This pdf contains:
Balk, D., M.R. Montgomery, G. McGranahan and M. Todd 2009. "Understanding the Impacts of Climate Change: Linking Satellite and Other Spatial Data with Population Data." Pp 206-217 in: *Population Dynamics and Climate Change,* edited by J.M.Guzmán, G. Martine, G. McGranahan, D. Schensul and C. Tacoli. New York: UNFPA; London: IIED.

The contents of the full book can be found here:
https://www.unfpa.org/public/publications/pid/4500

This pdf was downloaded from www.iied.org

13

Understanding the Impacts of Climate Change: Linking Satellite and Other Spatial Data with Population Data

Deborah Balk
Mark Montgomery
Gordon McGranahan
Megan Todd

Introduction

An understanding of the impact of global climate change requires knowledge of who lives where. Accumulating sufficient knowledge about the locations and characteristics of the people who will be disproportionately and negatively affected by climate change, and, in particular, to identify the most vulnerable groups at risk, is a non-trivial undertaking. While data sets on climate patterns and data sets on populations exist, no single data set provides a complete picture of individuals and the communities and environments in which they live, making a comprehensive understanding of the impact of climate change on populations difficult. A more complete understanding can only be achieved by combining data from different sources, a practice that is increasingly possible, but still poses many challenges. Some important advances have been made: Satellite data are increasingly available, demographic data are increasingly spatially rendered and environmental data are increasingly being collected and produced with interdisciplinary inquiry in mind.

Data integration between two data sets that share identifying units can be straightforward, but data inconsistency within and between places may not be trivial. National statistical offices collect and report information in many different ways (United Nations, 2009), making, for example, a comparison of educational attainment among the residents of the neighbouring states of Texas, United States, and Tamaulipas, Mexico, impossible. This is true despite the fact that the United States and Mexico are quite similar in terms of census data collection. The complexities increase when definitions differ among data sources and even more so when there is a need to use data with dissimilar reporting units. The challenges that arise when combining population data—whether from censuses or surveys—with environmental data useful for describing or predicting climate-change hazards—whether derived from satellites or other spatial analyses—is the focus of this chapter. International agencies, national governments and local planners all need to prepare for climate change, and, in order to understand its potential impact on population, data must be organized and analysed in a spatial framework. This is true regardless of whether people live in cities or villages, though emphasis in this chapter will be given to the challenges of urban area analysis.

Climate modelling and physical geography help identify where climate-change-induced hazards are likely to occur, but in order to assess the resulting risks, human settlements and activities must be located in relation to these hazards. For example, it is essential to know to what extent people live in areas where coastal flooding and extreme weather events are expected to worsen and to what extent agricultural production is located in areas where water availability is expected to decline. In order to reduce risks resulting from climate change, it is most important that this spatially integrated information be available and put to use locally, as the impacts of climate change will be borne on particular localities. However, because of the global nature of climate and the likelihood that an increasing share of impacts of climate change will be felt in Africa, Asia and Latin America, it is also crucial to understand these risks globally.

Hazards faced by urban settlements are particularly important, not simply because urban areas concentrate people and their economic activities, but also because future population growth and economic growth are expected to be concentrated in these locations. The 2007 revision of *World Urbanization Prospects* (United Nations, 2008) projects that during the first half of the 21st century, the world's urban population will grow by about 3.5 billion, while its rural population will decline by about 0.5 billion. Since the spatial distribution of urban settlements is different from that of rural settlements, urbanization will play a role in how the burden of risks associated with climate change will shift in the future. Most urbanization will occur in Africa and Asia, but it is important to know more precisely where this growth will occur. Countering the increased population density inherent in urbanization, urban areas are also expanding spatially, thus reducing urban density. This phenomenon is most advanced in North America's sprawling suburbs but is also occurring in most other parts of the world (Angel et al., 2005).

This chapter draws heavily on the research of its authors, though it aims to comment more generally on the complexities of integrating spatial and non-spatial data to address dynamic, contemporary concerns. Much more could be said by the research and planning community that undertakes this type of interdisciplinary data integration.

What is meant by integration of data in a spatial framework?

