MAPPING THE WORLD'S CITIES:
AN EXAMINATION OF GLOBAL URBAN MAPS AND THEIR IMPLICATIONS FOR CONSERVATION PLANNING

David Potere

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE OFFICE OF POPULATION RESEARCH

ADVISED BY PROFESSOR DOUGLAS MASSEY

January 2009

For Sheherazade.

ABSTRACT

Since 2000, eight groups from government and academia in both the EU and the US have created global maps that can be used to describe urban areas. Although most share common inputs, they differ by as much as an order of magnitude in their estimates of the total extent of the Earth's urban land (from 0.27 – 3.52 million sq. km). Based on a new Discrete Global Grid-based system for map comparison, we find inter-map correlations highest in North America ($\overline{r}$ = 0.83) and lowest in Asia ($\overline{r}$ = 0.66). We conduct the first global accuracy assessment focused on urban areas using a two-tiered approach that draws on high-resolution Google Earth imagery and medium-resolution Landsat-based maps. Across a wide range of accuracy measures, spatial scales, and world regions, the new MODIS 500 meter-resolution global urban map has the highest accuracy (aggregate accuracy rank of 1.0), followed by a thresholded version of the Global Impervious Surface Area map (aggregate accuracy rank of 2.7).

Based on an analysis of the MODIS 500 meter-resolution land cover map and the 2007 World Database of Protected Areas, we observe 19,600 sq. km of urban incursions or inholdings in protected areas (PAs); 77 percent within developed countries. A simple spatial model of urban expansion that is driven by several demographic and urban density scenarios predicts a near doubling of urban areas through 2050 (from 0.66 to 1.22 million sq. km). Assuming that PAs have no ability to deter urban encroachment, this urban expansion will threaten an additional 12,000 sq. km of PA, a 61 percent increase over observed incursions. Despite the relative inaccessibility of Africa's PA, by 2050 that region faces the largest potential PA losses: 2,700 sq. km, an increase of 392 percent. Overall, developing countries account for 76 percent of all potential urban PA incursions through 2050. High-growth demographic scenarios combined with declining urban densities yield global PA incursions twice the size of those derived from lower-growth, more compact futures. Our findings highlight the need to account for demographic pressure and urban planning when designing sustainable conservation strategies in the context of a rapidly urbanizing world.

ACKNOWLEDGEMENTS

I would like to first thank my family, whose constant support and understanding made this work possible. Thanks to my committee members, co-authors, and outside readers; your feedback and ideas were essential. Special thanks to Budhu Bhaduri and Eddie Bright from the Geographic Information Science and Technology team at Oak Ridge National Laboratory for hosting me over the course of two summers and providing considerable enthusiasm and resources over the past four years. Thanks also to Shlomo Angel for providing thoughtful comments and advice at all stages of this research. This work was made stronger through a series of consultations with experts across a wide range of fields, including: Alessandro Baccini, Dan Civco, Erle Ellis, Neal Feierabend, Mark Friedl, Kosuke Imai, Mutlu Ozdogan, Jason Parent, Kevin Sahr, Burt Singer, Chris Small, Will Turner, Ben Tuttle, and Alan Strahler. Thanks to Wayne Appleton, Chang Chung, Jennifer Curatola, Bill Guthe, Dawn Koffman, Wangyal Shawa, and Henry Umansky, for excellent computer and programming support. This research was made possible in part by support from the Department of Energy's Computational Science Graduate Fellowship and Princeton University's Office of Population Research.

CONTENTS

LIST OF FIGURES

CHAPTER ONE

CHAPTER TWO

CHAPTER THREE

APPENDIX A

LIST OF TABLES

CHAPTER THREE

APPENDIX A

ACRONYMS AND ABBREVIATIONS

AVHRR	Advanced Very High Resolution Radiometer Instrument (US)
CIESIN	Center for International Earth Science Information Network
DCW	Digital Chart of the World
DGG(s)	Discrete Global Grid(s)
DMSP-OLS	Defense Meteorological Satellite Program's Operational Line Scanner (US)
DOE	Department of Energy (US)
ENVISAT	Environmental Satellite (EU)
ESA	European Space Agency (EU)
EU	European Union
EC	European Commission
GDP	Gross Domestic Product
GE	Google Earth
GIS	Geographic Information System
GLC00	Global Landcover 2000
GLCC	Global Land Cover Characteristics Database (EU)
GLCF	Global Land Cover Facility
GLOBC	GlobCover (EU)
GRUMP	Global Rural-Urban Mapping Project
GTOPO-30	Global 30" arc-second Digital Elevation Model
HABITAT	United Nations Human Settlement Programme
HYDE3	History Database of the Global Environment Version 3
IGBP	International Geosphere-Biosphere Programme
IIASA	International Institute for Applied Systems Analysis
IMPSA	Global Impervious Surface Area Map

IPCC	Intergovernmental Panel on Climate Change
ISO	International Organization for Standardization
IUCN	International Union for Conservation Nature
JRC	European Joint Research Center (EU)
KM (km)	Kilometer(s)
LC	Land Cover
LCVT	Land Cover Validation Tool
LITES	Nighttime Lights Imagery from the DMSP-OLS Instruments (US)
LSCAN	LandScan 2005 (US)
M (m)	Meter(s)
MERIS	Medium Resolution Imaging Spectrometer (EU)
MMU	Minimum Mapping Unit
MOD1K	MODIS 1 km Urban Land Cover (US)
MOD500	MODIS 500 meter Urban Land Cover (US)
MODIS	Moderate Resolution Imaging Spectroradiometer (US)
NASA	National Aeronautics and Space Administration (US)
NOAA	National Oceanic and Atmospheric Administration (US)
NGDC	National Geophysical Data Center (US)
NMI	Normalized Mutual Information Coefficient
ODC	Other Developed Countries
OMA	Overall Map Accuracy
PA(s)	Protected Area(s)
PPP	Purchasing Power Parity
SPOT4-VEG	Satellite Pour l'Observation de la Terre – VEGETATION (EU)
SRES	Special Report on Emissions Scenarios (IPCC)
SRTM	Shuttle Radar Topography Mission

TSS	True Skill Statistic
UCONN	University of Connecticut, USA
UMD	University of Maryland, USA
UN	United Nations
US	United States
USA	United States of America
USGS	United States Geological Survey
UWISC	University of Wisconsin, USA
VMAP0	Vector Map Level Zero
WCMC	World Conservation Monitoring Center
WDPA	World Database of Protected Areas

CHAPTER ONE

A CRITICAL LOOK AT REPRESENTATIONS OF URBAN AREAS IN GLOBAL MAPS [1]

1.1 INTRODUCTION

Prior to 2000, if one asked the question, "How much of the Earth is urban, and where are these urban places located?" the only global map available was a digitized mosaic of maps and charts from the 1960s through the 1990s (Digital Chart of the World or Vector Map Level Zero; Danko 1992). That landscape has changed. International research groups from both government and academia have produced eight global-scale urban maps and two global-scale urban-related maps; seven of the eight urban maps were released since 2000. These new maps have been made possible by the availability of global coarse-resolution daytime and nighttime satellite remote sensing observations, frequent collection of census results and other urban-related map layers that are in a geographic information system (GIS) format, and increases in both computational power and the effectiveness of semi-automated mapping techniques. The new methods and data sources offer, for the first time, a global perspective on the spatial extent and distribution of urban land. Because this synoptic perspective is quite new and the community of map users is diverse and expanding, it is important that we step back from the limited perspective of individual mapping efforts and strengthen our understanding of the similarities and differences across the suite of available global urban maps. The aim of this chapter is to conduct that comparison by employing a wide array of map agreement measures across a range of spatial resolutions and

[1] **Chapter 1** is based on Potere and Schneider (2007, forthcoming).

geographic regions, and to provide the diverse community of global urban map users with the diagnostic tools to determine the map or set of maps most suited to their application.

We begin this chapter by exploring why global urban maps are an important scientific aim, summarizing briefly the methodologies and underlying source data for each of the eight existing global urban maps, and reviewing previous global comparative studies (**Section 1.1**). Although most of the maps fail to provide sufficient descriptions of their 'urban' land cover class, the class definitions that can be inferred from their various methodologies are quite heterogeneous. In the next section, we bring these maps into focus by creating a common analysis environment (**Section 1.2.A**) and exploring each map at three scales: the level of the city, the continental region, and the globe (**Section 1.2.B**). Our results suggest a surprisingly high degree of heterogeneity among the eight maps across all world regions and map resolutions, including an order of magnitude difference in total urban extent (Figure 1.1). We next explore the structure underlying these bulk differences by making all of the 28 possible inter-map comparisons (**Section 1.2.C**). We employ three different approaches to measuring map agreement: (1) pixel-by-pixel comparisons, (2) national-level comparisons of urban extent, and (3) multi-resolution comparisons of urban extent based on a global system of hexagonal cells. The inter-map differences that we uncover in **Section 1.2** point toward the need for a common urban taxonomy and a global map assessment based on medium- and high-resolution validation data.

A. Why global urban maps?

The density, spatial distribution, and physical characteristics of human settlement are important drivers of social and environmental change at multiple scales (Massey 2005). Based on national-level urban population projections, the UN now estimates that urban communities account for more than half of the human population (UN 2008a). Aside from some city-level statistics on the largest urban areas, these UN estimates do not reveal where

urbanites live and work at sub-national scales. Many countries track their urban populations with censuses, but these census counts cannot be readily transformed into estimates of urban *areal extent*. Conversion from urban population to urban land area is complicated by high variance in the density of urban settlement, and the variable size of the administrative districts that underlie census results. To estimate the amount of land devoted to urban areas, it is necessary to build spatially explicit human settlement maps that combine ground-based census data with satellite observations of the built environment. There exists a largely unmet need for a map of human settlement patterns that is global, assessed for accuracy, at high to moderate spatial detail or resolution, and regularly updated. Such a map is crucial on several fronts:

(1) Global urban maps can help us track contemporary urban expansion. We are living in an urbanizing[2] world. Two components underlie the conclusion that this urbanization will lead to an expansion of urban land: (a) despite cautions regarding the confidence that should be placed on more distant projections (Cohen 2004), the number of urban dwellers is on course to increase by one billion people over the next fifteen years and by nearly two billion through 2030 (UN 2008a); and (b) mounting evidence suggests that urban population density (population per unit of urban land area) is decreasing for a number of cities in both the developing and the developed world (Angel 2005; Schneider and Woodcock 2008). When combined, these two factors point to the potential for a rapid expansion of urban areas worldwide within the next twenty-five years. The magnitude of the demographic trend and its developing world character are evident in Figure 1.2b, which traces UN urban population estimates and projections by world region from 1950 to 2030. Figure 1.2b reveals that whereas the urban residents of the developing countries accounted

[2] The level of *urbanization* describes the fraction of a population living in urban areas. *Urban expansion* refers to an increase in the areal extent of urban land.

for only 42 percent of the global urban population in 1950, today they represent 71 percent of the total, and by 2030 they are projected to account for 79 percent.

Although we are confident that urbanization will continue throughout the developing world during the next quarter-century (Brockerhoff 2000; Montgomery et al. 2003), there is considerable uncertainty behind aspects of both the UN's pre-2005 estimates and their post-2005 projections (Figure 1.2b). Aside from the unknowns in effectively modeling urbanization (Montgomery et al. 2003), there are significant challenges in providing a consistent, operational definition of the term 'urban.' In a review of the UN's statistics on urban populations, Utzinger and Keiser (2006) point out that 228 member nations employ at least 10 categories of urban classification, drawing on various combinations of population size and density, administrative boundaries, and economic activities. Global urban maps produced using consistent methodologies that incorporate synoptic satellite observations could improve current census-based urbanization estimates and provide useful constraints for models of urban expansion and demographic trends.

(2) Global urban maps can improve our understanding of the influence of urban areas on the biosphere. As cities grow, the degree and complexity of their effects on ecosystems expand, prompting some to define a uniquely urban ecology (Kaye et al. 2006; Pickett et al. 2008). Despite their small area relative to the total land surface, urban locations are known to impact environmental systems across scales, including species diversity (Davies et al. 2006), microclimate (Oke 1982; Unger 2001), phenology and net primary production of urban-proximate vegetation (Milesi et al. 2003; Zhang et al. 2004), and global climate and biogeochemical systems (Calbo et al. 1998; Oleson et al. 2008). Relative to rural areas, cities appropriate a disproportionate share of the Earth's carrying capacity in terms of both resource inputs and waste sinks (Folke et al. 1997). The ecological footprint of urban areas extends beyond city boundaries, including the conversion of natural ecosystems, the loss of valuable agricultural land, fragmentation of natural habitats,

contamination of air, soil and water, increased water use and runoff, reduced biodiversity, and introduction of non-native species (Rees 1992; Pickett et al. 1997; El Araby 2002; Alberti 2005). In order to confront these environmental consequences of urbanization and urban expansion, we need to build maps of human settlement that are global, reliable, of adequate spatial resolution, and regularly updated.

(3) Global urban maps can help us adapt when the biosphere/geosphere "pushes back." We live on a dynamic planet that is undergoing global-scale change. Hazards such as volcanoes, earthquakes, tsunamis, hurricanes, cyclones, floods, and droughts may be more destructive in areas that have undergone urbanization or other geo-engineering efforts (Ehrlich 1991; Stren et al. 1992). The devastation of New Orleans during Hurricane Katrina is a good example; intense management of the Mississippi river destroyed wetlands and barrier islands south of the city that might have helped to weaken the storm before it moved inland (Bourne 2000; Travis 2005). Although both urban and rural areas are vulnerable to natural disasters, the scale of the immediate humanitarian needs and the magnitude of the recovery efforts are often higher within densely settled urban areas. The hazards imposed by climate-change add a new set of vulnerabilities which span longer time scales, including eustatic sea-level rise and significant shifts in precipitation and temperature regimes (IPCC 2007). The degree of damage inflicted by all of these hazards is strongly proximity-dependent, which is why it is essential that we have spatially explicit estimates of urban populations at scales appropriate to the full suite of potential hazards. Regularly updated global urban maps could serve as the basis for vulnerability assessments, improving both our preparations for and our responses to natural disasters.

(4) Global urban maps can help illuminate the connection between settlement characteristics and human health and well-being. For the first time in recorded history, urban environments influence the health and well-being of the majority of the world's population. In addition to many of the environmental issues mentioned above,

rapid population growth in cities of developing countries with a paucity of resources can lead to increased urban poverty (Wratten 1995; Brockerhoff and Brennan 1998), greater economic and social inequality (Massey 1996), and increased exposure and susceptibility to disease (Montgomery et al. 2003; Keiser et al. 2004). At the same time, sustainable urban expansion within a favorable development climate can lead to economic growth, higher incomes, technical innovation, more efficient land and energy use, better living conditions, cleaner water, and increased access to health care and education (Sassen 1994; Castells 1996; Quigley 1998; Van Vliet 2002). Overall, the only certainty with respect to urban growth is that it is going to occur rapidly for at least the next twenty-five years. The ultimate social, economic, and environmental impacts are complex and will depend on the soundness of our urban planning decisions (Burgess 2000; McGranahan and Satterthwaite 2000, 2002). To make informed decisions regarding the future of urban expansion, regional and national planners cannot rely on a piecemeal patchwork of city-scale maps. What is needed is a consistently-produced, regularly-updated, and accuracy-assessed global urban map.

B. Methodologies behind global urban maps

Since 1990, there have been several attempts to estimate the total urban extent of the Earth using national- and regional-scale tabular methods. The basic approach is to combine urban population estimates from census data with regional estimates of urban population density. Applying this technique to country-level urban population estimates from the United Nations (UN) and urban density estimates from unnamed sources, Douglas (1994) arrived at a global figure of 2.47 million sq. km of urban land circa-1985. Grubler (1990) used similar global values on urban populations in combination with a small collection of city-scale studies of the built environment to estimate a global total of 1.3 million sq. km for urban area circa-1990. Angel et al. (2005) used satellite imagery and census data to calculate regional urban population densities, and combined these densities with a global

dataset of city populations circa-2000 to arrive at an estimate of approximately 400,000 sq. km of urban land within cities of 100,000 people or more.

The first attempt to build a global land cover map based on satellite observations began with the Global Land Cover Characteristics Database (GLCC), a joint effort between the United States Geological Survey (USGS) and the European Joint Research Center (JRC), based on one year (1992-1993) of observations from the National Aeronautics and Space Administration (NASA) Advanced Very High Resolution Radiometer instrument (AVHRR) (Loveland et al. 2000). Because the AVHRR-based vegetation indices used by the GLCC were not particularly well suited to classifying non-vegetated areas, the GLCC group chose not to map urban areas (Loveland et al. 2000). Instead, GLCC rasterized the urban polygons from the Digital Chart of the World (DCW), which is a public-release version of the Vector Map Level Zero (Danko 1992). The same approach was used by the University of Maryland (UMD) group in their global AVHRR-based land cover map (Hansen et al. 2000). Both the GLCC and the UMD maps are not included here; instead, we include VMAP level 0, which is the common data source underlying their representations of urban areas.

Table 1.1 describes the eight global urban maps and two urban-related maps included in this chapter, together with the acronyms that will be used from this point forward. They are arranged in order of increasing global urban extent (from left to right in Figure 1.1 and top to bottom in Table 1.1), with the two urban-related maps listed in the last two rows. Each of these maps approaches urban land from a unique perspective (Table 1.2), employing methodologies that draw on a sometimes-overlapping pool of remote sensing imagery, ground-based census results, geographic information systems (GIS) data layers, and other global maps. Three of these maps — Vector Map level Zero (VMAP0), Global Landcover 2000 (GLC00), and GlobCover (GLOBC) — are general multi-class land cover maps that include an urban class. Three are binary (presence/absence) maps devoted entirely to urban land: Moderate Resolution Imaging Spectroradiometer (MODIS) Urban Land Cover 1 km

(MOD1K), MODIS Urban Land Cover 500 m (MOD500), and Global Rural-Urban Mapping Project (GRUMP).[3] Two maps, Global Impervious Surface Area Map (IMPSA) and History Database of the Global Environment (HYDE3), characterize urban land as a continuous variable: the fraction of impervious surface [4] and the fraction of urban land, respectively. The remaining two maps, LandScan 2005 (LSCAN) and Nighttime Lights (LITES), measure continuous variables closely associated with urban land: the ambient human population (defined as the population of a given area averaged over a 24-hour period) and the intensity of stable nighttime illumination, respectively. The majority of the maps are conducted at approximately 1 km resolution, with the exception of HYDE3 (9,270 m), MOD500 (463 m), and GLOBC (309 m).

These maps are beset by the problem of a consistent operational definition of urban areas. There is no generally accepted set of criteria for creating a binary urban-rural classification. This challenge is not unique to the land cover mapping community; as mentioned earlier, the UN Population Division faces the same problem in their national-level census estimates of urban populations. The map legends employed by our eight global urban maps reveal the same lack of consensus and precision regarding a definition of the urban land category, for example: VMAP0 uses the term "built-up," GLC00 and GLOBC make use of "artificial surfaces and associated areas," MOD1K and MOD500 employ "urban and built-up," and GRUMP utilizes "urban extent." Although not reflected in these simple legends, each of the eight global urban maps model urban areas based on a complex matrix of attributes, including: (1) remotely sensed daytime observation of impervious surfaces, (2) remotely sensed nighttime observation of artificial illumination, and (3) map layers

[3] Although both MODIS land cover maps include 17 land cover classes, the urban class is created in a wholly independent process from the other 16 land cover categories (Schneider et al. 2003, forthcoming).

[4] Impervious surfaces reduce the penetration of rainwater into the soil to levels below that of undisturbed land. Examples of impervious surfaces can involve many forms of human disturbance, including: paved and dirt roads, parking lots, buildings, airport runways, etc. Ridd (1995) introduced the term to the remote sensing literature.

characterizing the built environment or census data. Only one of these attributes is central to the urban models of every global urban map: artificial impervious surface area or built environment. The built environment is visible in the daytime satellite imagery that underlies the VMAP0, GLC00, GLOBC, MOD500, MOD1K, and LSCAN maps. Although GRUMP, HYDE3, and IMPSA do not directly include daytime imagery in their models, they rely heavily on maps that do (VMAP0, GLC00 and LSCAN). For the remainder of this chapter, the term 'urban' refers the human built environment. For the six binary raster maps considered here, the 'urban' label means that a pixel is more than 50 percent built-up. For the two continuous maps, the actual amount of built-up area is described by the 'percent-urban' label.

