CIESIN's Experience in Mapping Population and Poverty

Alex de Sherbinin and Susana Adamo
Center for International Earth Science Information Network (CIESIN), Columbia University

Introduction

It has been agreed that the Sustainable Development Goals (SDGs) should be monitored using disaggregated data. For example, the Open Working Group report states: *In order to monitor the implementation of the SDGs, it will be important to improve the availability of and access to data and statistics disaggregated by income, gender, age, race, ethnicity, migratory status, disability, geographic location and other characteristics relevant in national contexts to support the monitoring of the implementation of the SDGs.* [1]

This recommendation has been echoed in subsequent treatment of the SDGs, including the Secretary General's Synthesis Report which notes: *Mechanisms to review the implementation of goals will be needed, and the availability of and access to data would need to be improved, including the disaggregation of information by gender, age, race, ethnicity, migratory status, disability, geographic location, and other characteristics relevant to national contexts*.[2]

This brief note presents and reflects on CIESIN experiences in compiling global subnational demographic and poverty data sets for use in measuring progress towards the Millennium Development Goals (MDGs) and now for the Sustainable Development Goals (SDGs), and provides recommendations for how to strengthen the demographic evidence base needed for attainment of the SDGs.

Demographic Data for the MDGs

CIESIN was called on by Jeffrey Sachs in 2003 to be the "mapping arm" of the Millennium Development Project (MDP), a two year process to develop the evidence base necessary for the achievement of the MDGs. CIESIN worked most closely with the Poverty and Hunger task forces, providing maps for reports and later producing data to aid in site selection for the villages that were included in the Millennium Project (test beds for MDP recommendations).

Under this project, in collaboration with the World Bank, CIESIN also developed a poverty atlas (CIESIN 2006) and a collection of spatial poverty data for dissemination. These included small area estimate data on poverty headcounts for 26 countries (CIESIN 2005a), a global map of infant mortality rates (a measure of extreme poverty) (CIESIN 2005b), and a global map of the percentage of children underweight (CIESIN 2005c). The small area estimates were developed by the World Bank and country partners, while the two latter data sets were developed by CIESIN based on statistically representative subnational regions of varying sizes from Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), vital statistics and other country sources. These were the first globally complete subnational maps of the spatial dimension of poverty (Fig. 1); an update of the infant mortality rate grid for circa 2015 is in preparation.

CIESIN used the data to better understand a number of spatial factors that may influence the distribution of poverty such as distance from roads and ports, high elevation zones, rainfall levels and dryland ecosystems. For example, de Sherbinin (2009) examined geographic and biophysical correlates of hunger in Africa, and Balk et al. (2004 and 2005) conducted analyses using DHS cluster level data for West Africa to understand the impact of distance from urban areas on infant mortality, and for all developing regions to understand the relation of environmental variables to child malnutrition. CIESIN population grids were also used with age structure data by country to develop global grids of vulnerable populations. Data from the small area estimates collection were more actively used in the development of country interventions.

Although this was a high payoff activity, and it made use of existing CIESIN experience in matching geostatistical data with boundaries, it was challenging and the spatial resolution of the subnational units varied substantially.

For some countries, e.g. Brazil and Mexico, admin3 level data were available. For most other low income countries, inputs were on the order of admin1 or even aggregates thereof.[3]

