The Distribution of People and the Dimension of Place: Methodologies to Improve the Global Estimation of Urban Extents

Deborah Balk[1]
Francesca Pozzi[1]
Gregory Yetman[1]
Uwe Deichmann[2]
Andy Nelson[3]

[1]CIESIN, Columbia University
contact: dbalk@ciesin.columbia.edu

[2] The World Bank
[3] University of Leeds

8 DECEMBER 2004

Acknowledgments

The study has been undertaken by Columbia University's Earth Institute's Center for International Earth Science Information Network (CIESIN), in collaboration with the International Food and Policy Research Institute (IFPRI), the World Bank, and the International Center for Tropical Agriculture (CIAT) with primary support to CIESIN from National Aeronautics and Space Administration under Contract NAS5-03117, International Food Policy Research Center (IFPRI) under Contract 2001-X046COL, and World Conservation Monitoring Center (WCMC) for the Millennium Ecosystem Assessment under Contract CU-02225301 Adam Storeygard, Greg Booma, and Bridget Anderson, were hugely helpful in pulling the final version of this paper together. We also thank Marc Levy, Gordon McGranahan, Stan Wood, Jordan Chamberlin and Kate Sebastian who asked a continuous stream of good questions to keep us on track. Additionally, we thank to Ron Rindfuss and Mark Montgomery for comments on early drafts of this paper.

We are indebted to the many individuals for their participation in the development of the associated data products: Edwin Adkins, Sonya Ahamed, Bridget Anderson, Melanie Brickman, Greg Booma, Jordan Chamberlin, Betty Hsiu-tsiu Chang, Jessica Forrest, Christine Grimando, Robby Gutmann, Erika Helms, Glenn Hyman, Malanding Jaiteh, German Lema, Marc Levy, Rosalba Lopez, Lisa Lukang, Melissa Neuman, James "Cuz" Potter, Kate Sebastian, Chris Small, Adam Storeygard, Maarten Tromp, Katya Vasilaky, Keelia Wright, Stan Wood, and Xiaoshi Xing.

Abstract

What is known about the urban world is largely derived from local knowledge. This paper showcases substantial efforts at new data integration with existing technologies to develop a new suite of global datasets on urban population and extents. These new databases far surpass past efforts to construct a systematic global database of urban areas by combining census and satellite data and methods of analysis, in an integrated geospatial framework. The resulting data allow for inquiry into analysis of urban issues and population by environmental and other ecological characteristics in novel ways. This paper focuses on the methodologies employed to construct these new datasets. Summary results regarding population distribution at continent- and global-levels are also given, as well as suggestions for future research.

INTRODUCTION

Human settlements occupy a relatively small fraction of Earth's surface area but their extent and distribution have significant impacts on their surroundings, both from an environmental and a socio-economic perspective. By 2007, it is estimated that over half of the world’s population will reside in urban areas (United Nations, 2002). Despite increasing knowledge about the characteristics of urbanization, little is known about its spatial dimensions. For example, only guesswork has provided prior estimates on the share of the world’s inhabited land area that is urbanized, or on the classification of the world’s population by city-sizes other than the very largest ones (UN, 2002; UN 2002b). Even when cities are tallied by their population sizes and types (such as agglomerations), little effort has gone into systematically capturing the spatial dimension of these places.

There are also increasing demands for greater specificity in defining the impacts of agricultural change and development, particularly with regard to the likely impacts of policy, technology, and institutional changes on poverty (Wood *et al*. 1999). It is not only the growth of urban areas, but also the interconnection between urban and rural areas that is important to understand these impacts. Improved knowledge of the spatial distribution of urban and rural population is extremely important for assessing socioeconomic, demographic and environmental change in urban and rural areas.

In order to understand and study the impacts of urbanization, population and physical factors need to be made available as detailed, spatially disaggregated data and reduced to comparable scales. Although there is ample research on urban growth as separate geographic and demographic phenomenon, there is little research and no dataset in which these parameters are integrated. This study proposes a new methodology to foster this integration. That methodology is the focus of this paper, along with the discussion of some analytical results and suggestions for future research.

BACKGROUND

While many environmental data are available already as spatial datasets, demographic data tend to be collected by administrative units and therefore require some form of spatial allocation to convert irregularly shaped census units to globally or regionally consistent population grids. Several researchers and institutions in recent years have used new methods and data to map the global distribution of human population. The first major effort to generate a consistent global georeferenced population dataset was the Gridded Population of the World (GPW, Balk and Yetman, 2004), produced at the National Center for Geographic Information Analysis (NCGIA) in 1995 (Tobler *et al.*, 1997), and updated by CIESIN in 2000 (Deichmann *et al.*, 2001). The inputs to the GPW dataset are solely administrative boundary data and population estimates associated with those administrative units. Other efforts then followed, generally incorporating satellite data and other ancillary data. In the LandScan database, for example, developed by the Oak Ridge National Laboratory, "sub-national census counts are apportioned to each grid cell based on likelihood coefficients, which are derived from proximity to roads, slope, land cover, night-time lights and other data sets, (Dobson et al, 2000). A third effort by UNEP and partners includes an "accessibility" model, whereby access to roads is used as a means to reallocate population (UNEP, 1996a, 1996b, 2000, 2004, and CIAT, 2003).

Each of these approaches has strengths and weakness (see Dooley, 2004): GPW uses a simple areal weighting scheme for reallocation and the best possible census and administrative data available, but does not model population distribution within the administrative units. Its output resolution is 2.5 arc minutes. LandScan, conversely, does not use very high-resolution population input data, but uses an extensive model to reallocate people on a 30 arc second grid. Since some of the data used to reallocate persons may be outcomes of interest (e.g., distance to roads, or land cover), LandScan must be used with caution in studies involving environmental outcomes. GPW attempts to represent decennial population counts, whereas LandScan attempts to capture ambient—or at risk—population. The Accessibility model builds on the GPW tradition, but takes into account road networks and populated places in the reallocation of population. Unlike LandScan, only roads and populated places are used, and there is no explicit effort to capture the ambient quality of the LandScan approach. Like GPW, its output resolution is also moderate (2.5 arc minutes). The study presented here forwards a new methodology to extend the initial GPW efforts, to improve output resolution, and to overcome some of the use limitations of LandScan, by accounting explicitly for urban areas.

In the process of reallocating population to urban areas, it is necessary to first construct a spatial database of those areas. To accomplish that, satellite data were a necessary additional input. The Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS) "night-time lights" dataset has been increasingly used to estimate aspects of human activity at the global level. Satellite imaging of stable anthropogenic lights provides an accurate, economical and straightforward way to map the global distribution of urban areas, and several studies of DMSP-OLS stable night lights have shown encouraging agreement between temporally stable lighted areas and various definitions of urban extent. The stable lights for the 1994-1995 time period have been produced for most of the Americas, Europe, Asia and Northern Africa (Elvidge *et al.*, 1997b) and have been used for a variety of applications. As with the population databases, there have been some relevant uses of the night-time lights data. For example, Sutton *et al.* (1997) examined the potential use of the stable lights data to revise estimates of the population of urban areas around the world; Imhoff *et al.* (1997a, 1997b) used the stable lights to estimate the extent of land areas withdrawn from agricultural production; and Elvidge *et al.* (1997b, 1997c) found that the area lit from the stable lights of individual countries is highly correlated to the Gross Domestic Product. More recent efforts include a pilot study to map urban land cover by fusing the night-time lights dataset with GPW and a MODIS-derived land cover classification (Schneider *et al.* 2003), and Pozzi *et al*. (2003), which maps global urban population by integrating GPW and the night-time lights. None of these efforts, however, attempts to merge the lights directly with city-level census data to derive population estimates of urban areas. Thus, using GPW as a base, in 2000, CIESIN, IFPRI, the World Bank and CIAT began the multi-year effort of the Global Rural Urban Mapping Project (GRUMP). This effort aimed not only to construct an improved population grid, but also to construct a globally consistent database of urban areas.

The methods presented in this paper are based on the premise that data may be combined from several disparate data streams: census (or census-type) inputs on the population size (of settlements and administrative areas), with associated names; and two key pieces of geographic information, latitude and longitude of settlements, and boundaries for administrative areas and urban extents (the latter being identified using the night-time lights and ancillary geographic datasets). The resulting dataset consists of three distinct products: a human settlements database, an urban extent database, and an urban-rural population grid or surface.

METHODOLOGY

This project aimed to produce three databases that could be used in an of themselves or in combination with each other. These databases are: 1) Human Settlements points, 2) an Urban Extent 'Mask' and 3) an population grid or surface where population is reallocated in urban areas. Here we describe the methodology—and underlying data—used in the development of these datasets.

Human Settlements

Although there are many gazetteers listing populated places, few of these contain population estimates for those named places. Similarly, databases of city population estimates rarely include geographic information, such as the latitude and longitude coordinates let alone area or other spatial information about each urban area. Several collections, such as the UN *Demographic Yearbook* (UN, 2002b) or the UN's *World Urbanization Prospects* (UN 2002a) include coordinates, and type of urban area, for places of 100,000 and 750,000 persons, respectively.