Table 1.2 lays out the complex relationships between the global maps and their input data, and serves as a reference point for a more detailed discussion of their methodologies. VMAP0 is the earliest global map that includes an urban class. It is a GIS-product created by the United States (US) through digitizing a large collection of roughly 1:1,000,000 scale maps and navigational charts starting in the 1960's. The Digital Chart of the World (DCW) map was based on a 1992-version of VMAP0 (Danko 1992). The urban polygons of VMAP0 often trace the outer edge of urban areas, without delineating interior patches of non-urban land. These urban polygons are not labeled with the date of the underlying map from which they were extracted, and are sometimes poorly geo-located. Because VMAP0 was designed to support the needs of the US government, it also contains a higher level of detail within regions of national interest to the US (e.g. Russia). Nevertheless, because VMAP0 is a global dataset and a conservative estimate of urban land area, it is used as part of the input stream for GLC00, GRUMP, and HYDE3. LSCAN relies on higher resolution VMAP products (levels one and two), which are at scales of 1:250,000, are based on more recent imagery, and are not publically available.

The five columns on the left side of Table 1.2 describe maps that draw on a full year of coarse-resolution satellite imagery for their primary input data: MOD500, MOD1K, GLC00, GLOBC, and IMPSA. MOD500 and MOD1K are derived from the supervised classification of bi-monthly multispectral daytime data spanning the year 2001 from the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument aboard NASA's *Terra* and *Aqua* satellite platforms (Schneider et al. 2003, forthcoming). MOD500 is based on newly released Collection Five 463 m-resolution nadir bi-directional reflectance factor-adjusted data (NBARs), and MOD1K is based on the earlier Collection Four 927 m-resolution NBARs data (Schaaf et al. 2002). Both MOD1K and MOD500 employ manual interpretation of medium-resolution (28.5 m) Landsat imagery and other ancillary data sets to construct training sites for supervised classifications. GLC00 is also built using one year of approximately 1 km resolution data from 2001, drawing on the Satellite Pour l'Observation de la Terre (SPOT4)-VEGETATION instrument. The GLC00 map was completed by separate teams in eighteen production regions, each relying primarily on unsupervised classification methods (Bartholome and Belward 2005). The MOD1K and GLC00 maps incorporate LITES data to constrain their classifications (1996-97 and 1994-95 LITES respectively), while MOD500 is the first global map to rely entirely on daytime multi-spectral observations.

The GLOBC map employs 309 m-resolution data spanning May 2005 – April 2006 from the Medium Resolution Imaging Spectrometer (MERIS) instrument aboard the ENVISAT platform. In general, the GLOBC methodology exploits unsupervised approaches similar to those of GLC00. The exception for GLOBC is Australia, where it appears that the road network and other detailed map layers were included during processing. Over several large regions and countries, GLOBC has relied on the earlier GLC00 map for the urban class, including South America, much of Western Asia, much of Africa, and all of India and Japan. GLOBC v2 was released subsequent to the completion of this chapter and is therefore not

included here (although it is included in **Chapter 2**). This new version of GLOBC includes fewer fill values and 7.3 percent more urban land. Finally, IMPSA models impervious surface area globally by employing LSCAN 2004 and radiance-calibrated LITES 2000-01 data (Elvidge et al. 2007). The IMPSA model was trained using a medium-resolution Landsat map of impervious surface area for the US produced by the USGS as part of the National Land Cover Data set project (Vogelmann 2001; Yang et al. 2003).

The remaining two urban maps in Table 1.2 (HYDE3 and GRUMP) employ a combination of remote sensing and ground-based inputs. GRUMP integrates VMAP0, thresholded 1994-95 LITES, census data, and a variety of ancillary GIS data sets such as city gazetteers (CIESIN 2004). HYDE3 combines LSCAN 2005 population density with 2003 UN national urban population estimates, city gazetteers, and assumptions about mean urban population densities to estimate the fraction of urban land cover (Goldewijk 2001, 2005). One additional global land cover map that also contains a representation of urban areas is GeoCover Land Cover (LC; not listed in Table 1.2). GeoCover LC is a supervised land cover classification of the GeoCover image archive, a near-global collection of circa-1990 and 2000 Landsat imagery (Tucker et al. 2004). Although the underlying 28.5 m-resolution imagery is freely available (Global Land Cover Facility 2007), the land cover classification is a commercial product produced under NASA contract. GeoCover LC includes a class for "urban / built-up, developed areas greater than 60 m wide." Where available, GeoCover LC is an input into the LandScan product stream, but is not included in this analysis because of data quality and availability issues.[5]

Two additional products included in this study, LITES and LSCAN, map urban-related variables. LSCAN models the ambient human population using GeoCover LC,

[5] A global analysis of GeoCover LC revealed that the imagery is often obscured by clouds in the tropics, the map relies on outside datasets for coverage of the US, and land cover data is missing for regions above 60° N, much of Europe (except for a land-water mask), all of Australia and New Zealand, and much of Central Asia. In addition, distinct differences in urban mapping methodologies are apparent at Landsat scene boundaries, particularly in Asia. Finally, the high cost of GeoCover LC data makes it impractical for global-scale academic research.

MOD1K, VMAP level one and above, GIS census products, Landsat data, and high-resolution imagery (1 to 5 m). Although LSCAN was originally conceived during work related to LITES (Sutton et al. 1997; Dobson et al. 2000), the current LSCAN versions do not rely on LITES data (Bhaduri et al. 2002). LSCAN is an ongoing project, and each year the LSCAN team releases a new global map based on the latest census estimates from the International Programs Center at the US Census Bureau and new urban imagery and map layers.

The LITES map is created by the Earth Observation Group at the National Geophysical Data Center (NGDC) using data from the Defense Meteorological Satellite Program's Operational Line Scanner (DMSP-OLS), a nighttime imaging satellite that has a primary mission of monitoring cloud cover by moonlight. NGDC models the average illumination intensity of human settlements and activities for all cloud-free observations within a given year by compositing many individual images at 2.2 km resolution, later resampled to 1 km (Elvidge et al. 2001; NGDC 2007). The DMSP-OLS instrument applies a variable gain during flight, and the gain settings are only monitored on a select number of orbits. Because of these variable gain settings, it is not possible to convert standard LITES composites into radiance values. However, the NGDC group also produces a smaller number of radiance-calibrated composites using fixed-gain imagery (Elvidge et al. 2007). Whether or not the composites are radiance-calibrated, thus far, the relationship between LITES and urban areas appears to be complex and context-dependent (Imhoff et al. 1997; Henderson et al. 2003; Schneider et al. 2003; Small et al. 2005), which is why the IMPSA team draws on LSCAN as well as LITES data for their urban model.

C. Previous studies

Urbanized areas make up a small portion of the Earth's total land area, and the urban class is rare when compared to land cover types such as forest, grassland, and savanna. This minority status has led to scant urban-specific comparisons of global land cover maps. This

study, together with a companion article (Potere and Schneider 2007; forthcoming), is the first systematic global comparison focused on urban maps. Although there has been a recent increase in the number of global-scale land cover comparison projects, these assessments have either failed to include the urban class, considered only areal extent, included only a subset of currently available urban maps, or inadvertently compared multiple maps where the urban layer is actually derived directly from VMAP0 (Hansen and Reed 2000; Giri et al. 2005; Jung et al. 2006; Mayaux et al. 2006; McCallum et al. 2006; See and Fritz 2006).

At regional and city scales, there have been a number of comparative studies that include some of the urban maps considered here. Schneider et al. (2003) compared VMAP0, LITES and a MODIS-derived urban map for cities in North America, uncovering variable amounts of underestimation of urban land by VMAP0 and systematic overestimation of urban land by the 1994-95 nighttime lights data (thresholded) when compared against more recent medium-resolution maps of urban land derived from Landsat imagery. Tatem et al. (2005) compared a map of urban areas in Kenya to five global scale urban maps, and unsurprisingly reported that a medium-resolution Kenya map based on Landsat and Radarsat imagery is of higher accuracy than the coarse-resolution global maps. Finally, Small et al. (2005) assessed thresholded 1994-95 and 2000 LITES data against Landsat-based maps from a global sample of 17 cities, confirming that no single LITES threshold is suitable for mapping urban land.

1.2 RESEARCH METHODS AND RESULTS

A. Map preparation

Not all of the global urban maps share a common geospatial model (see Table 1.1, rightmost column). For our analytical environment we selected a geographic projection, 30" arc-second raster (approx. 0.86 sq. km cells at the equator), and the WGS-84 datum. LITES,

LSCAN, GRUMP, and IMPSA required no modification. HYDE3 was down-sampled to 30" arc-second cells from 5 arc-minute cells (approx. 9.27 km wide at the equator). The MODIS land cover group provided a geographic reprojection of their native sinusoidal 1 km and 500 m products.[6] For the 463 m-resolution MODIS product and the 309 m-resolution GlobCover map, we aligned the 15" arc-second and 10" arc-second grids (respectively) with the common 30" arc-second grid. Although GLC00 is in a geographic projection, the pixel size was resampled from 32.14" arc-seconds to 30" arc-seconds. Finally, the VMAP0 product is in a vector format, and was converted to raster. All of the modified products were checked against their native counterparts at various stages and biases were determined to be negligible throughout the map preparation process.

To assess the maps across continental, sub-continental (regional), and national scales, it is necessary to have a coarse-resolution map of international boundaries. Unfortunately, few options exist at 1 km spatial resolution with a consistent land and water boundary. We opted to use the international boundary grid produced and distributed by the LSCAN program, as it is at 30" arc-second resolution and is updated annually. To account for significant differences in how each map delineated land and water, we created custom international border files for each product. In this step, land pixels from each map that fell outside of the LSCAN land-water mask were retained as land and assigned to the country of their nearest land neighbor in the LSCAN country boundary file. This procedure prevents the land water mask from eliminating some of the coastal areas within several of the maps, which can result in serious under-estimation of urban land.

We then cross-walked the LSCAN country names (derived from the US Census Bureau) to their counterparts in the system employed by the UN Statistics Division and the International Organization for Standardization (UN 2008b). These UN country codes are

[6] For simplicity, we refer to these maps as MOD500 and MOD1K, but their actual resolutions are 463 m and 927 m, respectively.

accompanied by a twenty-region scheme for delineating world regions. The regional scheme employed throughout the rest of this chapter is a slightly modified version of this UN regional scheme (Figure 1.2). These modifications include: (1) reassignment of Sudan from the Northern Africa region to the West Africa region in order to maintain Sudan's traditional association with the super-region of Sub-Saharan Africa; (2) combining the Australia and New Zealand region with the North America region on the basis of demographic and economic similarities (e.g. moderate levels of projected population growth and in-migration); (3) merging Japan with the Western Europe region due to Japan's similarly high levels of contemporary urbanization and projected population decline; and (4) combining the remaining UN regions into ten super-regions, the majority of which are related to the nine-region scheme employed by the UN Human Settlement Programme (UN-HABITAT; Angel et al. 2005). Figure 1.2 depicts these ten regions and describes the UN's urban population projections for each region through 2030.

The global GLC-2000 v1.1 map indicates that virtually no urban areas are present in southern China. To correct for these and other urban omissions, we relied on recently updated regional GLC 2000 maps. These updates are still based on circa-2000 imagery, but represent refinements in the regional protocols by the GLC 2000 production regions for Europe and Asia. By including these improved regional maps, the GLC00 map used in this analysis contains 9 percent (28,000 sq. km) more urban area than the original global map. The following regions received additional urban area: Southern Europe (9,000 sq. km), Northern Europe (6,800 sq. km), East Asia (4,700 sq. km), South East Asia (4,500 sq. km), and South Central Asia (2,300 sq. km).[7]

[7] MOD1K also includes a small region of urban-omission, just visible as a thin stripe over southern Europe in Figure 2.7g, which was not corrected.

B. Portraits of global urban maps across scales—from cities to continents

(1) City-scale descriptions. Figure 1.3 depicts the eight urban maps and two urban-related maps for Beijing and Tianjin, China (top row), Mumbai, India (second row), Paris, France, (third row), Moscow, Russia, (fourth row), and Cairo, Egypt (bottom row). The first two columns are for the urban-related maps: Nighttime Lights for one year spanning 2000-01, and LandScan 2005. The global urban maps begin with the third column (from left to right): VMAP0, GLC00, GLOBC, HYDE3, IMPSA, MOD500, MOD1K, and GRUMP. Although only HYDE3 and IMPSA map urban land as a continuous variable (fraction of a pixel classified as urban), for the purpose of visualization we create continuous variables for the binary (urban/rural) maps by aggregating the maps from 30" arc-second pixels to 1.5' arc-minute cells (approx. 3 km resolution at the equator). Each new 1.5' arc-minute cell is shaded from zero percent urban (blue) to fully urban (red) using a linear scale. The same legend is used for IMPSA and HYDE3, but those pixels remain at their native resolution (30" arc-seconds and 5' arc-minutes, respectively). All of the city subsets are approximately 150 km by 150 km and oriented north-up.

Even a quick glance at the VMAP0 column confirms the often-dated nature of this map, particularly in the countries of the developing world; significant urban expansion has occurred in all of these cities since the charts and maps that underlie VMAP0 were digitized. In the GLC00 column, the heterogeneous impact of the distributed mapping strategy is apparent; the Europe mapping team clearly enforced a different standard for the urban classification of the Paris metropolitan area than the China team did for Beijing and Tianjin. The close relationship between the GLOBC and GLC00 research teams is evident in Mumbai, where GLOBC relied completely on the GLC00 map. The important role of both LITES and LSCAN is clear when one considers the shape of IMPSA's Cairo map, which includes both the distinctive fan shape of the Nile delta and the urban core of the city center. The two MODIS maps are quite similar to IMPSA in overall form, but are somewhat more

expansive. Comparing GRUMP to LITES reveals the important role of the thresholded nighttime lights data in the GRUMP mapping process. Note that several urban clusters in the LITES imagery for both Beijing-Tianjin and Cairo do not appear in GRUMP because the LITES imagery in Figure 1.3 is for 2000-01, while GRUMP relied on 1994-95 LITES. GRUMP's use of buffers surrounding gazetteer city points is also noticeable in the cluster of circular urban patches in the lower left corner of the GRUMP Beijing map. Overall, these examples from Figure 1.3 demonstrate the diversity of the eight global urban maps, and the challenges of inter-map comparison.

We continue this city-level comparison by examining the size distribution of urban 'patches,' where patches are defined as contiguous areas of urban pixels (Figure 1.4).[8] This analysis provides information on the effective minimum mapping unit used in each map, as well as some indication of whether urban area is distributed in large clusters (e.g. extensive cities such as Chicago), or in a large number of small-sized patches (e.g. small towns and villages, one to two sq. km in size). The absolute number of urban patches varies widely across data sets (in descending order): MOD500 (140,000), GLOBC (88,000), MOD1K (54,000), VMAP0 (32,000), GLC00 (22,000), GRUMP (21,000), and HYDE3 (17,000). The number of patches is influenced by the resolution at which each map is produced; the two maps conducted at resolutions finer than 30" arc-seconds (GLOBC and MOD500) have higher numbers of patches than the rest, while HYDE3 (5' arc-minutes resolution) naturally has the fewest patches. IMPSA is not included here or in Figure 1.4 because of the difficulty of delineating individual patches from a continuous (percent impervious surface) 1 km-resolution map.

The most striking feature of Figure 1.4 is the strong log-linear decay of patch size from 2 to 1,000 sq. km for all seven maps. Although this trend is consistent with the long-recognized rank-size rule for city population sizes (Berry 1961; Nishiyama et al. 2008), it is

[8] Patches were described using a rooks-case rule for adjacency (eight neighbors for each pixel).

interesting how marked the trend is among global maps which are based on such a diverse set of methods. On the far left of Figure 1.4 (smallest patch sizes), the double-peaks for the four imagery-based distributions (GLC00, GLOBC, MOD1K, MOD500) correspond to patches that contain only one or two pixels. Because the geographic projection used for map comparison is not equal-area, single pixels in each map vary in size from 0.86 sq. km at the equator to 0.42 sq. km at the southern tip of Greenland. The overall shape of the GRUMP distribution is quite distinct, and it reflects the larger size of most GRUMP patches relative to all other global urban maps. Yet for the majority of city sizes, GRUMP adheres to the pattern of roughly log-linear decay.

(2) Regional-scale descriptions. Figure 1.5 and Table 1.3 present the total areal extent of urban land for each of the eight urban maps at the scale of world regions. The thickness of each horizontal bar in Figure 1.5 is scaled to reflect the relative amount of urban land within each of the eight maps. These horizontal bars are divided into colored sections representing the relative distribution of urban land within each of the ten world regions (the ten-color scheme is the same as in Figure 1.2). Because both axes are scaled proportionately, it is possible to make direct comparisons between segments and rows. The most obvious feature of Figure 1.5 and Table 1.3 is the previously-noted order of magnitude difference in total urban area, revealed by the pronounced differences in the thickness of the eight horizontal bars.

The pattern of regional banding within each map (measured on the horizontal axis) reveals large differences at the scale of world regions (Figure 1.5). VMAP0, GLOBC, and IMPSA have some of the most unusual inter-regional banding patterns. The results for VMAP0 reflect the cold-war legacy of this data source, as shown in the wide band for Eastern Europe. The US agencies responsible for VMAP0 likely focused mapping efforts on this region, causing VMAP0's relative amount of Russian urban land to be far above that of any other map (Table 1.3). GLOBC has by far the widest East Asian band, nearly the same width

as North America, Australia, and New Zealand. GLOBC (version 1) is also unusual in that it attributes more urban land to Australia than to the entire US (Table 1.3); this may be linked to the decision to use road networks in mapping Australian urban areas. Overall, IMPSA's regional bands are the most distinct from the rest. In relative terms, IMPSA has the least amount of urban area in the North America, Australia, and New Zealand region (dark green), and the most in Asia (red, salmon, and gray). This difference can be explained by IMPSA's reliance on LSCAN; because LSCAN tracks population density, it favors the two demographic billionaires of India and China.

(3) Global-scale descriptions. Figure 1.1 presents the total urban extent of the eight maps at the global scale. The urban extents vary by an order of magnitude, from 0.3 to 3.5 million sq. km. For scale, the difference in the GRUMP and VMAP0 totals is approximately equal to the land surface area of India. It is a challenge to move beyond the plots of Figures 1.1 and 1.5 and to display 30" arc-second resolution maps at a global scale with sufficient detail for a meaningful comparison. At a reasonable resolution of 300 dots-per-inch, the 40,000 x 20,000 raster (column x row) would be roughly 3.4 x 1.7 m in size. Even at this resolution, the urban class would be difficult to discern since urban areas occupy at most one to three percent of the Earth's 140 million sq. km of land. One way to deal with these issues is through the use of Geodesic Discrete Global Grids (DGGs) (Sahr et al. 2003). DGGs are a class of equal-area, uniformly distributed partitions of the Earth's surface. The hexagonal partitions provide a scheme for assessing the urban area of a given map that is independent of arbitrarily defined political boundaries or world regions. Moreover, the hexagons are constant in size (unlike the varying raster cell sizes of a geographic projection), and constant in cell shape (unlike the deformities of sinusoidal and other equal-area global projections). By computing the square root of the facet area, we can also derive a rough estimate of the effective spatial resolution of the DGGs across different hexagonal sizes (Small and Cohen 2004). For illustration, in Figure 1.6 we employ five

resolutions of DGGs with hexagonal cell sizes ranging from 800 to 70,000 sq. km (29 to 264 km resolution).

By aggregating the 30" arc-second pixels of our global urban maps up to the 51 km-resolution DGG facets (yellow facets in Figure 1.6; 2,591 sq. km in size), we have created a series of maps that provide a comprehensive view of urban areas at a global scale (Figures 1.7a-h). In all of these maps, royal blue areas represent completely urban-free facets. Of the eight maps, IMPSA (Figure 1.7d) has by far the fewest facets that are completely free of urban land, while GLC00 (Figure 1.7b) has the most. From these global-scale visualizations, regional trends are apparent: VMAP0 (Figure 1.7a) portrays Eastern Europe as a more extensive urban network than Western Europe, GLC00 depicts a world in which Europe and the Eastern US are the only urban-dominated regions and Africa and East Asia are almost completely urban-free, and GLOBC is unique in depicting Australia as an urban system on par with Western Europe or North America. Overall, HYDE3 has the most uniform view of the distribution of urban land, while IMPSA portrays a massive portion of the world's surface area as very low-fraction urban (the light blue regions of the IMPSA map represent DGG facets that are less than 0.05 percent urban).

