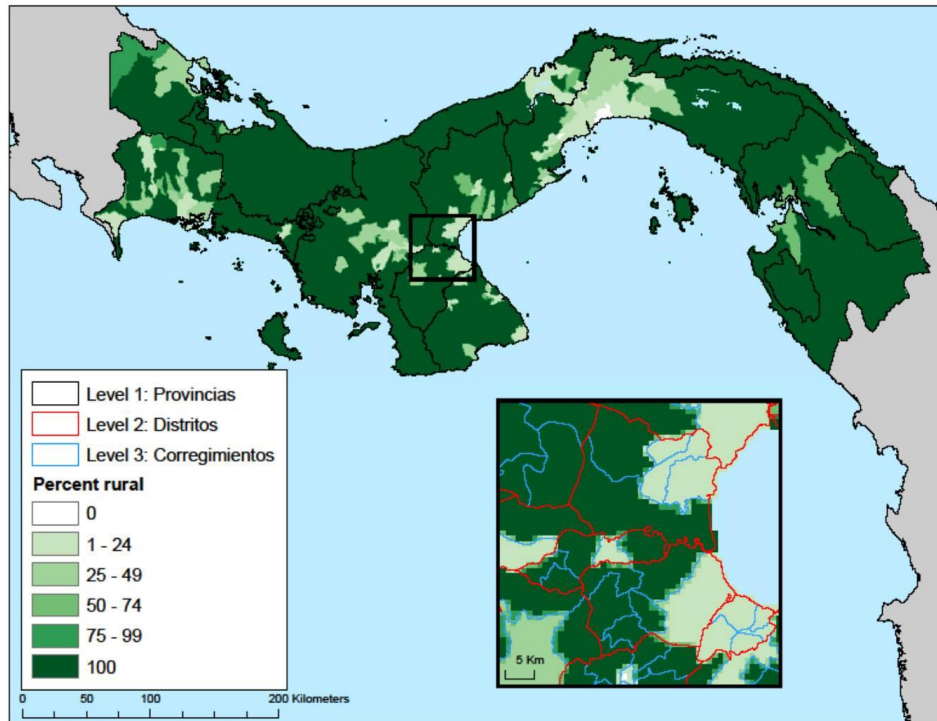
Demand to include demographic information in global population grids

GPWv4 will include detailed global grids for:

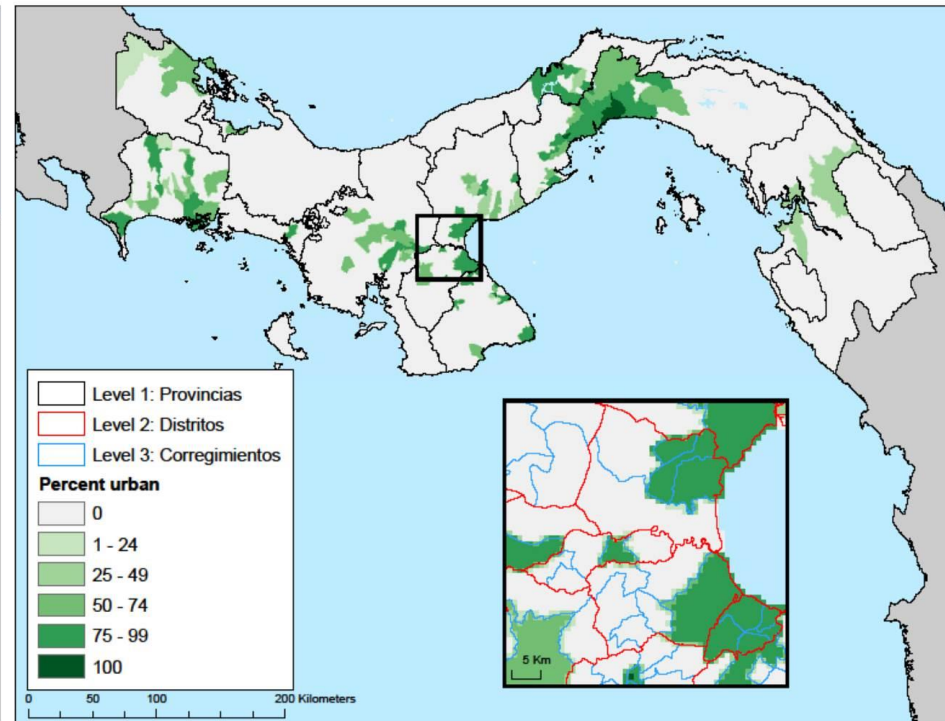
- Sex
- Age 5-year age groups
- Urban/Rural status