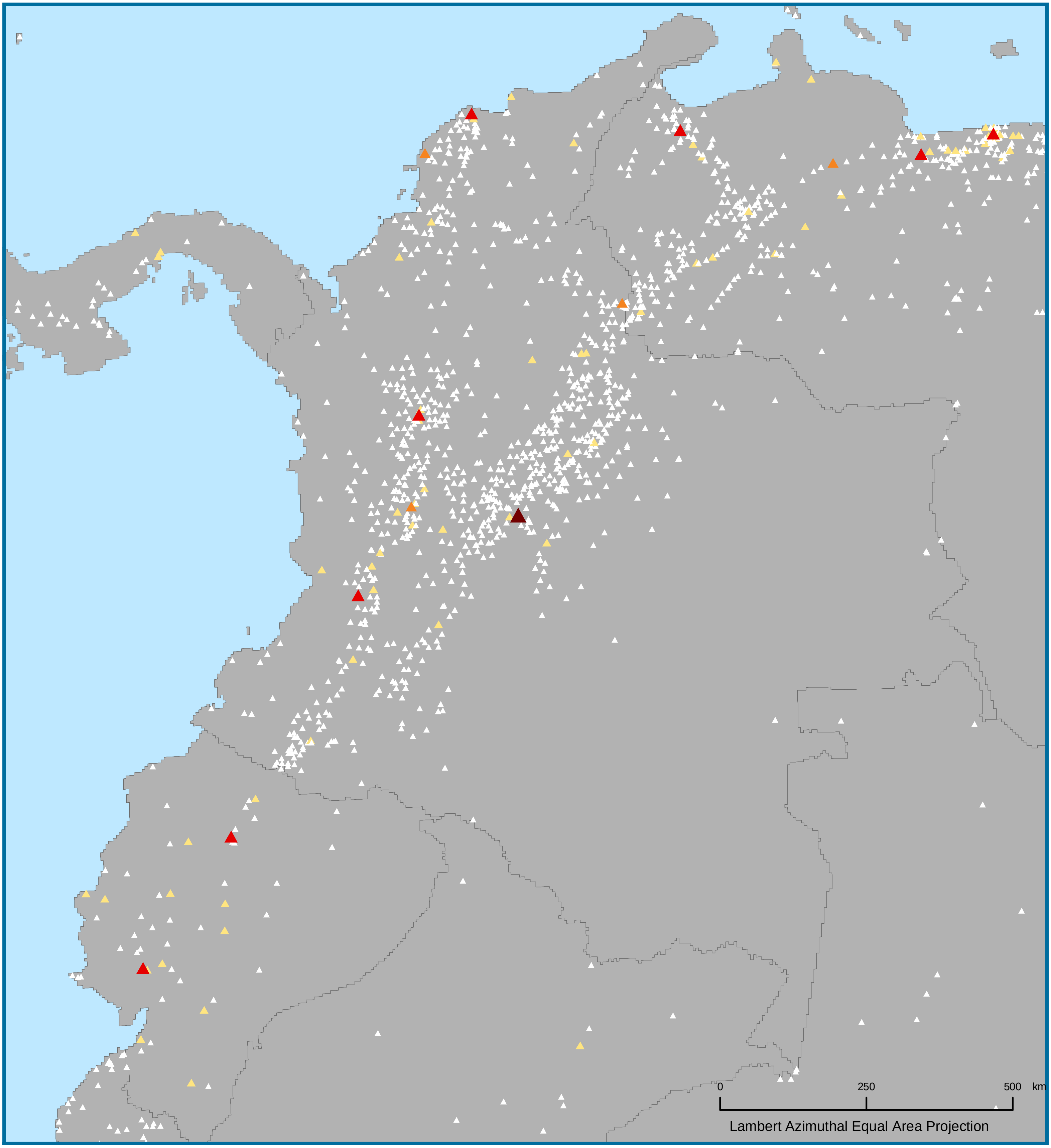
The GRUMP human settlements database is a global database of cities and towns of 1,000 persons or more, where each settlement is spatially represented as a point, and has associated tabular information on its population and data sources. Population data were gathered primarily from official statistical offices (census data) and secondarily from other web sources, such as Gazetteer and CityPop[1], or from specific individual databases when official statistical databases were not available. Based on the data available and applying UN growth rates, we estimated population in 1990, 1995, and 2000. In some cases, the records for cities and town included latitude and longitude coordinates. For those where coordinates were not available we matched the settlement name and administrative units with the National Imagery and Mapping Agency (NIMA) database of populated places. Although we automated the matching of places to coordinates found in the NIMA database, this process still required considerable validation with other sources, and sorting through multiple options (i.e., NIMA often provides several, slightly varying sets of coordinates to match a single place name in a given administrative unit). Nonetheless, we did not have means to consistently validate the positional accuracy of the NIMA coordinates for all cities, and therefore some of the cities and towns might not be accurately geolocated. Table 1 shows the distribution of data sources, while figure 1 shows the settlement points database in a portion of South America, by population size.

[1] Gazetteer (www.gazetteer.de) and CityPop (www.citypop.de) are two publicly available datasets which contain collections of information from varying sources, not always specified.

Table 1. Distribution of population data sources by continent.

Source	Asia	Africa	Europe	North America	South America	Oceania	World	Percentage
Census	9,666	2,525	6,641	27,493	4,889	451	51,665	73.2
World Gazetteer	2,210	561	4,799	243	74	168	8,055	11.4
CityPop	1,363	319	3,364	443	304	179	5,972	8.5
Others	7	737	0	119	4,002	1	4,866	6.9
Total	13,246	4,142	14,804	28,298	9,269	799	70,558	100.0

Note: “Census” include also data from Statistical Yearbooks, Statistics from State Departments and on-line Statistical Offices. “Others” include data from AFCities, ASCities, the World Bank, CIA Factbooks and CELADE (Latin American and Caribbean Demographic Centre).



Settlements by Population Size, 2000

<100K | 100 to 500K | 500K to 1 Million | 1 to 5 Million | 5 Million +

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), 2004. Global Rural-Urban Mapping Project (GRUMP): Urban Extents. Palisades, NY: CIESIN, Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw.

Urban Extent 'Mask'

While much research has been undertaken to describe the extents, landscapes, and changes of local urban areas (Seto, 200X, Weeks et al. 200X) and applied to a variety of environmental and ecological changes (e.g., Tatem and Hay 2004), none has been undertaken in a systematic way at the global level. Efforts to use moderate-resolution vegetation-detecting optical satellite imagery prove too costly, and inconsistent, to detect built-up areas globally (Small, 200x, Tatem *et al*. 2004). While other satellite data, such as new radar data from the SRTM, holds promise for the detection of built-up areas globally (Ngheim et al., 2001), the effort has yet to be attempted.

The GRUMP Urban Extent Mask attempts to somewhat crudely represent the extents associated to the human settlements. In particular we now describe the sources of the physical extents of the settlements and the methodology to assign population from the point database to the areal extents.

The physical extent of settlements has been derived largely from NOAA's the Night-time lights dataset (Elvidge et al., XXXX), which produced a composite data set of stable "city" lights using time series data from the DMSP-OLS for the period 1 October 1994 to 30 April 1995, where the pixel values are measurements of the frequency with which lights were observed, normalized by the total number of cloud-free observations. Additionally, we used one other global-scale dataset: Digital Chart of the World's Populated Places (DCW): an ESRI product originally developed for the US Defense Mapping Agency (DMA) using DMA data and currently available at 1:1,000,000 scale (1993 version). The "populated places" coverage is available for most countries and contains depictions of the urbanized areas (built-up areas) of the world that are represented as polygons at 1:1,000,000 scale.

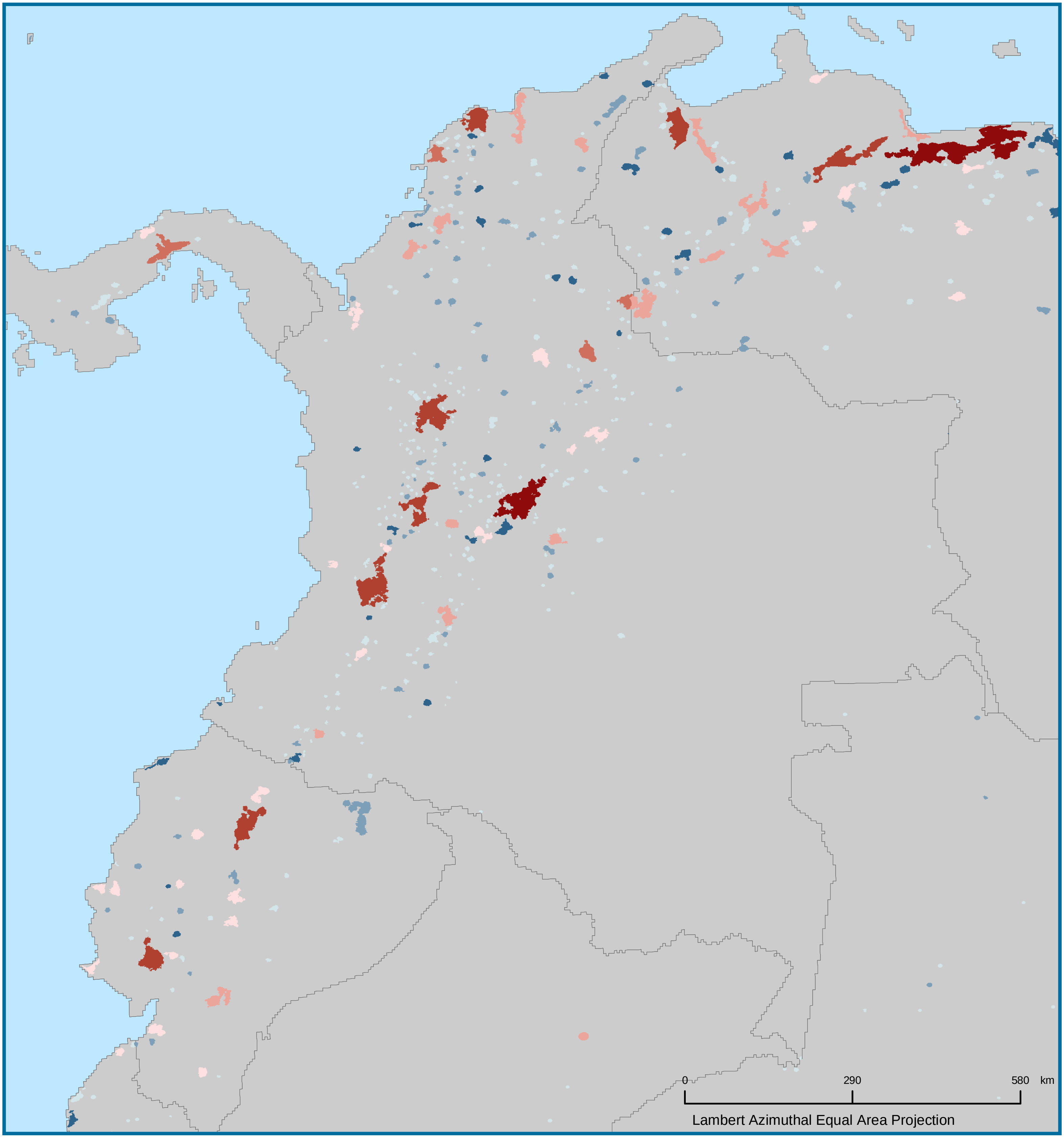
Additionally, in Africa and Latin America, two other sources of data were used to supplement the lights and DCW datasets. Tactical Pilotage Charts (TPC): standard charts produced by the Australian Defense Imagery and Geospatial Organization, at a scale of 1:500,000, originally designed to provide an intermediate scale translation of cultural and terrain features for pilots/navigators flying at very low altitudes where used for Africa. Each chart contains information on cultural, drainage/hydrography, relief, distinctive vegetation, roads, sand ridges, power lines, and topographical features. Settlements are reported both as polygons and points. Polygons and points were digitized for a number of countries, especially where lights and DCW data did not adequately delimit urban areas.

