El uso de imágenes satelitales para cartografía demográfica, con el ejemplo de Haití

The use of satellite images for demographic mapping, with the example of Haiti[*]

Joshua Comenetz[†]

Key words: satellite images, demographic mapping, Haiti

Resumen

El suministro oportuno de mapas unidos con información demográfica subnacional es crítico para mejorar la eficacia de los esfuerzos humanitarios durante y después de los desastres naturales y otras crisis. Adquirir los mapas y datos después de un evento dura demasiado tiempo, especialmente cuando un evento afecta a más que un país. A pesar de la demanda, una base de mapas mundial unidos con datos actualizados de población no está disponible. La Oficina del Censo de los EE.UU. está construyendo un sitio del Internet para ayudar satisfacer esta necesidad. El terremoto reciente en Haití proporcionó una prueba de los productos de la Oficina del Censo. El sitio del Internet estuvo en construcción con una apertura pública programada para el otoño del 2010; pero debido a éste evento, el sitio se hizo disponible en sólo 24 horas después del terremoto. El sitio contiene un visor de datos (no se requiere ningún SIG) con archivos descargables para análisis espacial. El visor incluye herramientas que producen estimados para polígonos definidos por el usuario y pistas de huracán (zonas de búfer cerca de una línea). Se puede ver la demanda para este servicio por la respuesta positiva de una variedad de usuarios.

---

[*] Paper presented at the IV Congress of the Asociación Latinoamericana de Población, ALAP, Havana, Cuba, 16 to19 November, 2010.

[†] Population Division, U.S. Census Bureau, Washington, DC 20233, USA (Joshua.Comenetz@census.gov).

El uso de imágenes satelitales para cartografía demográfica, con el ejemplo de Haití

The use of satellite images for demographic mapping, with the example of Haiti*

Joshua Comenetz†

Introduction‡

The U.S. Census Bureau has a long history of producing annual population estimates by age and sex for every country and territory in the world. Through the populations at risk initiative, the Census Bureau is beginning to provide current international, subnational demographic data and maps online, to aid in humanitarian response to disasters and crises worldwide.

Over the past two decades, digital demographic data, boundary maps in GIS formats, and satellite images have increasingly become available. Advances in data accuracy and spatial precision can be achieved by incorporating census data and digital base maps at lower administrative levels, application of improved subnational demographic projection methods, and through use of satellite images to map populated places. Image data can also assist in developing demographic estimates for countries that have no recent census data or that restrict public access to demographic data. This is of particular value in developing countries where migration and other population changes mean that census results become outdated rapidly. Provided via an internet site, the populations at risk database will help improve the efficiency of response to climatic disasters, famines, and other natural or human-induced crises through provision of demographic estimates that are more geographically and temporally accurate than has previously been possible. The Census Bureau's first online assistance of this type occurred just after the January 2010 earthquake in Haiti, as described below.

Goals

The primary purpose of the international populations at risk project is to produce and provide geospatial data and maps that can be used to help save lives during and after disasters. This Census Bureau effort responds to the U.S. National Research Council's (NRC's) call for better data and maps in support of disaster response. A recent NRC report included the following recommendations (NRC 2007, pp. 152-55):

> Develop a template of minimum acceptable population and other geospatial data sets that are required by disaster responders. The data sets should be updated frequently (at least

* Paper presented at the IV Congress of the Asociación Latinoamericana de Población, ALAP, Havana, Cuba, 16 to19 November, 2010.

† Population Division, U.S. Census Bureau, Washington, DC 20233, USA (Joshua.Comenetz@census.gov).

‡ This paper is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, technical, or operational issues are those of the author and not necessarily those of the U.S. Census Bureau.

mid-decade if not more frequently) and include digital census enumeration units and other census maps in digital form.

Support should be given to test the accuracy of estimates of size and distribution of populations based on remotely sensed imagery, particularly in rural and urban areas of countries with spatially, demographically, and temporally inadequate census data.

The U.S. Census Bureau should be given greater responsibility for understanding populations at risk and should be funded to…support the U.S. government in international disaster response and humanitarian assistance activities. The U.S. Census Bureau should also have an active research program in using and developing these tools and methods, including remotely sensed imagery and field surveys.