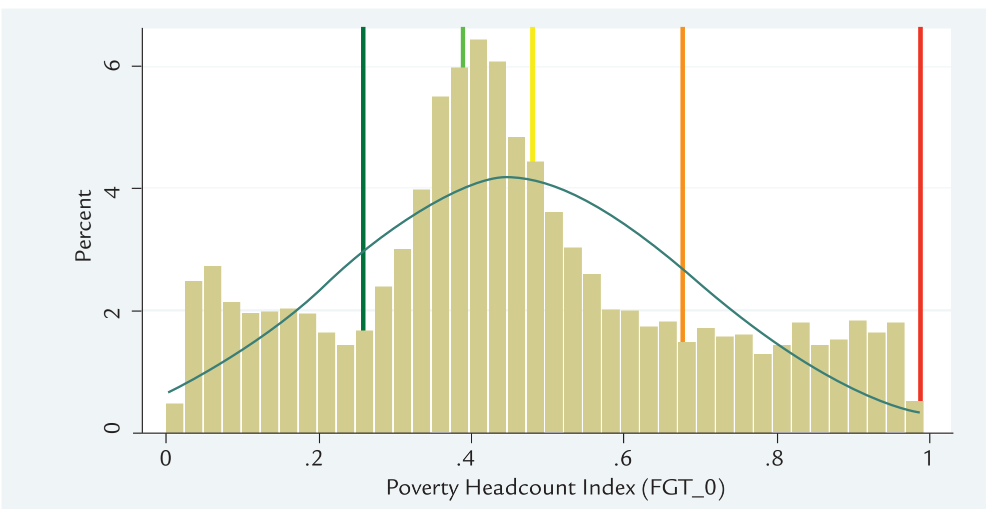
The urban and rural population and land at risk of sea level rise has been described elsewhere (Chapter 5; McGranahan et al., 2007). Here an example using poverty mapping data rather than population data is drawn on to demonstrate that population is not the only attribute that can be described in spatial terms. For example, one might wish to know the number of poor people living at risk of sea level rise—

i.e., within a 10-metre rise in elevation that is contiguous to the sea coast (the low elevation coastal zone or LECZ)—because the poor are expected to be more vulnerable to the untoward effects of climate change and the least able to adapt (Hardoy and Pandiella, 2009; Hardoy et al., 2001).

Figure 13.1 shows the frequency distribution of poverty among districts (third-order administrative units) in Viet Nam (Muñiz et al., 2008). Along the x-axis is the proportion of each district that is poor (ranging from 0 [no one] to 1 [everyone]). The y-axis indicates the percentage of districts that have each level of poverty. These data are derived from the World Bank's Small Area Estimation (SAE) of Poverty (Elbers et al., 2003, 2005; Minot, 2000; Minot et al., 2003), and though these units correspond to spatial boundaries for the administrative units, they were originally reported in a table (not shown). By matching this table of poverty attributes with corresponding spatial boundaries, these units can be rendered spatially, as shown in Figure 13.2 (Muniz et al., 2008). Without the map, it is not possible to recognize the spatial pattern of high poverty (shown in red) that is concentrated in central and northern Viet Nam. This type of data integration is relatively straightforward since the same administrative units are used as reporting units for both the table and the spatial boundaries. Even so, this type of integration requires common codes (or names) in each data set with which to link attributes (e.g., poverty rate) with spatial information. This linking is a simple function in geospatial and other statistical software packages, without which this task would be cumbersome and error-prone.

Additional spatial data integration in Viet Nam not only shows that some units are coastal, but also allows the land area and population at risk of seaward haz-

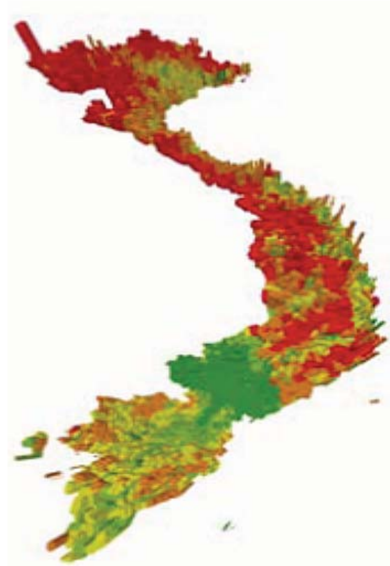
Figure 13.1: Frequency Distribution of Per Cent Poor in Each District, Viet Nam



Note: The coloured lines from green to red indicate country-specific quintiles of poverty rates by district (third-order administrative unit).

Source: Muñiz et al., 2008.

