Overview of Recent CIESIN Spatial Data Activities

Kytt MacManus & Alex de Sherbinin

UNGIWG 2014 Plenary
Friday May 16, 2014
CIESIN Background

- CIESIN founded in 1989 as a non-profit consortium based in Michigan

- Center within the Earth Institute since 1998 based at the Lamont campus in Rockland County, NY

- About 45 professional staff from the social and natural sciences, information technology & data management plus students, postdocs, interns and visiting scholars (from France, China, Japan, Sweden, …)

- Teaching at Columbia, with courses the schools of
  - International and Public Affairs
  - Public Health
  - Continuing Ed
  - Columbia College (undergrads)
Selected CIESIN Projects

- Socioeconomic Data and Applications Center (SEDAC)
- Millennium Villages Project
- African Soil Information Service (AfSIS)
- Health Effects and Geochemistry of Arsenic and Manganese (NIEHS Superfund Research Program)
- African and Latin American Resilience to Climate Change (ARCC) Project (USAID)
- Planning for Resilience in East Africa through Policy, Adaptation, Research, and Economic Development (PREPARED) (USAID)
- Consortium for Climate Risk in the Urban Northeast (NOAA RISA)
- Terra Populus, NSF DataNet project
- Global Forum for Geography and Statistics
- Global Exposure Database for the Global Earthquake Model
Africa Soil Information Service

Data and Map Portal
Map products and tools, methods, and data for digital soil mapping and soil spectral analysis

New Release: Soil Property Maps of Africa at 1 km²

Remote Sensing Soil Covariates
- AfricaGrids.net Data Sets
- SRTM Data

Soil Profile Databases
- ICRAF-ISRIC Soil Spectral Library
- Africa Soil Profile Database - 12,000+ records

Field Data Collection
- New Release: Diagnostic Field Trial Database
- Land Degradation Surveillance Framework

Video
Why is soil important?

The CEO of the ATA talks with...
Interview with Khalid Bomba

In the Spotlight
Are large-scale commercial agricultural expansions possible in Nigeria?

News and Updates
- Article: Addressing nutrient
Terra Populus: 
A Global Population/Environment Data Network

- 5-year NSF DataNet project led by Minnesota Population Center
- CIESIN helping to integrate land cover, climate, and other environmental data with census microdata
- Building on IPUMS International holdings (859 million records across 65 countries)

[Logo of Terra Populus] [MPC Minnesota Population Center] [ICPSR Inter-University Consortium for Political and Social Research] [IPUMS International]

Center for International Earth Science Information Network
EARTH INSTITUTE | COLUMBIA UNIVERSITY
Global Exposure Database for the Global Earthquake Model (GED4GEM)

- People and their distribution in three time intervals (day, night, transit)
- Building counts classed by building type based on their structural characteristics
- 30 x 30 arc-second grid cells
Global Forum for Geography and Statistics (GFGS)

A global task force initiative by EFGS www.efgs.info

Voluntary professional network for merging geography and statistics

Focus on global issues and cooperation between national statistical institutes/mapping agencies and research institutions

Current Members

ARGENTINA FINLAND KOSOVO SLOVENIA
AUSTRIA FRANCE LATVIA SPAIN
BRAZIL GEORGIA LITHUANIA SWEDEN
BULGARIA GERMANY LUXEMBOURG SWITZERLAND
CROATIA GREECE MEXICO THE NETHERLANDS
CZECH REPUBLIC HUNGARY NORWAY THE REPUBLIC OF MACEDONIA
DENMARK ICELAND POLAND TURKEY
ECUADOR INDONESIA PORTUGAL UNITED KINGDOM
ESTONIA IRELAND ROMANIA UNITED STATES OF AMERICA
EUROSTAT ITALY SLOVAKIA ÅLAND
NASA Socioeconomic Data and Applications Center (SEDAC)

SEDAC’s mission is to develop and operate applications that support the integration of socioeconomic and earth science data and to serve as an “information gateway” between the earth sciences and social sciences.

http://sedac.ciesin.columbia.edu/
Key SEDAC Global Environmental & Socioeconomic Datasets