Timely provision of good subnational data can improve the efficiency of humanitarian relief efforts during and after disasters. The goal of the populations at risk database is to help solve problems ranging from basic questions—"How many tents should we load on the plane?"—to longer-term issues such as provision of age-appropriate health care to evacuees. Better targeting of services may improve both the quality of response and allocation of scarce financial resources.

Background

For more than 60 years, the U.S. Census Bureau has been a source of international demographic estimates at the national and subnational levels. Annual national estimates and projections (data for entire countries) are included in the agency's International Data Base (U.S. Census Bureau 2010a). Subnational estimates can be linked to digital maps, allowing for the creation of international demographic maps. Some of these data and maps serve as a base for the LandScan international gridded population database (www.ornl.gov/sci/landscan), created by the Oak Ridge National Laboratory.

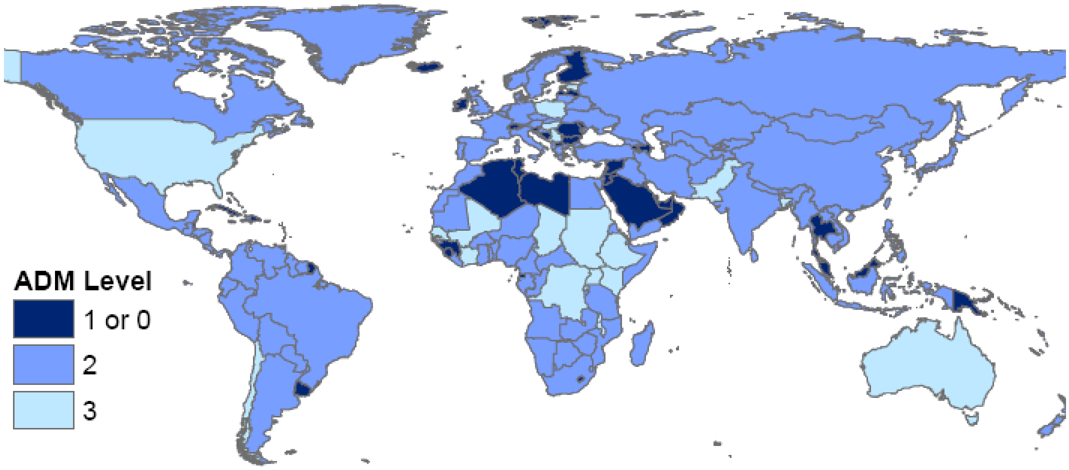
Data

The Census Bureau's digital map archive covers every country and territory in the world. Figure 1 shows the smallest administrative unit (ADM) currently held in ArcGIS shapefile format for each country. The goal is to obtain both data and maps at the lowest (smallest) available ADM level. ADM levels are defined as follows:

- ADM0 – Countries and territories.
- ADM1 – States, provinces, or equivalent.
- ADM2 – Districts, counties, or equivalent.
- ADM3 – Municipalities, townships, or equivalent.
- ADM4 and below – Villages or other small units.

Most of the world is covered in the archive at the ADM2 level, but even at ADM2 there are gaps in the Middle East and other areas. On the other hand, ADM3 maps have been obtained for more than twenty countries. Separate maps are obtained for each ADM level wherever possible. While smaller ADMs can usually be aggregated to form higher-level ADM units, this is not always the case (examples include South Africa and New Zealand). Also, boundary revisions can occur at any time, making it difficult to map change over time (Comenetz and Caviedes 2002). Data sources include national statistical agencies, international organizations such as the United Nations and World Health Organization, the Department of State and other U.S. government agencies, online collections maintained by academic and other non-governmental organizations, and digitization of paper maps by Census Bureau staff.

Figure 1. Smallest administrative level held in the U.S. Census Bureau digital map archive. There are no ADM3s in some U.S. states.