Addition of gridded census variables

% Rural population, 2010, Panama

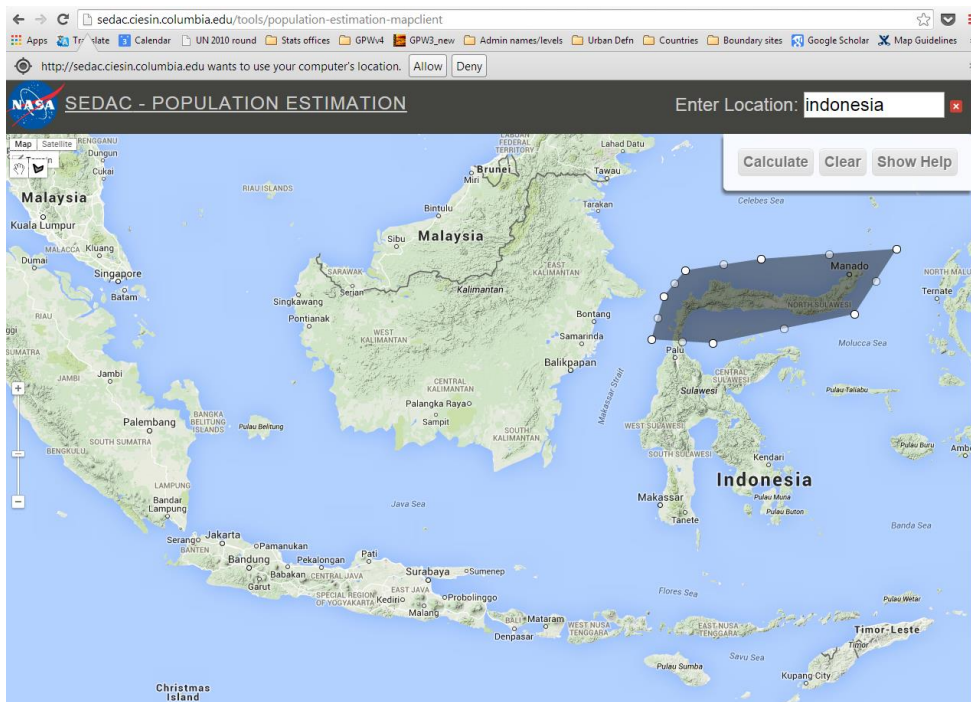


% Urban population, 2010, Panama



SEDAC Population Estimation Service

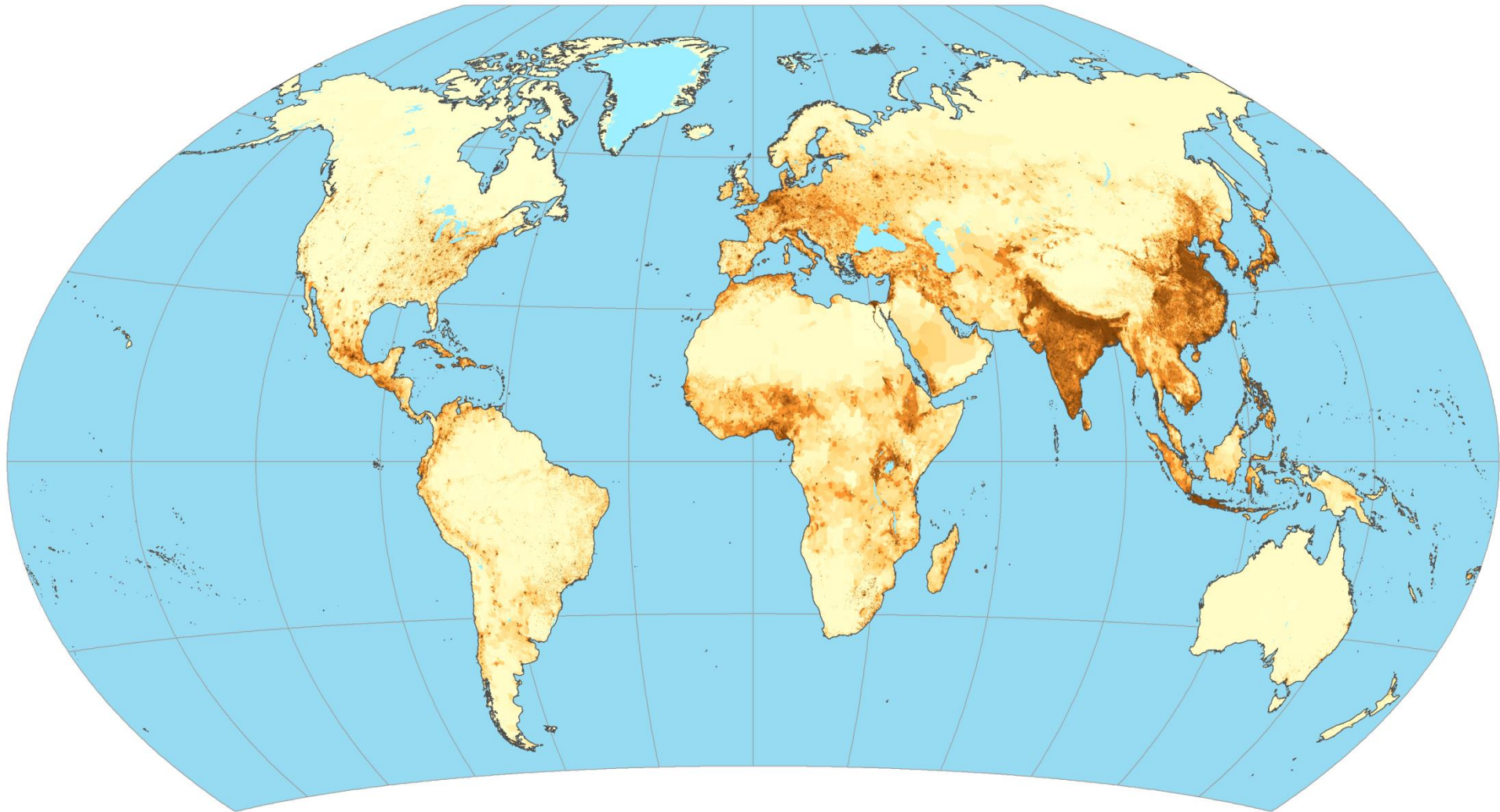
- Web-based tool for estimating population totals and related statistics from a user-defined region
- Uses GPWv3 dataset for the year 2005; will be updated with GPWv4



Population Estimates:

Estimated Population 2005	3,555,146
Cell Count	2,236
Minimum	0
Maximum	125,298
Mean	432
Standard Deviation	2,325
Estimated Land Area	39,729 km ²
Mean Unit Area	89 km ²
Quality Warning	0

GPWv4: Population Density



Additional Modeling: GHSL

- Integration with GHSL via a joint project with JRC
 - Details to follow in 2016