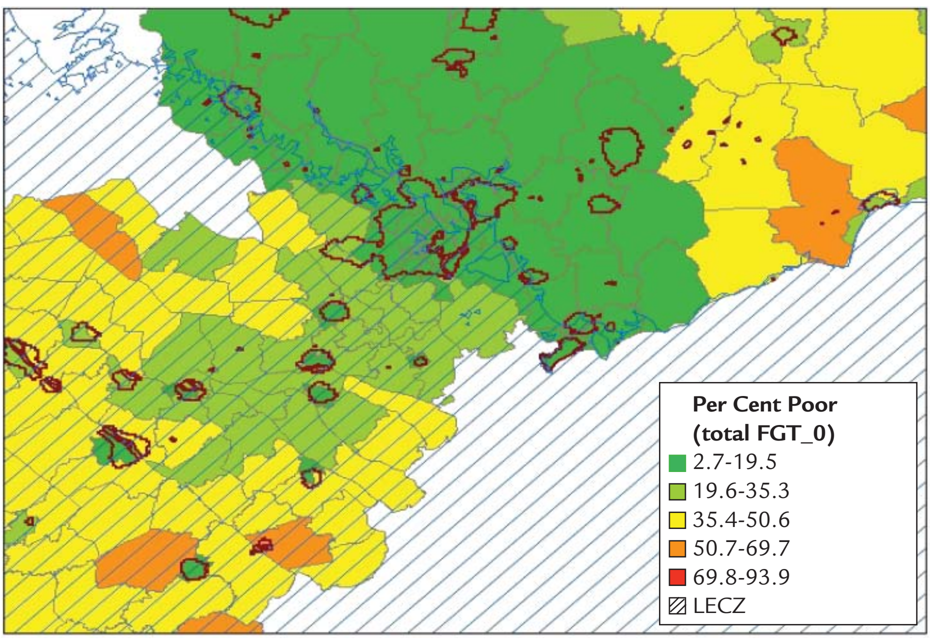
Figure 13.2: Poor in Each District, Viet Nam



Note: The shading from green to red represents country-specific quintiles of poverty rates by district (third-order administrative unit), and the extrusion represents numbers of poor persons (expressed in tens of thousands).

Data source: Minot, 2000.

Figure 13.3: Per Cent Poor, LECZ, and Urban Footprints, Viet Nam



Data sources: Minot, 2000; Muñiz et al., 2008; and McGranahan et al., 2007.

ards to be calculated according to a set of systematic characteristics accounting for coastal proximity and elevation. Figure 13.3 shows the data layers: poverty level in green to orange hues, the LECZ in blue hatching (McGranahan et al., 2007) and the urban footprint, derived from satellite imagery of night-time lights (from the Global Rural-Urban Mapping Project [GRUMP]), outlined in brown (described below and in Balk, 2009). In this view, the LECZ covers a large portion of the land area, illustrating that, in this low-lying delta, it would be a huge underestimate to define the

Table 13.1: Estimates of Urban Poor at Risk of Climate Change Coastal Hazards, Viet Nam

	% Poor	Number of Poor	Number of 1 km cells
Non-LECZ City	26.6	342,030	79
Cities with any land area within the LECZ			
LECZ City, Non-LECZ Land	20.30	413,623	36
LECZ City, LECZ Land	28.0	2,112,987	131

Data sources: Minot, 2000; Muniz et al., 2008; McGranahan et al., 2007.

vulnerable population as only those residing in districts bordering the sea coast. Figure 13.3 also makes it possible to distinguish urban population and land (within the GRUMP urban footprints) from the rural (outside the footprints).