The Census Bureau's subnational demographic holdings similarly include population data (total population, and age/sex breakdowns typically in 5- or 10-year bands) for the entire world. In more than 90 percent of countries or territories, covering over 95 percent of the world's population, these data are derived from national censuses. Census data are essential for creation and maintenance of a populations at risk database because censuses usually cover all of a country's territory, allow for "continuity of statistics from census to census," and provide "detail…about population sub-groups in local areas" (NRC 2002, p. 7). Where national census data are unavailable or outdated, official estimates or other sources are used—for example, in countries affected by war (Afghanistan, Angola) or where no census has been conducted in more than a decade (Congo-Kinshasa, Greenland).

In 2008, the Census Bureau began to plan for expansion of the data archive to include more variables. Important variables—components of the "minimally essential data set" (NRC 2007, 89-96) for populations at risk projects—vary according to type of disaster and recovery stage but always include total population disaggregated by age and sex. Table 1 outlines other variables that are potentially useful in disaster response and that are actually available for many countries.

Table 1. Selected geo-demographic variables for a populations at risk database.

Type of Event	Some Key Variables
Short-term disasters (e.g., hurricanes, floods, earthquakes, tsunamis)	• Baseline populations of small administrative areas • Vulnerable populations – elderly, lower income, migrants (especially those lacking knowledge of national language) • Housing characteristics – e.g., wall and roof materials, source of drinking water, age of housing
Longer-term disasters (e.g., climate variation, sea level rise, famine)	• Population density, urban and rural population • Growth projection by coastal administrative unit • Occupation • Elevation above sea level (McGranahan *et al.* 2007)
Disease	• Populations most vulnerable to disease – children, elderly, lower income • Health status, e.g., HIV infection increases vulnerability to other diseases
War and genocide	• Baseline populations of administrative areas – total population, age, sex • Cultural variables – ethnicity, language • Household size and composition for estimating housing needs

Data Problems

To provide geo-demographic information that is as accurate and current as possible, both archives—demographic data and digital maps—require revision and updating as new census results, demographic estimates, and administrative boundary maps are released by national governments. The following issues complicate this process:

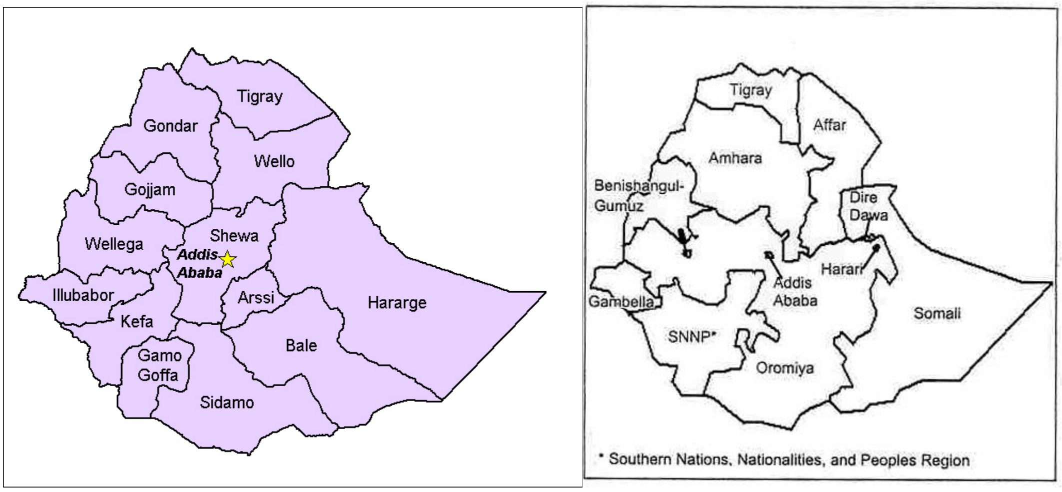
- Demographic data and maps are often produced by different agencies or released separately (it is typically more difficult to obtain administrative maps than numerical data). If only one of these two elements is available, it is not possible to *map* population. Also, governments may release data and maps at different ADM levels. For example, ADM2 data cannot be mapped if the best-available shapefile contains only ADM1 boundaries.

- Countries around the world choose to split, combine, or completely redraw administrative areas. For example, Ethiopia (Figure 2) and South Africa redrew provincial and lower-level boundaries in the early 1990s; India and Nigeria periodically create new states; and new counties are occasionally created in the United States.