In Figure 1.7i, the amount of urban area for each grid cell is averaged across all eight maps. The total extent of this mean urban map is 864,000 sq. km. Because GRUMP's urban area is more extensive than any of the other five maps, GRUMP has strongly influenced this mean urban map, as is evident in a comparison of Figures 1.7h and 1.7i. In the mean map, the largest blocks of intensely urban areas include the Eastern US, Western Europe, India, Eastern China and Japan, and to a lesser degree the south-eastern coasts of South America and Australia. The largest contiguous urban-free areas are the Sahara desert, interior Australia, Siberia, Mongolia, Northern Canada and Greenland, the Rhub al Khali of Saudi Arabia, and to a lesser degree the rainforests of South America and the Kalahari Desert (were it included in this analysis, Antarctica would be the largest urban-free landmass).

C. Inter-map comparisons

(1) Per-pixel comparison. Contingency tables, sometimes referred to as confusion matrices, are a common way to conduct inter-map comparisons (Table 1.4a). These tables are produced by overlaying a set of maps and estimating the areas of agreement and disagreement pixel-by-pixel. We constructed contingency matrices for all eight global urban maps (Table 1.5a). Each entry describes the fraction of urban pixels in *map A* (shown in the rows) also mapped as urban in *map B* (shown in the columns). For example, although VMAP0 is a part of the input data streams for both GLC00 and GRUMP (Table 1.2), only 41 and 76 percent of VMAP0's urban pixels are also mapped as urban in GLC00 and GRUMP, respectively. Some portion of this could be due to differences in geo-registration; relative to the other global maps, GLC00 appears to be shifted 1'30" eastward (three pixels) and 30" arc-seconds (one pixel) southward over parts of North America, Central America, and the Caribbean. It is not possible to correct for this shift because the geo-registration problem appears to be non-constant within the northwest hemisphere. Similar registration problems are not apparent for GLC00 in any other region of the world, or for any other urban map examined in this chapter. A thorough evaluation of geo-registration is beyond the scope of this chapter, but is addressed in **Chapter 2**.

The GLC00 example highlights one of the fundamental problems of conducting a global map comparison at the pixel-level: a shift of a single pixel can significantly reduce map agreement. Another issue that arises when interpreting these simple contingency tables is the order of magnitude variance in total map area (Figure 1.1). For instance, 60 percent of GLC00's urban pixels are also mapped as urban in MOD1K, while only 26 percent of MOD1K urban areas are mapped in GLC00. This is not surprising when one considers that the urban extent of MOD1K is 2.4 times the size of GLC00. Even if every MOD1K urban pixel agreed with GLC00, that would only represent 42 percent of MOD1K urban land.

There are several statistics based on contingency tables that have long been used to confront the challenges of assessing map agreement against random chance and for maps of varying extent (Congalton and Green 1999). There remains considerable debate in both the remote sensing and the spatial ecology literatures concerning the selection of optimal map agreement measures (Fielding and Bell 1997; Manel et al. 2001; McPherson et al. 2004; Allouche et al. 2006; Lobo et al. 2007; Liu et al. 2007; Foody 2006, 2007), yet there is broad consensus that statistics based on the entire contingency table are welcome improvements on the basic accuracy, sensitivity, and specificity measures (Table 1.4b). In our comparison, we follow the recommendations of Foody (2006, 2007) and rely on several agreement measures across a wide range of spatial scales. We choose Cohen's Kappa (Congalton and Green 1999), the true skill statistic (TSS; Allouche et al. 2006), and the normalized mutual information coefficient (NMI; Forbes 1995) for their consideration of both omission and commission errors and their reliance on the entire contingency table (Table 1.4b).

Cohen's Kappa statistic (κ) is designed to measure the strength of agreement, taking into account the potential for chance agreement (Cohen 1960; Monserud and Leemans 1992; Goldewijk and Ramankutty 2004). Landis and Koch (1977) suggest that Cohen's κ values of 0.00-0.20 can be considered 'slight' agreement, values of 0.21-0.40 as 'fair,' and 0.41-0.60 as 'moderate' agreement; however, they caution that these are only intended as useful benchmarks. In Table 1.5b, we report global-scale Kappa values for the 28 map pairs. The overall impression is that these maps are quite distinct on the per-pixel level, as shown by the 'fair' level of the mean Kappa score across all of the map pairs ($\overline{\kappa}$ = 0.26). Only two map pairs achieve 'moderate' agreement, and these pairs were created by related research groups: MOD1K and MOD500 (κ = 0.48), and GLC00 and GLOBC (κ = 0.47). Of the 17 'fair' Kappa values, the strongest agreements are between GLC00 and VMAP0, and GLC00 and the two MODIS maps. This result is also expected, considering that VMAP0 is an input into the GLC00 product and that MOD1K and GLC00 are both derived primarily from coarse-

resolution daytime imagery. Overall, nine map pairs had Kappa values indicating 'slight' agreement, and IMPSA or GRUMP were involved in each of these pairings, signaling that both maps are quite distinct on a per-pixel scale.

In order to understand how map agreement varies in space, we conducted these same per-pixel comparisons at a regional scale. We began by creating contingency tables and associated map agreement statistics for all of the inter-map comparisons (28 pairs) for each of the ten world regions. We use box-plots in Figure 1.8 to display the medians and inter-quartile ranges of these 28 agreement measures for each region (color-coded to correspond to Figure 1.2). We report Cohen's Kappa, NMI, and TSS measures, where complete agreement is 1.0 for each measure. The dominant trend in Figure 1.8 is that of declining overall agreement as one moves from the affluent regions of North America and Europe (green bars on left) and towards the developing countries of Asia (red, salmon, and gray bars on right). This does not imply that the maps are any less accurate in these lesser-developed regions, merely that there is less inter-map agreement. Across all three agreement metrics, the overall pattern points towards three super-regions: relatively high agreement in the North America, Australia and New Zealand region (also including Western Europe in the TSS measure), relatively low levels of agreement in most Asian regions (South-Central Asia, East Asia, and Southeast Asia and the Pacific Islands), and medium levels of agreement for the remaining regions (Eastern Europe, Central and South America and the Caribbean, Africa, and Western Asia). This trend aligns well with our more complete understanding of cities of the developed world, and the relative challenges of describing the complex and rapidly changing cities of the developing world.

(2) Country-level comparison. By comparing the total extent of urban area within each country, we can approach the issue of map agreement without some of the limitations of pixel-by-pixel comparisons. Large area aggregations using national borders also reduce the importance of any persistent geo-registration problems. We first present our

country-level results by ranking the ten countries with the largest urban extents for each of the eight global urban maps (Table 1.6a). There is considerable variation in these rank lists, with only USA, India, Russia, and China appearing in all of the top ten lists. For three maps, IMPSA, MOD500, and GRUMP, the five countries with the largest mapped urban extents also have the five largest urban populations according to the UN Population Division. For IMPSA, this correlation extends further; IMPSA's top ten rankings closely mirror their order by urban population. This is not surprising considering the important role that demographic attributes play in the methodologies of IMPSA and GRUMP.

We estimate the relationship between the country-level urban extents of each map-pair by using Kendall's Tau (τ), which provides a quantitative assessment of the correlation among the urban area rankings for all 223 countries within the global datasets (Kendall 1938) (Table 1.6b). The overall impression from this measure is one of far greater agreement than that of the per-pixel comparisons (Tables 1.5a-b), with $\overline{\tau}$ = 0.73 and a range of 0.64 - 0.86 (top 8 rows, Table 1.6b); all correlations are highly significant. This country-level analysis reveals that international comparisons of the *relative* areal extent of urban land are far less sensitive to the selection of a particular global urban map than the per-pixel results imply. Of course, the large *absolute* differences between these maps remain important across all aggregation scales. As in the earlier analysis of global Kappa values, the map-pairs with the highest Tau values are either produced by closely related research teams (GLC00-GLOBC: τ = 0.86, MOD1K-MOD500: τ = 0.81) or rely on similar demographic data (HYDE3-GRUMP: τ = 0.81). Of the map-pairs with the eleven lowest Tau values (0.64 – 0.70), GLOBC or GLC00 are involved in ten, indicating that these two maps are very distinct at the national scale.

(3) Multi-resolution hexagonal comparison. As mentioned previously, any comparison of the spatial pattern of urban land in these eight global urban maps is made more challenging by the rarity of the urban class, the very large differences in estimates of

the total urban area for each map, and the potential problems of geo-registration. Although using national borders can help alleviate these problems, the distribution of country sizes varies widely and it is not possible to explore sub-national inter-map differences with this approach. Here, we again turn to Discrete Global Grids (DGGs) to create a global equal-area partitioning system with minimal shape distortion over the entire Earth's surface. For each hexagonal DGG facet, we estimate the percentage of urban coverage for each of the eight global maps. Because the resulting aggregations contain continuous values, we can test for linear correlations. We carry out these tests in Figure 1.9 for each map combination (28 pairs) across all five DGG resolutions (29 to 264 km). We restrict our Pearson correlation coefficient (r) estimates to hexagonal facets that contain urban land in at least one map. Although this multi-resolution global map comparison is similar to the methods used by Goldewijk and Ramankutty (2004) and Greyner et al. (2006), both studies employed a series of traditional rectangular raster grids in a geographic projection. While the raster approach is more straightforward than DGGs, the drawback is that the size and shape of individual cells varies widely with latitude. This effect creates non-uniform sampling, with sparser sampling in tropical versus temperate regions.

The highest overall inter-map correlation curves are for the closely related MOD500-MOD1K maps and for MOD500-HYDE3, a correlation that is more difficult to explain. The correlations among the VMAP0-GLC00-HYDE3 group (upper left of Figure 1.9) are also quite high (at or about 0.8 for all resolutions). This result is understandable considering that VMAP0 was an important input for GLC00 (in some regions, the sole input), and that HYDE3 was the only map to draw on both GLC00 and VMAP0. The lowest overall correlation curve is for VMAP0-GLOBC ($r \leq 0.6$ for all resolutions). In general, GLOBC (version 1) has the lowest correlation curves of any of the maps; GLOBC is involved in six of the eleven curves where $r \leq 0.7$. When one compares IMPSA and MOD1K with either VMAP0 or GLC00, the resulting correlation curves are also quite low ($r \leq 0.7$). Considering

that neither MOD1K nor IMPSA drew on VMAP0, and given the aforementioned similarity between VMAP0 and GLC00, these low correlations are also to be expected.

The slopes of the correlation curves are more difficult to interpret, but may be related to the size distribution and density of urban patches within each urban map and their relationship to the various hexagonal facet sizes. The 28 curves in Figure 1.9 can be divided into three categories: twelve cases of flat curves (no significant change in correlation with change in grid resolution), eight cases of upward sloping curves (increasing correlation with coarser grid resolution), and eight cases of downward sloping curves (decreasing correlation with coarser grid resolution). All but one of the flat, resolution-independent curves involve MOD1K (4 cases) or MOD500 (three cases). The downward sloping curves are more evenly divided, but IMPSA (six cases) and VMAP0 (five cases) are most frequently involved. Interestingly, nearly all of the positive-sloped curves (seven of eight cases) involve GRUMP. This result is likely tied to the unique size distribution of the GRUMP urban patches relative to all of the other maps (Figure 1.4). At the finest DGG resolution of 29 km, each facet contains roughly 900 sq. km of land. GRUMP has by far the most urban patches that are of sufficient size to saturate these facets, potentially reducing inter-map correlations at this finest hexagon grid cell resolution. This effect gradually fades as the pixel sizes become large relative to the mean urban patch size of GRUMP (resolutions above 150 km).

The DGG aggregates are also an effective tool for uncovering regional patterns in the inter-map correlations. The box-plots of Figure 1.10 chart the distributions (median, inter-quartile range, and outliers) of the Pearson correlation coefficients for the 28 global urban map pairs across all five DGG resolutions and all ten world regions. Each region has five box-plots showing correlations from the finest to the coarsest DGG (29 to 264 km, left to right). The colors correspond to the world regional scheme used throughout this analysis (Figure 1.2). From Figure 1.10, it is clear that the North America, Australia, and New Zealand region (left-most box-plot group) has by far the strongest inter-map correlations ($\overline{r}$

= 0.83), indicating the highest agreement in terms of the intra-regional distribution of urban land. By comparison, the Asian regions (four right-most box-plot groups) have the lowest inter-map correlations ($\overline{r}$ = 0.66), and the remaining regions are intermediate ($\overline{r}$ = 0.79).[9]

The eight maps have the most dispersed distribution of correlations in the East Asia region (salmon color, including China). Most of the variance in this region is caused by very low inter-map correlations for all the comparisons that include VMAP0 or GLC00 ($\overline{r}$ = 0.36 and $\overline{r}$ = 0.44, respectively) relative to the other six maps ($\overline{r}$ > 0.60). As previously discussed, this result is an artifact of GLC00's significant urban omissions over East Asia. VMAP0's under-estimation is likely tied to China's exceptionally rapid urban expansion over the past twenty years and the older dates of many of the maps used to create VMAP0.

1.3 DISCUSSION AND CONCLUSIONS

As both recipients of ecosystem services and modifiers of ecosystem processes, humanity is an important part of the biosphere. In order to better understand the role humans play in ongoing global change processes, there is a need for a global, accuracy-assessed, moderate resolution, and regularly updated map of contemporary human settlement. Global urban maps can make an important contribution here, by accounting for the residences of more than half of the human population. To meet this need, eight international groups from government and academia have created global maps that can be used to describe contemporary urban areas (Vector Map Level Zero, Global Landcover 2000, GlobCover, History Database of the Global Environment, Global Impervious Surface Area, MODIS Urban Land Cover 1 km, MODIS Urban Land Cover 500 m, and Global Rural-Urban Mapping Project). Such maps are of significant interest to a wide range of users, including: regional and national planners, disaster management specialists, humanitarian and

[9] When single maps are held out of the box-plots from Figure 10, the same overall regional pattern persists. The one exception is GLOBC; when GLOBC is held out, the mean North American correlation increases to 0.91, and the other regions remain relatively unchanged.

development aid coordinators, epidemiologists and public health officials, demographers, economists, conservation biologists, climatologists, and urban ecologists.

Despite the considerable resources allocated to the task of creating global urban maps, this chapter has revealed that the eight maps differ by as much as an order of magnitude in their estimates of the total areal extent of the Earth's urban land (0.27 – 3.52 million sq. km). Differences in these eight maps persist at the scale of urban patches, countries, and world regions. To better visualize these maps and to conduct a quantitative map comparison, we employed a hexagonal system of Discrete Global Grids (DGGs). An analysis of the spatial distribution of urban land based on these DGGs across a wide range of spatial resolutions (29-264 km) has revealed that inter-map correlations are highest in North America ($\overline{r}$ = 0.83), lowest in Asia ($\overline{r}$ = 0.66), and intermediate in Europe, South and Central America, and Sub-Saharan Africa ($\overline{r}$ = 0.79). It seems that large inter-map differences are driven by a combination of several factors, including differences in the timing of map construction, differences in map resolution and class enumeration, and fundamental differences in each group's approach to urban land. Scale and resolution are of concern because many of these maps combine coarse-resolution inputs with binary classifications (urban / non-urban classes). Because urban land does not occur in neat 30" arc-second square blocks, there are considerable problems inherent in any attempt to infer total areal extent from coarse-resolution binary classifications (Latifovic and Olthof 2004; Ozdogan and Woodcock 2005).

Of the aforementioned sources of inter-map variance, perhaps the most important is a fundamental divergence in each group's approach to defining and modeling urban land. The problem of creating a meaningful and workable characterization of 'urban' is not trivial. In the absence of a clear set of definitions, each group constructs an implicit model of urban extent that can be inferred from their methodologies. The eight global urban maps that emerge are sensitive to many attributes commonly associated with urban areas, including

high population density, extensive built-environment, nighttime illumination, and proximity to transportation infrastructure. The degree to which any one of these attributes contributes to an urban classification is likely regionally dependent and is not specified by any of the makers of these maps. We can do better if progress can be made towards synchronizing legends and negotiating a common set of multi-faceted urban definitions that draw on both the demographic and the physical dimensions of urbanization.

The only way to truly understand the source and impact of the large inter-map differences is to conduct a global map accuracy assessment, the subject of **Chapter 2**. In the longer-term, an improved understanding of each map's strengths and weaknesses in both space and time will facilitate the construction of a suite of composite urban maps that are tailored to particular user-groups. We could address the specific needs of users by first devising a group-specific definition of urban areas (either ordinal or categorical), and next fusing components of existing global urban maps and urban-related maps that best meet the requirements of those definitions.

The human built environment is complex, and no single binary definition of the urban / rural divide is likely to satisfy the demands of a growing user community. With more than two billion new urban residents due to arrive within the next quarter-century, there can be little doubt that a new series of urban maps is on the way. Whether those maps are more successful at capturing the complexities of urban areas depends on how effectively we engage the multi-disciplinary community of map users and map producers to resolve the fundamental issue of urban taxonomy and agree upon a common map assessment regime.

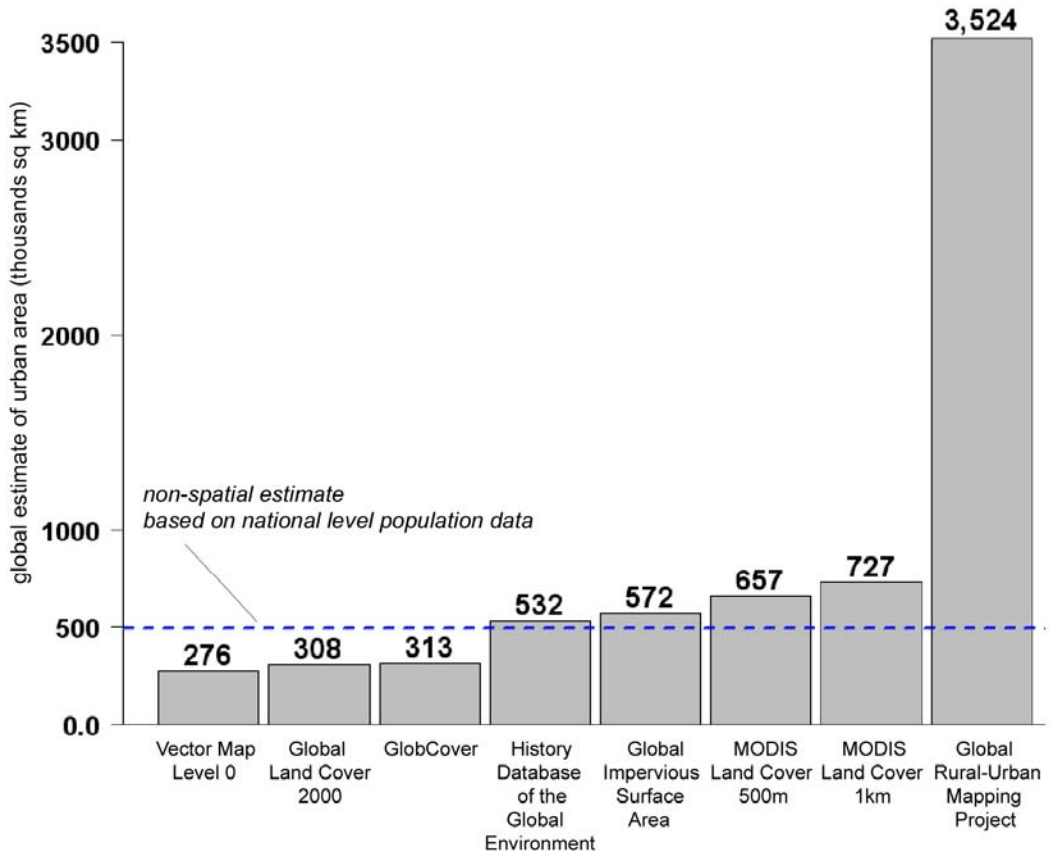
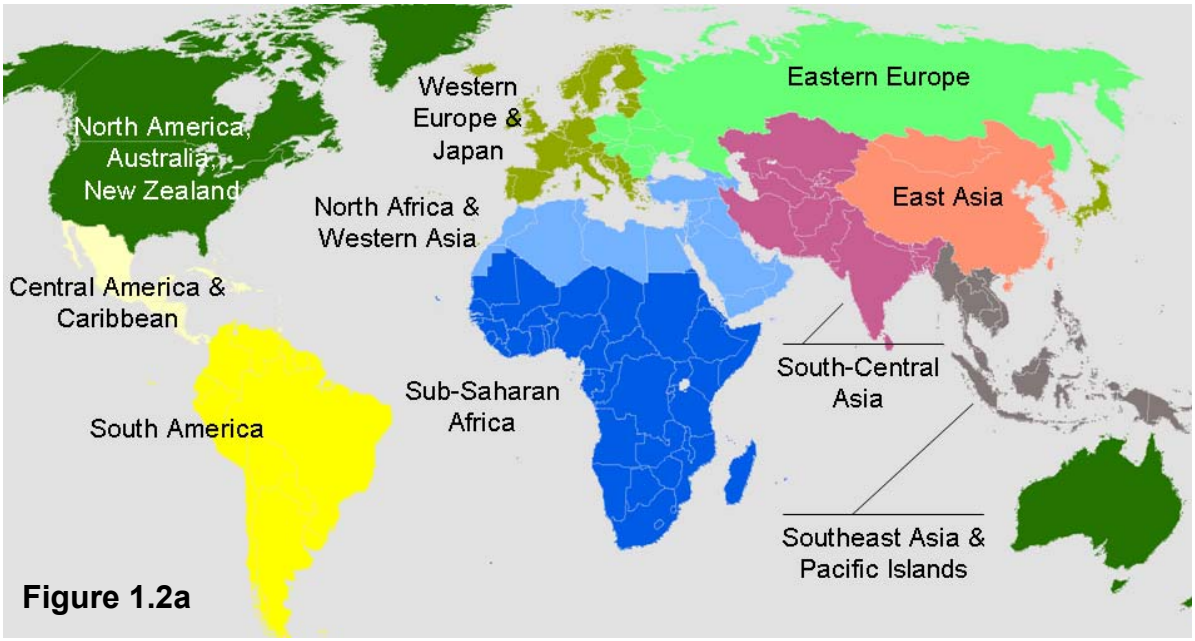


Figure 1.1 Total extent for eight global urban maps. Global extents of urban lands in thousands of sq. km: Vector Map Level Zero, Global Land Cover 2000 v1.1, GlobCover, History Database of the Global Environment v3, Nighttime Lights-based Global Impervious Surface Area beta product, Moderate Resolution Imaging Spectroradiometer (MODIS) 500m Urban Land Cover v5 2001, MODIS 1 km Urban Land Cover v4 2001, and the Global Rural-Urban Mapping Project version alpha. Above each bar is the estimate of global extent in thousands of sq. km. The dotted line is an estimate of urban area based on national-level urban population statistics (UN 2005) and regional-level urban population densities for the year 2000 from Angel et al. (2005). The GlobCover map described here is the version one release; the urban areas in version two have a total area of 336,000 sq. km.