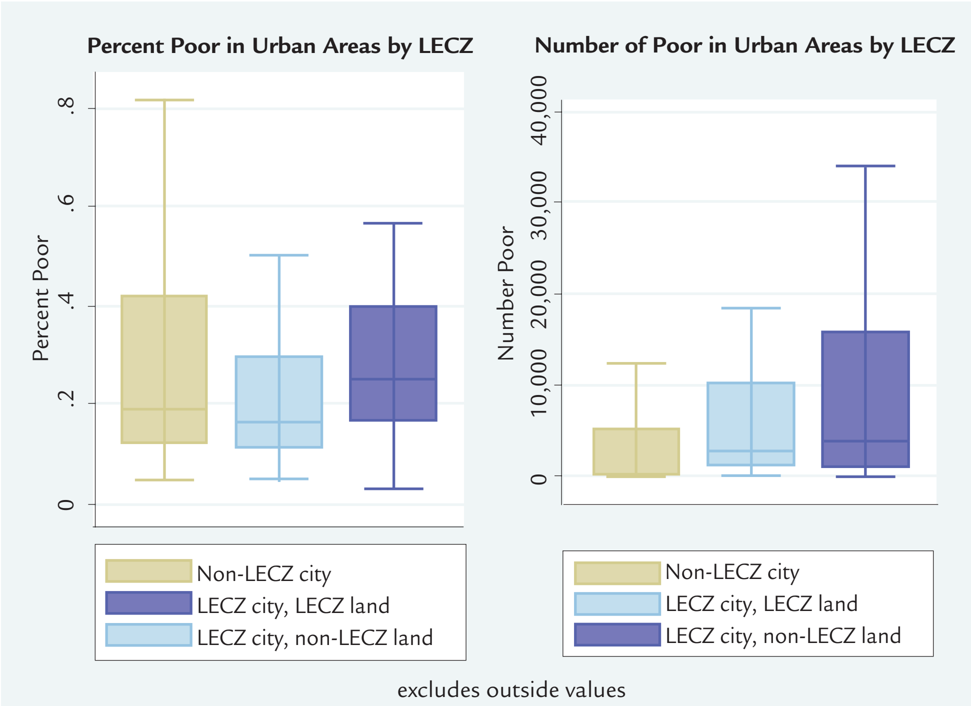
From this integration, it is possible to summarize populations at risk, as in Table 13.1 and Figure 13.4. Table 13.1 summarizes three types of exposures: persons who live in cities entirely outside the LECZ, persons who live outside the LECZ but in cities with some land area in the LECZ and urbanpersons living in the LECZ. Vulnerability—here expressed as poverty—is indicated two ways: by the proportion of the exposed population that is poor and by the total number of exposed persons who are poor. Though the poverty rates are similar among residents of the LECZ and residents of non-LECZ cities, those living in the LECZ outnumber poor people living in non-LECZ cities by five to one. Figure 13.4 shows this graphically. Only data integration in a spatial framework makes these estimates possible.

Table 13.2: Comparison of Spatial and Tabular Approaches to Estimating Urban Population Distribution: Population Density in the Urban Low-elevation Coastal Zone

Country	Average Resolution of Underlying Census Data (km)	Distribution of Urban Areas	Urban LECZ Population ÷ Urban LECZ Land: Tabular (National Aggregate) Estimates	Urban LECZ Population ÷ Urban LECZ Land: Spatially-derived Estimates	Ratio of Tabular to Spatial Estimates
Viet Nam	8	Most in LECZ	4,489	3,317	1.4
Philippines	14	Many in the LECZ	13,284	3,636	3.7
Brazil	29	Biggest ones in the LECZ	11,523	939	12.3
South Africa	1	Most not in the LECZ	92,413	1,463	63.2
Congo, Democratic Republic of the	184	Most not in the LECZ	146,533	102	1,431.2

Data sources: Minot, 2000; Muniz et al., 2008; McGranahan et al., 2007.

Figure 13.4: Percent Poor and Number of Poor in Urban Areas, Viet Nam



Data sources: Minot, 2000; Muñiz et al., 2008; McGranahan et al., 2007.

Data integration challenges in general

Estimates that are derived strictly from national-level aggregates and that do not formally integrate demographic and environmental data using a spatial framework are likely to produce highly inaccurate estimates. Neither cities nor persons (regardless of whether they are urban or rural residents) are uniformly distributed across national territories. New analyses have identified some patterns of city and population distribution vis-à-vis geographic characteristics. In a global study, McGranahan et al. (2007) found that LECZs are disproportionately urban compared to other ecozones such as drylands. Further, they found that 75 per cent of all countries have their largest city in the LECZ. In Table 13.2, two estimates of population density in urban areas in the LECZ are examined for five countries. Estimates based on national-level aggregates that are expressed only as tables (tabular estimates) are compared to estimates based on overlaid spatial data comprising administrative, night-time lights and settlement attribute data (see below and Balk, 2009, for more on GRUMP methodology). The five countries in Table 13.2 vary in the spatial resolution of their census data; low numbers indicate many smaller census units. Viet Nam is an example of a country with most of its urban areas (and much of its land area) in the LECZ.