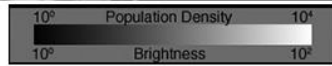
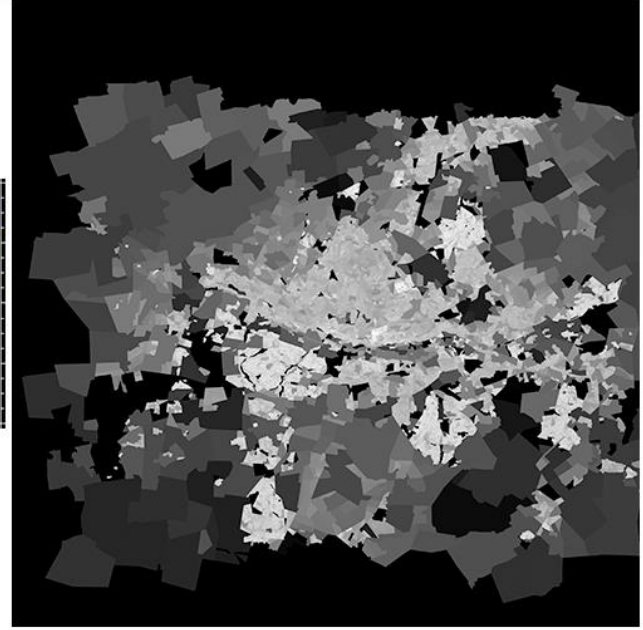
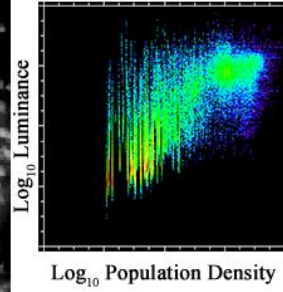
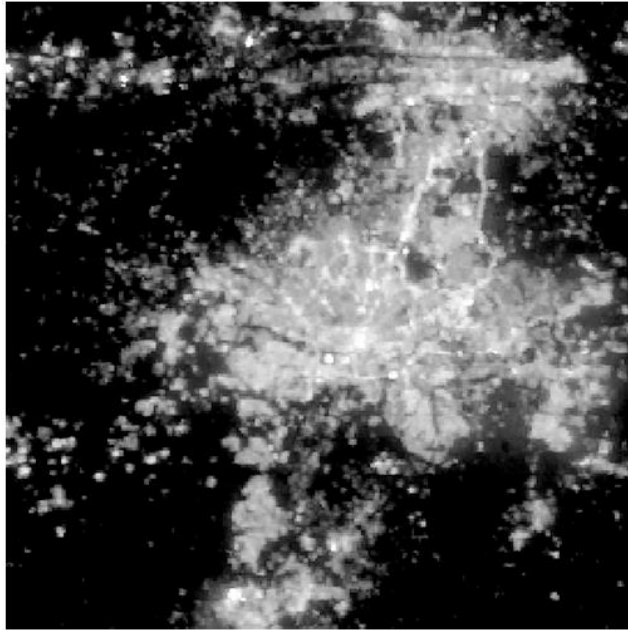
Additional Modeling: VIIRS

- Day-Night band from the joint NASA-JAXA Suomi NPP satellite
- Detects lights with greater sensitivity than DMSP at a finer resolution (~300m)
- Working with beta lights composites
 - Final annual composites of stable lights scheduled for October 2015
- Relationship between high resolution census data (population density) and emitted lights

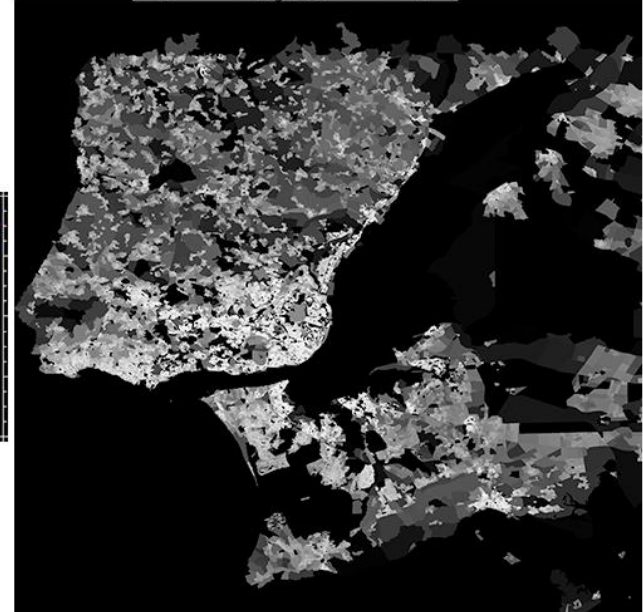
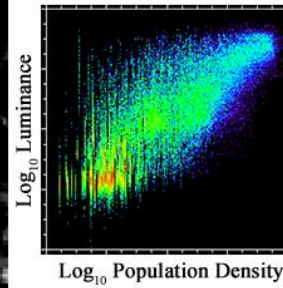
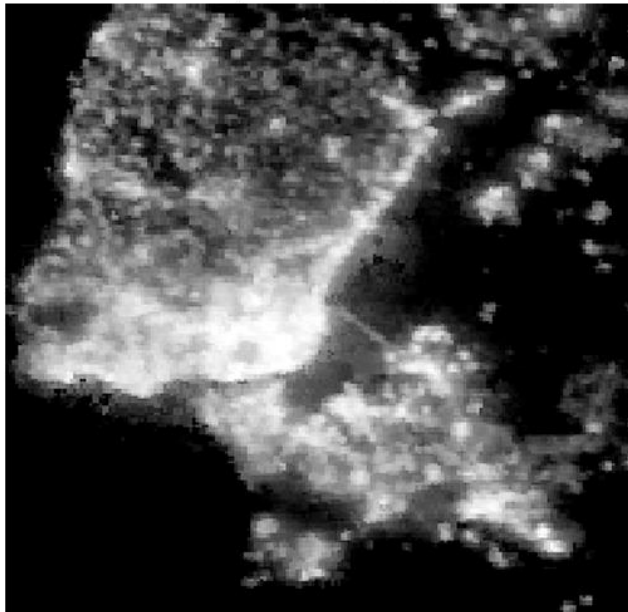
Northeast Corridor