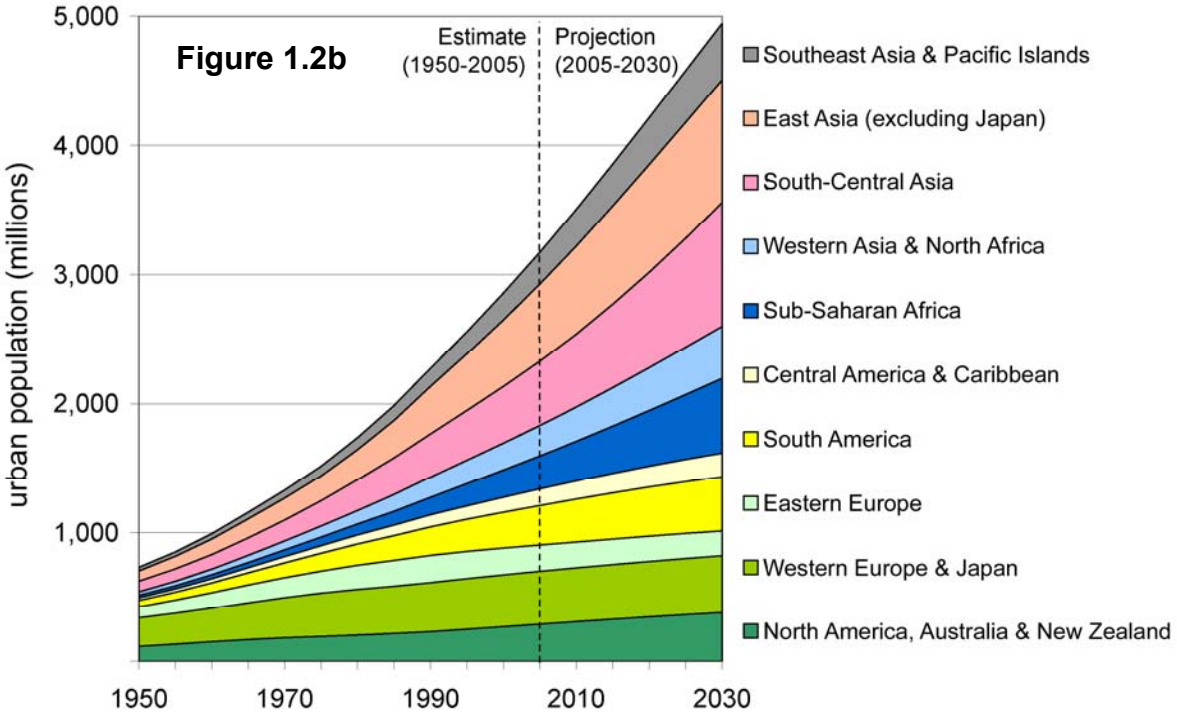


Figure 1.2 Ten world regions. The regions used in this study are modified from the UN regional scheme. The color codes in this map match the legend in Figures 1.5, 1.8, and 1.10. Figure 1.2b tracks the UN Population Division estimates and projections for urban populations by region from 1950-2030 (UN 2005).

Code	Map (Citation)	Producer (Source)	Projec -tion	Map type	Resolu -tion	Extent (sq. km)
VMAP0	Vector Map Level Zero (Danko 1992)	National Geospatial-Intelligence Agency (US) *(http://geoengine.nga.mil)*	geog.	thematic	1:1mil.	276,000
GLC00	Global Land Cover 2000 v1.1 (Bartholome et al. 2005)	European Commission Joint Research Center *(http://www-gvm.jrc.it/glc2000)*	geog. (32")	thematic (22 class)	988 m	308,000
GLOBC	GlobCover (Arino et al. 2007; ESA 2008)	European Commission Joint Research Center *(ftp://uranus.esrin.esa.int/pub/globcover_v2)*	geog. (10")	thematic (22 class)	309 m	*V1*: 313,000 *V2*: 336,000
HYDE3	History Database of the Global Envt. v3 (Goldewijk 2005)	Netherlands Environmental Assessment Agency *(http://www.mnp.nl/hyde)*	geog. (5')	continuous (pct. urban)	9,270 m	532,000
IMPSA	Global Impervious Surface Area (Elvidge et al. 2007)	National Geophysical Data Center (US-NOAA) *(http://www.ngdc.noaa.gov/dmsp)*	geog. (30")	continuous (pct. imp. surf.)	927 m	572,000
MOD500	MODIS Urban Land Cover 500m (Schneider et al. forthcoming)	U. of Wisconsin and Boston U. (US-NASA) *(http://www.sage.wisc.edu)*	sinusoid -al	binary (urban / rural)	463 m	657,000
MOD1K	MODIS Urban Land Cover 1km (Schneider et al. 2003)	Boston U., Dept. of Geog. & Envt. (US-NASA) *(http://www-modis.bu.edu/landcover)*	sinusoid -al	binary (urban / rural)	927 m	727,000
GRUMP	Global Rural-Urban Mapping Project, alpha (CIESIN 2004)	Earth Institute at Columbia University *(http://sedac.ciesin.columbia.edu)*	geog. (30")	binary (urban / rural)	927 m	3,524,000
LITES	Nighttime Lights v2 (Elvidge et al. 2001)	National Geophysical Data Center (US-NOAA) *(http://www.ngdc.noaa.gov/dmsp)*	geog. (30")	continuous (nighttime illum. intensity)	927 m	NA
LSCAN	LandScan 2005 (Bhaduri et al. 2002)	Oak Ridge National Laboratory (US-DOE) *(http://www.ornl.gov/sci/landscan)*	geog. (30")	continuous (ambient human population)	927 m	NA

Abbreviations: DOE, Department of Energy; DMSP-OLS, Defense Meteorological Satellite Program-Operational Line Scanner; EC, European Commission; MODIS, Moderate Resolution Imaging Spectroradiometer; NASA, National Aeronautics and Space Administration; NOAA, National Oceanographic and Atmospheric Administration; UN, United Nations; NA, not applicable.

Table 1.1 Eight global urban maps. Maps are listed in order of increasing global urban extent followed by the two urban-related maps (bottom two rows). For maps in a geographic projection, the resolution column describes pixel size at the equator.

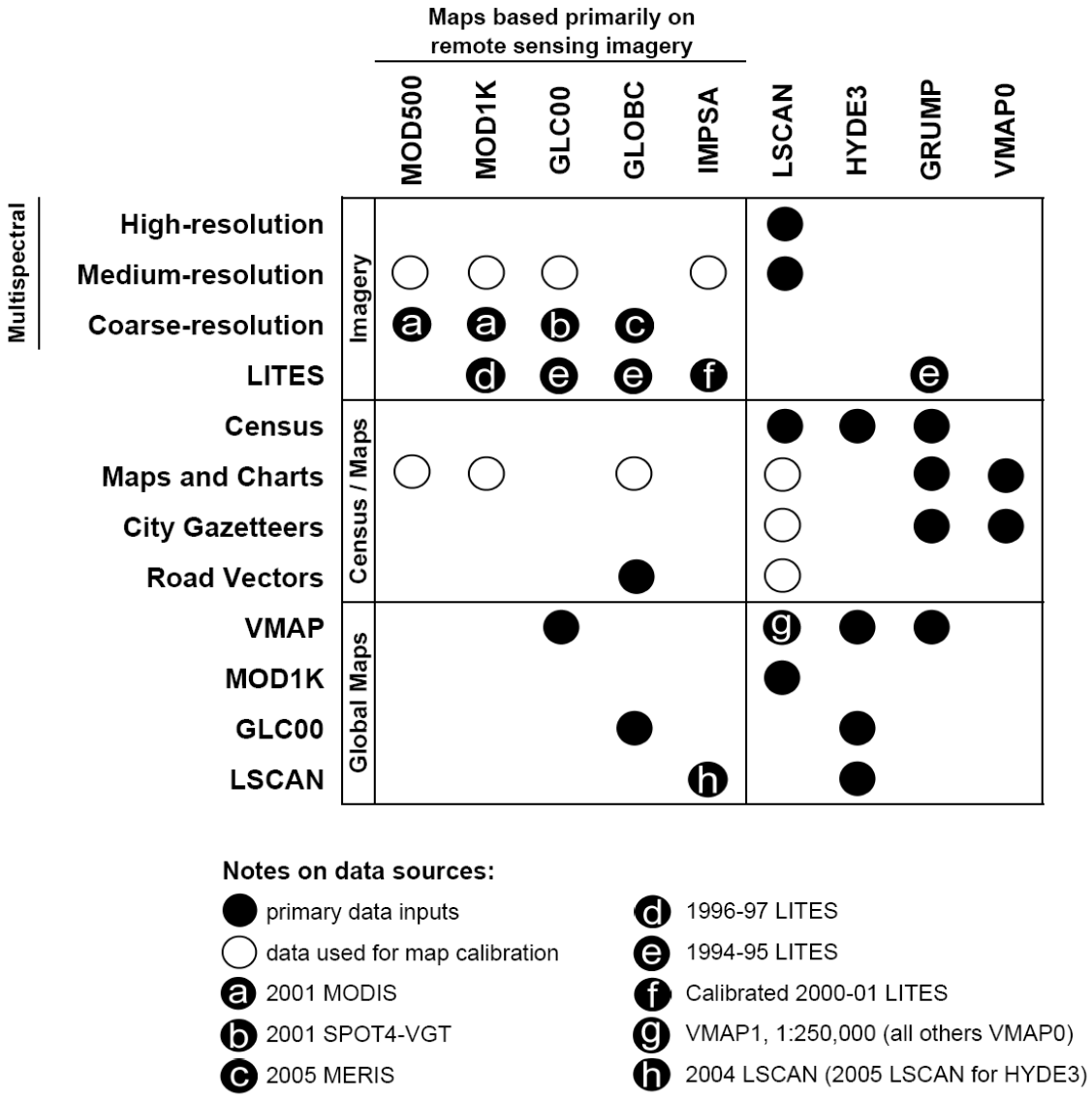


Table 1.2 Major inputs used to produce global urban maps. The five left-most columns are primarily based on expert and semi-automated classification of satellite imagery, and the four right-most columns employ a combination of imagery, census data, and vector maps and charts. The bottom four rows (global maps section) indicate data sharing among global urban maps. Abbreviations correspond to those employed in Table 1.1. In note (e), pending more detailed documentation from the GLOBC mapping group, we assume that GLOBC draws on the same 1994-95 LITES imagery that was used by the GLC00 mapping team.

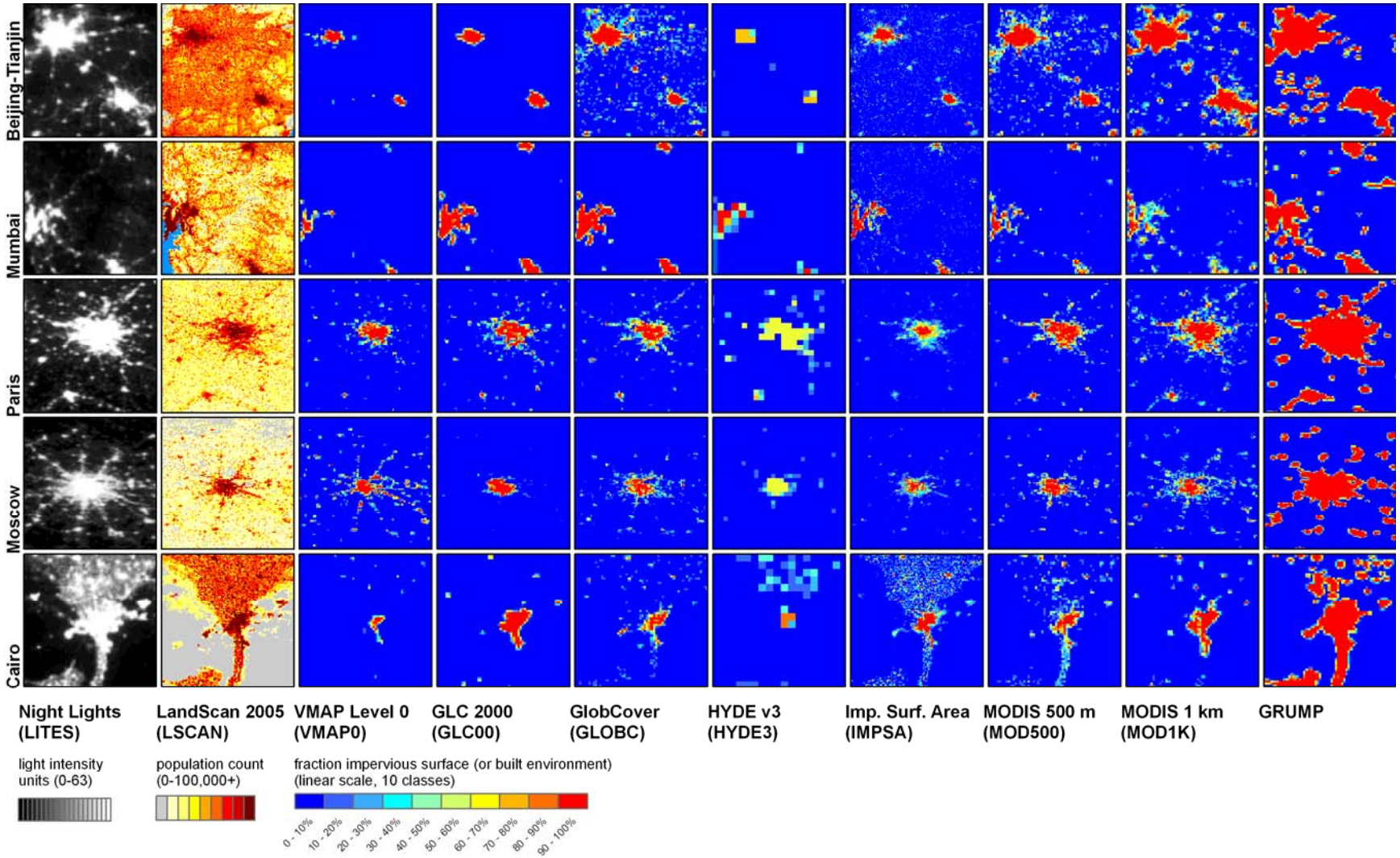

Beijing-Tianjin
Mumbai
Paris
Moscow
Cairo
Night Lights (LITES)
LandScan 2005 (LSCAN)
VMAP Level 0 (VMAP0)
GLC 2000 (GLC00)
GlobCover (GLOBC)
HYDE v3 (HYDE3)
Imp. Surf. Area (IMPSA)
MODIS 500 m (MOD500)
MODIS 1 km (MOD1K)
GRUMP
light intensity units (0-63)
population count (0-100,000+)
fraction impervious surface (or built environment) (linear scale, 10 classes)
0 - 10%
10 - 20%
20 - 30%
30 - 40%
40 - 50%
50 - 60%
60 - 70%
70 - 80%
80 - 90%
90 - 100%

Figure 1.3 (previous page) Global urban portraits of five world cities. The eight global urban maps and two urban-related maps for Beijing-Tianjin, China (top row), Mumbai, India (second row), Paris, France (third row), Moscow, Russia (fourth row), and Cairo, Egypt (bottom row). The first two columns are Nighttime Lights for 2000-01 and LandScan 2005. The eight global urban maps begin with the Vector Map Level Zero (VMAP0), in the third column and proceed from left to right in order of increasing total extent. Each subset is approximately 150 x 150 km and all are oriented north-up. LITES, LSCAN and IMPSA are at full 30" arc-second resolution, HYDE3 is at 5' arc-minutes (native), and all of the remaining maps have been aggregated from 30" arc-seconds to 1.5' arc-minutes for the purpose of display and comparison. This aggregation effectively converts their legends from binary (urban/rural) to continuous (percentage urban) to remain consistent with the two native-continuous maps (IMPSA and HYDE3).

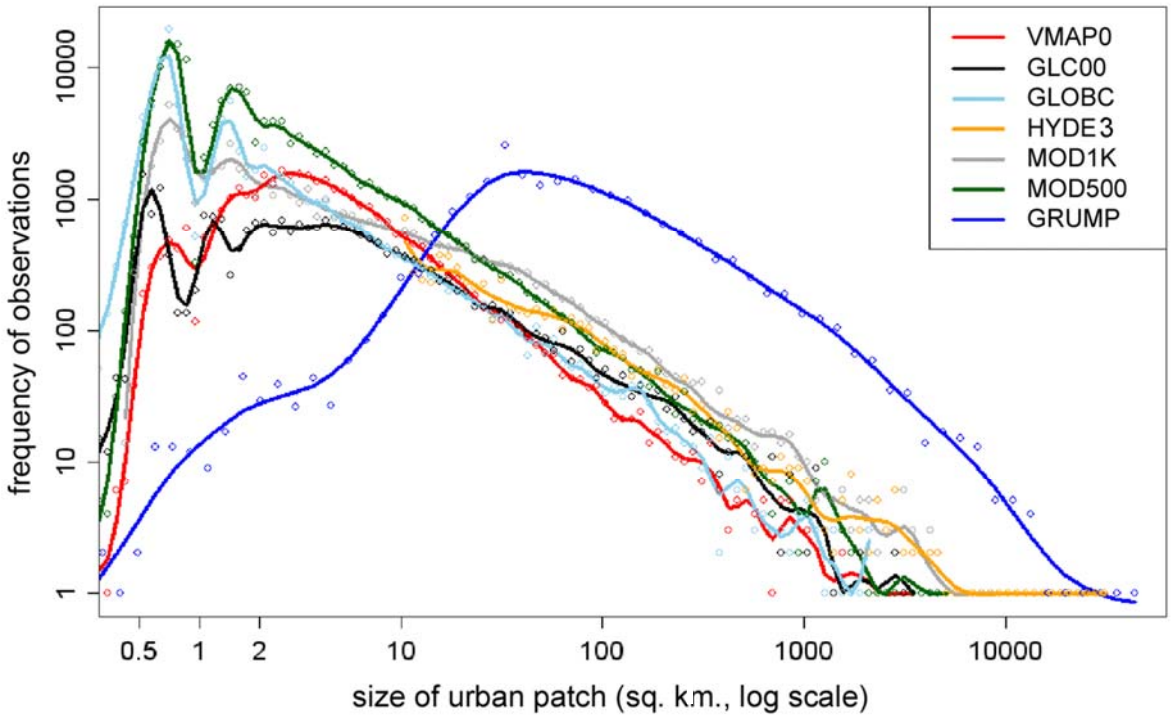


Figure 1.4 Distribution of urban patch sizes. The frequency of urban patch sizes (log-log scale) for each map (excluding IMPSA). Observations are indicated with hollow circles and the solid line is a fitted spline. HYDE3 is plotted starting at 10 sq. km because of the coarse resolution of HYDE3 pixels (5' arc-minutes).