Scale and Resolution

In order to produce robust urban estimates, data must be spatial and must be of sufficient resolution and scale. Recent efforts have produced basic descriptions of the population distribution of urban areas in a spatial framework (Balk, 2009; Montgomery and Balk, forthcoming). These efforts are an important departure from prior approaches: Population estimates for cities can now be rendered in physical space. However, there is currently no spatial database that allows for the estimation of changes in urban population and area at a global scale. While moderate- and high-resolution data permit change estimation at the scale of a city or a handful of cities, much work in methodology, data processing and validation remains to be done before a globally consistent, spatial-temporal view of urban areas exists (Small, 2005).

While spatial data are necessary, they may not be sufficient: The properties of the spatial data matter. Especially when evaluating spatially-specific urban population data with respect to environmental data, resolution must be at a scale that is appropriate for urban-area-level analysis; that is, the unit of analysis must be fine enough to adequately capture variation within and around the urban area. Currently, the resolution of most geophysical data—such as the historical climate record and future climate predictions, as well as many disaster databases—is much coarser than that of the city, preventing meaningful analysis of these geophysical data at the city and sub-city scales. Just like national-level aggregates, coarse spatial data misleadingly distribute place-specific characteristics over a too-large

Table 13.3: Resolution of Selected Spatial Data and Size of Average Urban Areas

East-West Arcs		Distance per side (km)	Area (km^2)
		at equator	
5 degrees	Climate models		
1 degree		111.32	12,392.1
0.5 degee	Rainfall, precipitation models	55.66	3,098.0
	Average urban area, 1 million + persons		**1,650.0**
5 minutes (0.083°)	Gridded Populatin of the World (GPW) v1	9.30	86.5
	Average urban area, <1 million persons		**70.0**
2.5 minute (0.042°)	GPW v3	4.65	21.6
30 arc-sec (0.0083°)	GRUMP, SRTM, Ecozones	0.93	0.9
1 arc-sec (0.000278°)	Landsat	0.03	0.0009
	Quiksat, Ikonos	0.001	0.000001

Note: Average urban area size is determined by GRUMP for cities in Africa, Asia and South America. Distance and area calculated at the equator.

region. If the true characteristic is variable across the areal unit of analysis (the 'cell' in geophysical parlance), coarsely resolved data will mask intra-cell variation.

Table 13.3 shows the resolution of some key environmental data and the average size of urban areas in Africa, Asia and South America. Urban localities are much smaller than, and would be subsumed by, the cells of many spatial data sources. This means that no variation in these environmental data would be observed within an urban area, a presumption that is clearly false. High- and moderate-resolution satellite imagery can often provide a smaller-celled view into urban areas, but these images typically show vegetation and other features rather than climate data. Appropriate use of integrated data critically depends on issues of scale and precision.

Satellite data: More than just a pretty picture?

Satellite data may be the most objective means by which to systematically identify urban areas (Potere and Schneider, 2009). At a local or regional scale, these data may be used to identify change in built-up area or land use. Inexperienced users of satellite imagery must remember that without expertise an image is not much more than a pretty picture. It takes considerable knowledge to evaluate, interpret and classify satellite data (Small, 2005). With expert evaluation, satellites can reveal a great deal about vegetation, permanent lights and built-up areas, but these data are not ready to use outside of a spatial framework. It usually takes interdisciplinary teams of researchers to translate satellite data into outputs that can be integrated with census-type population data (Small and Cohen, 2004).

Satellite data have some shortcomings. The data never provide the city names, identifying codes or statistical reporting concepts (e.g., city proper, urban agglomeration) that are commonly attached to population censuses. This may not be an important shortcoming when studying a single location, but it is a significant limitation when working across many localities or at a regional or global scale. Satellite data are costly to process, and many types are prone to cloud cover, which obscures the features of interest. Additionally, analysis is more subjective than the typical social scientist is used to.

Satellite data, however, also have unique strengths. Unlike surveys, censuses or even administrative boundary data, for which the cost of collection tends to be borne by countries, the cost of satellite imagery is borne in large part by the data collector. (Sometimes, a portion of that charge is passed on to data users as fees.) The data may be supplied by international experts or their space-borne technology. This means even countries with limited resources can be studied with high-quality data.