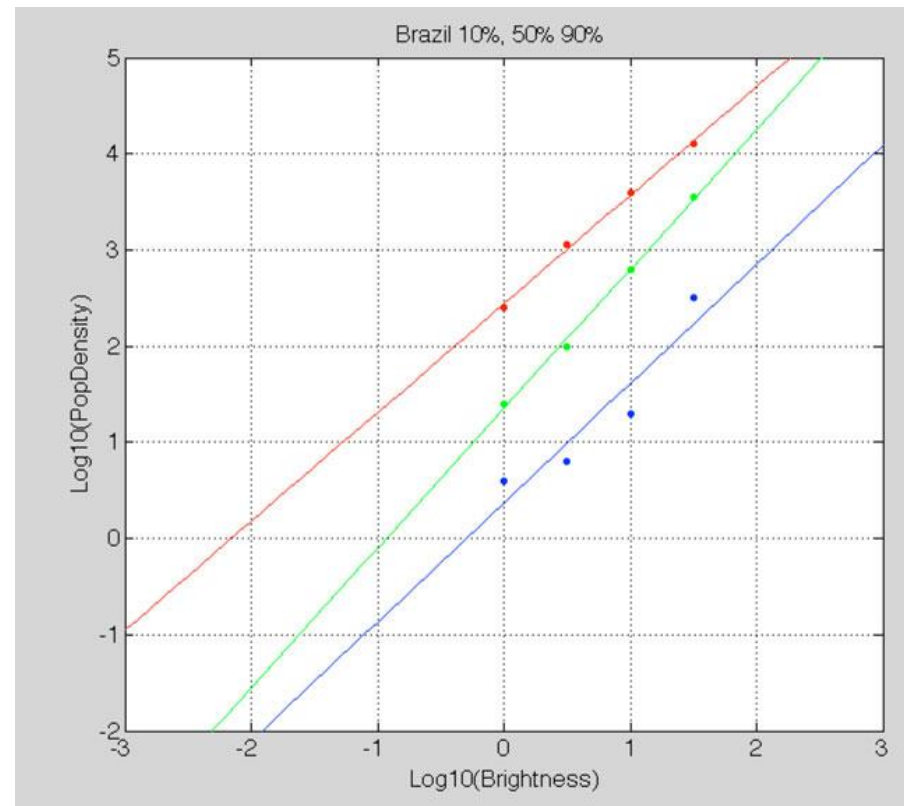
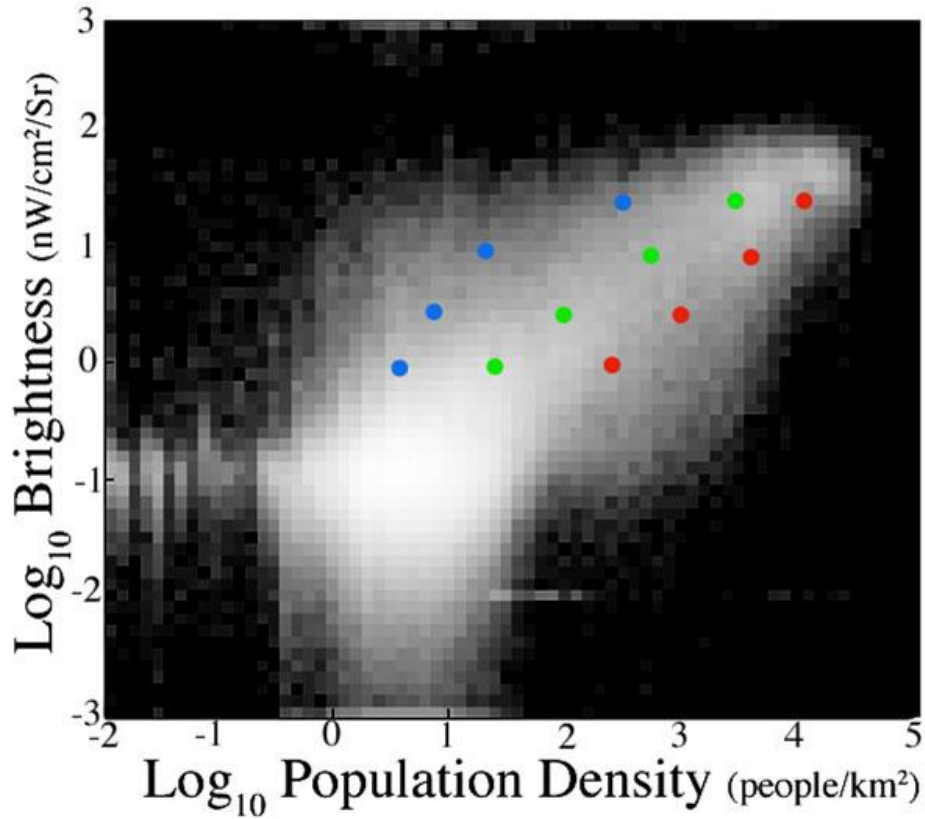
Johannesburg