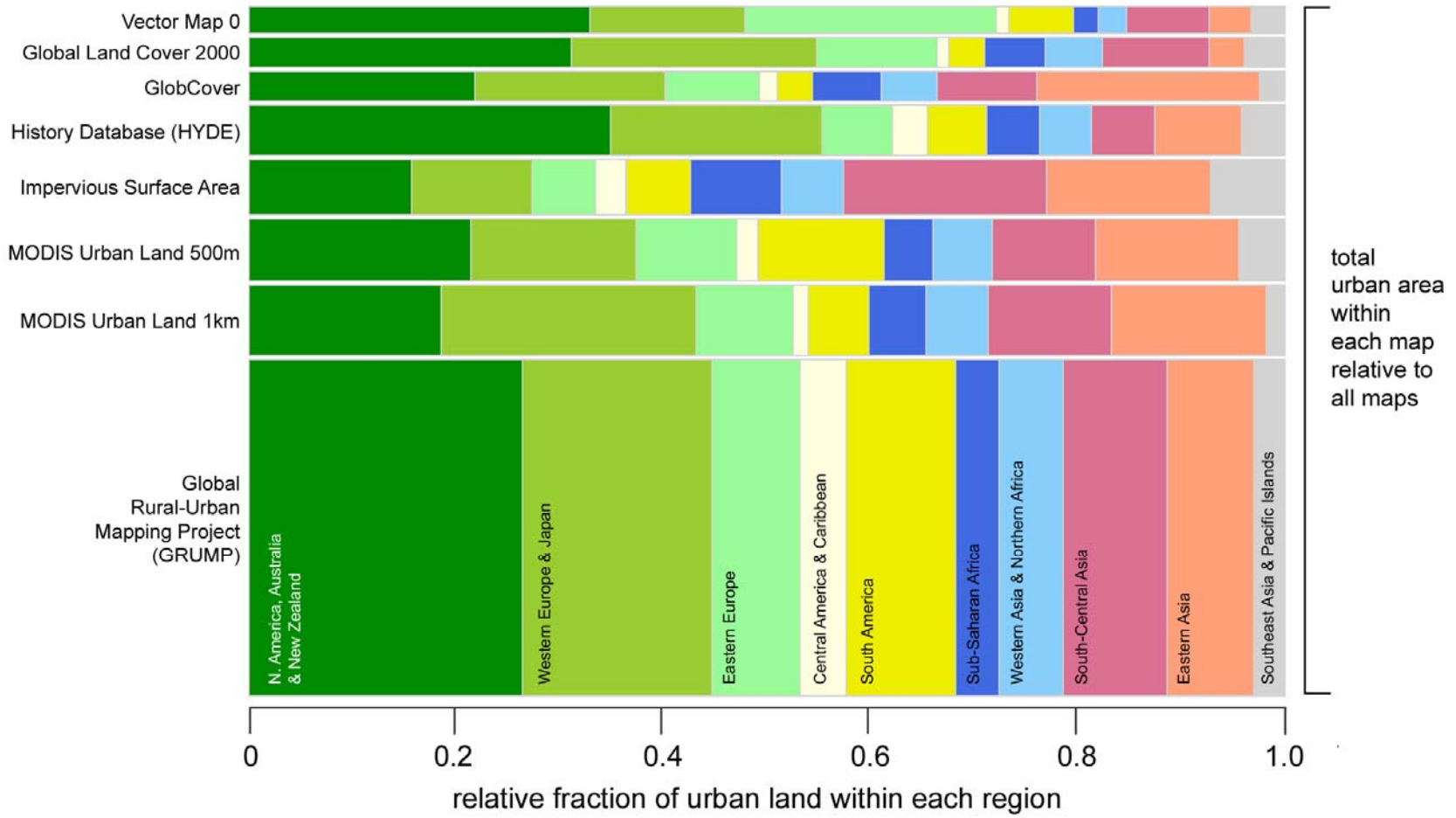


Figure 1.5 Total urban extent by continental region. The distribution of urban land per continental region for eight global urban maps: VMAP level 0, GLC-2000 v1.1, GlobCover, HYDE v3, IMPSA, MODIS 500 m, MODIS 1 km, and GRUMP alpha. The thickness of the horizontal bars is scaled to reflect the relative amount of urban land within each of the eight maps. The horizontal bars are divided into sections representing the relative distribution of urban land area within each of the ten regions (the color scheme is the same as in Figure 1.2a).

Region	VMAP0	GLC00	GLOBC	HYDE3	IMPSA	MOD500	MOD1K	GRUMP
North America, Australia & N. Zealand	***91,123***	***95,982***	***68,409***	***186,303***	***89,963***	***140,506***	***134,764***	***933,537***
USA	79,607	80,482	26,434	161,041	75,488	121,200	117,504	755,881
Canada	4,201	6,054	6,021	12,400	11,300	8,704	7,895	132,472
Australia & New Zealand	7,315	9,446	35,954	12,845	3,161	10,602	9,366	45,027
Western Europe & Japan	***41,327***	***72,761***	***57,602***	***108,435***	***66,469***	***105,449***	***178,887***	***641,608***
Western Europe	15,789	27,372	18,726	56,509	20,855	47,855	53,386	180,709
Northern Europe	16,083	27,413	13,855	18,788	13,157	12,585	21,829	159,799
Southern Europe	4,102	16,133	23,183	19,625	19,250	25,460	50,127	196,262
Japan	5,352	1,842	1,838	13,513	13,207	19,548	53,545	104,839
Eastern Europe	***67,056***	***35,937***	***28,232***	***35,796***	***34,540***	***63,494***	***68,487***	***301,596***
Russia	40,760	15,044	15,374	16,321	20,073	26,504	37,731	188,346
Eastern Europe (excluding Russia)	26,296	20,893	12,858	19,475	14,467	36,990	30,757	113,250
Central America & Caribbean	***3,466***	***3,468***	***5,322***	***17,802***	***17,581***	***13,099***	***10,274***	***154,951***
Central America	2,534	2,309	4,146	12,254	14,245	8,896	8,863	122,462
Caribbean	933	1,159	1,176	5,548	3,335	4,203	1,411	32,490
South America	***17,074***	***10,731***	***10,801***	***30,499***	***35,382***	***82,242***	***42,876***	***374,942***
Brazil	10,113	5,025	5,007	17,021	17,938	39,989	19,254	189,286
South America (excluding Brazil)	6,962	5,706	5,794	13,478	17,445	42,253	23,622	185,655
Sub-Saharan Africa	***6,828***	***17,937***	***20,458***	***27,201***	***49,788***	***31,053***	***39,621***	***144,996***
East Africa	1,585	3,286	4,835	6,181	17,887	5,006	10,136	32,310
West Africa (plus Sudan)	1,483	5,378	6,879	10,448	20,430	12,788	15,468	45,967
Southern Africa	2,459	7,883	7,775	7,717	4,965	8,486	10,482	49,977
Middle Africa	1,301	1,390	970	2,856	6,506	4,772	3,535	16,741
Western Asia & North Africa	***7,433***	***16,905***	***17,285***	***27,114***	***34,492***	***37,782***	***44,039***	***222,113***
West Asia	5,732	6,108	6,684	20,396	21,527	27,347	29,393	145,247
Northern Africa (excluding Sudan)	1,701	10,797	10,601	6,717	12,966	10,435	14,645	76,866
South Central Asia	***22,026***	***31,680***	***29,690***	***32,327***	***112,296***	***64,973***	***86,298***	***350,383***
India	8,160	21,288	20,904	17,020	76,244	31,334	30,857	204,676
South Central Asia (excluding India)	13,866	10,392	8,786	15,307	36,052	33,639	55,440	145,708
East Asia	***11,063***	***10,788***	***67,562***	***44,634***	***90,059***	***90,966***	***109,100***	***297,692***
China	9,579	10,012	65,263	39,547	82,301	81,120	88,977	261,920
East Asia (excluding Japan & China)	1,484	776	2,299	5,087	7,758	9,846	20,123	35,772
Southeast Asia & Pacific Islands	***8,981***	***11,819***	***7,416***	***21,874***	***40,933***	***29,197***	***12,597***	***102,290***
Southeast Asia	8,866	11,714	7,308	21,560	40,252	28,911	12,522	97,440
Melanesia	97	102	104	216	553	285	71	3,454
Micronesia	4	3	4	40	68	0	4	688
Polynesia	13	0	0	58	60	0	0	707
Total (sq. km)	***276,377***	***308,007***	***312,779***	***531,985***	***571,504***	***658,760***	***726,943***	***3,524,109***

Table 1.3 Regional urban extent for the global maps. The areal extent of each of the eight global urban maps (in sq. km) for each of the ten regions (modified from the UN regional designations) and select countries. The North American total includes Greenland, Bermuda, and St. Pierre and Miquelon. The GLOBC column describes GlobCover version 1.

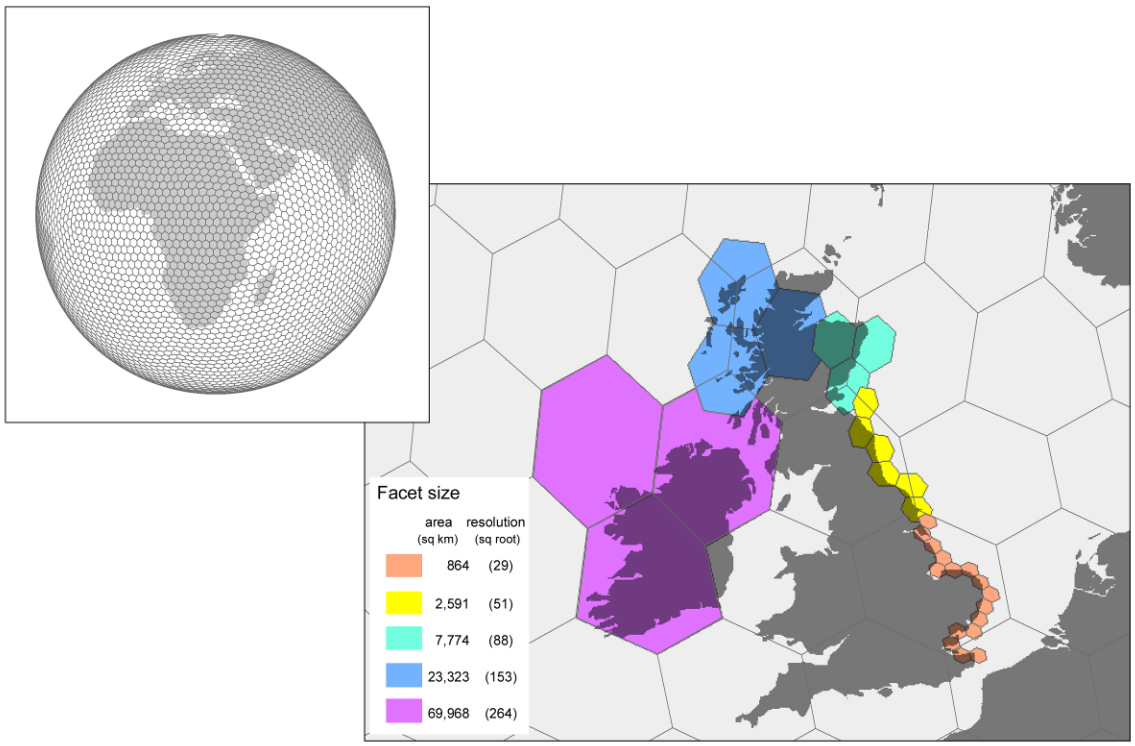
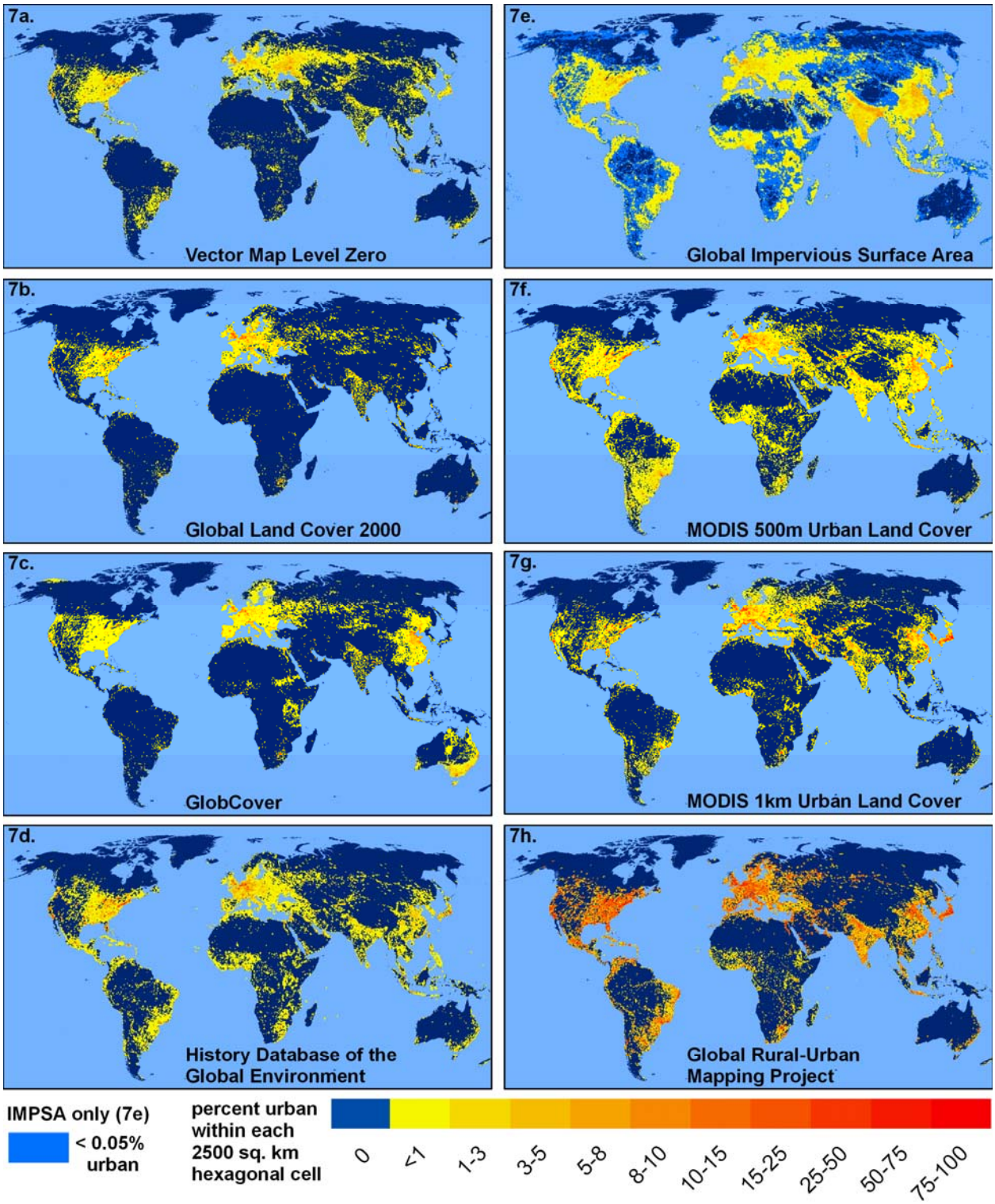


Figure 1.6 Five levels of a discrete global grid. In order to conduct inter-map comparisons across a wide-range of spatial scales, we employed a series of five DGGs with equal-area facets of between 800 and 70,000 sq. km in total area. The DGG displayed globally in the upper left figure is at the coarsest facet size of 70,000 sq. km, which correspond to the purple facets in the map of the United Kingdom. The map legend also shows the effective spatial resolution for each facet size, estimated by the square root of the facet area (from 29 to 264 km).


7a.
Vector Map Level Zero
7e.
Global Impervious Surface Area
7b.
Global Land Cover 2000
7f.
MODIS 500m Urban Land Cover
7c.
GlobCover
7g.
MODIS 1km Urban Land Cover
7d.
History Database of the Global Environment
7h.
Global Rural-Urban Mapping Project
IMPSA only (7e)
< 0.05% urban
percent urban within each 2500 sq. km hexagonal cell
0
<1
1-3
3-5
5-8
8-10
10-15
15-25
25-50
50-75
75-100

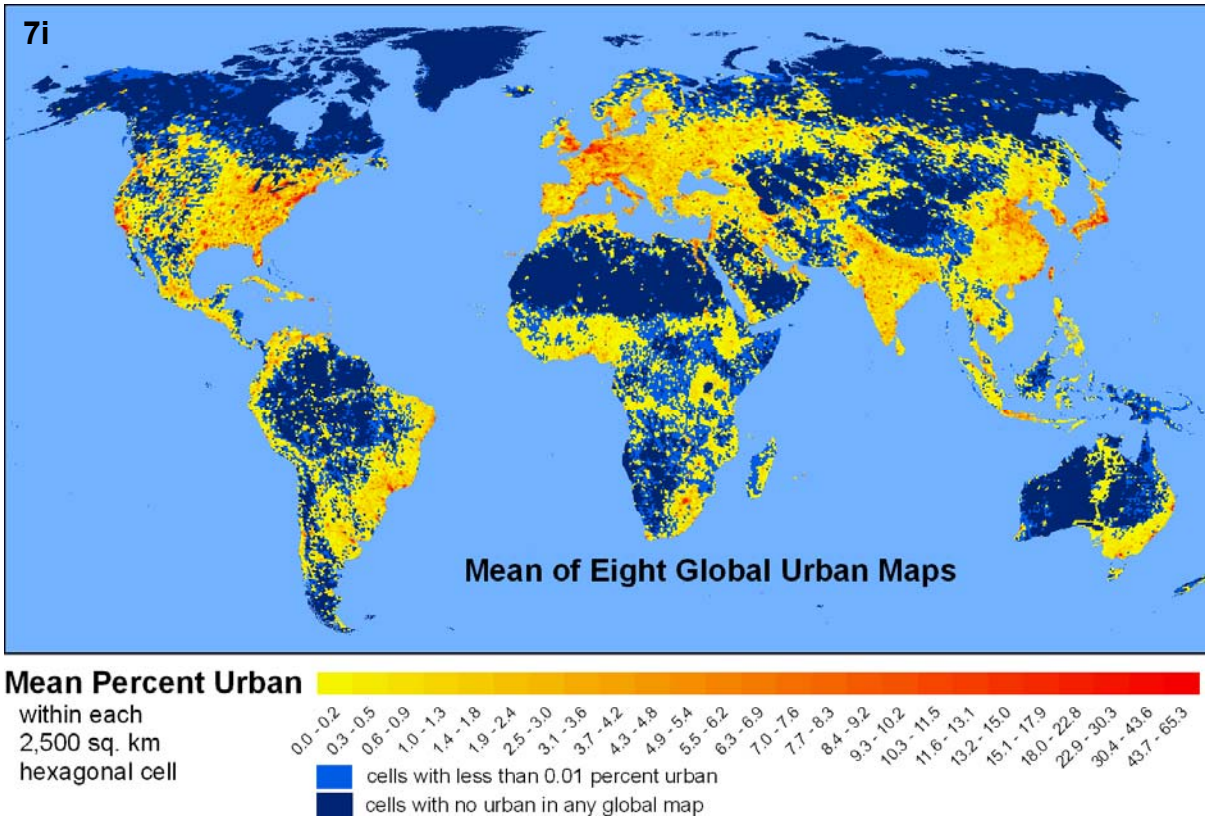


Figure 1.7 (including previous page) Global aggregations of urban area. Figures (a) through (h) depict the percentage of urban land per facet for all eight global urban maps, aggregated to a discrete global grid with hexagonal facets 2,591 sq. km in area (effective resolution of 51 km). The mean map, (i), is the amount of urban area for each grid cell averaged across all eight maps. The total extent of this mean urban map is approximately 864,000 sq. km. For figures (a) through (i), royal blue shades indicate the absence of urban land. The light-blue shading in (e) indicates IMPSA hexagons that contain less than 0.05 percent urban land, and in (i) hexagons that contain less than 0.01 percent urban land.

a. Typical contingency table

Total number of cells: **$n = a + b + c + d$**		Validation Data: Presence (urban)	Validation Data: Absence (non-urban)
Data Under Review (global urban maps)	Presence (urban)	***a***	***b***
	Absence (non-urban)	***c***	***d***

b. Several contingency table-based measures of agreement.