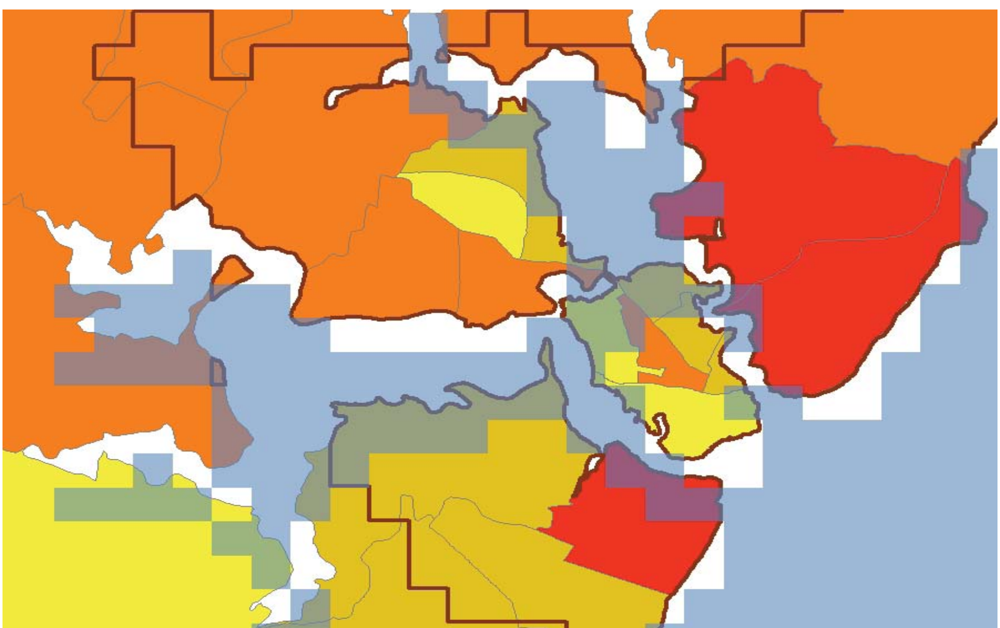
Although some satellite data can be a time series, few global studies of integrated data are. In Chapter 5 of this volume, a study is described wherein an integrated data set linking satellite-derived urban footprints to names and population values is constructed, but it does not have time-varying spatial data for cities. Though much has been learned from data integration, those lessons have not resulted in an automated process for continued future integration. Integration of disparate data sets is—and should be expected to remain for the foreseeable future—a highly labour-intensive enterprise, even with considerable programming aids.

Demographic data: More than just head counts?

The global demographic record is built from a variety of data sources: national censuses, vital registration data and household survey data. Sometimes censuses or even vital registrations are rendered at a fine sub-city resolution, but surveys typically are not. Survey data tend to be relied on mostly in countries with weak census or vital registration systems. How to best piece together these multiple and fluid data types to build a fuller record for city (and sub-city) demographic data at a global scale is an open question.

The demographic record has its own shortcomings (see Montgomery et al., 2003). There is no globally consistent or systematic set of demographic estimates for the world's cities, except for the most populous cities and those large enough to be comprised of standard census reporting units (such as counties or districts). For most cities of the world, there are no data on age distribution, fertility, mortality or migration. Even when this type of information is available at the city scale, it is rare that it also exists for neighbourhoods within cities (Weeks et al., 2007). In some cases, urban estimates exist that are aggregated to the national or first-order subnational units, but this record cannot be translated to specific cities or even classes of cities based on their population size.

Figure 13.5: Mismatch Example, LECZ and Per Cent Poor, Kenya



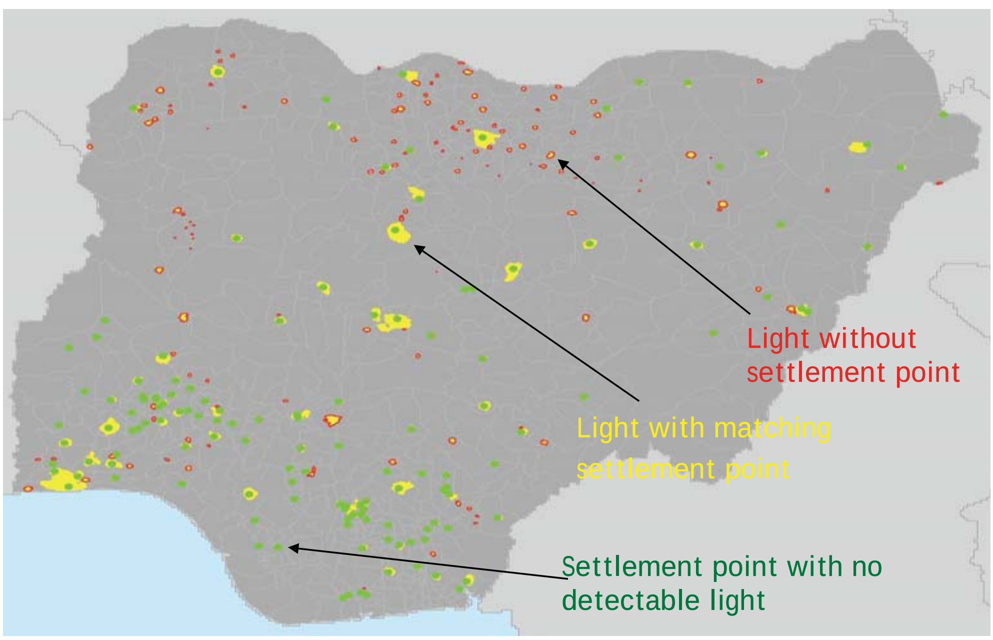
Data sources: CIESIN et al., 2008; Ndeng'e, 2003; McGranahan et al., 2007.