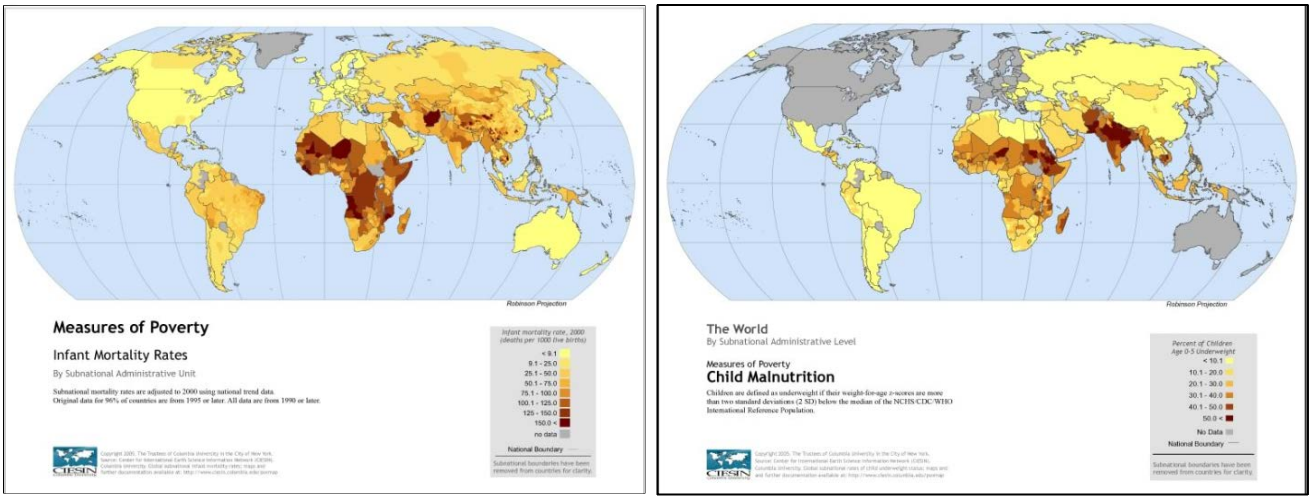


Fig. 1: Global Infant Mortality Rates and Rates of Child Malnutrition (weight-for-age)

Demographic Data for the SDGs

As mentioned above, demographic data has a critical role in enabling the monitoring of the 2015 SDGs.[4] Because the SDGs are intended to cover all countries, and not just low income countries, complete global coverage is important. The impacts of climate change and natural disasters do not respect levels of wealth.

National population censuses provide an essential common foundation for looking at the distribution of populations by age and sex groupings and, by embedding such measurements in a spatial framework, it becomes possible to create adaptable data frameworks that can support SDG monitoring. Spatial representation of census information creates flexible aggregation possibilities (e.g. number of elderly women living in high-disaster risk zones). CIESIN's *Gridded Population of the World* (GPW) data set is relevant for SDG monitoring and reporting. GPW displays the global distribution (counts and densities) of the human population on a continuous surface (Fig 2). Since its first version in 1995, the essential data inputs to GPW have been population census tables and cartography as these provide complete and exclusive geographic coverage of a country's population at a given point in time. While recognizing the uncertainties inherent in a "census only" approach to spatial allocation of population, for many applications there is utility in having an unmodeled population surface. For example, these grids may provide a "ground truth" for remote sensing or heavily modeled products, or may be used in analyses where heavily modeled population would be inappropriate.

In order to extend their usefulness for social applications, there is a need to add population characteristics to global population grids. Thus, version 4 of GPW (GPWv4) will include three additional variables from the 2010 round of censuses: 1) sex, 2) age by five-year age groups, and 3) urban/rural status. GPWv4 is available now by request, and will be released via the NASA Socioeconomic Data and Applications Center (SEDAC) later this Fall. The other products will be released progressively over the next 12 months. Where possible, the variables will be cross-tabulated, resulting in a consistent global gridded population data set with detailed estimates of age, sex, and urban/rural distribution within each country (Doxsey-Whitfield et al. 2015). Spatial databases that quantify the distribution of sex, age cohorts and other groupings across geographic areas provide crucial framework data for the purpose of SDG monitoring and reporting (e.g., Fig 3 and Fig 4). It is worth noting that even with simple variables, such as age cohorts, global census comparability is not as high as desired, and more work will be needed to facilitate comparable reporting across categories.