Lisbon

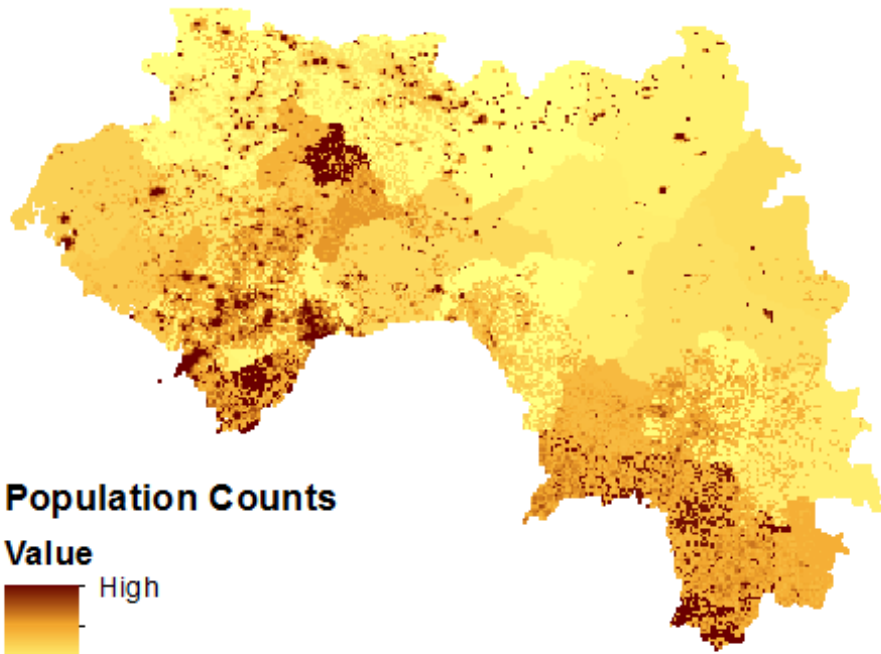


Population-Light Relationship

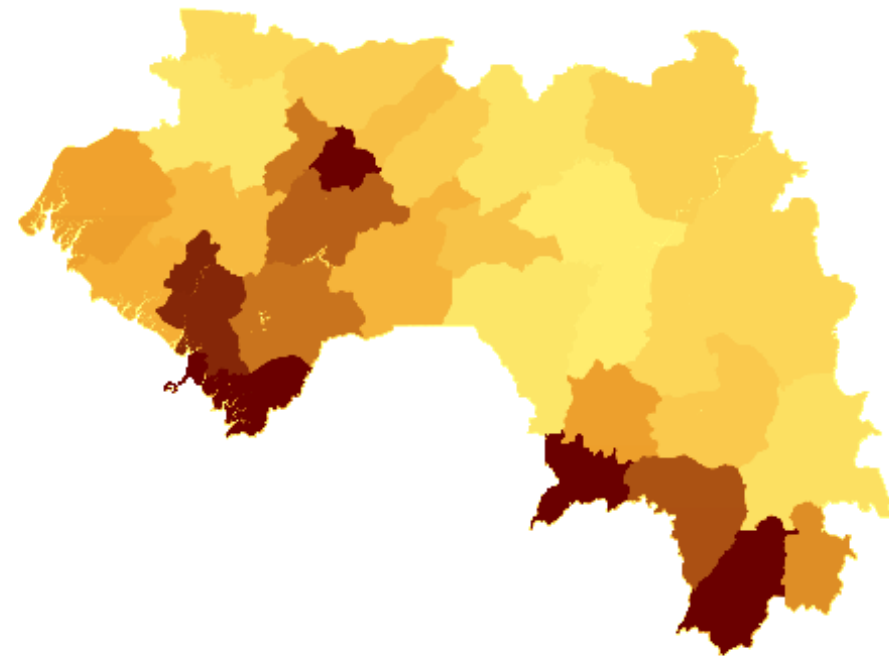


Test Case: Guinea

VIIRS Reallocation

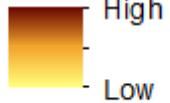


Proportional Allocation



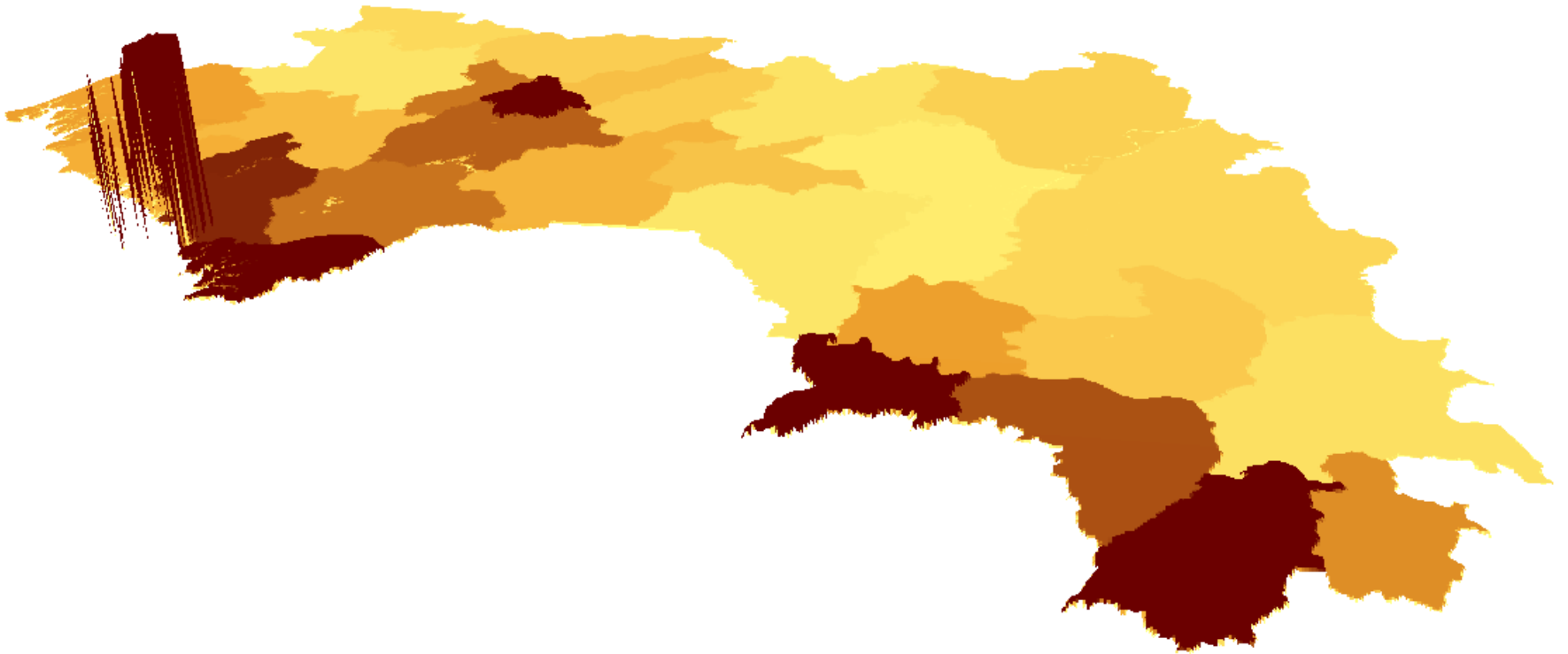