Issues that arise upon integration

The precision and accuracy of the various data layers matter, and small differences may be amplified when different data sources are integrated. Figure 13.5 shows how small differences in the precision of the administrative boundary-based coast line of Kenya (data initially supplied by the Kenyan National Statistical Office [Ndeng'e, 2003]) and the LECZ layer (from Shuttle-Radar Topography Mission Digital Elevation data) result in evident gaps between the land (yellow-red hues depicting the distribution of poverty) and sea (in blue). White space shows the mismatch. This can result in the mischaracterization of the population at risk of coastal flooding and other seaward hazards. Whether one or both data sources are inaccurate is a matter yet to be determined. All data integration is at risk of this type of mismatch. Even within countries, different data users might modify boundary data to suit their needs. Some agencies wish to include water bodies in jurisdictional boundaries while others wish to omit them. Sometimes there is agreement on how to reconcile multiple sets of boundaries, but often there is not.

To study urbanization, the Global Rural-Urban Mapping Project (see Balk, 2009) uses night-time lights satellite data as a proxy for urban areas, combined with population settlement data. While night-time lights are the most systematic urban footprint, it is evident that, in some locations at least, they are an imperfect proxy for urban areas. Lights can be seen where there is no identifiable settlement, and some settlements have no corresponding light. Figure 13.6 shows these mismatches. GRUMP accepts the latter type of location and estimates a settlement size based on other known settlements, but disregards lights without points, as these are believed

Figure 13.6: Night-time Lights, Nigeria



Data source: CIESIN et al., 2008.

to be unlikely to be true settlements (e.g., lights produced by oil flares in oil-producing countries are not permanent human settlements). It is, however, possible that some lights without points were omitted from the national statistical reporting of settlements for other reasons (such as political ones) and that the lights are a means for challenging census results. Either way, it is integration that opens up these important inquiries for each analyst to adjudicate for him- or herself.

Every data integration effort will require some subjective judgement on the part of researchers. In addition to the examples above, the authors of this chapter have faced several puzzles. For example, if two settlements are identified with near-identical geographic coordinates, and have names that match but for one letter, are they the same city or different cities? If a settlement falls outside an observed light, should that settlement be considered part of the same urban area as the light, given the spatial measurement error of the lights (about 3km)? Should settlements falling within 3km of a light be assigned to the light? (Perhaps a band of 3km of the lights on both sides of the light boundary should be treated as less sure matches.) If a light appears very near the border of two countries, is the precision of the country boundary great enough to definitively place the light in one country? Does a many-to-one match between settlements and urban areas make sense if the settlements are meant to represent a relatively large metropolitan area? The answer to these dilemmas may depend on the purpose of the study.

Most important, when subjective data integration decisions are made, a transparent record of these decisions and the reasoning behind them must be kept, and an effort must be made to develop a systematic approach for analogous issues. Both of these concerns can be addressed by programming the data integration process in statistical software such as Stata or SAS. Python is particularly useful for working with spatial data, as it can create output easily read by map making software such as ArcGIS. Careful, well-documented programming is crucial to a study's repeatability and transparency, a point that cannot be made strongly enough.

Conclusions

It is not known precisely where climate change will occur. But to prepare for those changes, both climate-change and social-science researchers need to adopt a spatial framework of analysis that is attentive to current and future population concentrations in urban areas. The integration of these data is essential to understanding the risks that populations face from climate change.