While GPWv4 incorporates significantly higher resolution census inputs than prior versions (12.5 million input units in GPWv4 vs ~400,000 in GPWv3), there are still significant portions of the world in which the spatial resolution of input units is suboptimal. The average input unit resolution for very high development regions is

944 sq. km, whereas the low and medium human development countries have an average input resolution of 3,518 sq. km and 4,700 sq. km, respectively. Partly to overcome these limitations, CIESIN developed a modeled population surface, the Global Rural Urban Mapping Project (GRUMP) (CIESIN 2011), that reallocated GPWv3 population counts using the Defense Meterological Satellite Program's nighttime lights data sets. Currently, CIESIN is working with partners at the Joint Research Center (JRC) to allocate GPWv4 data using the JRC's Global Human Settlements Layer derived from Landsat imagery, and is experimenting with a lightly modeled population distribution grid combining GPWv4 and VIIRS day-night band data using the relationship between population density and luminosity in countries within regions for which we have very high resolution population data.

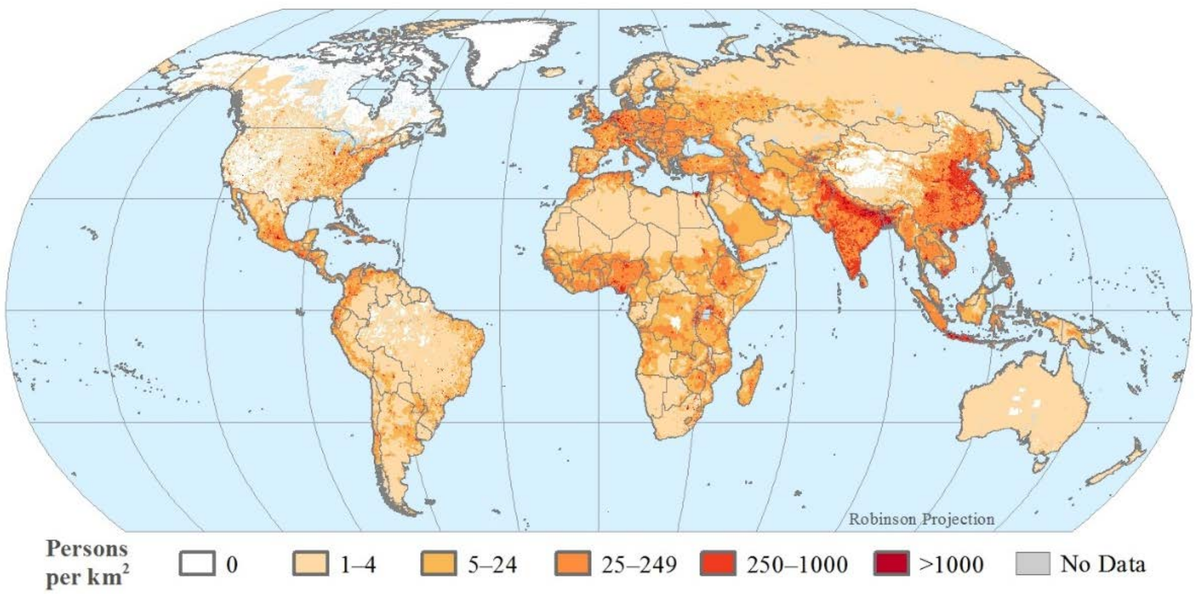
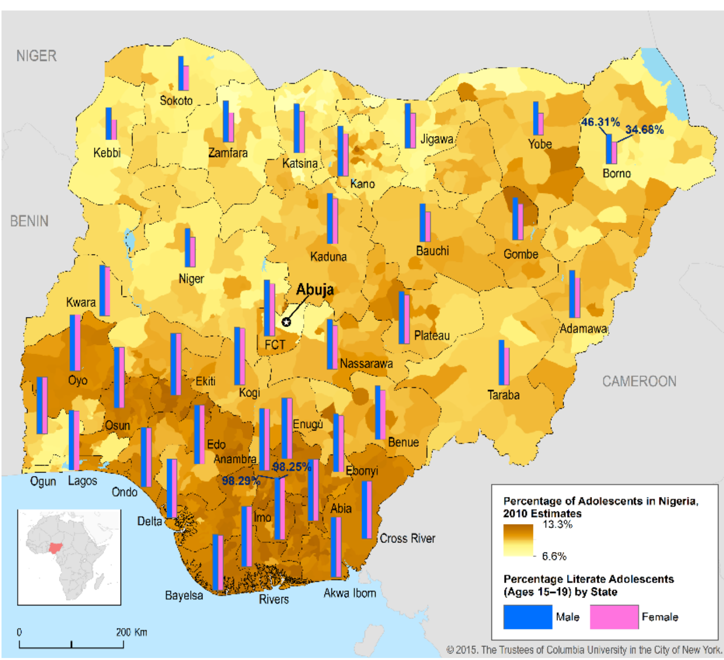