All the sources of urban extent were combined in order to obtain the maximum possible coverage for each country. The night-time lights were used as a baseline (due to its global coverage), and then polygons that did not intersect any existing light were added from other sources. Therefore, the total number of urban polygons in each country is the number of lighted areas plus all the other polygons identifying settlements that were not already identified by the lights. These polygons are characterized only by the basic geographic attributes, such as area and perimeter and do not have population attribute data. To create the Urban Mask from all the different sources we developed a hierarchical process, as follows.

First, we assign population from the points to the settlement extents, based on a 3km buffer distance.[2] If multiple points are present, as in the case of an urban agglomeration, the sum of the population of all points is assigned to the polygon. The name of the most populous city within the buffer is assigned to the polygon. Then, we estimate areal extents for points without polygons, based on a relationship between population size and areal extents for the points with known parameters. This relationship is derived from a logarithmic regression that predicts the expected geographic size of a place, given its population size. Where the number of observations is greater than 20 known relationships, we use country-specific regressions, otherwise regional regressions were used where regions were constructed according to the UN Statistics Division (UNSD) sub-continental regional codes. Based on these area values, we create circles, centered on the points. Finally, we add these newly created polygons to the existing ones to create a complete urban extent coverage for each country. Figure 2 shows the extent of urban places (as identified by the urban mask) in a portion of South America.

There are two main problems that arise when using the night-time lights dataset as a baseline to identify urban extents: the insufficient detection of small settlements that are not frequently illuminated and the blooming effect. While the first one has been reduced by using ancillary data, such as DCW and TPCs polygons, the blooming effect still remains unsolved. The blooming effect is an overestimation of the real extents of urban areas, and is believed to be dependant on the intrinsic characteristics of the sensor (Elvidge *et al.*, 2004). For a more detailed discussion and examples of the blooming effect, see the section under Discussion.

[2] The selection of the 3km buffer was based on the instrinic error associated with the night-lights of 2.7 km (Elvidge, YEAR).



Urban Extents

Urban places > 100k		Urban places 5k-100k	
	100,085 - 250,000		5,009 - 25,000
	250,001 - 500,000		25,001 - 50,000
	500,001 - 1 M		50,001 - 100,000
	1 M - 5 M		
	> 5 M		countries

SEDAC

CIESIN
Columbia University

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), 2004. Global Rural-Urban Mapping Project (GRUMP): Urban Extents. Palisades, NY: CIESIN, Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw.

Population grid with urban reallocation

The GRUMP population grid is a 30-arc second population distribution raster dataset that was developed by combining population data from the census administrative units and from the urban extent mask. To create the population surface, we developed a mass-conserving algorithm called GRUMPe (Global Rural Urban Mapping Programme) that reallocates people into urban areas, within each administrative unit. In particular we used data inputs from two vector sources: (1) Administrative polygons, containing the total population for each administrative unit; (2) Urban areas, containing the urban population for each area.

These two data sets are combined in such as way that an intermediate (polygon) data set representing the urban and rural areas, but which does not assign populations into those areas, is produced.[3] This intermediate dataset is then passed to GRUMPe, a stand-alone model written in C, that assigns population to each new polygon and labels it as rural or urban. Typically, the algorithm works on a country-by-country basis and uses the following pieces of information: The size and population of each urban area, denoted by a unique urban area identifier, the size and population of each administrative area denoted by a unique administrative identifier, the size of the intersect areas where the urban and administrative areas overlap, and the UN national estimates for the percentage of the population in urban and rural areas (UN, 2002).

The goal of the algorithm is to reallocate the total population in each administrative unit into rural and urban areas while reflecting the UN national rural-urban percentage estimates closely as possible. The algorithm was designed to have few constraints and to make the constraints simple and reflect common sense. There are 6 constraints in total: (1) The total population (urban + rural) within any given administrative units remains constant; (2) The urban population density in any given administrative unit must be greater than the rural population density in that administrative unit; (3) The rural population density in any given administrative unit cannot be lower than a national minimum rural population density threshold; (4) The rural population density in any given administrative unit cannot be greater than a national maximum rural population density threshold; (5) The urban population density in any given administrative unit cannot not be greater than a national maximum urban population density threshold; (6) The urban population density in any given administrative unit cannot not be lower than a national minimum urban population density threshold.

The algorithm works on each administrative unit in turn, and checks the urban and rural populations within that administrative unit against constraints 2 to 5. If any of the constraints are