- Censuses are relatively infrequent—usually a decade or more apart (UN 2010). Most population estimates are based on censuses, and therefore calculated using the ADMs in force at the time that the most recent census was conducted. If boundaries are redrawn after a census, it can be difficult to produce accurate estimates of the populations of new ADMs unless the new ADMs are defined in terms of pre-existing smaller units. For example, if a country redraws ADM1 boundaries but keeps ADM2 boundaries constant, one can calculate the new ADM1 populations; otherwise, this may not be possible.

- Data and maps made available by governments can be withdrawn or revised at any time. Frequent online searches must be conducted, and personal contacts maintained with agencies that produce data, to ensure that all new materials are added to the Census Bureau archive as soon as possible after release or revision.

Figure 2. Ethiopian provincial boundaries in effect at time of 1984 census (left) and 1994 census (right). The newer boundaries were drawn along ethnic lines.



The above issues require continuous update and revision of the data and maps. All available data and maps are collected, but data and maps with matching geography frequently must be obtained at different times or from different sources. At best, an online populations at risk database of demographic maps reflects the lowest common denominator in both archives—that is, the smallest and most recently delineated ADMs for which both data and digital boundaries are available.

Remote Sensing Methods

The goals of satellite-imagery-based population mapping are as follows:

- Map population distribution at finer scales than is possible with census data (Tatem *et al.* 2007). This is especially useful where countries do not gather or release data for smaller administrative units. Current Census Bureau projects are using image analysis to map built-up areas, and then distributing population to one-hectare pixels (100-meter square grid cells).

- Provide information about the geographic distribution of population where no census has been taken recently, or where census data are no longer reliable. Examples include Somalia, Afghanistan, and Lebanon. Image-based mapping also has the potential to highlight possible quality problems with national censuses by revealing geographic regions where there are large differences between intercensal population change as measured by censuses and by image analysis.

- Improve the quality of intercensal population estimates. Few countries take censuses more frequently than once per decade (UN 2010). Image mapping could help validate demographic projections produced by standard mathematical methods by revealing, for example, rapid growth in urban areas. Image mapping is unlikely to be of as much use in areas of population decline, because the built environment will remain; however, this will not pose a serious problem in the medium term because populations are still growing in most countries.

To map population distribution at pixel scale, the Census Bureau is using a combination of demographic data from national censuses and built-up area maps derived from satellite imagery. Built-up areas are mapped by finding human-created hard or "impervious" surfaces such as roads and roofs in satellite imagery, using image processing and GIS software. The Census Bureau's methodology for classifying impervious land cover features (Azar *et al.* forthcoming) was largely adapted from methods developed for the Multi-Resolution Land Characteristics Consortium's National Land Cover Data (Yang *et al*. 2003, Homer *et al.* 2007) by the U.S. Geological Survey's Center for Earth Resources Observation and Science (USGS EROS). USGS EROS used a classification and regression tree approach, which is increasingly common in image classification. In this approach, multiple small training sites derived from high-resolution imagery (e.g., Quickbird or IKONOS, with a pixel size of 4m or less) are used to train a regression tree that is then used to estimate impervious surfaces over larger areas using lower resolution imagery such as Landsat or ASTER (pixel size of 28.5m or 15m, respectively).

The use of a mix of imagery types is an efficient and cost-effective way to map entire countries because it takes advantage of the benefits of each kind of imagery. High-resolution imagery has the advantage that it shows the built environment clearly. It is therefore possible to classify each pixel as impervious or pervious, and to omit non-human impervious surfaces such as bare rock from the "impervious" category. However, each scene (individual image) covers a comparatively small area—meaning that processing enough images for a national map becomes time-consuming—and it is also difficult or impossible to obtain complete coverage for entire countries. Landsat and ASTER imagery provide more complete coverage of the earth and require fewer scenes to create a national map, so processing time is reduced.

The key question is how well image-based maps of impervious surface correlate with population distribution. Comprehensive testing of the procedures requires:

- Census data, linked to boundary maps of the smallest available ADMs in several countries. The procedure must be tested in regions with different environmental characteristics (e.g., desert, humid tropical, mountainous) and urban planning and building traditions, because varying environments and construction styles are likely to produce different patterns of impervious land cover.