- **Population Grids**
  - Gridded Population of the World, Version 3 (GPW)
  - Global Rural-Urban Mapping Project (GRUMP)
  - US Census Grids

- **Poverty Maps**
  - Global Distribution of Poverty
  - Global Child Malnutrition

- **Climate Related**
  - Low Elevation Coastal Zone
  - Ramsar Wetlands at Risk to Sea Level Rise
  - Population, Landscape, And Climate Estimates (PLACE)
  - Global Natural Disaster Hotspots

- **Environmental Indicators**
  - Environmental Sustainability Index (ESI)
  - Environmental Performance Index (EPI)
  - Human Appropriation of Net Primary Productivity
  - Human Footprint/Last of the Wild
  - Natural Resource Management Index (NRMI)

- **Infrastructure**
  - Global Roads Open Access Data Set (gROADS)
  - Global Reservoirs and Dams (GRanD)
  - Nuclear Power Plants
Global Roads Open Access Data Set (gROADS)

Work carried out under the CODATA Global Roads Data Development Task Group

Future collaborations planned with:
- UN Cartographic Division
- Open Street Map
Global Rural-Urban Mapping Project, v.1

- Three products in one:
  - A population count/density grid (30 arc-second resolution)
  - An urban extents grid
  - Settlements points (all settlements >5k population)

Low Elevation Coastal Zone

Population Density (per km²)

Low Elevation Coastal Zone (m)

© 2013. The Trustees of Columbia University in the City of New York.
columbia.edu/data/set/lec2-urban-rural-population-land-area-estimates-v2

The Low Elevation Coastal Zone (LECZ) Urban-Rural Population and Land Area Estimates Version 2 (data set provides continent-level and country-level estimates of land area and urban, rural, and total population for 202 statistical areas (countries and other UN-recognized territories). Population inputs were derived from Gridded Rural-Urban Mapping Project, version 1 (GRUMPv1)). Elevation data were derived from the Shuttle Radar Topographic Mission (SRTM) 90-meter data set.
SEDAC Population Estimation Service

- WPS-accessible service that returns an estimate of 2005 population for a user-defined polygon
- SEDAC Client recently upgraded to support tablets, phones
- CHANGE viewer also supports queries

http://climatechangehumanhealth.org/changeviewer/

Mobile: http://sedac.ciesin.org/tools/population-estimation-mapclient/m
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)
Uniform Distribution

• GPW is a minimally modeled product.
  – Does not incorporate ancillary data
  – Uniformly distributes population based on land area

• The accuracy of GPW pixel estimates is directly related to the size of the input census units
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><strong>Number of Input Units</strong></td>
<td>19,000</td>
<td>127,000</td>
<td>c. 400,000</td>
<td>&gt; 13,000,000</td>
</tr>
<tr>
<td>(subnational geographic units)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Grid Resolution</strong></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><strong>Census variables</strong></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>
Panama, GPW4 Boundaries
Panama, Population Distribution Grids 2010
GPW4 Total Population Progress as of 5–13–14
Data Acquisition Status as of April 23, 2014: Age

- **Acquired-143 Countries**
- **Not Acquired-98 Countries**
Data Integration Difficulties

• GIS Data from non-Census Source
  Some common problems:
  – Data sources might be from different points in time
  – Might refer to the same unit by different names
  – GIS data might not capture changes in boundaries over time that are usually identified in the census, and therefore require editing

• Significant effort is needed to reconcile Census areas with those present in the GIS data and to assign common identifiers

• An effort should be made by national statistical offices to coordinate and cooperate with national geographic data managers to produce integrated products that are both consistent and accurate.

• As changes occur to internal area boundaries, those changes should be tracked and versioned in the geographic data.
SEDAC Future Investments

• We have an interest in distributing more global infrastructure data, moving beyond roads, reservoirs & dams, and nuclear power plants to include electrical grids, other power plants, waste water treatment facilities, building footprint and type, etc…

• We will be focusing on data relevant to climate impact, vulnerability, and adaptation assessment

• We are investing in linking population data to urban “footprints” to better understand the interplay of population dynamics and urban expansion