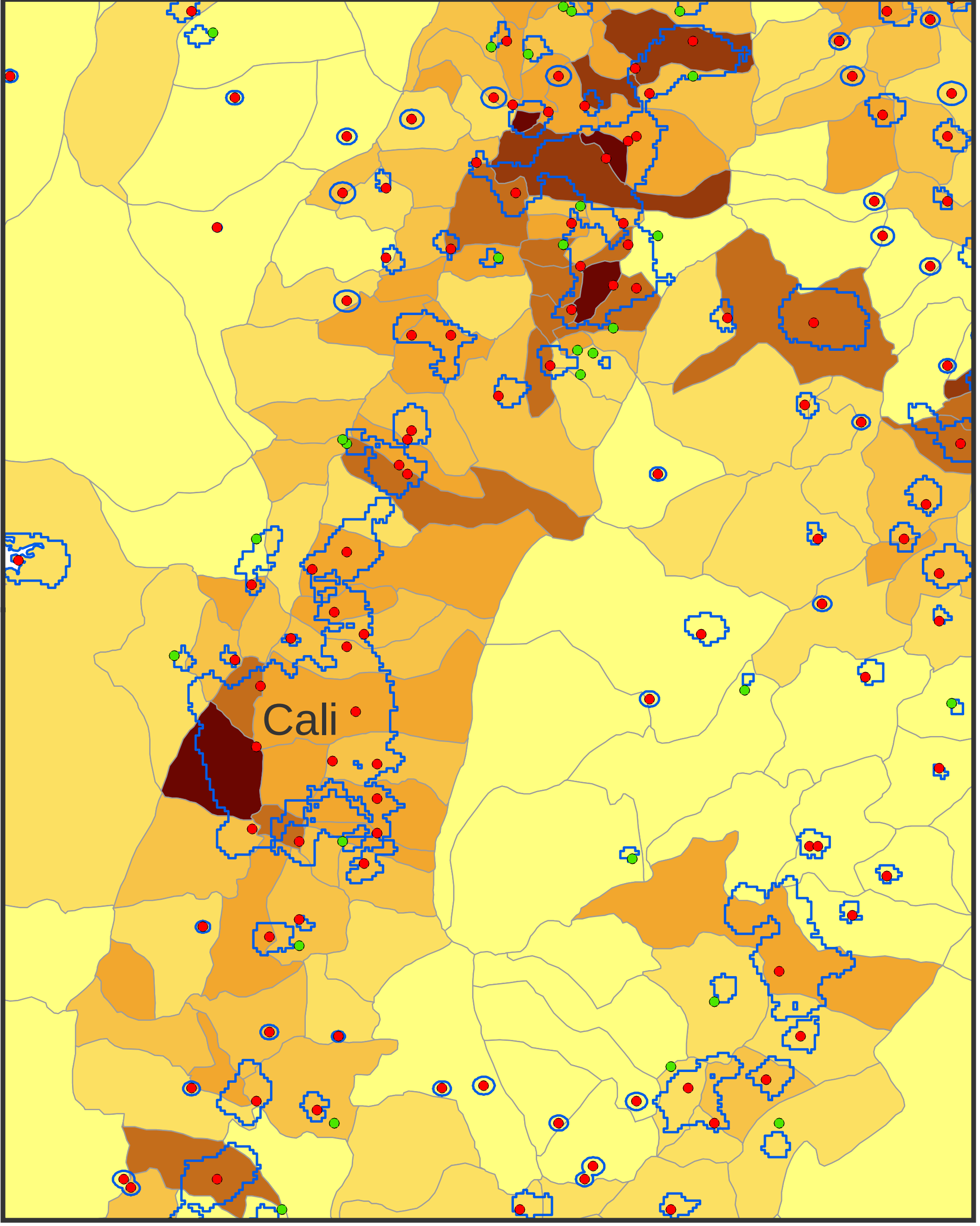
[3] Using the ArcInfo command IDENTITY

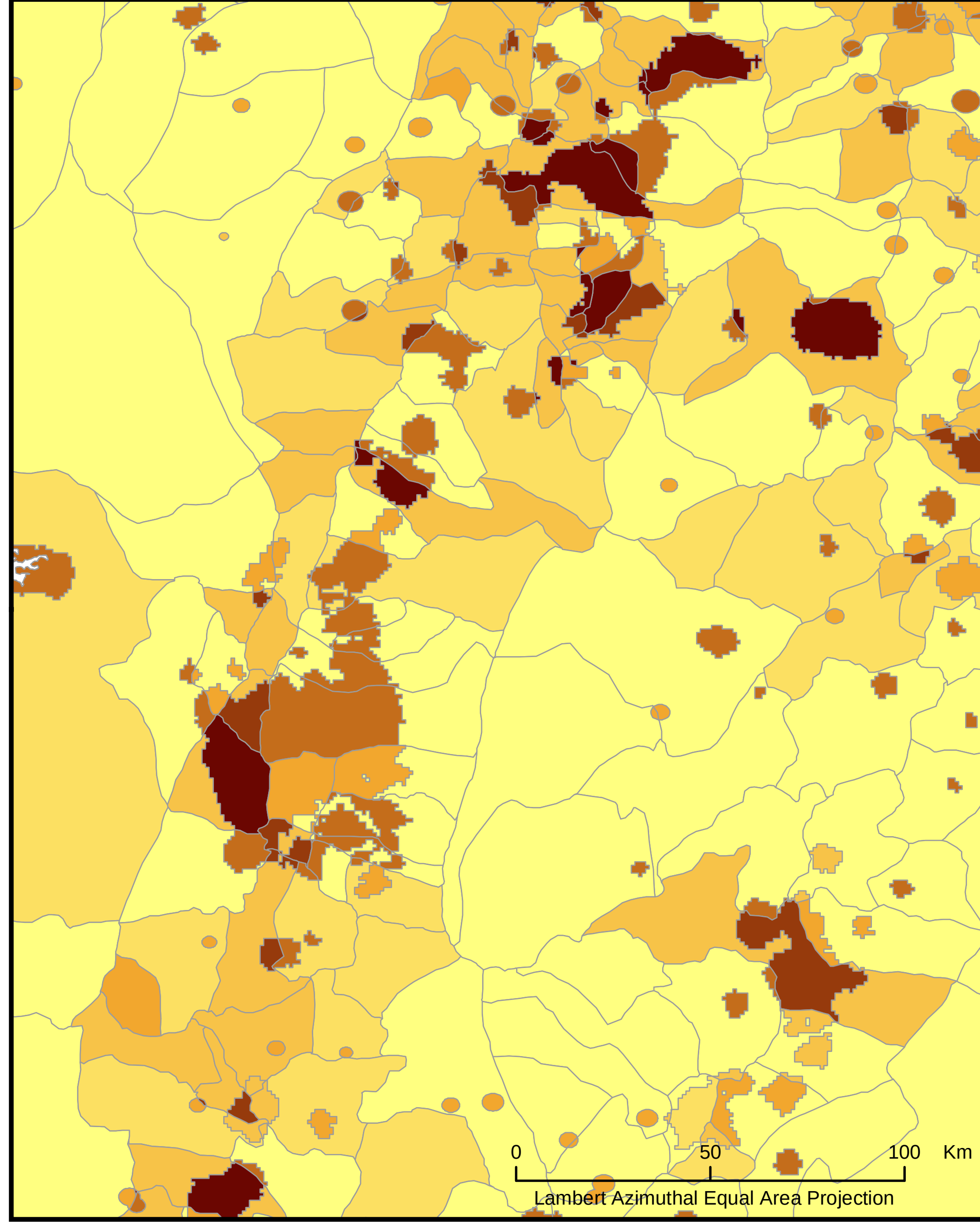
not met, then the rural and/or urban populations are adjusted literately to meet them while ensuring that constraint 1 is met. These constraints and the national population density thresholds are controlled by parameters that are passed to the algorithm. If no parameters are specified then the algorithm will assign fixed values that have been empirically determined to be good first estimates.

The adjustment in population is trivial when there are no or one urban area per administrative unit, and where the urban area lies wholly within the administrative unit. It becomes increasingly complex however when there are more than one urban area, and urban areas overlap more than one administrative area (e.g., Cali, Colombia), and large urban areas contain more than one administrative area (e.g., Quito, Ecuador). All of these are common situations, and may require successive iterations to meet all the constraints. The algorithm can also be run on a region-by-region basis (such as states or other first-level administrative units), such that the national constraints (3 to 6) now become regional constraints and will better reflect the state-level variation in rural/urban population percentages in large countries like the USA. This approach was employed for most of the largest countries or countries with very large numbers of administrative units (e.g., South Africa).

The resulting map, Figure 3 (on the following page)—a close-up of Cali, Colombia—shows the data before and after running GRUMPe. Note how, where urban areas are present in a given administrative unit, the density of the GPW administrative units decreases after GRUMPe because people are reallocated into their respective urban areas. The final results from each country are then compared to the UN urban population estimates. Although the UN totals are useful as a benchmark, they are only that. Not only have recent studies shown the uncertainty associated with UN urban estimates (NRC, 2003), there are many reasons why our estimates may differ considerably from that of the UN's. For example, our data stream may have included many more small settlements, including those below the urban threshold either given by the country, or implied by the region, in which case we would expect the comparison between percentages of the population living in urban areas to be quite different between the two. We estimate that in X% of the countries, we had *a priori* reasons to expect much different outcomes from the UN estimates (mostly but not always for the better), and in another Y% for them to match rather closely because our data streams matched closely those which they also report. In the remainder of the countries, we had no information either way to predict the closely to those estimates.

The final stage is to convert the output coverage from GRUMPe into a grid, at 30 arc-seconds resolution.





Administrative units, urban extents, and points

GRUMP output

Population Density 2000

25 50 100 250 500 1000 1001+

Points inside the urban extents

Points within a 3 km buffer of the lights

Urban extents

SEDAC

CIESIN
Columbia University

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), 2004. Global Rural-Urban Mapping Project (GRUMP): Urban Extents. Palisades, NY: CIESIN, Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw.

RESULTS

In addition to validating the data set (described in the next section), we initially constructed continent- and global-level summaries of the number, size and density of settlements and urban extents. Table 2 shows the characteristics of the population data for the three data products, including summary statistics for the year 2000. The minimum population size of 1 in the human settlements database corresponds to a few cities in the US census that are declining in population. While we included all settlements in the points database, we only included those with a population greater than 1,000 people in the urban mask. This explains the minimum population size of the urban extents. Even though not included in the table, we note here that the minimum areal size is less than 1 km^2 (several places across the world) while the largest is that of Tokyo (which is a continuum that extends south to Osaka) with more than 30,000 km^2.

Table 2. Characteristics of the three data product.

Data Product	Measure	Africa	Asia	Europe	Oceania	North America	South America	Global	Min pop size	Max pop size
Settlement Points	Total number of points	4,142	13,246	14,804	799	28,298	9,269	70,558	1 (various, USA)	18,326,722 (Mexico City)
Urban Extents	Total number of polygons	2,778	10,123	4,666	242	1,970	4,356	24,135	1,000 (various, USA)	72,786,683 (Tokyo)
Population Grid	Total number of administrative areas	109,120	88,782	91,086	2,153	74,421	10,919	376,481	1 (various)	10,434,252 (Brazil)

Note: as the GRUMP input are the GPW administrative units, we report that number as input of the urban-rural surface grid.

An interesting example that shows the differences in data collection and spatial representation of the cities between the UN view and the GRUMP approach is that of Tokyo. Although the UN reports Tokyo as the largest city, with more than 26 million people in 2000 (UN, 2002b), Tokyo is not the largest city in the settlements database, due to different data collection systems (the census data reports population for the city proper, while the UN might include some other neighboring settlements). Consistently with the UN numbers, Tokyo is the largest city when we look at the urban extents. In this case, though, the total population is much greater than what reported by the UN, because the lights extend far beyond the administrative boundaries of Tokyo. In particular the Tokyo light covers an area of more than 30,000 km^2 and includes more than 500 other settlements, making it the largest urbanized area. Although, in the population grid this does not pose a problem, and while most of the land area around Tokyo is likely urban, this is almost certainly an upper bound estimate of even an agglomerated view of the Tokyo metropolitan area.

Figure 4. shows a continental breakdown of the number of settlements by population size, in the database (noting the truncated display for North America). This figure speaks both to the availability of data and to the world's settlement patterns. There is much more information on smaller settlements for the more developed or urbanized regions of the world, i.e., in North and South America, Europe, and Oceania. While it is probably true that in these continents, as well as globally, there are many more small settlements than larger ones, the smaller ones require greater institutional capacity on the part of national statistical offices to track and disseminate information about. Thus, in Asia and Oceania, the GRUMP databases rely on less information about settlements below 20,000 persons. Cautious users, therefore, might want to apply a city-size threshold of 20,000 persons, if they are interested in making comparisons across countries that do not reflect data collection as much as settlement patterns.

Figure 4. Settlements distribution by continent and population size.