- National built-up area maps of the same countries, derived from satellite images obtained as near to the census dates as possible.

- Information on the location of urban areas in each country. The population-built-up area relationship will vary not only among world regions but also within countries, because in poorer countries, for example, rural roads are less likely to be paved than city roads.

With the above in hand for a country, the next step is to aggregate Landsat pixels to the smallest available ADM boundaries. It is then possible to chart the correlation between impervious area and census-derived population at the administrative unit level. The goal is to calibrate image maps using census data, making it possible to estimate population in areas where images but not census data are available. Among the questions to be answered are:

- How well does the distribution of built-up areas correlate with population patterns?

- How does the built-up area-population relationship vary regionally (among major world environmental and cultural regions) or among countries?

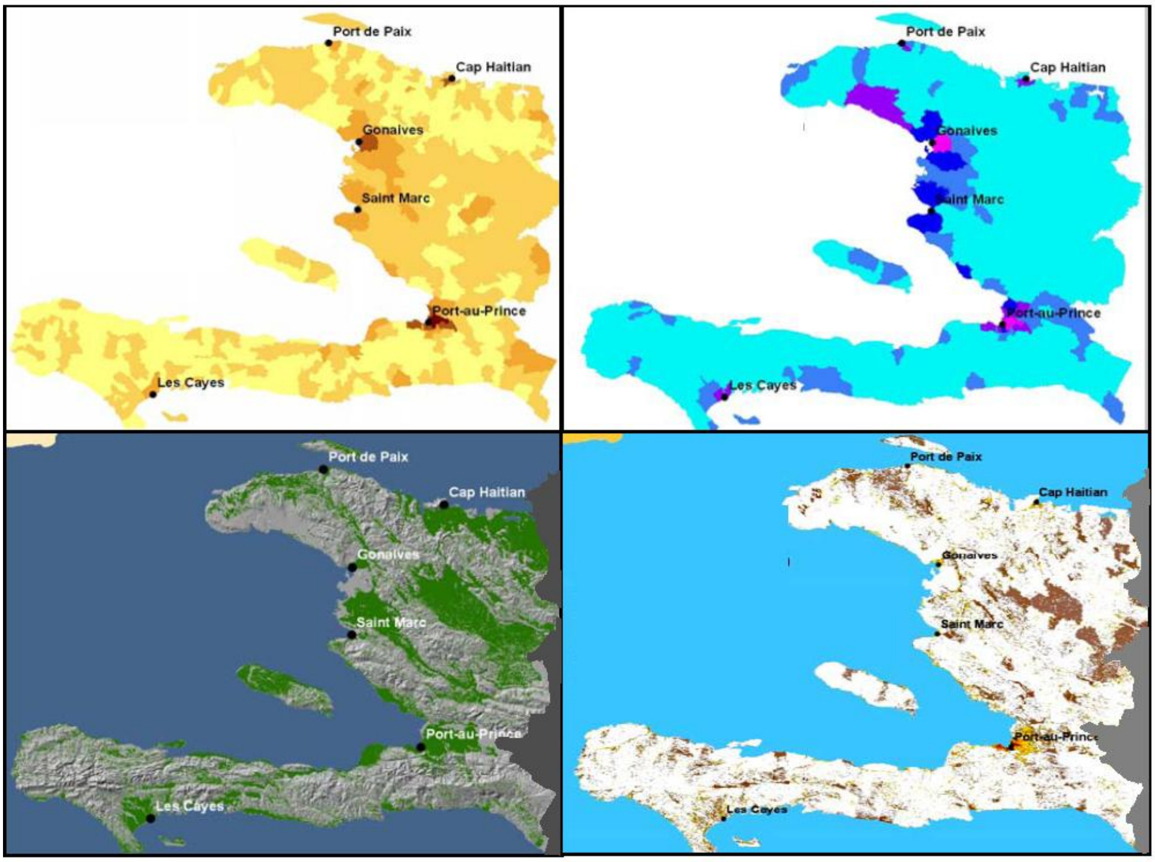
- Can the built-up area-population relationship in countries with census data be used to estimate population distribution in countries with no recent census data?

- How does the built-up area-population relationship vary within countries, or within the urban and rural areas of each country?