Measure	Formula
Overall Map Accuracy (OMA)	$\frac{a+d}{n}$
Specificity (1–Commission)	$\frac{d}{b+d}$
Sensitivity (1–Omission)	$\frac{a}{a+c}$
Cohen's Kappa (Kappa; Cohen 1960)	$\frac{\left(\frac{a+d}{n}\right) - \frac{(a+b)(a+c)+(c+d)(d+b)}{n^2}}{1 - \frac{(a+b)(a+c)+(c+d)(d+b)}{n^2}}$
True Skill Statistic (TSS; Allouche et al. 2006)	sensitivity + specificity – 1, or $\frac{ad-bc}{(a+c)(b+d)}$
Normalized Mutual Information (NMI; Forbes 1995)	$1 - \frac{-a\,ln(a) - b\,ln(b) - c\,ln(c) - d\,ln(d) + (a+b)\,ln(a+b) + (c+d)\,ln(c+d)}{n\,ln(n) - ((a+c)\,ln(a+c) + (b+d)\,ln(b+d))}$

Table 1.4 Contingency tables and map agreement measures. A typical contingency table for a two-class map comparison (a), and the six map agreement measures employed in this chapter (b). Each measure is based on the elements of the contingency table (a).

	VMAP0	GLC00	GLOBC	HYDE3	IMPSA	MOD500	MOD1K	GRUMP
VMAP0	-	0.41	0.28	0.37	0.26	0.57	0.54	0.76
GLC00	0.37	-	0.47	0.40	0.27	0.55	0.60	0.89
GLOBC	0.25	0.47	-	0.29	0.24	0.49	0.52	0.72
HYDE3	0.19	0.23	0.17	-	0.18	0.34	0.38	0.74
IMPSA	0.12	0.15	0.13	0.17	-	0.25	0.26	0.46
MOD500	0.24	0.26	0.24	0.28	0.22	-	0.51	0.75
MOD1K	0.21	0.26	0.22	0.28	0.20	0.45	-	0.80
GRUMP	0.06	0.08	0.06	0.11	0.07	0.14	0.16	-

Table 1.5a Summary of contingency tables for each map pair. This table lists the degree of urban agreement in the contingency tables for each of the 28 pair-wise map comparisons. Each entry represents the percent of urban area in Map A (shown in the rows) that is also mapped as urban in Map B (shown in the columns).

	VMAP0	GLC00	GLOBC	HYDE3	IMPSA	MOD500	MOD1K	GRUMP
VMAP0	-							
GLC00	0.39	-						
GLOBC	0.26	0.47	-					
HYDE3	0.25	0.29	0.21	-				
IMPSA	0.16	0.19	0.17	0.17	-			
MOD500	0.34	0.35	0.32	0.30	0.23	-		
MOD1K	0.30	0.36	0.31	0.31	0.22	0.48	-	
GRUMP	0.11	0.14	0.11	0.19	0.12	0.23	0.27	-

Table 1.5b Summary of Cohen's Kappa values for each map pair. For each combination of the eight maps, this table lists Cohen's Kappa (κ) statistic, a measure of overall map agreement designed to adjust for the amount of agreement that one could expect due to chance alone. Kappa values of 1.0 indicate perfect agreement.

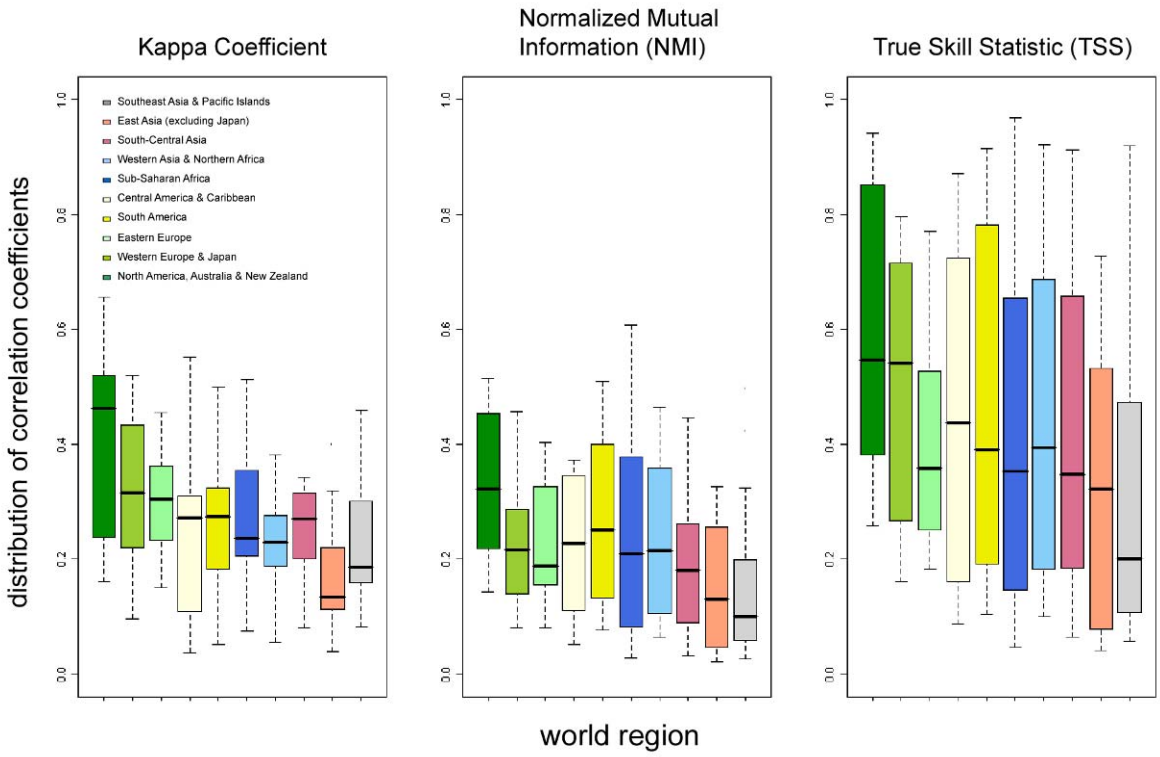


Figure 1.8 Regional box-plots of agreement metrics. Agreement metrics based on contingency tables for all 28 inter-map comparisons, grouped by world region (the color scheme matches that of Figure 1.2a). The plots depict (from left to right) Cohen's Kappa, Normalized Mutual Information (NMI), and True Skill Statistic (TSS). All of these measures approach one when the maps are in complete agreement.

	China (2)	India (4)	USA (1)	Brazil (9)	Russia (10)	Indonesia	Japan (3)	Mexico	Germany (5)	UK (7)	Pakistan	France (6)	Iran	Italy (8)	Ukraine	Argentina	Spain	Canada	South Africa	Australia
VMAP0	6	8	1	5	2		10		7	4					3					9
GLC00	6	2	1		4				5	3		10		8					9	7
GLOBC	1	4	3		5				8	7				6			10		9	2
HYDE3	2	6	1	5	7		8		3			4						9		10
IMPSA	1	2	3	5	4	6	7	9			10							8		
MOD500	2	4	1	3	5		7		6			10		9		8				
MOD1K	2	5	1	10	4		3		9			8	7	6						
GRUMP	2	3	1	4	5		7	8				9		10				6		

Table 1.6a Country-level ranks of urban extent. Ranks for the ten countries with the most urban area within each of the eight global urban maps. The countries are ordered left to right according to their rank in the UN's estimate of the total urban population for the year 2000 (UN Statistics Division 2008). In parenthesis is the country-rank for the top ten highest per capita gross domestic products (purchasing power parity) in the year 2000 (World Bank 2008).

	VMAP0	GLC00	GLOBC	HYDE3	IMPSA	MOD500	MOD1K	GRUMP
VMAP0	-							
GLC00	0.69	-						
GLOBC	0.64	0.86	-					
HYDE3	0.74	0.69	0.65	-				
IMPSA	0.69	0.65	0.65	0.78	-			
MOD500	0.79	0.69	0.67	0.78	0.75	-		
MOD1K	0.75	0.72	0.70	0.78	0.76	0.81	-	
GRUMP	0.76	0.71	0.67	0.81	0.77	0.79	0.79	-

Table 1.6b Country-level urban extent rank correlations. The rank correlation matrix (Kendall's Tau, τ) for a sample of 223 countries.

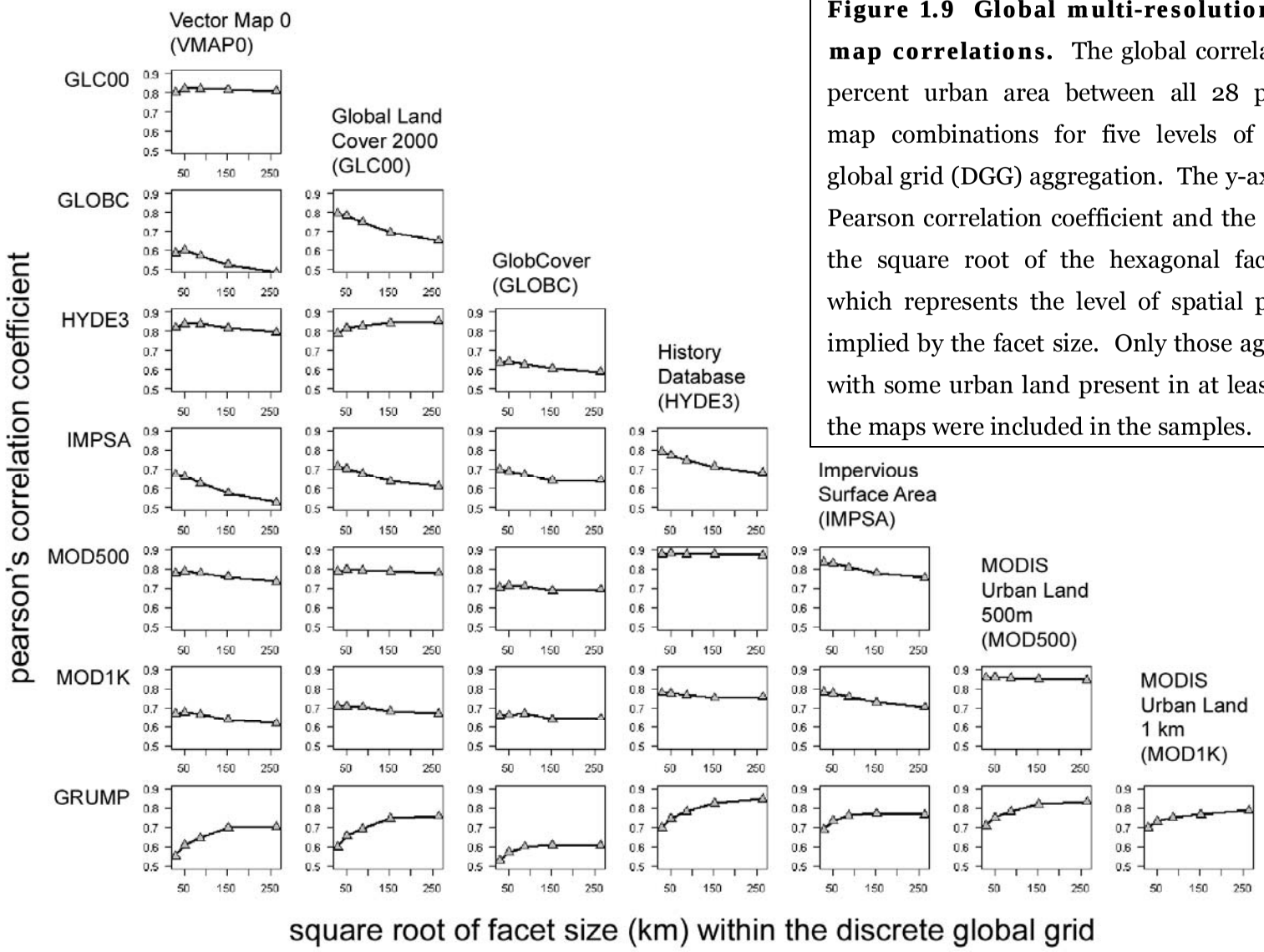


Figure 1.9 Global multi-resolution inter-map correlations. The global correlations in percent urban area between all 28 pair-wise map combinations for five levels of discrete global grid (DGG) aggregation. The y-axis is the Pearson correlation coefficient and the x-axis is the square root of the hexagonal facet area, which represents the level of spatial precision implied by the facet size. Only those aggregates with some urban land present in at least one of the maps were included in the samples.

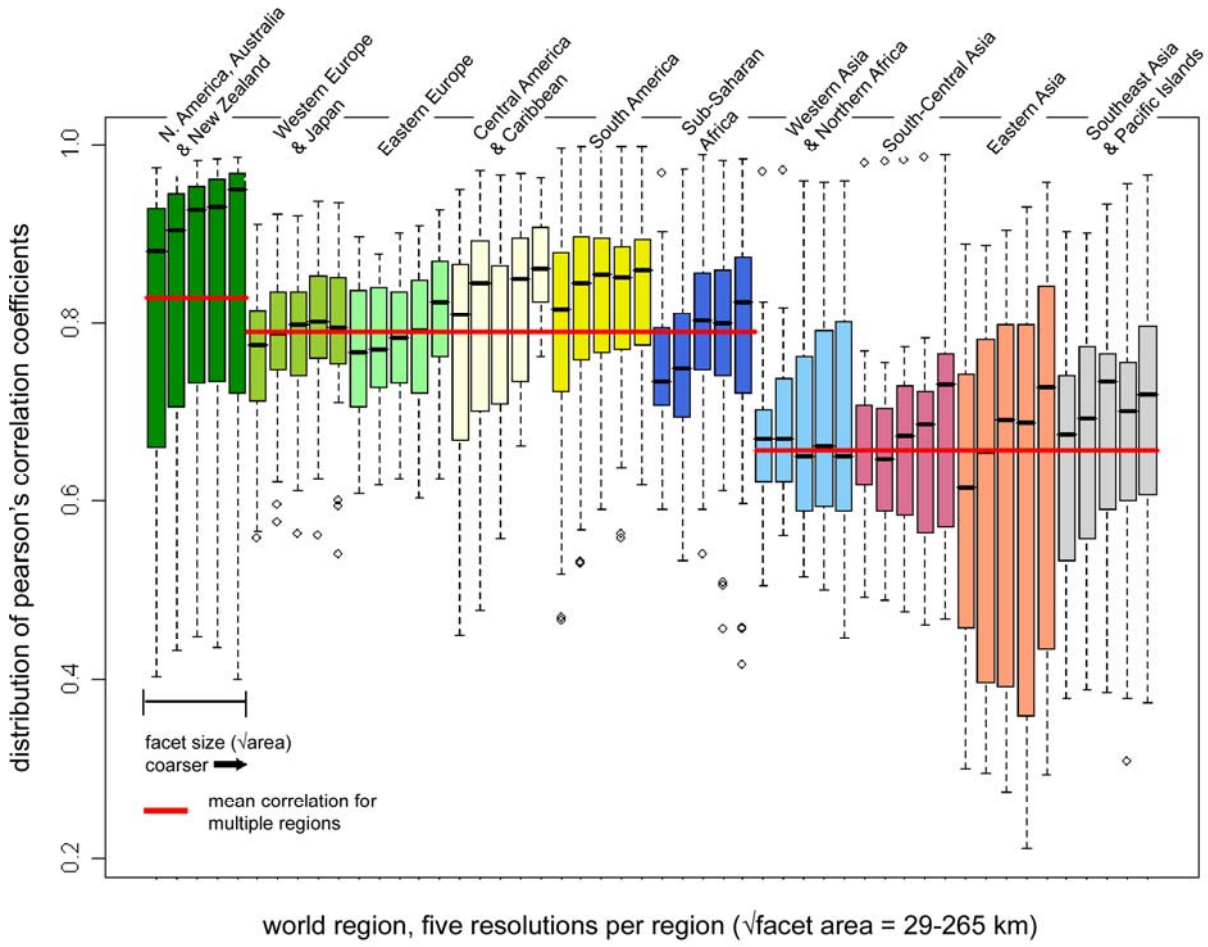


Figure 1.10 Regional inter-map correlations. These box-plots capture the 28 pair-wise Pearson correlations between global maps at five resolutions for each of ten world regions. For instance, the five left-most box-plots (dark green) represent the inter-map correlations for the North America, Australia, and New Zealand region at five discrete global grid (DGG) resolutions (finest to coarsest, 29–264 km, left to right). For this region, the median inter-map correlation (black horizontal bars) improves as the hexagonal grid resolution becomes coarser. Each of the red horizontal bars represents a mean correlation for a group of regions: the top bar ($\overline{r}$ = 0.83) is for North America, Australia, and New Zealand, the middle bar ($\overline{r}$ = 0.79) is for Europe and Japan, South and Central America, and Sub-Saharan Africa, and the bottom bar ($\overline{r}$ = 0.66) is for Asian regions.

CHAPTER TWO

AN EMPIRICAL ASSESSMENT OF EIGHT GLOBAL URBAN MAPS

2.1 INTRODUCTION

A. Overview

We live in an increasingly urbanized world, with more than half of the world's population living in cities and nearly two billion new urbanites projected to arrive in the next twenty-five years (UN 2008). The residences and related infrastructure necessary to sustain more than three billion urban dwellers already require an estimated one to three percent of the Earth's land surface (Potere and Schneider 2007). Urbanization is an important component of regional and global environmental change (Foley et al. 2005; Turner et al. 2007; Ellis and Ramankutty 2008), and has significant implications for both environmental systems (Grimm et al. 2008) and human health and well-being (Dye 2008; Montgomery 2008). Global to regional scale analyses of the connections between urbanized areas and natural and anthropogenic processes have only just begun, and much work remains as we work toward a more sustainable urban future. To this end, high quality, regularly updated information on the patterns and processes within the urban environment — including maps that monitor location and extent — is essential (Potere and Schneider 2007).

In response to the need for global maps of land cover and urban areas, international groups from academia, government, and industry have developed eight global-scale maps that can be used to describe urban areas (Table 1.1; map acronyms will be used from here

forward). Seven of these maps have been released since 2000 as a result of advances in satellite remote sensing, geographic information systems, and semi-automated mapping techniques. Although these eight maps often rely on overlapping input data sets and methodologies (Table 1.2), there are order of magnitude differences in the global extent of urban areas portrayed by these maps: from 0.3 million sq. km for VMAP0 to 3.5 million sq. km for GRUMP (Table 1.1). Because these differences persist across geographic scales and world regions, two important questions arise regarding map accuracy: (1) which maps correspond most closely to built-up patterns observed on the ground?, and (2) how does each mapping group define the 'urbanized' portion of the human-modified landscape? (Potere and Schneider 2007; forthcoming). In the face of such extreme map variability, an assessment of the accuracy of these urban maps is long overdue.

In this chapter, our goal is to establish the accuracy of the coarse-resolution global urban maps by comparing them to a new, random-stratified global assessment data set derived from two collections of medium-resolution (28.5 m) Landsat-based city maps (Angel et al. 2005; Schneider and Woodcock 2008) and a new high-resolution, web-based Google Earth (GE) imagery tool (**Appendix A**). We conduct a two-tiered empirical assessment, first establishing the accuracy of the Landsat-based city maps using GE imagery (tier one), and then using the city-scale Landsat-based maps to quantify the accuracy of the eight global urban maps under review (tier two). We identify MOD500, a global map based on 463 m-resolution imagery from the MODIS sensors, as the map with the fewest omitted cities, the highest city-size correlation, and the highest overall map agreement measures. For several of these measures, MOD500 is followed closely by IMPSA, a global map of impervious surface area based on nighttime lights imagery and human population density from LandScan. In the conclusion, we offer several potential user communities our best-use recommendations for each global map.