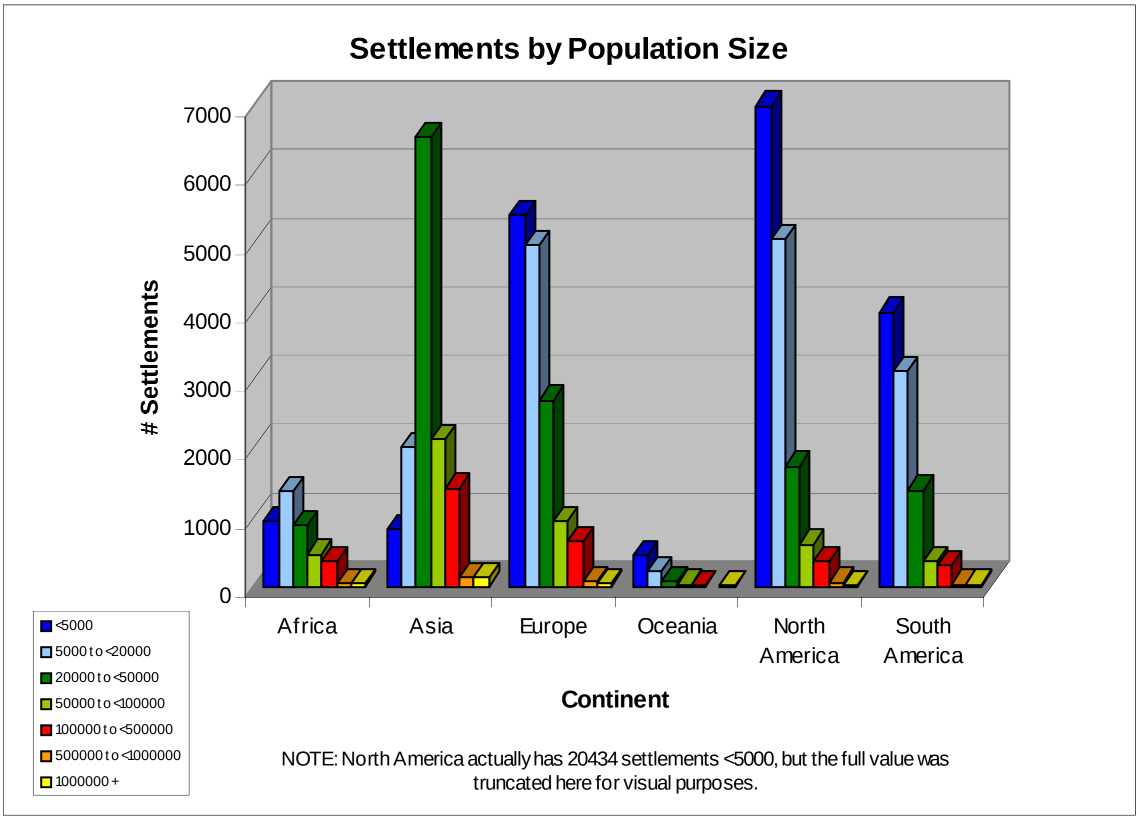


Table 3 shows the distribution of the world's population by city size classes, with associated population densities. While there is general agreement in the population totals, and overall proportions urban, GRUMP estimates that 6.7 rather than 3.7 of the urban dwellers reside in the world's largest megacities. It estimates close to 24,000 urban areas of 5,000 persons or more, in the year 2000. As expected, the geographically largest cities are the most dense with decreasing densities as city size declines.

Table 3. Distribution of the world's population by size class of settlements, in 2000, as estimated by the UN and GRUMP.

Distribution of the world's population by size class of settlement, 2000						
	UN	GRUMP			UN	GRUMP
Size class of urban settlement (number of inhabitants)	**Total Population (000s)**		***Number of settlements***	***Population Density***	**% of Total**	**Population**
Total	***794,000***	***791,821***		***27***		
Urban area	295,000	303,784	2,778	1278	37.2	38.4
10 million or more	0	32,049	2	2208	0.0	4.0
5 million to 10 million	23,000	21,414	3	1948	2.9	2.7
1 million to 5 million	64,000	86,015	40	2529	8.1	10.9
500,000 to 1 million	26,000	31,657	45	2004	3.3	4.0
under 500,000	181,000	132,649	2,688		22.8	16.8
100,000 to 500,000		75,173	365	1357		*9.5*
50,000 to 100,000		22,797	322	863		*2.9*
20,000 to 500,000		20,336	649	580		*2.6*
5,000 to 20,000		14,343	1352	316		*1.8*
Rural area	498,000	488,037	--	17	62.7	61.6

Note: The total number of settlements is different from that in Table 2, as the numbers in this table are calculated after the GRUMPe and the gridding processes by class of settlements. For technical reasons, settlements just above or below 5,000 might be classified slightly differently in the pre-GRUMPe mask than in the post-GRUMP mask and associated 1 km population surface.

A second analytic objective was to summarize urban patterns —both in terms of population and land areas-- by ecosystems. Table 4 shows the power of these new data, when integrated with other geographic data. Here the urban extent mask and the gridded population surface are overlaid with ecosystem boundaries from the Millennium Ecosystem Assessment (ref). This table, prepared for an assessment of urban systems (McGranahan et al., forthcoming) shows that coastal and island systems tend to be the most densely populated, followed by systems with water and other agricultural resources—namely, cultivated and inland water systems—but that in coastal areas, land area is disproportionately urban. Two systems—coastal and cultivated—also sustain high rural population densities. Forested and mountain ecosystems, which sustain the same total population as coastal ecosystems are much less urban, and thus sustain much lower population densities, even its urban areas.

These data suggest that roughly 3% of the earth's surface is occupied by urban areas, the majority of which concentrated in coastal and cultivated environments. This is somewhat greater than the often-cited suggestion of 1-2% of land area (ref). It is noteworthy that the although the highest share of urban land area is in the coastal zone (10.2%) is coupled with the highest share of urban population (64%), cultivated systems have 6.8% of their land area in cities, perhaps somewhat surprisingly. These areas are somewhat less urban (44.9%) other ecosystems, such has islands and water, with smaller shares of land going to urban area. Although many cultivated areas explicated omit urban centers from their ecosystem, smaller settlements, and the periphery of large settlements are commonly found in cultivated zones.

DISCUSSION

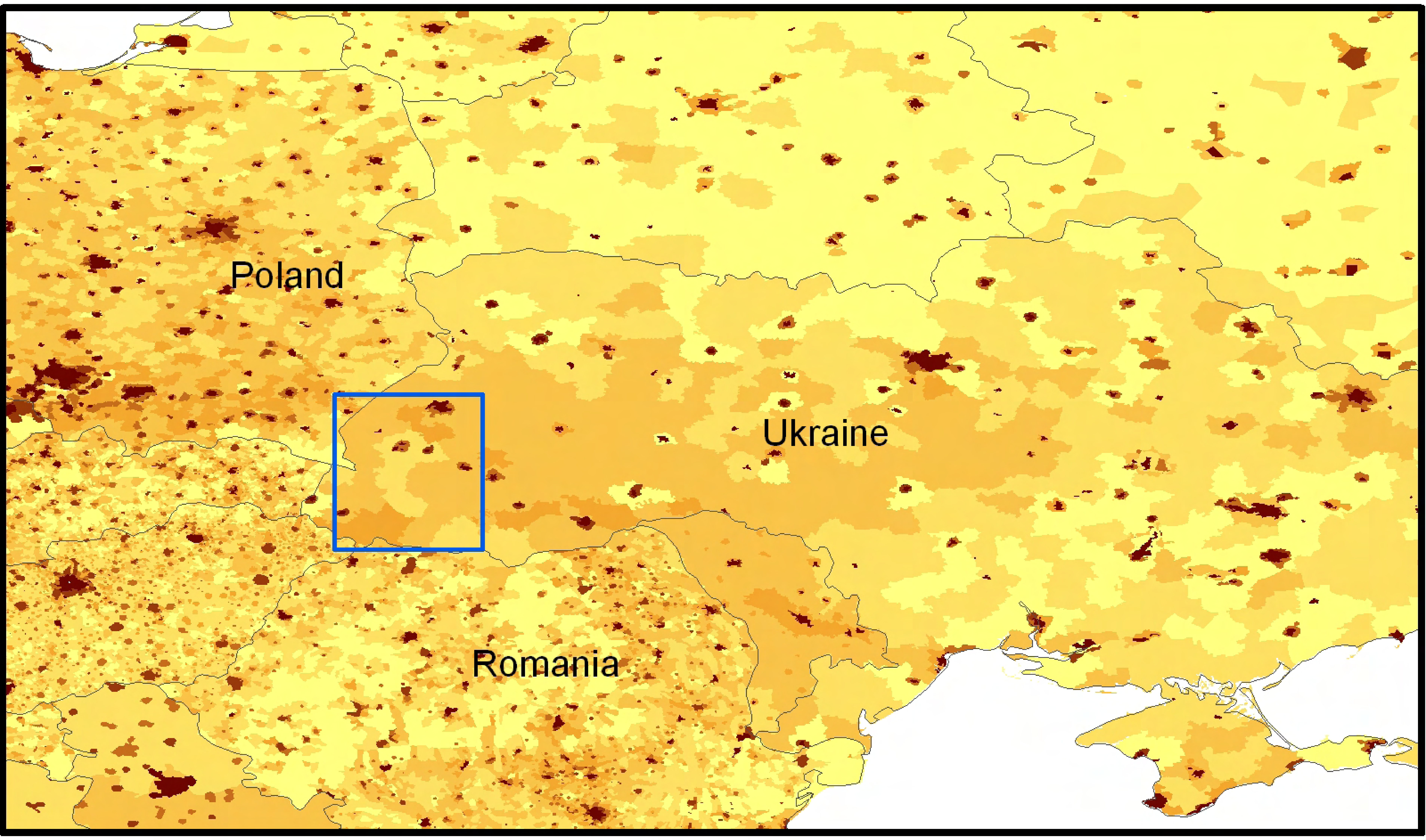
Advantages and disadvantages of the GRUMP methodology