Understanding the construction of integrated data is essential for data users, even for those users uninvolved in the integration process. Development and planning efforts for improvements in urban drainage or sanitation, for example, require both spatial and population data; so does projecting where migration will swell the populations of towns and cities that lie in the path of risk. National economic strategists need to be made aware of the implications of locating special economic zones in sensitive areas and of promoting coastal development in what will become environmentally risky sites. Secondary data users, such as developers and planners, must inform themselves of methodological decisions made in the construction of any integrated data sets, because these decisions will impact the interpretability—and conclusions—of analyses resulting from these data.

References:

Angel, S., et al. 2005. "The Dynamics of Global Urban Expansion." Washington, D. C.: Transport and Urban Development Department, The World Bank.

Balk, D. 2009. "More than a Name: Why is Global Urban Population Mapping a GRUMPy Proposition?" Pp. 145-161 in: *Global Mapping of Human Settlement: Experiences, Data Sets, and Prospects,* edited by P. Gamba and M. Herold 2009. Boca Raton, London and New York, N.Y.: Taylor and Francis Group.

Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IFPRI); The World Bank; and Centro Internacional de Agricultura Tropical (CIAT). 2008. Global Rural-Urban Mapping Project (GRUMP), Alpha Version. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Data available in beta version at http://sedac.ciesin.columbia.edu/gpw.

Elbers, C., J. O. Lanjouw, and P. Lanjouw. 2005. "Imputed Welfare Estimates in Regression Analysis." *Journal of Economic Geography* 5(1): 101–118.

______. 2003. "Micro-level Estimation of Poverty and Inequality." *Econometrica* 71(1): 355–364.

Hardoy, J., and G. Pandiella. 2009. "Urban Poverty and Vulnerability to Climate Change in Latin America." *Environment and Urbanization* 21(1): 203–224.

Hardoy, J., D. Mitlin, and D. Satterthwaite. 2001. *Environmental Problems in an Urbanizing World*. London: Earthscan.

McGranahan, G., D. Balk, and B. Anderson. 2007. "The Rising Tide: Assessing the Risks of Climate Change to Human Settlements in Low-elevation Coastal Zones." *Environment and Urbanization* 19(1): 17–37.

Minot, N. 2000. "Generating Disaggregated Poverty Maps: An Application to Vietnam." *World Development* 28(2): 319-331.

Minot, N., B. Baulch, and M. Epprecht. 2003. "Poverty and Inequality in Vietnam: Spatial Patterns and Geographic Determinants." Washington, D. C.: International Food Policy Research Institute.

Montgomery, M. R., and D. Balk. Forthcoming. "The Urban Transition in Developing Countries: Demography Meets Geography." In: *Global Urbanization in the 21st Century*, edited by E. Birch and S. Wachter. Philadelphia, PA: University of Pennsylvania Press.

Montgomery, M. R., et al. (eds.) 2004. *Cities Transformed: Demographic Change and Its Implications in the Developing World*. Washington, D.C.: National Academies Press.

Muñiz, M., et al. 2008. "Charting a New Course in the Analysis of Poverty and Inequality: Contributions from a Spatial Framework." Revision of a paper presented at the Annual Meeting of the Population Association of America, Los Angeles, CA, 30 March-1 April 2006.

Ndeng'e, G. 2003. *Geographic Dimensions of Well-being in Kenya: Where are the Poor: From Districts to Locations*. Nairobi, Kenya: Central Bureau of Statistics.

Potere, D., and A. Schneider. 2009. "Comparison of Global Urban Maps." Ch. 13 in: *Global Mapping of Human Settlement: Experiences, Data Sets, and Prospects,* edited by P. Gamba and M. Herold. 2009. Boca Raton, FL, London and New York, N.Y.: Taylor and Francis Group.

Small, C. 2005. "Global Analysis of Urban Reflectance." *International Journal of Remote Sensing* 26(4): 661-681.

Small, C., and J. E. Cohen. 2004. "Continental Physiography, Climate and the Global Distribution of Human Population." *Current Anthropology* 45(2): 269-277.

United Nations. 2009. *Handbook on Geospatial Infrastructure in Support of Census Activities*. Studies in Methods Series F. No. 103. New York: Statistics Division, Department of Economic and Social Affairs, United Nations.

United Nations. 2008. *World Urbanization Prospects: The 2007 Revision*. Population Database. New York, N. Y.: Population Division, Department of Economic and Social Affairs, United Nations.

Weeks, J. R., et al. 2007. "Can You Spot a Neighborhood From the Air: Defining Neighborhood Structure in Accra, Ghana." *GeoJournal* 69(1-2): 9-22.