- To what extent does the built-up area-population relationship vary over time? Such changes could occur through increasing urban sprawl or changes in building methods.

If there is regional consistency in the built-up area-population relationship, then it will be possible to apply the results from countries with census data to neighbors with none—for example, results from Ethiopia or Kenya could be used to map population in Somalia. Temporal consistency in the impervious area-population relationship will determine whether imagery can be used to map population change between censuses, though it seems unlikely that the relationship could change radically in a decade or two.

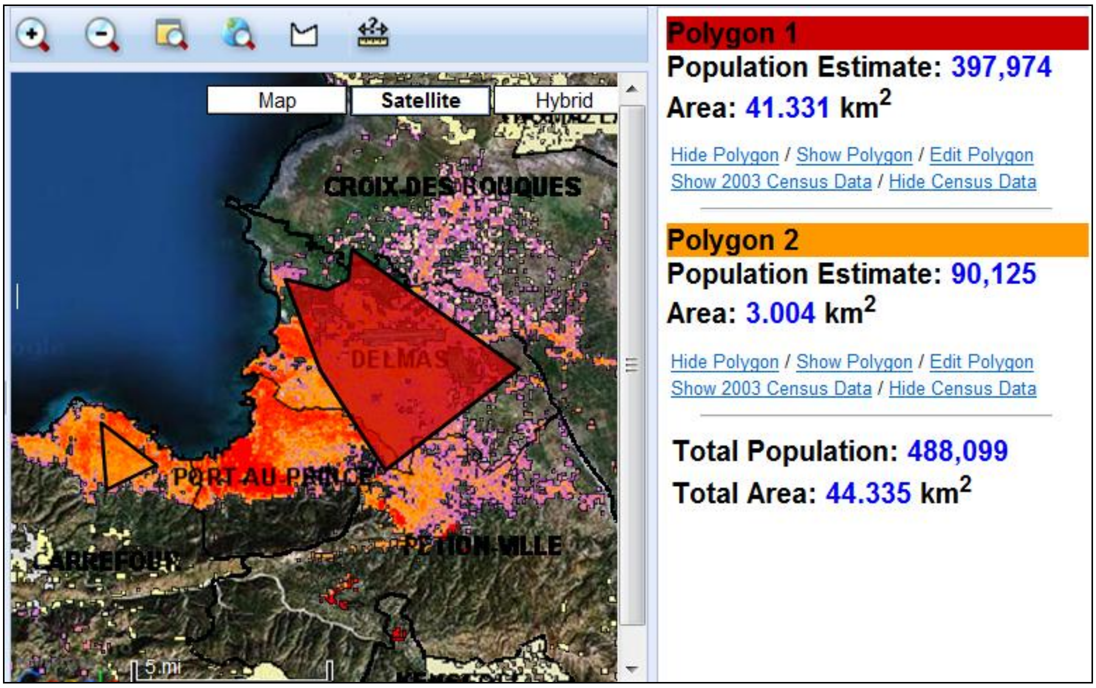
Figure 3. Mapping population in Haiti, based on image processing and census data.



Example: Haiti

The methods described above have been used in a study of Haiti. The illustrations in Figure 3 illustrate the process. First, data from the 2003 census were mapped at ADM3 level (the 560 *sections communales*) (Figure 3, upper left, with darker colors indicating greater population). Next, satellite imagery was used to map human-created impervious surface by ADM3 (Figure 3, upper right, with darker colors indicating more impervious area). To obtain pixel-level population estimates, people were only allowed to be assigned to the areas shown in green in Figure 3, lower left: the remaining areas contained steep slopes, water, or were otherwise assumed uninhabitable. The final population distribution map is shown in Figure 3, lower right, with rural or sparsely-populated areas in brown and denser or urban population in orange or yellow.

Figure 4. Haiti online gridded population mapping website, with polygon tool. Users can obtain population estimates for multiple custom polygons.