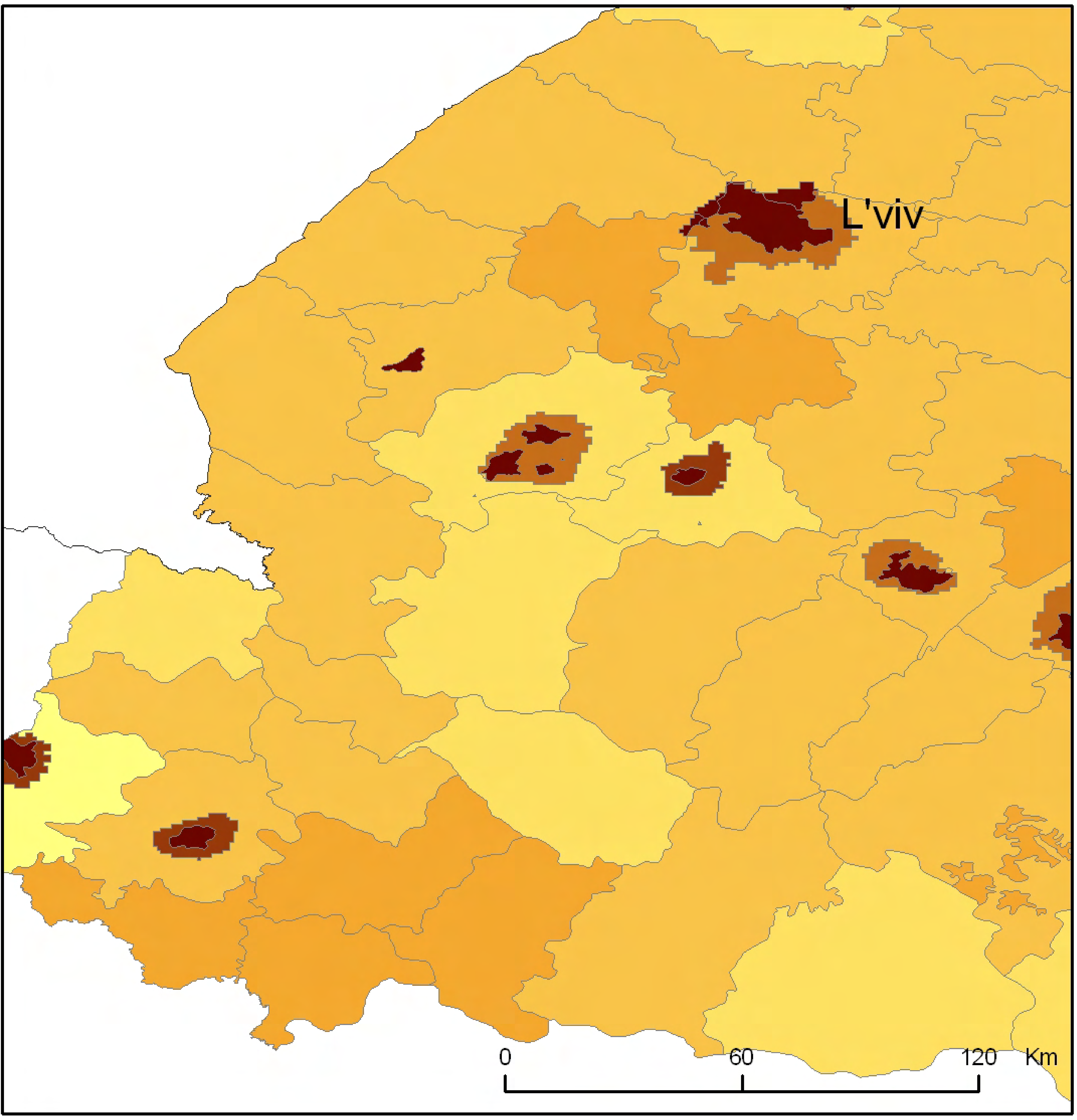
The methodology presented has both advantages and disadvantages compared to existing datasets, such as GPW, LandScan and other approaches. As compared to the approaches discussed in Pozzi et al. (2003), Schneider et al. (2003), the Accessibility model (UNEP et al., 2004), and to LandScan (Dobson et al., 2000), this method has the advantage of *using* population data for settlements from census data, rather than *predicting* population density based on probability coefficients or lighted areas. Therefore, we have an independent and more or less reliable measure of population. Further, this methodology makes use of other GIS data to identify urban areas, compensating for the small settlements in poor countries that are not detected by the night-time lights. We know that the lights dataset has two main problems: the blooming effect (see below) and the insufficient detection of small settlements that are not frequently illuminated. While there is not yet a method to improve upon these two elements at a global scale, using ancillary data to identify small settlements has proved useful in several countries in Africa. As compared to GPW, although this method produces a model surface, as opposed to a more heuristic one from GPW, it allows for improvements not only in resolution but also in the positional accuracy of human population distribution.

As for GRUMPe—the mechanism through which the modeling occurred—it proved to be a good tool to refine GPW in countries where administrative data is coarse. Although the administrative data in Colombia is relatively good, the size of the units is such that the reallocation works very well. As shown in Figure 5 there are cases of relatively large administrative units with one or two cities within, and we can clearly see how the GRUMPe assigns people to the urban areas, decreasing the density of the remaining administrative unit.

Then, there are cases, like Cali, where the urban areas identified by the lights expand over several administrative units. Also in this case, the reallocation process outputs a population distribution that is more consistent with the notion that people tend to be more concentrated into the urban areas rather than uniformly distributed across an administrative unit. GRUMPe also proved to be an effective way to compensate for the blooming effect in some countries where administrative data include city boundaries. In this case the total population allocated to the light is larger than the census value for the same city, but, given the type of administrative data, the result is a structure that has a more densely populated urban core at the center of the light, surrounded by a lower-density outskirt, as shown in Figure 5.

Figure 5 (see following page). Grump output in Eastern Europe, and close -up view of western Ukraine, showing the effect of the grumping process where administrative data include city boundaries smaller than the urban extents derived from the lights.





GRUMP Output

Population Density 2000

- 1 - 25
- 26 - 50
- 51 - 100
- 101 - 250
- 251 - 500
- 501 - 1000
- 1001+

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), 2004. Global Rural-Urban Mapping Project (GRUMP): Urban Extents. Palisades, NY: CIESIN, Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw.

We also have found some instances where GRUMPe does not work as effectively as in other cases. First, in countries such as Malawi, where GPW is very detailed, there are more than 9,000 administrative units, but only about 40 urban extents, most of them very small. Therefore GRUMPe does not provide any additional information on population distribution. This is not so much of a problem, rather than noteworthy for consistency sake, since most countries do not have such high- resolution data. Another example of the GRUMPe limitations is related to small, populated islands, like several islands in the Caribbean or in the Pacific. These islands might have one or two isolated small urban centers, but, due to the blooming effect, appear all lit. In this case, the reallocation of the population into urban areas and rural areas is not very effective, as the urban areas identified by the lights could cover the entire islands, even though the urban population is only a fraction of the total population, and the administrative data is generally good. Fortunately, inputs in the underlying administrative and population data for GPW version 3 have improved substantially for more than half of the world's island nations (Balk and Yetman, 2004). Where the administrative data are poor in the sense that they attempt to approximate urban centers, but do so inadequately (e.g., the construction of small triangular shapes to represent urban centers in the former Soviet republics) and the lights data are moderate to poor, GRUMPe may assign too high a population density value to such a small area. In this case, its general assignment is correct, but the extent is limited both by the lights and the administrative data's shortcomings.

In sum, GRUMPe performs moderately well. Where administrative data and the lights data are good (x% of cases), GRUMPe does not perform very well. However, in these instances, there is less imperative for it to work well; it is only a problem in the event that it introduces error or degrades the data quality, both open questions at this point. Where administrative data are moderate or poor, and the lights (and more intensive substitutes) are moderate to good, GRUMPe performs very well (y% of cases). Where both the administrative data and the lights (or its substitutes) are poor GRUMPe just does not have much to work with (z% of cases). As is generally the case, there are no perfect substitutes for good data.

Blooming Effect

The blooming effect is an overestimation of the real extents of urban areas, and is believed to be dependant on the intrinsic characteristics of the sensor (Elvidge *et al.*, 2004). Early efforts to threshold the lights globally, such as by Imhoff and colleagues (1997), are inappropriate in this context because their study sample was small and concentrated in a well-light region of the United States. Figure 6 shows examples of the extents of the urban areas of Quito, Ecuador, and Lagos and Ibadan, Nigeria, as detected by Landsat and by the night-time lights. Note how much larger the night-time lights areas are. For Quito in particular, the area estimated by the Landsat polygons is 187 km^2, while the area covered by the light is 1690 km^2. A more detailed comparison of lighted area with built area estimates from Landsat imagery of 17 cities worldwide (Small et al., forthcoming 2005) shows that lighted areas are consistently larger than even maximum estimates of built areas for almost all cities in every light dataset. Thresholds >90% can often reconcile lighted area with built area in the 94/95 dataset but there is not one threshold that works for a majority of the 17 cities considered. Moreover, such a high threshold would result in the loss of several small settlements that are not frequently illuminated. The same study shows that, even though a 10% threshold could reduce the blooming effect without significantly attenuating many individual small settlements for the 1994/1995 dataset, this detection frequency threshold does not provide a globally consistent basis for reconciling lighted areas to urban extent. For this reason, and because detection is more of a concern than blooming for the global database, we did not apply any thresholds to the lights.

Further, while blooming is noted to be a problem, it is probably much less of one for the production of a global population distribution grid because that redistribution is to go from even much coarser administrative units to these urban areas. Thus, the direction of the reallocation we argue is a vast overall improvement in the database. Furthermore, for the largest cities, where blooming is probably greatest, there tend to be better sub-urban administrative units, so that the population within the extents will show the detail of the underlying detail. Nevertheless, future work should continue to determine the possibility of reducing the lights, as appropriate, so that the blooming effect is minimized.

A new generation of annual global OLS nighttime lights are currently in production for the 1992-2003 time period. The new products will report the average visible band digital numbers (DN) of lights, but will not be radiometrically calibrated. NGDC plans to cross calibrate each of the annual products, but in relative sense, not absolute. The Visible Infrared Imaging Radiometer Suite (VIIRS) instrument will continue the low light imaging measurements of the OLS, with substantial improvements in calibration, spatial resolution and levels of quantifications. It is anticipated that the nighttime lights products derived from VIIRS data will be superior to those possible from the OLS.