B. Remote sensing of urban areas

When present, elements of the human built environment leave at least some physical trace in virtually all terrestrial Earth observation data. Table 2.1 outlines some of the strengths and weaknesses of satellite imagery that can be used to monitor the urban environment at high-, medium-, and coarse-spatial resolutions. From Table 2.1, the central trade-off when selecting satellite imagery is apparent: the higher spatial resolution imagery that is essential for detailed analysis of urban features (top row) is gathered infrequently, contains less spectral information and is often prohibitively expensive relative to coarse- or moderate-resolution imagery. Accordingly, when regional to global data coverage is desired, the only viable options are medium- to coarse-resolution data. To varying degrees, all global mapping efforts that rely on remotely sensed data employ coarse-resolution imagery (Table 1.2), in part because of the launch of a series of second-generation coarse-resolution Earth observation satellites (MODIS, MERIS, SPOT4-VGT; Table 1.2) with improved radiometric quality, higher temporal resolution, and higher geodetic accuracy than earlier sensors (e.g. AVHRR). Although the pixel size of the data from these sensors is relatively coarse (250-1,000 m), the high temporal frequency (daily to weekly) and geospatial and spectral precision of these data can be exploited to discriminate urban features from surrounding land cover types.

Despite recent advances in the acquisition and processing of satellite data at regional to global scales, the accuracy of maps created from these data is constrained by the problem of sub-pixel mixing (Small 2005) and the uncertainty inherent in semi-automated classification. The characteristic scale of features from the built environment is roughly 20 m (Jensen and Cowen 1999; Small 2003), and thus, even the newest coarse-resolution sensors (250-300 m resolution) create very large fractions of mixed pixels. This sub-pixel mixing, in turn, leads to a high degree of local map variability, and points to the need for careful, systematic accuracy assessment of any coarse-resolution map.

The accuracy limitations of coarse-resolution maps are somewhat mitigated in medium-resolution sensors (e.g. Landsat) because the smaller pixels (28.5 m) much more closely approximate the 10-20 m characteristic scale of urban features (Small 2003). Compared with coarse-resolution urban mapping, there is a much longer history of using medium-resolution imagery for building maps of cities (Jensen and Toll 1982; Treitz et al. 1992; Longley and Mesev 2000). At spatial resolutions of less than 10 m, the problem of mixed pixels is less pronounced and many individual elements of the built environment are immediately evident without semi-automated classification tools. Large-area mapping from high-resolution imagery is not yet possible because of the high cost of both imagery acquisition and interpretation. Nevertheless, because these maps have high accuracy and precision, they are often ideal reference data for medium- to coarse-resolution map assessment.

C. Previous assessments of global urban maps

This work and a related analysis associated with the release of the MOD500 map (Schneider et al. forthcoming) are the first global-scale investigations of map accuracy aimed specifically at urban areas. This research builds on our initial comparative assessment that identified a low degree of inter-map agreement for six global urban maps and two urban-related maps (Potere and Schneider 2007, forthcoming; **Chapter 1**). Of the research teams identified in Table 1.1, only two have conducted global-scale validation campaigns: MODIS and GLC 2000. The MODIS land cover team (2003) produced a validation of the 927 m MODIS Collection 3 product, but this land cover map pre-dated the Schneider et al. (2003) urban layer (introduced in the Collection 4 map) and instead relied on VMAP0 to depict urban areas. Mayaux et al. (2006) constructed a global, stratified random sample of test sites using manual interpretation of medium-resolution (28.5 m) Landsat imagery to asses GLC00, yet this effort did not produce enough urban sites to generate accuracy results for

that class. Similar sample-size limitations prevented Herold et al. (2008) from drawing urban-specific conclusions in their assessment of four global land cover maps. There are several projects that consider inter-map differences across global maps, but these either do not address map accuracy (Potere and Schneider 2007) or are focused on non-urban maps and land cover classes (Hansen et al. 2000; Giri et al. 2005; Jung et al. 2006; McCallum et al. 2006; See and Fritz 2006). In addition, several region-specific investigations of urban maps have been performed (Schneider 2003; Tatem et al. 2005), but these are of limited utility in the context of a global-scale assessment.

2.2 DEFINING AN 'URBAN' LAND-USE CATEGORY

In any assessment of map accuracy it is important to begin with a clear definition of the map attributes under review. Because our focus is on the map accuracy of urban areas, our assessment includes only two classes: urban areas and non-urban areas. For the purposes of this assessment, we define 'urban' based on physical attributes: urban areas are places that are dominated by the built environment, where 'built environment' encompasses all non-vegetative human-constructed elements, including roads, buildings, runways, industrial facilities, etc., and 'dominated' implies coverage greater than or equal to fifty-percent of a given landscape unit (in this case, a single pixel). Within the remote sensing literature the term 'impervious surface' is often used interchangeably with 'built environment' (Ridd 1995), but we prefer the more direct term 'built environment' because of ambiguity and uncertainty regarding the impervious surface concept (Small 2002).

In addition to the physical attributes associated with urban land-use, there are a wide range of social, economic, and demographic indicators that are often used to delineate urban areas. Although these definitions are often useful for country-level studies, they carry two principal limitations when extended to the problem of global urban mapping: (1) there is no consensus on how to construct a consistent definition of urban areas based on any single set

of social, economic or demographic attributes (Montgomery et al. 2003, Keiser et al. 2004), and (2) the quality and quantity of spatially-explicit demographic and economic data vary enormously both across countries and within national borders. For many developing countries, the necessary spatially-explicit socio-economic data are virtually non-existent (Van Oosterom 2005). In order to create a consistent global map of urban areas, a physically-based approach is currently the best alternative. Moreover, the synoptic perspective of available remote sensing data (as discussed above) make this type of mapping possible.

To illustrate the concept of a 'physical' definition of urban land, Figure 2.1 presents a simple Venn diagram of a land-use classification system composed of three primary classes: the built environment (gray), the cultivated environment (yellow), and the natural environment (green). These land-use categories correspond loosely to those employed by Foley et al. (2005) in their model of land-use transitions. All landscapes can be described by their membership in one or more of these classes. The red oval in Figure 2.1 represents urban areas, a subset of the built environment which includes: (a) areas of contiguous built land, (b and d) areas of interspersed built and cultivated or natural land, and (c) areas that contain a mosaic of built, cultivated, and natural land. This classification system is sensitive to the effects of scale: a reduction in the size of the minimum mapping unit (MMU) generally reduces the proportion of landscape elements (e.g. pixels) that are classified as mixtures of the three land-use categories (Ozdogan and Woodcock 2006).

Figure 2.1 is also a useful tool for describing the differences between a hard-classified or thematic map (where only one land-use label is allowed for each pixel) and a soft-classified or continuous-value map (where each pixel is assigned fractional membership in a land-use category). For the six global urban maps that represent urban areas as a binary urban / non-urban thematic class, our 'urban' definition implies that urban pixels are those where the built environment spans the majority (50 percent or greater) of the sub-pixel

space. Although this assumption is implicit, these thematic maps do not contain any information regarding the proportion of urban coverage within each pixel. Provided a given map has not missed any urban land, these binary pixels can be regarded as an upper-bound on the urban built environment (red oval, Figure 2.1). In contrast, soft-classified maps attempt to characterize the entire built environment, but can only represent urban areas when thresholds are applied (e.g. pixels with a majority of built environment are labeled 'urban'). In **Sections 2.4.C and 2.4.D** we discuss the importance of exploring a range of thresholds for assessing the accuracy of IMPSA and HYDE3, the two continuous-value maps in this study.

Our physical definition of urban areas as landscapes dominated by the built environment aligns well with the eight global maps under review. In Table 1.2, they are arranged from left to right in roughly decreasing order of their dependence on remotely sensed imagery, although to varying degrees all of these maps employ a classification strategy that relies in part on daytime or nighttime satellite data. MOD500 and MOD1K are the only maps that are completely independent of ancillary data (i.e. sources other than image data) for their primary classification inputs (Table 1.2). GLOBC and GLC00 also rely primarily on daytime imagery; data derived from the MERIS and SPOT4-VGT satellites, respectively. The remaining four maps draw directly on imagery through nighttime lights data (IMPSA and GRUMP), or indirectly through the use of other image-based global maps (HYDE3). The exception is VMAP0, which was produced primarily through the analysis of navigational charts and maps. The methodologies of each map are discussed in more detail in the various map-specific source documents (column 1, Table 1.1), Potere and Schneider (2007, forthcoming), and **Section 1.1.B**.

2.3 DATA SETS

In addition to the eight global coarse-resolution urban maps under review, we employ five distinct data sets for this assessment (Table 2.2); examples of each of these are shown for Moscow, Russia (Figure 2.2). These data sets include: the *Angel-gazetteer*, a global index of 3,949 cities from Angel et al. (2005); *GE-validation sites*, a set of more than 10,000 high-resolution Google Earth sites rated by expert analysts (green hexagons in Figure 2.2); *GE-boundaries*, a group of 107 outer city boundaries constructed using Google Earth; *UCONN-maps*, a collection of 120 city maps constructed using Landsat imagery as part of a joint University of Connecticut and World Bank study (Angel et al. 2005; red pixels in Figure 2.2), and *UWISC-maps*, a similar collection of 24 maps assembled in a University of Wisconsin project (Schneider and Woodcock 2008; also red pixels in Figure 2.2). Each of these data sets is discussed in detail in the following sections.

A. City gazetteers and the global sampling scheme

Because of the small extent of urban areas relative to other land uses, it is difficult to conduct an efficient global assessment based on a completely random sample. Instead, it is necessary to construct a sample based on areas where urban landscapes are known to exist. Such a sample can be drawn from a global gazetteer—a point dataset of known city locations. There are at least three circa-2000 global gazetteers: (1) the United Nations (UN) Population Division's global gazetteer of cities with populations greater than 100,000 (UN 2008c), (2) the Columbia University Center for International Earth Science Information Network's human settlement point database (CIESIN 2004), and (3) the Angel et al. (2005) "universe of cities" with populations greater than 100,000. We avoid the UN gazetteer because it is of low reliability for cities of less than 750,000 inhabitants (pers. comm. UN Population Division). Both the CIESIN and Angel et al. gazetteers have city size distributions that are in

agreement, but we chose the Angel-gazetteer to maintain continuity with the 120-city *UCONN-maps,* which were selected using the Angel-gazetteer as the sampling framework.

Because the Angel-gazetteer includes only cities with populations greater than 100,000, our assessment can make no claims about the accuracy of the global urban maps for cities smaller than this size. To understand how many people may reside in these small cities, towns, and villages, we can compare the circa-2000 UN estimate for urban population (2.86 billion; UN 2005) with the total population in the Angel-gazetteer (2.12 billion). In total, 25.9 percent of the UN's total urban population is unaccounted for in the Angel-gazetteer (similar omissions are present in the CIESIN and UN gazetteers). This missing population may dwell in cities with populations larger than the gazetteers record, in cities that are missing from the gazetteers, or in cities that are smaller than 100,000 inhabitants.

Figure 2.3 depicts the distribution of cities by city population size versus country income level (a, b) and world region versus country income level (c, d). As is clear from Figure (a), 79.4 percent of all cities have populations smaller than 500,000, and this size distribution is relatively stable across country income divisions. Rather than create a sample that maintains this small-city focus, we opted to follow the methods of Angel at al. (2005), and evenly sample each income/size division (b). For each of the sixteen combinations of city size and country income, we include 16 cities except for the large city division (populations greater than 4.2 million), which we sample fully. When we estimate sample means in **Section 2.4**, we weight our city estimates so that the global means reflect the distribution of cities in Figure 2.3a. From (c), it is apparent that most of the cities within moderate to low-income countries (dark bars) are located in Africa and Asia, while most of the cities in wealthy countries are in Europe and Other Developed Countries (ODC). Comparing (c) to (d) reveals that our expanded 247-city UCONN/UWISC/GE-boundary sample follows a similar pattern, and is thus relatively unbiased by region and income. The map of assessment cities (Figure 2.4) confirms the global distribution of our sample.

B. High-resolution Google Earth assessment data

In 2006, Google Earth reported that their high-resolution image archive spanned twenty percent of the Earth's landmass and more than a third of the human population (Google 2006). This contemporary high-resolution archive represents a significant, rapidly expanding, cost-free and largely unexploited resource for mapping urban areas (**Appendix A**). The QuickBird satellite, a high-resolution commercial imaging satellite launched in 2001 (0.6 m pan-chromatic, 2.4 m multispectral data), is the primary source for high-resolution GE imagery. The GE images have been degraded from their native format: the infrared band has been removed, they have undergone undocumented and non-uniform spectral alterations, and they are at a coarser spatial resolution (approximately 3-5 m). Nevertheless, an expert analyst (and often a lay person) can readily discern individual landscape elements and the presence of the built environment (Figure 2.2d). Crucially, the positional accuracy of the GE images is adequate for assessing coarse-resolution maps. In an assessment of 436 control points in the vicinity of Angel et al. (2005) cities, on average the GE points varied by only 29.4 m from the corresponding Landsat GeoCover control points (standard deviation of 26.6 m; **Appendix A**). This positional error is significantly less than the spatial resolution of the global urban maps, making it accurate enough for use in our assessment. In **Section 2.4.A**, we use a GE tool to gather 10,000 *GE-validation sites* for analysis of the *UWISC / UCONN-maps* (tier one).

GE imagery also serves as the basis for constructing our *GE-boundaries* dataset. These boundaries serve as one component in a simple analysis of the presence/absence of urban areas within each global urban map (**Section 2.4.B**). To determine which cities to include in the *GE-boundaries* dataset, we constructed a random-stratified sample from the *Angel-gazetteer* based on the divisions from **Section 2.3.A**. To delineate urban boundaries in GE, we used GE polygon drawing tools to outline the outer perimeter of each city and converted these 107 polygons into a layer within our geographic information system (GIS)

database. These 107 GE boundaries were relatively easy to assemble because we did not actually map the built environment in detail; rather, we traced the approximate outer edge of each city. In areas of continuous conurbation like Tokyo and Los Angeles, we simply created polygons that enclosed large portions of the cities (larger than 25 sq. km).

C. Medium-resolution Landsat-based urban maps

Both the *UCONN-maps* and *UWISC-maps* are medium-resolution (28.5 m) city maps created with circa-2000 six-band Landsat imagery through a combination of classification, spatial filtering, and hand editing (Figure 2.2). The mean acquisition year of the Landsat images is 2001, with a range of 1999-2003. Angel et al. (2005) conducted an accuracy assessment of 10 percent of their 120-city sample and reported a mean overall map accuracy of 89.2 percent. Schneider and Woodcock (2008) assessed the accuracy of 10 cities (roughly half of their sample), and reported a range of overall map accuracies from 84 to 97 percent. To further strengthen confidence in these maps as a source of reference data for this assessment, we conduct a comprehensive tier-one accuracy analysis using GE validation sites in **Section 2.4.A.**

Because the UCONN and UWISC maps were created by separate teams of analysts, it is important to compare the four cities that are present in both map collections: Alexandria, Egypt; Madrid, Spain; Warsaw, Poland; and Guadalajara, Mexico. A per-pixel map comparison between the *UCONN-maps* and *UWISC-maps* conducted at the full 28.5 m-resolution revealed an overall agreement across the four map-pairs ranging from 91.7 (Alexandria, Egypt) to 96.1 percent (Guadalajara, Mexico). Given that Guadalajara was the only city-pair with the same Landsat acquisition dates, it makes sense that this city has the highest overall accuracy (the Landsat images underlying the other three city pairs are separated by 1.5-2.5 years). Overall, the four independently mapped cities agree well, providing an indication of the robustness of both groups' urban mapping strategies.

D. Coarse-resolution global urban maps

Although we have already discussed the basic characteristics and methodologies of the eight global urban maps under review (**Section 2.2**; Table 1.1; Table 1.2), it is important to note that the majority of these maps have been modified somewhat from their native data formats. To facilitate a consistent assessment methodology, we selected a common geographic projection, WGS-84 datum, and either 30", 15", or 10" arc-second cells (the 15" and 10" arc-second cells are for MOD500 and GLOBC, respectively). VMAP0, GLC00, HYDE3 had to be altered as described in Potere and Schneider (2007, forthcoming) and **Section 1.2.A**; these modified maps were compared with their native counterparts at various stages in the transformation process and differences were negligible. The common geospatial environment allowed us to perform: (a) rigorous geospatial alignment checks, (b) standardized assessment methods, and (c) consistent division along international borders using methods from Potere and Schneider (2007, forthcoming) and **Section 1.2.A**. Since **Chapter 1** was completed, the GLOBC group has released a new version (v2) of their global map. The urban areas in GLOBC v2 are 7.4 percent larger than those of the original version and include fewer 'fill' values. For this chapter, we employ the new version of GLOBC (v2).

Aside from the eight maps in this study, there are two additional global maps from Table 1.1 that describe features closely related to urban areas: LSCAN and LITES. In Potere and Schneider (2007), we conducted a basic threshold analysis of both LSCAN and LITES. Although applying thresholds to these continuous maps of human population density and nighttime illumination intensity (respectively) can yield estimates that correlate with urban areas, we have chosen not to include LITES or LSCAN in this study. The reasons for this are twofold: (1) they are not designed for urban mapping, and (2) they are both inputs to IMPSA, a map that is part of this study.

2.4 METHODS AND RESULTS

We begin by assessing the 140 medium-resolution Landsat-based city maps with more than 10,000 high-resolution Google Earth validation sites (tier one). We then use these medium-resolution maps to assess the eight global urban maps (tier two). In this second stage, we test for the presence / absence of cities (hereafter *urban omission rates*), compare estimates of city size, and compute a suite of overall map agreement measures.

A. Tier one accuracy assessment

To augment the assessments of Angel et al. (2005) and Schneider and Woodcock (2008), we conducted an independent validation of the UCONN and UWISC maps using Google Earth (GE) imagery. The horizontal positional accuracy of GE imagery (29.4 ± 26.6 m; **Appendix A**) is not sufficient for a pixel-based assessment, therefore we opted for a site-based analysis. To define the sites, we overlaid an equal-area hexagonal discrete global grid (DGG) with a facet size of 0.132 sq. km (ISEA3H-18; Sahr 2003) on the Earth's surface, and selected those hexagons that fell within the boundaries of the UWISC / UCONN maps. After eliminating hexagons that contained more than 30 percent water, we selected a stratified-random sample of 100 assessment sites for each city (green hexagons in Figures 2.2a-d), where the stratification ensured that our sample contained a roughly equal portion of sites with low, medium, and high fractional coverage of urban pixels. This process produced 9,423 cloud-free hexagonal assessment sites distributed across 102 cities in the UWISC / UCONN map collection.

To facilitate rapid assessment of such a large number of sites we designed a web-based Land Cover Validation Tool (LCVT; **Appendix A**). This tool provided a way for individual analysts to quickly label each site on a zero to five ordinal scale: (0) cloud cover or a lack of high-resolution coverage, (1) no appreciable built environment, (2) majority non-built environment, (3) an even mix of built environment and natural or cultivated

environment, (4) majority built-environment, and (5) virtually all built-environment. For 800 of the GE sites, multiple analysts completed the rating procedure, which allowed us to calculate inter-analyst agreement. Based on these estimates, we opted to exclude mixed sites (class three) and collapse the four remaining classes into two simple groups: 'majority non-urban' (classes one and two) and 'majority urban' (classes four and five). For these 7,403 'non-urban' and 'urban' sites, the inter-analyst agreement rates are 80.0 and 85.3 percent, respectively.

The tier one accuracy assessment results are summarized in Table 2.3; the overall accuracy for this sample is 87.1 ± 0.4 percent and the area-adjusted user's accuracy is 83.6 percent. The user's accuracy for the urban class is quite high, indicating that a portion of the Landsat map that is labeled as 'urban' will also appear as urbanized in the high-resolution imagery 91.0 percent of the time. The producer's accuracy for urban areas is slightly lower, indicating that urbanized areas within our sample were correctly identified 89.3 percent of the time. For this assessment, both of these measures are important because we want to be certain that the UWISC/UCONN map collection is neither missing urban land nor mis-labeling non-urban areas as urban land. Since the user's and producer's accuracies from this tier one assessment are high across both classes (82.8-91.0 percent), we feel confident that the Landsat-based maps are suitable for use in our tier two assessment.

It is important to note that on average, the Landsat images were acquired five years earlier than the GE images. In the span of five years, it is possible that sites were converted from non-urban land cover types to urbanized areas. This source of spurious urban omission errors makes the 82.8 percent non-urban producer's error rate from Table 2.3 a conservative estimate. Because the full UWISC / UCONN map collection contains a much higher fraction of non-urban hexagons (90.1 percent), an underestimation of this non-urban producer's error would significantly reduce the area-adjusted user's accuracy. The problem of spurious omission errors is likely to be less important for cases where the entire site is

completely covered by urban features (class five) or cases where urban features are not at all present (class one). Based on a contingency table for just these 2,211 hexagons (not shown), the area-adjusted user's accuracy is 97.3 percent and the overall map accuracy is 97.0 ± 0.4 percent. Overall, our GE analysis confirms the findings of Schneider and Woodcock (2008) and Angel et al. (2005); we can be confident that the UWISC / UCONN data accurately portray urban features according to our physical definition from **Section 2.2**.