Population Counts

Value



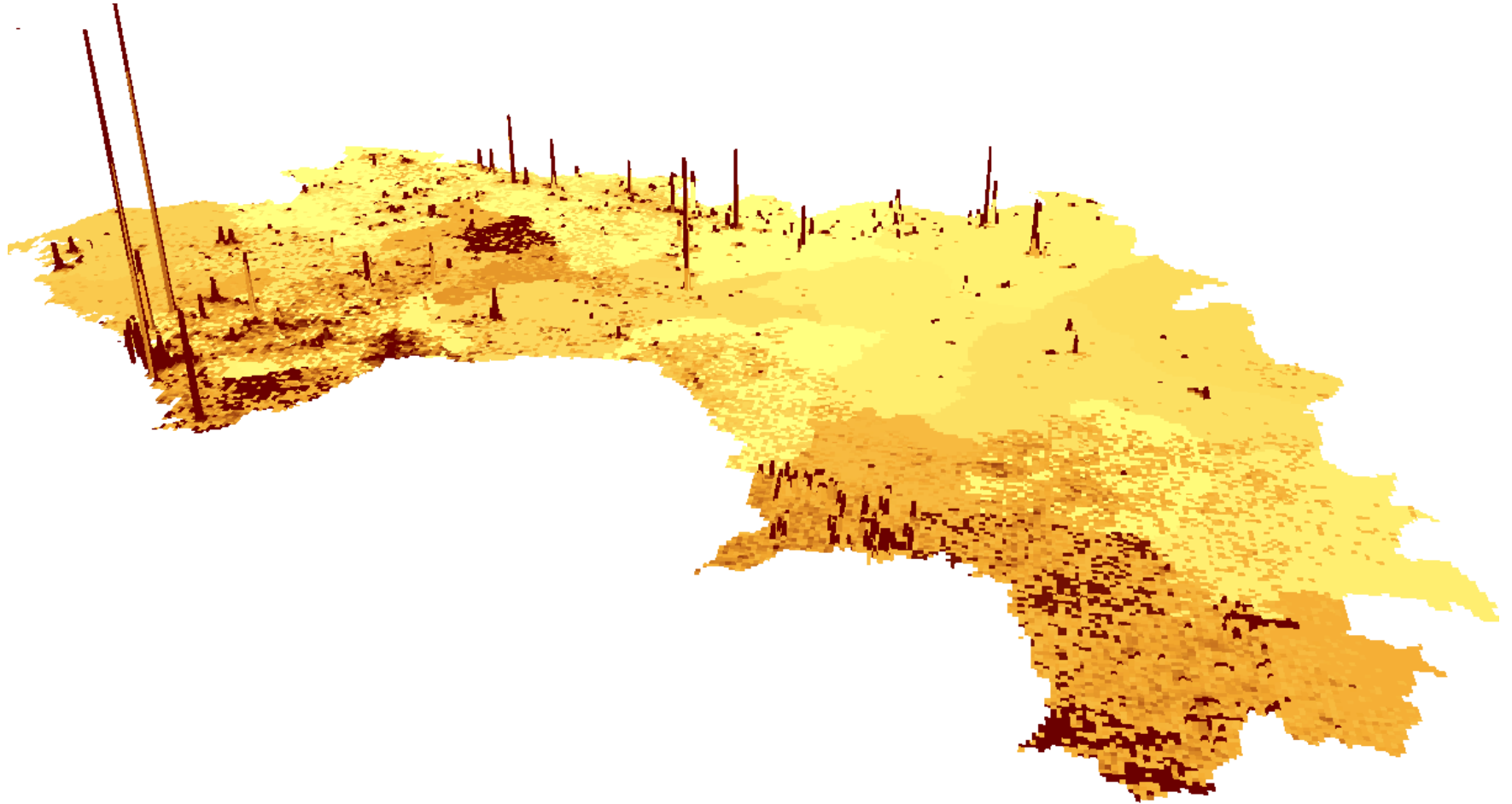
Proportional Allocation

Height Proportional to Population Counts



VIIRS Reallocation

Height Proportional to Population Counts



GPW is a census-based global population grid that is:

- Free and open source
- Well-documented
- Transparent
- Highly used
- Widely applicable

How to access data:

- GPWv3: <http://sedac.ciesin.columbia.edu/data/collection/gpw-v3>
- GPWv4: Email gpw@ciesin.columbia.edu to request alpha data

Greg Yetman

gyetman@ciesin.columbia.edu