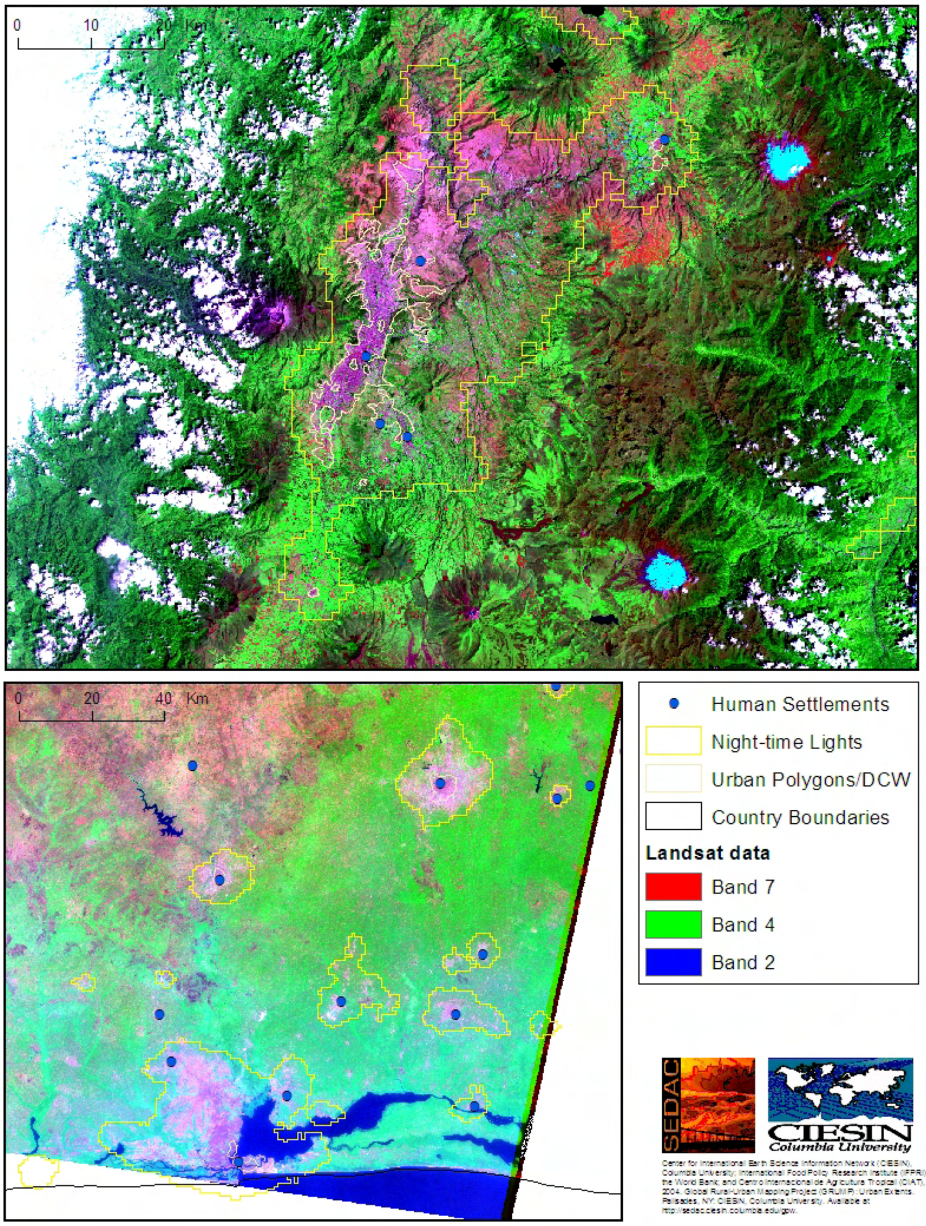


Figure 6. Night-time lights and urban polygons overlayed to Landsat scenes for Quito, Ecuador (top) and southern Nigeria (bottom). Urban areas are easily identifiable in purple. In the Quito image, the urban polygons are derived from digitization of urban areas from Landsat data, as part of the TREES project, while the polygons in the Nigerian map are DCW polygons. Note how in both cases the urban polygons are much smaller than the night-time lights estimates. In particular, the TREES polygons in the Quito urban area sum to 187 km2, while the area covered by the light is 1690 km2. The blooming is also apparent in Nigeria, where all cities in the Landsat image appear smaller than what detected by the lights, even though in this case the cities seem to bloom less that Quito. Also note how the DCW polygon underestimate the urban area of Ibadan (the large one in the northern portion of the image) while the lights provide a more accurate spatial representation of the city.

VALIDATION

When producing these datasets, we faced two main problems: first the mismatch between the lights and the points, which leads to 1) the removal of several lights, and to 2) the creation of circles for the points; and second the blooming effect and the overestimation of the urban extents. The first type of problem is primarily dependant on population data collection and on coordinates availability. In some cases, the census or gazetteer just have poor points data to begin with. In some other cases, for a given settlement, the NIMA matching produces several inconsistent results and we have no means to validate the geolocation accuracy of one over the other, resulting in the removal of that given point from the database. The result is that often we have lights with no settlements (that will be removed from the Urban Mask) and settlements characterized by an artificially circular shape.

Lost Lights

To assess the extent of the “lost lights” we selected all the lights with no population and calculated the proportion of these unpopulated lights over the total number of lights for each country. …

Circles

To come…

Overestimation

As detailed in Elvidge *et al*. (2004), we know the lights tend to overestimate the actual extents of the urban areas. Since we use the areas of the light along with population data to estimate the areas of the settlements and to construct circles, it is very important that we also try to assess the extent of this overestimation. As previously mentioned, Small et al. (forthcoming 2005) show that lighted areas are consistently larger than even the maximum estimates of built areas for almost all cities considered, and that only thresholding at very high frequency will reconcile lighted areas with built areas. To assess the extent of the overestimation, we used areal estimates from polygons derived from high resolution satellite data and compared those to lighted areas for corresponding settlements. In particular, we used the polygons from the TREES project for a sample of countries in Latin America.

The TREES Project (Tropical Ecosystem Environment Observation by Satellites) was set up in the early 1990s, by the European Commission for global humid tropical forest monitoring, using medium-resolution satellite data (in particular SPOT and Landsat). Several sites in Latin America were identified according to the high (hot spot) and low range of deforestation. Satellite data for these sites were classified based on visual interpretation and in accordance to the CORINE land use/land cover classification system and supported by ancillary data (GET REF.). Only polygons greater than 50 hectares were digitized at a scale of 1:100 000. We used the polygons classified as “urban” for 13 sites in Columbia, Ecuador and Peru, processed by CIAT. We found that the average ratio of area estimated by the TREES polygons to that of the lights is about 5%. It also appears that this proportion is not dependant on the size of the light or the TREES polygons.

As we use DCW polygons on a global scale to supplement the lack of lights, we also compared the areal extents of DCW polygons with those of the lights in selected areas worldwide. In this case the ratio of DCW area to the area of the lights is about 7%. This should not be taken as an absolute ratio, because it was calculated only on a sample of 20 countries, and the DCW polygons are very inconsistent, in that they are very accurate in some countries (especially some European countries, where the ration can be as high as 30%), and very much inaccurate in others (especially African and Asian countries, where the ratio is often around 1%).

This type of analysis gives us an indication of how large the lights are compared to other polygons, but it cannot be used as a sound method to shrink the lights. In fact we need to keep in mind that sometimes other polygons tend to underestimate urban areas, and can be either inconsistent across different countries or digitized according to standards of minimum areal extents.

CONCLUSIONS

One of the main objectives of this project was not only to construct an improved population grid that systematically accounts for urban centers, but also to construct a globally consistent database of those urban areas. As several studies show, there have been several attempts to map or model population distribution, but few of them account explicitly for urban areas, or attempt to merge the lights directly with city-level census data to derive population estimates of urban areas. The methodologies detailed here take a comprehensive and systematic approach to combine several data streams into estimates of urban extents and population distribution.

If we look at the three separate products of the Global Rural Urban Mapping Project (GRUMP) we can draw the following conclusions: Data gathered from the census or census-like sources seems to provide considerably more detailed information about population distribution for settlements under 500,000 people than the UN estimates. The Population Division actually collects information for smaller cities, but does not do so systematically. The Statistics Division collects some—less systematic—information for places of 100,000 or more, but neither they nor the Population Division attempt to collect data below 100,000 persons. We show here that nearly 20% of the world's urban population lives in cities of these sizes, thus GRUMP substantially contributes by amassing these data.

Spatial information about cities is also very important and it is a piece of information that is missing entirely from the UN statistics. Despite the caveats about using the night-time lights dataset as a baseline for delineating urban extents (i.e., the blooming effect, etc), this project represent the first attempt to systematically map urban areas at the global level, by using satellite and geographic data without any prediction modeling. The Urban Extents Mask could certainly be considered the first global population dataset to explicitly include the extents associated to human settlements.

The GRUMPe algorithm also presents some limitations (where administrative data and lights both very good or fairly poor, the performance of GRUMPe is suboptimal), but overall it seems to perform well, especially where administrative data are moderate or poor, and the lights (and more intensive substitutes) are moderate to good. Although the algorithm might still need some refinement to further improve the reallocation process, the methodology is overall conceptually sound and allowed us to improve the positional accuracy of human population distribution.

Improved knowledge of the spatial distribution of urban and rural population is extremely important for assessing socioeconomic, demographic and environmental change in urban and rural areas. These products allow one to easily estimate the percentage of urban people living in a given environment, or a given country or region, based on an improved spatial distribution of population at a relatively fine scale (30 arc-second, nominally 1 km at the equator). As shown here, these data provide the first systematic assessment of the world's urban land area (nearly 3%), and how distributions by ecosystems differ dramatically. Coastal zones are the most urban of all systems, and sustain the highest population densities, not only in the urban areas, but in the rural ones as well.

These data can also be used to provide much greater information on distances to urban areas, where the places can be classified according to information about their population and geographic size. Much attention has been made on the importance of moving toward an urban continuum (NRC, 2003, Woods, 2003) and these datasets are a first big step in moving in that direction. These data have already been used in studies of mortality (Balk et al., 2004) and hunger (Balk et al., 2004) in Africa, by combining these data with household survey data. Other early uses of the data have been to use the population surfaces in assessing malaria risk (Hay et al, 200X) and dimensions of rurality (Chomitz et al., 200x). Continued uses will assist in contributing to the dialogue on how best to collect, interpret, and process information to maximize flexible and creative new uses of these data.