Fig 2. CIESIN/SEDAC's Gridded Population of the World version 4 (GPWv4), 2010 population density[5]



	State	% Literate, Male Adolescents	% Literate, Female Adolescents	Male/Female Ratio of Literate Adolescents
5 States with lowest total literacy	Borno	46.3	34.7	1.3
	Kebbi	51.1	31.8	1.6
	Yobe	52.2	34.7	1.5
	Sokoto	54.5	39.4	1.4
	Niger	60.6	47.2	1.3
5 States with highest total literacy	Osun	96.6	96.1	1.0
	Ekiti	97.6	97.8	1.0
	Anambra	97.8	97.6	1.0
	Abia	97.7	97.9	1.0
	Imo	98.3	98.3	1.0
	NIGERIA	82.8	76.7	1.1

Source: National Population Commission of Nigeria. 2009

Fig 3. Adolescent literacy in Nigeria: *The map displays a GPWv4 grid of the proportion of adolescent population (age 15-19) by Local Government Area in Nigeria, ca 2010. Superimposed are bar charts of the percentage of literate adolescents within the same age group by state and sex, from the 2006 population census*[6]*. The table compares the states with the lowest and highest total adolescent literacy rates, highlighting the intra-country differences in the sex ratio of adolescent literacy.*

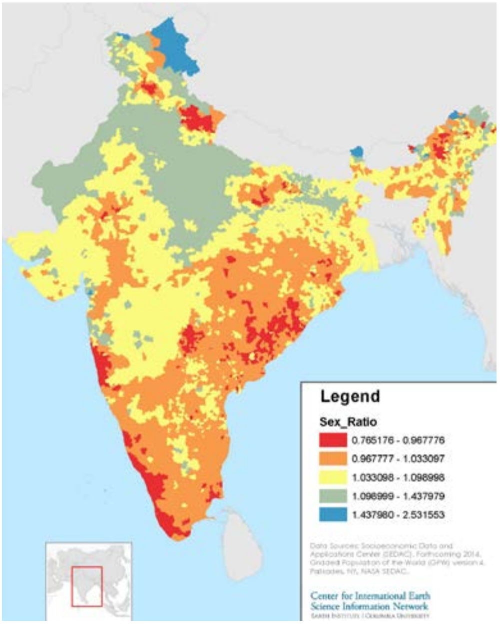


Fig 4. Estimated sex ratio in India (2010)

Observations and Recommendations

Spatial data frameworks support multi-scale reporting across different administrative levels as well as across non-administrative units such as watersheds, climate zones, and disaster risk zones. To do so requires reporting data at the highest resolution possible without violating confidentiality concerns. Yet there are challenges with the resolution at which several large, demographically important countries release their data. For India, admin3 (Tehsil level) data are released inconsistently across states. For Indonesia, the admin4 level data (Desa/Village) are only available at the price of USD 15,000. These are not isolated cases.