Haiti Earthquake Response and Website Tools

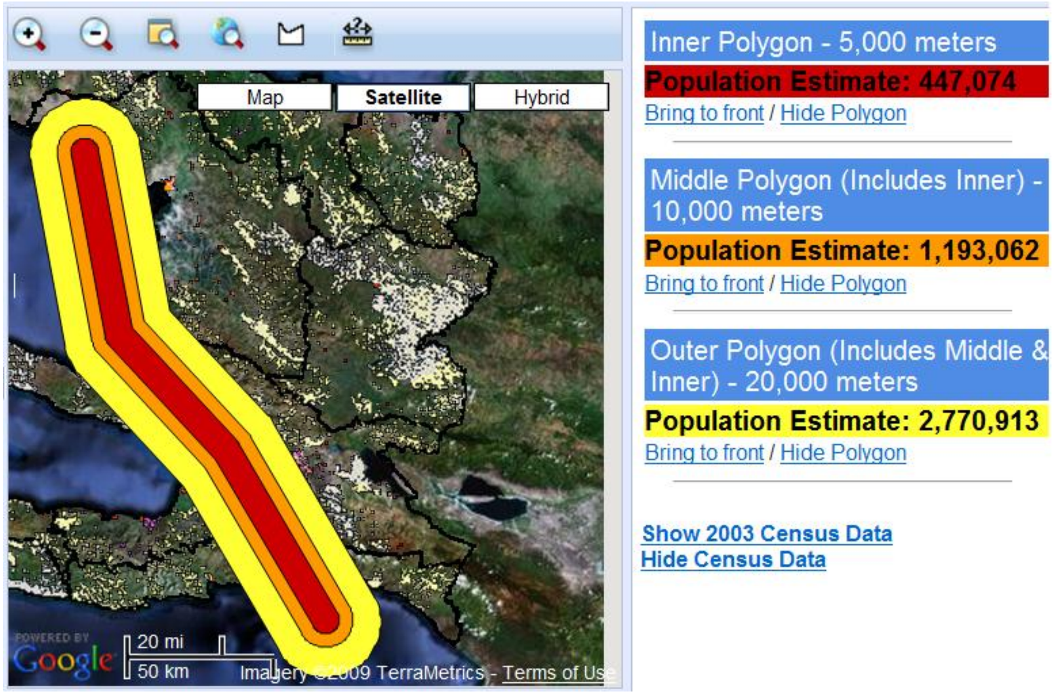
The Census Bureau, in partnership with the National Geospatial-Intelligence Agency (NGA), began constructing a website for free dissemination of subnational demographic data and maps in 2009. The goal was to release the site to the public in summer 2010. Fortunately, website development was ahead of schedule when a 7.0-magnitude earthquake struck Haiti in January 2010. The Census Bureau and NGA were able to open the site (U.S. Census Bureau 2010b) within 24 hours after the earthquake. It contains a data viewer (Figure 4) that provides access to census and 100-meter gridded population data without need for GIS software, and also a download page that provides the complete gridded dataset and census data in GIS format for more complex applications.

Tools available in the Haiti data viewer, or in development for the future, include:

- Custom polygon selection, allowing users to obtain population estimates for any area of interest (Figure 4). In the future, it may be possible to enhance this tool by allowing users to upload polygons. This would allow someone to, for example, obtain the boundary of a flooded area from one source and obtain a population estimate, all without using GIS software.

- Hurricane track selection, providing population estimates for buffered areas around a user-drawn line (Figure 5). This tool will also help with other disasters that affect a roughly linear area, such as river or coastal flooding. Other potential uses include estimating population along roads or borders.

Figure 5. Haiti online gridded population mapping website, with hurricane track tool. The tool provides population estimates for buffers around a user-drawn line .



- Earthquake point selection, for estimates of populations in circular buffers around earthquake epicenters. This will also be useful for other disasters originating at point locations, such as volcanic eruptions, and for estimating population within a user-defined radius of any point, such as a proposed hospital or shopping center.

- Street-level polygon selection, which permits users to draw custom polygons following roads or other features shown in publicly-available online mapping applications such as Google Maps. One can, for example, outline and obtain a population estimate for a city neighborhood.

The value of and level of demand for detailed, online international demographic data tied to maps has been clearly demonstrated by the traffic to the Haiti site from government agencies, United Nations offices, non-governmental organizations, universities, and businesses. The data sets and maps were downloaded more than 700 times in the weeks after the earthquake, and positive comments on this new Census Bureau service were received from numerous users, including humanitarian relief providers planning aid distribution.