The main challenge of this methodology, is the complex and time-consuming procedure that goes from collecting and processing the census data, to combining the city population with the spatial information about the settlements and finally to reallocating people from the administrative units into the urban centers. Some of that complexity could be reduced as institutional capacity increases in the production and distribution of urban data, as has already happened for administrative data over the past 10-15 years (see Balk and Yetman, 2004). Further gains may be made by establishing international guidelines on the definition and correspondence between metropolitan areas of different types (see Champion and Hugo, 2003). In hindsight, such guidelines would make an invaluable contribution in reducing the processing time, but also increasing the accuracy of the underlying point data, upon which both the extent mask and population surface are based. Finally, even though we estimated population for three time periods (1990, 1995, and 2000), users need to remember that the lights refer to one point in time only (the 1994/1995 time period), therefore it would not be advisable to use these extents for any analysis of change in spatial parameters.

The three GRUMP datasets (CIESIN et al., 2004a, 2004b, 2004c) are available freely at: *http://beta.sedac.ciesin.columbia.edu/gpw*. We also welcome users who notice such errors to report them to *info@ciesin.columbia.edu*.

REFERENCES

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

Balk, D. and Yetman, G. 2004. "The Global Distribution Of Population: Evaluating The Gains In Resolution Refinement", Draft Documentation for GPW v3, downloadable from http://beta.sedac.ciesin.columbia.edu/gpw.

Center for International Earth Science Information Network (CIESIN), Columbia University; and Centro Internacional de Agricultura Tropical (CIAT), (2004). *Gridded Population of the World (GPW), Version 3.* Palisades, NY: Columbia University. Available at http://beta.sedac.ciesin.columbia.edu/gpw.

Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IPFRI), the World Bank; and Centro Internacional de Agricultura Tropical (CIAT), 2004a. Global Rural-Urban Mapping Project (GRUMP): Settlement Points. Palisades, NY: CIESIN, Columbia University. Available at: http://beta.sedac.ciesin.columbia.edu/gpw .

Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IPFRI), the World Bank; and Centro Internacional de Agricultura Tropical (CIAT), 2004b. Global Rural-Urban Mapping Project (GRUMP): Urban Extents. Palisades, NY: CIESIN, Columbia University. Available at: http://beta.sedac.ciesin.columbia.edu/gpw .

Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IPFRI), the World Bank; and Centro Internacional de Agricultura Tropical (CIAT), 2004c. Global Rural-Urban Mapping Project (GRUMP): Gridded Population of the World, version 3, with Urban Reallocation (GPW-UR). Palisades, NY: CIESIN, Columbia University. Available at: http://beta.sedac.ciesin.columbia.edu/gpw .

Champion, Anothony and Graeme Hugo (eds). 2003. *New Forms of Urbanisation: Beyond the Urban-Rural Dichotomy* IUSSP [Get complete reference]

Chomitz, K., P. Buys, and T. S. Thomas, 2004, "Quantifying The Rural-Urban Gradient In Latin America And The Caribbean," paper revised: *7 October 2004*

Deichmann, U., Balk, D. and Yetman, G., (2001). Transforming Population Data for Interdisciplinary Usages: From Census to Grid. Unpublished manuscript available on-line at: http://sedac.ciesin.columbia.edu/plue/gpw/GPWdocumentation.pdf

Dobson, J. E., E. A. Bright, P. R. Coleman, R.C. Durfee, B. A. Worley (2000), LandScan: A Global Population Database for Estimating Populations at Risk, *Photogrammetric Engineering and Remote Sensing*. 66(7): 849-857.

Elvidge, C.D., Baugh, K.E., Kihn, E.A., Kroehl, H.W., Davis, E.R. (1997a). Mapping City Lights with Nighttime Data from the DMSP Operational Linescan System.*Photogrammetric Engineering and Remote Sensing, 63*(6): 727-734.

Elvidge, C.D., Baugh, K.E., Kihn, E.A., Kroehl, H.W., Davis, E.R. and Davis, C.W. (1997b). Relation Between Satellite Observed Visible-Near Infrared Emissions, Population, Economic Activity and Electric Power Consumption, *International Journal of Remote Sensing, 18*(6): 1373-1379.

Elvidge, C.D., Baugh, K.E., Hobson, V.R., Kihn, E.A., Kroehl, H.W., Davis, E.R. and Cocero, D. (1997c). Satellite Inventory of Human Settlements Using Nocturnal Radiation emissions: a Contribution for the Global Toolchest. *Global Change Biology, 3*: 87-395.

Elvidge, C.D., Safran, J., Nelson, I.L., Tuttle, B.T., Hobson, V.R., Baugh, K.E., Dietz, J.B., Erwin, E.H., 2004, Area and position accuracy of DMSP nighttime lights data. Chapter 20 in

Remote Sensing and GIS Accuracy Assessment (Lunetta R.S. and Lyon, J.G. editors), CRC Press, p. 281-292.

Hay, S. I. Noor, A. M. Nelson, A. Tatem, A. J. (2004 *submitted*) The Precision of Large-Area Population Maps. American Journal of Epidemiology xx(x) xxxx-xxxx.

Imhoff, M.L., Lawrence, W.T., Elvidge, C.D., Paul, T., Levine, E., Privalsky, M.V. and Brown, V. (1997a). Using Nighttime DMSP/OLS Images of City Lights to Estimate the Impacts of Urban Land Use on Soil Resources in the United States. *Remote Sensing of the Environment, 59*: 105-117.

Imhoff, M.L., Lawrence, W.T., Stutzer, D.C. and Elvidge, C.D. (1997b). A Technique for Using Composite DMSP/OLS "City Lights" Data to Map Urban Areas. *Remote Sensing of the Environment*, *61*: 361-370.

Imhoff, M.L., W.T. Lawrence, D.C. Stutzer, and C.D. Elvidge, (1997). A technique for using composite DMSP/OLS "city lights" satellite data to map urban area, *Remote Sensing of Environment*, *61* (3), 361-370.

International Center for Tropical Agriculture (CIAT), (2003). "Digital maps for the analysis of temporal and spatial changes in population of Latin America and the Caribbean". URL http://gisweb.ciat.cgiar.org/population/, Accessed on 12th October 2004

McGrahanan et al., 2004 (draft). "Urban Systems," Millennium Ecosystem Assessment, Trends and Conditions Assessment.

National Research Council. 2003. *Cities Transformed*: *Demographic Change and Its Implications in the Developing World.* Panel on Urban Population Dynamics, M.R. Montgomery, R. Stren, B. Cohen, and HE Reed, eds., Committee on Population, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Nghiem, S. D. Balk, C. Small., U. Deichmann, A. Wannebo, R. Blom, P. Sutton, G. Yetman, R. Chen, E. Rodriguez, B. Houshmand and G. Naumann. Nov. 2001, "Global Infrastructure Mapping: The Potential for SRTM to Break New Ground", JPL White Paper, available at: http://www.ciesin.columbia.edu/whitepaperfinal.pdf.

NIMA, (gnswww.nima.mil/geonames/GNS/index.jsp).

Pozzi, F., C. Small and G. Yetman (2003). Modeling the Distribution of Human Population with Night-time Satellite Imagery and Gridded Population of the World. *Earth Observation Magazine, Vol. 12, No. 4:* 24-30

Schneider, A., M.A. Friedl, D.K. McIver, and C.E. Woodcock (2003). Mapping Urban Areas by Fusing Multiple Sources of Coarse Resolution Remotely Sensed Data. *Photogrammetric Engineering and Remote Sensing, 69* (12):1377-1386.

Seto, K.

Small, C., F. Pozzi and C.D. Elvidge (2004, or forthcoming 2005??). Spatial Analysis of Global Urban Extents from the DMSP-OLS Night Lights, submitted to Remote Sensing of the Environment.

Sutton, P., D. Roberts, C. Elvidge, and K. Baugh, (2001). Census from Heaven: An estimate of the global human population using night-time satellite imagery, *International Journal of Remote Sensing*, *22* (16), 3061-3076.

Tatem A.J, Hay SI., (2004) "Measuring urbanization pattern and extent for malaria research: a review of remote sensing approaches" *J Urban Health*. 81(3):363-76.

Tatem, A.J. Noor, A.M. Hay, S.I. (2004) "Defining approaches to settlement mapping for public health management in Kenya using medium spatial resolution satellite imagery", *Remote Sensing of Environment*, 93(1-2):42-52.

Tobler, Waldo, Uwe Deichmann, Jon Gottsegen and Kelly Maloy. 1997. "World Population in a Grid of Spherical Quadrilaterals," *International Journal of Population Geography*, 3:203-225.

United Nations, 2002a. *World Urbanization Prospects, The 2001 Revision.* United Nations: Department of Economic and Social Affairs, Population Division, New York.

United Nations, 2002b. *Demographic Yearbook.* United Nations: New York.

United Nations Environment Programme (UNEP), (1996a). "Asia Population Distribution Database", URL http://grid2.cr.usgs.gov/globalpop/asia/index.php3, Accessed on 12th October 2004.

United Nations Environment Programme (UNEP), (1996b). "Global Population Distribution", URL http://grid2.cr.usgs.gov/globalpop/1-degree/description.php3, Accessed on 12th October 2004.

United Nations Environment Programme (UNEP) International Center for Tropical Agriculture (CIAT), World Resources Institute (WRI), (2000). "Latin Amercia Population Distribution Database", URL http://grid2.cr.usgs.gov/globalpop/lac/index.php3, Accessed on 12th October 2004.

United Nations Environment Programme (UNEP); Center for International Earth Science Information Network (CIESIN), Columbia University (2004). "Africa Population Distribution Database", URL http://grid2.cr.usgs.gov/globalpop/africa/index.php3, Accessed on 12th October 2004.

Weeks, John.

Wood, S., K. Sebastian, F. Nachtergaele, D. Nielson, and A. Dai, 1999 "Spatial Aspects of the Design of Targeting of Agricultural Development Strategies," Environmental and Production Technology Discussion Paper, #44, International Food Policy Research Institute, Washington, DC: May.

Woods, Robert. 2003. Urban-Rural Mortality Differentials: An Unresolved Debate. *Population and Development Review*, March 2003, *29*(1), pp. 29-46.