Based on two decades of experience in working with census and survey data to create gridded products, CIESIN recommends the following:

- Disseminate data freely for at most the cost of reproduction: this supports research, discovery, and information flows that can promote policies that reduce poverty
- Report population counts at enumeration area level and all other census variables at census tract or smaller census geographies: this facilitates a whole range of population-based analyses important to the SDGs
- Include common identifying codes for the tabular population counts and census geographies to allow for seamless and accurate data integration: this would reduce the time needed to compile spatial population data and increase the time for analyses
- Make the census geography available to the public in a digital format: too many countries do not disseminate spatial data files associated with their census results
- Document changes in administrative areas from one census round to the next: this is vital for tracking progress towards SDGs over time
- Report all ages in 1-year age groups: having 1-year age groups would allow for grouping the age data as needed, for example as denominators for education statistics or for calculation of infant and child mortality and malnutrition rates
- Do not truncate age reporting over a certain age: many countries group everyone over age 60 or 70 rather than reporting all age groups in one or five year intervals up to age 100; with increasing longevity and heterogeneity across the elderly populations over age 60 it is important to disaggregate
- Report infant and child mortality disaggregated by sex at the highest resolution reporting units possible: this facilitates tracking of sex-differentiated development across space

For more information, contact the authors at adesherbinin@ciesin.columbia.edu or sadamo@ciesin.columbia.edu, or visit http://ciesin.columbia.edu.

References

Balk D, Pullum T, Storeygard A, Greenwell F, Neuman M. 2004. A Spatial Analysis of Childhood Mortality in West Africa. *Population, Space and Place*. 10: 175–216.

Balk D, Storeygard A, Levy M, Gaskell J, Sharma M, Flor R. 2005. Child hunger in the developing world: An analysis of environmental and social correlates. *Food Policy*. 30: 584-611.

CIESIN, IFPRI, The World Bank, and CIAT. 2011. Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Population Density Grid. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://sedac.ciesin.columbia.edu/data/set/grump-v1-population-density

CIESIN. 2006. *Where the Poor Are: An Atlas of Poverty*. Palisades NY: CIESIN, Columbia University.

CIESIN. 2005a. *Poverty Mapping Project: Small Area Estimates of Poverty and Inequality*. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H49P2ZKM

CIESIN. 2005b. *Poverty Mapping Project: Global Subnational Infant Mortality Rates.* Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4PZ56R2.

CIESIN. 2005c. *Poverty Mapping Project: Global Subnational Prevalence of Child Malnutrition.* Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4K64G12.

de Sherbinin, A. 2009. "The Biophysical and Geographical Correlates of Child Malnutrition in Africa" *Population, Space and Place*, 15. http://dx.doi.org/10.1002/psp.599.

Doxsey-Whitfield, E., K. MacManus, S. Adamo. L. Pistolesi, J. Squires, O. Borkovska and S. Baptista. Taking advantage of the improved availability of census data: a first look at the Gridded Population of the World, version 4 (GPWv4). *Papers in Applied Geography,* 1(3). http://dx.doi.org/10.1080/23754931.2015.1014272

---

[1]United Nations General Assembly, Report of the Open Working Group of the General Assembly on Sustainable Development Goals. A/68/970 12 August 2014. https://sustainabledevelopment.un.org/owg.html

[2] United Nations, The Road to Dignity by 2030: Ending Poverty, Transforming All Lives and Protecting the Planet. Synthesis Report of the Secretary General on the Post-2015 Agenda, 4 December 2014. http://www.un.org/disabilities/documents/reports/SG_Synthesis_Report_Road_to_Dignity_by_2030.pdf

[3] Administrative levels referred to in this document are as follows: admin1 = state or province, admin2 = country, district or *municipio*, admin3 = town or village.

[4] IUSSP. 2014. *Defining and successfully accomplishing the Data Revolution: the perspective of demographers*. http://iussp.org/sites/default/files/Data_Revolution_Demographers_IUSSP.pdf

[5] Gridded Population of the World, Version 4 (GPWv4) can be obtained on request to ciesin.info@ciesin.columbia.edu.

[6] National Population Commission of Nigeria. 2009. 2006 Population and Housing Census of the Federal Republic of Nigeria. National and State Population and Housing Tables: Priority Tables, Vol. I. Abuja, Nigeria: National Population Commission. http://www.population.gov.ng/images/Priority%20Tables%20Volume%20I-update.pdf