Conclusion and Next Steps

Methods for mapping population distribution using a combination of satellite imagery, image processing and GIS software, administrative boundary maps, and census data can improve the resolution and currency of demographic maps and databases. With online distribution, highly detailed population maps can be available to anyone with an internet connection when disasters happen, for pre-disaster planning, or for other educational, governmental, charitable, or commercial purposes. Creation of the products and websites described in this paper is time-consuming so cannot be done ad-hoc as disasters happen. By coincidence, the Census Bureau completed its study of Haiti just in time for the results to be of use in humanitarian response to a major natural disaster, but the best way to ensure that data are available for future events is to work toward the goal of a global populations at risk database.

As more countries are mapped, it will also be possible to see how much the relationship between built-up areas and population distribution varies around the world. Demographers and geographers will benefit from having access to digital population data for user-defined areas such as city neighborhoods, without being limited by the ADM levels at which national statistical agencies release data. They will also have another tool for assessing census data quality: how well do population patterns predicted by built-up area mapping match what is stated in census tables?

As noted above, the remote sensing methods must be tested in several diverse regions (with different demographic patterns, environmental characteristics, and topography) in order to refine them for general use. While it will be necessary to study several more countries to further improve the methods and determine whether they work well in the broadest range of environmental and demographic conditions, the success of the Haiti project gives reason for optimism. The goal is to cover as much of the developing world as possible, with a focus on areas most prone to disasters and most in need of external aid in the aftermath of disasters.

Acknowledgements

The efforts of Population Division staff—Theresa Andrews, Derek Azar, Tom Fitzwater, Jordan Graesser, and Robert Leddy—were essential to completion of this paper.

References

Azar, D., Graesser, J., Engstrom, R., Comenetz, J., Leddy, R.M., Schechtman, N.G., Andrews, T. (Forthcoming). Spatial refinement of census population distribution using remotely sensed estimates of impervious surfaces in Haiti. *International Journal of Remote Sensing*.

Comenetz, J., and Caviedes, C. (2002). Climate variability, political crises, and historical population displacements in Ethiopia. *Global Environmental Change B: Environmental Hazards*, 4, 113-127.

Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., Herold, N., McKerrow, A., VanDriel, J.N., and Wickham, J. (2007). Completion of the 2001 National Land Cover Database for the Conterminous United States. *Photogrammetric Engineering and Remote Sensing*, 73, 337-341.

McGranahan, G., Balk, D., and Anderson, B. (2007). The rising tide: assessing the risks of climate change and human settlements in low elevation coastal zones. *Environment and Urbanization*, 19, 17-37.

NRC (National Research Council). (2002). *Down to Earth: Geographical Information for Sustainable Development in Africa*. Washington, DC: National Academies Press.

NRC (National Research Council). (2007). *Tools and Methods for Estimating Populations at Risk*. Washington, DC: National Academies Press.

Tatem, A.J., Noor, A.M., von Hagen, C., Di Gregorio, A., and Hay, S.I. (2007). High Resolution Population Maps for Low Income Nations: Combining Land Cover and Census in East Africa. *PLoS ONE*, 2(12), e1298. Available at: www.plosone.org/article/info:doi/10.1371/ journal.pone.0001298 [accessed April 2010].

UN (United Nations). (2010). *Census dates for all countries*. Available at: unstats.un.org/unsd/demographic/sources/census/censusdates.htm [accessed April 2010].

U.S. Census Bureau. (2010a). *International Data Base*. Available at: www.census.gov/ipc/www/idb [accessed September 2010].

U.S. Census Bureau. (2010b). Haiti Earthquake: U.S. Census Bureau Population Data [online] (January 2010 release). Available at: https://geoint-online.nga.mil/community/haitiearthquake/default.aspx [accessed September 2010].

Yang, L., Huang, C., Homer, C.G., Wylie, B. and Coan, M.J. (2003). An approach for mapping large-area impervious surfaces: synergistic use of Landsat-7 ETM+ and high spatial resolution imagery. *Canadian Journal of Remote Sensing*, 29, 230-240.