B. Cities omitted by the global maps

A primary requirement for any useful global urban map is its ability to depict the location of all major cities, such as those listed in a global city gazetteer. Because of time and cost limitations, it is not feasible to test for the presence / absence of each of the nearly 4,000 cities in the Angel-gazetteer. Instead, we opted to test the 247 cities in the combined *UCONN / UWISC / GE-boundary* data sets (Table 2.2, columns 3-5). If a given global map contained no urban pixels for a sample city, that city was considered omitted by that map. From this information, we estimated omission rates for all eight maps (Table 2.4) globally, and by city size, country income, and world-region. In order to establish that our findings were not dependent on our definition of an 'omitted' city, we created the same omission rates using slightly stricter definitions that required more than 5 or 10 sq. km of urban land for a city to be considered 'present' in a global map. We found that the inter-map patterns from Table 2.4 were unchanged.

Overall, MOD500 and IMPSA omit no cities in our 247-city sample, GLC00 and GLOBC have the highest omission rates (12.1 and 6.1 percent, respectively), and the remaining maps have intermediate omission rates (0.4 to 3.2 percent). Table 2.4 also allows us to track the omission rates for each map by city size, income class, and world region (same breakpoints as in Figure 2.3). From the mean omission rate column, it is clear that there are strong gradients in omission rates by both city size and country income. In

general, smaller cities and cities from low-income countries are missed by the global maps at rates several times higher than the omission rates for larger cities in high-income countries. Most global maps follow the overall trend of declining omission rates with respect to increasing city size and country income, with the exception of GRUMP, which has all of its omissions in moderate-low-income countries (10.4 percent omission rate), and VMAP0, which has the bulk of its omissions in moderate-income countries (9.4 percent omission rate). There is a similarly marked trend by world-region; the results for Europe and ODC reveal a mean omission rate of only 0.2 percent, a full order of magnitude smaller than the next best region (Africa, 2.8 percent). This is to be expected given the fact that Europe and ODC contain a much larger fraction of wealthy countries and large cities than any of the other regions. South-Central and Western Asia (including India) and Southeast and East Asia (including China) stand out as the regions with the largest omission rates (5.5 and 5.8 percent, respectively).

Looking more closely at the class-level omission rates by region (Table 2.4), two important trends stand out. First, it is evident that GLOBC relies on GLC00 maps for much of the Latin American and Caribbean region: this is apparent in the matching GLOBC/GLC00 omission rates (6.7 percent). Yet GLOBC, the newer of the two maps, does much better than GLC00 for Africa (3.7 versus 14.8 percent) and Southeast and East Asia (4.0 versus 24.0 percent), reflecting the updated MERIS imagery and analysis used for GLOBC in these regions. Second, the rates for the GRUMP and VMAP0 map also display unusual patterns with respect to world regions. GRUMP has all of its omissions in Southeast and East Asia (14.0 percent omitted), while VMAP0 has a very high omission rate for Latin America and the Caribbean (11.1 percent).

The omitted cities from Table 2.4 contain a surprising number of very large urban populations. To visualize these results, we identify the cities in our sample with populations greater than one million that were omitted by a global urban map (Table 2.5); dark circles

indicate cities completely omitted and open circles correspond to cities mapped with areas of less than five sq. km. For a city with more than 1 million residents, a total areal extent of five sq. km is unreasonably low, and should be considered only marginally better than an outright omission. Of the eight global urban maps, only IMPSA and MOD500 manage to include all 95 of the cities in the sample with populations greater than 1 million, also mapping all of these cities at greater than five sq. km. From Table 2.5, it is also clear that the vast majority of the large-city omissions occur in the East Asia region, specifically China. Although GLOBC represents a significant improvement over GLC00, Table 2.5 illustrates that both GLOBC and GLC00 omit a significant number of very large cities (8 and 2 full omissions, and 13 and 7 omissions with less than 5 sq. km extents, respectively).

C. City size estimates

With the results of the city omission rates in mind, we turn now to comparing the areal extent of the medium-resolution UWISC / UCONN maps against the extent mapped by each of the coarse-resolution global maps (Figures 2.5, 2.6, Table 2.6). First, the box-plots in Figure 2.5 describe the distributions of the ratios of city size derived from the global maps to city size derived from the Landsat-based maps (reference) for all of the cities not omitted in **Section 2.4.B**. For these plots, we use the full 28.5 m-resolution Landsat imagery to estimate the reference city size, and the continuous-values for HYDE3 and IMPSA extents, respectively. In this raw, full-resolution size comparison, MOD500 most closely approximates the city-scale extent of the built environment, followed by MOD1K. With the exception of GRUMP, all of the remaining urban maps systematically underestimate city size. The inter-quartile range (width of the shaded boxes) is narrowest for MOD500, indicating good agreement across all the cities in the sample, and particularly wide for VMAP0, GLC00, and GLOBC. When we aggregate the Landsat-based reference maps to the appropriate coarse resolutions (10”, 15”, or 30” arc-seconds) and threshold HYDE3 and

IMPSA at 50 and 20 percent, respectively (selection of these thresholds is discussed in detail in **Section 2.4.D**), these relative inter-map trends remain unchanged.

We continue an exploration of the ability of each global urban map to correctly estimate the physical size of cities according to the UCONN / UWISC reference maps by estimating the correlation between the two (Figure 2.6). The trends first observed in Figure 2.5 are also evident in the dispersion patterns of the scatter-plots and the location of the best-fit lines. GRUMP clearly overestimates city size, which is an artifact of the extensive nature of this map. VMAP0, on the other hand, underestimates city size, most likely because it is less current. MOD500 appears to have the best fit to reference city size, reflecting excellent size-agreement at the scale of individual cities. The ordinary least-squares (OLS) lines for both the raw (blue) and aggregated (red) reference data are quite similar, indicating that these relationships are robust with respect to scale. We examined these scatter-plots for patterns with respect to country income level and world-region, and found no clear trends. There are no cities which are consistently outliers across all eight maps.

Table 2.6 provides additional information on the coefficients, slopes, and adjusted R^2 values for regressions of city extents based on Landsat-based reference data as a function of city extents based on global urban maps. These regressions are similar to the blue OLS lines from Figure 2.6, except they are linear rather than logarithmic. For each regression, we include only those cities not omitted in **Section 2.4.B**. Because the standardized residuals from these regressions are much higher for cities at the extremes of the size-distribution, we exclude all cities with reference-map (Landsat-based) extents less than 100 sq. km or more than 2,000 sq. km (sample sizes were larger than 100 for all cities; column one of Table 2.6). As demonstrated graphically in Figure 2.6, MOD500 has the least unexplained variance in its estimates of city size (highest R^2 of 0.90), while GRUMP, HYDE3, MOD1K, and IMPSA have intermediate R^2 values (0.60-0.75). VMAP0, GLC00, and GLOBC have the most unexplained city size variance (lowest R^2 values of 0.56, 0.36, and 0.30, respectively). The

OLS regression results in Table 2.6 serve an additional function: they can be used to correct city-size estimates made from global urban maps. All of the slopes in Table 2.6 are highly significant (p-value < 0.01) and the intercepts are significant (p-value < 0.03) for all of the maps except GRUMP (p-value 0.66). These correction factors are most effective for cities whose global map-based extent is within the range included in the OLS regression (Table 2.6, column two).

D. Map agreement measures

To compliment our suite of accuracy measures, we conduct a pixel-based assessment comparing each global urban map to each city in the UCONN / UWISC collection. For this analysis, we begin by creating individual 2 x 2 contingency matrices for the urban and non-urban classes for each city in the sample; Table 1.4a is an illustrative example. As in **Section 2.4.C**, we use a simple 50 percent rule to aggregate the Landsat-based maps to the appropriate resolution prior to comparison. Translating the information in this simple contingency table into a robust measure of map agreement can take many forms (Table 1.4b; Congalton and Green 1999), and there is no clear consensus on the optimal approach either within or across disciplines (Fielding and Bell 1997; Manel et al. 2001; McPherson et al. 2004; Allouche et al. 2006; Foody 2006, 2007; Lobo et al. 2007; Liu et al. 2007). We align our assessment with Foody's recommendations (2006; 2007) by relying on several agreement measures. We report Cohen's Kappa (Table 1.4b; Cohen 1960) because of its long history of use in the remote sensing literature and the ability to estimate confidence intervals (Fleiss et al. 1969), and we also estimate the True Skill Statistic (TSS; Allouche et al. 2006) and Normalized Mutual Information (NMI; Forbes 1995). Landis and Koch's (1977) recommendations for interpreting Kappa values are a useful guide: (0.00 - 0.20) can be considered 'slight' agreement, (0.21 - 0.40) 'fair,' (0.41 - 0.60) 'moderate', (0.61 - 0.80) 'substantial,' and (> 0.81) 'almost perfect' agreement.

In assessing map agreement at the pixel level, the two continuous maps require special attention. HYDE3 and IMPSA represent the sub-pixel fraction of urban and impervious surface areas, and must be re-classified into binary 'urban / non-urban' maps before they can be compared to the UCONN / UWISC reference data. Simple 50 percent thresholds do not necessarily yield the highest overall accuracy measures. To select optimal thresholds for IMPSA and HYDE3, we estimate Kappa, TSS, and NMI values for cities in the UCONN / UWISC collection across the full range of potential thresholds (5 to 95 percent in 5 percent increments). Based on this analysis, a threshold of 50 percent urban for HYDE3 and 20 percent impervious surface for IMPSA yields global maps with the highest accuracies across all measures. For the remainder of **Section 2.4.D** (and previously in **Section 2.4.C**), we threshold these maps accordingly.

A second concern that is particularly relevant to pixel-based accuracy assessment is co-registration. Because we evaluate agreement on a pixel-by-pixel basis, a shift of only a single pixel in the geo-registration of a global map can result in significant differences in agreement statistics. This sensitivity represents both a challenge and an opportunity; although we conducted tests based on cardinal and inter-cardinal shifts of the global maps for **Sections 2.4.B –2.4.C**, only in this section is the response sensitive enough to allow robust testing of geo-registration. To optimize the map agreement results, we found that two maps required global shifts in their horizontal alignment: we shifted IMPSA one pixel (30" arc-seconds) southeast and MOD500 one pixel (15" arc-seconds) northwest.

The five accuracy measures in Table 2.7 and the red bars in Figure 2.7 are constructed using weighted global contingency matrices, where the weights are based on the income/size scheme from **Section 2.3.A**. The sensitivity and specificity measures in Table 2.7 are of limited utility because of their strong correlations with the total global extent of each map; each column is arranged in order of increasing global extent (top to bottom), and specificity declines roughly linearly while sensitivity increases in a similar fashion. VMAP0's

high specificity score of 0.996 is not surprising given the dated nature of that map and its very small global extent, likewise for GRUMP's high sensitivity score of 0.91 and its extremely large global extent. Because of its limited range (0.960 – 0.977, with the exception of GRUMP, 0.839), the overall map accuracy measure is also of little use in discerning inter-map differences. While overall accuracy is the most often citied measure in the remote sensing literature, the two overall agreement measures, TSS and Kappa, demonstrate a much wider range across our maps and are much less correlated with overall global extent. Here, the scores range from 0.38 to 0.75 for TSS and 0.22 to 0.63 for Kappa. From Figure 2.7 and Table 2.7, it is clear that IMPSA and MOD500 have significantly higher Kappa distributions indicating 'substantial' agreement with the Landsat assessment data, GRUMP has consistently lower Kappa values indicating 'slight' agreement, and VMAP0, GLC00, GLOBC, HYDE3, and MOD1K have intermediate values indicating 'fair' to 'moderate' agreement. The un-weighted box-plots of Figure 2.7 capture the 140 city-level UCONN / UWISC Kappa measurements for each of the global urban maps; these follow the basic trends of global weighted Kappa values (95 percent confidence bounds based on methods in Fleiss et al. 1969), indicating that the inter-map patterns are robust with respect to weighting schemes.

From Table 2.7, it is clear that the overall pattern of map agreement persists across measures (the same is true for NMI, although it is not shown here). Looking across all measures, MOD500 stands out as the closest to the Landsat-based reference data, GRUMP, MOD1K, and IMPSA are all of intermediate agreement, and VMAP0, GLC00, GLOBC, and HYDE3 have the lowest overall measures of agreement. GRUMP is an exception worth examining closer. Because it consistently overestimates the size of cities (Figures 2.5 and 2.6), GRUMP is much less likely to commit urban omission errors, leading to very high sensitivity, which is an important component of TSS (Table 1.4b). Yet the same overestimates force GRUMP to make a large number of urban commission errors, drawing

down specificity, overall accuracy, and Kappa. Although GRUMP is grouped with the MODIS maps based on its TSS value, GRUMP is probably more properly placed with the lower-tier maps because of its high commission rates and low overall map accuracy. It is also interesting to note that GLOBC has lower or equivalent Kappa and TSS values relative to VMAP0, although VMAP0 is an older map; clearly urban expansion since the 1990's release of VMAP0 was not significant enough to offset the accuracy limitations of GLOBC.

Table 2.8 depicts the un-weighted regional Kappa and TSS values for each global map. From the eight-map means (bottom row), there is a similar inter-map pattern as was clear for city omission rates (**Section 2.4.B**) and inter-map comparisons (**Chapter 1**): Europe and the ODC and Latin America and the Caribbean have the highest TSS and Kappa values, Asia has the lowest, and Africa is intermediate. For most regions, the inter-map pattern from Table 2.7 remains intact (green values are regional maxima and red values are regional minima): MOD500 and IMPSA have the highest Kappa and TSS values for all five regions and GRUMP has the lowest Kappa values. In reporting these regional map agreement measures, it is important to recall that the two continuous maps (IMPSA and HYDE3) have been thresholded for optimal performance at a global level; sets of regional thresholds may exist that would improve regional agreement measures for these two maps.

2.5 CONCLUSIONS

Monitoring the world's cities during the intense urbanization processes of the next twenty-five years is essential to a wide range of research agendas in both the social and physical sciences (**Section 1.1.A**). Within the context of this monitoring effort, it is crucial that we constrain the accuracy of global urban maps. We began this two-tiered accuracy assessment by re-examining the collection of UCONN and UWISC Landsat-based maps that form the basis for the majority of our accuracy measures (tier one). Based on a comparison with more than 10,000 validation sites constructed with high-resolution imagery from a new

Google Earth validation tool, we found the map accuracy of the UCONN and UWISC maps to be both high (upper-80 to lower-90 percent overall map agreement) and consistent with previous assessments (Angel et al. 2005; Schneider and Woodcock 2008).

In the second-tier of our assessment, we leveraged these medium-resolution city maps for three separate investigations of accuracy (**Sections 2.4.B, 2.4.C, 2.4.D**). The ranks in Table 2.9 summarize the city omission (column one), size correlation (R^2; column two), and map agreement (Cohen's Kappa; column three) for each of the eight global urban maps. In all four columns of Table 2.9, lower ranks indicate better accuracies. Overall, MOD500 has the lowest mean rank (1.0), followed by IMPSA (2.7) and MOD1K (3.0). Although there is no single, best method for combining disparate accuracy measures, these overall rankings are robust across a wide range of averaging schemes and accuracy measures.

In the following, we offer map users our wise-use recommendations, summarize some general methods for improving the accuracy of the maps, and discuss the road ahead for urban map accuracy assessment.

A. Wise-use recommendations

Clearly all map users would prefer to select the map that most closely corresponds to the patterns of built environment observed on the ground. MODIS and IMPSA are the maps that best meet this objective (row 1, Table 2.10). However, global map users often have a complex set of priorities, and map accuracy is not always their most pressing concern. It is possible to imagine several scenarios in which other attributes would take precedence (rows 2-7, Table 2.10):

(1) Highest resolution: Users who are conducting comparisons with medium-resolution regional to global-scale data may require the global map with the highest spatial resolution. Some examples of projects where resolution could be an important constraint

include: examinations of the global protected area system that draw on moderate-resolution maps of the world's parks, transportation studies that rely on detailed vector maps of road networks, and public health projects that include geo-coded point data from field surveys. GLOBC is the highest-resolution global urban map, but offers very low map accuracy, and in some cases the urban areas mapped in GLOBC are fill values from coarser-resolution maps (ESA 2008). If resolution is the primary concern, MOD500 offers the best compromise of accuracy and detail.

(2) Continuous-value urban land cover: Some users may need a more sensitive measure of urban areas than the binary, presence/absence model that is most common in global maps. With a continuous-value urban map, these users could disaggregate urban areas into several discrete classes of urban density. IMPSA and HYDE3 are currently the only maps to offer continuous-value estimates of impervious surface and urban areas, respectively. Of the two, IMPSA is the clear choice because of its superior map accuracy and spatial resolution. An alternative with equivalent or superior accuracy would be to aggregate MOD500 to 30" arc-second resolution (or coarser), creating a continuous-value, percent-urban map from the binary MOD500 map.

(3) Most recent map: In the case of disaster response and planning, there is a premium placed on maps that are current. GLOBC claims to be the most recent global urban map because it is based on 2005 observations from the MERIS satellite, but when GLOBC encounters consistent cloud-cover, it defaults to 'fill' values from a wide range of older maps (ESA 2008). Given these limitations, IMPSA is the global map with acceptable accuracy that is based on the most recent input data (2004 LandScan and 2000-2001 calibrated nighttime lights). An update to the MOD500 map based on 2005 data is in planning, and this map would represent an alternative for users who desire a current map that also has high accuracy and sub-kilometer spatial resolution.

B. Methods for improving map accuracy

In the course of this assessment, we have arrived at several procedures for improving the accuracy of the eight global maps under review:

(1) Map alignment for IMPSA and MOD500: Because of the sensitivity of the pixel-based measures in **Section 2.4.D** to small shifts in map alignment, we recommend a 30" arc-second (one pixel) southeast shift in the global alignment of IMPSA and a 15" arc-second (one pixel) northwest shift in MOD500 relative to their original projections.

(2) Thresholding for HYDE3 and IMPSA: We identified optimal thresholds for converting the continuous maps to binary urban / non-urban maps. Based on the map agreement measures (**Section 2.4.D**), thresholds of 50 and 20 percent maximize the accuracy of HYDE3 and IMPSA, respectively.

(3) Calibrating estimates of city size: The size-correlation analysis from **Section 2.4.C** allows for the correction of city-size estimates based on our reference data. Table 2.6 contains the coefficients and intercepts for calibrations of each global urban map. As noted in **Section 2.4.C**, these measures are best applied to cities with total extents within the range of the city sizes used as the basis for each regression.

C. The path ahead

An important area that was not explored in our analysis is the accuracy of human settlement maps for places with populations smaller than 100,000 (**Section 2.3.A**). Such an assessment would require the collection of new validation data from high-resolution imagery archives and fieldwork, and the expansion of the sample frame to include cities outside the scope of the Angel et al. (2005) gazetteer. Such an undertaking is of considerable importance because a quarter of all urban residents may live in such smaller cities and the regional to global extent of the rural built environment remains almost entirely unconstrained.

Our strategy covers a wide range of desirable map qualities and our assessment data sets can be readily updated using freely available imagery. In particular, the Google Earth land cover validation tool will be of use in providing a steady stream of updated high-resolution imagery for validating urban land cover maps. In addition to using GE imagery for assessing medium-resolution maps, future assessments could use the high-resolution imagery to directly assess individual coarse-resolution pixels on a regional scale. The repeatability of this assessment is essential; with the world's urban population projected to grow by nearly two billion through 2030, a steady stream of new global urban maps is on the way and that stream will require a standardized map assessment model that is tailored to the rare but critical urban land-use category.

	Cost	Image Extent	Revisit Times	Spectral Properties
High-resolution (< 10 m)	typically high	local (~10 x 10 km)	typically monthly to annually (weekly with pointing)	visible, SWIR
Medium-resolution (10-100 m)	low to medium	regional (~100 x 100 km)	typically monthly to bi-annually (16-18 days for some locations)	visible, SWIR, MIR, thermal
Moderate- to coarse-resolution (> 100 m)	typically low	global (~1,000 x 1,000 km)	24 to 48 hours (often composited to 8 or 16 days)	visible, SWIR, MIR, thermal

Abbreviations: 'SWIR' short-wave infrared, 'MIR' middle-infrared.

Table 2.1 Image sources for mapping urban areas. A comparison of the strengths and weaknesses of the imagery available for mapping urban areas and the built environment, focusing on the trade-offs between spatial resolution (rows) and cost, coverage, revisit times, and spectral properties (columns).

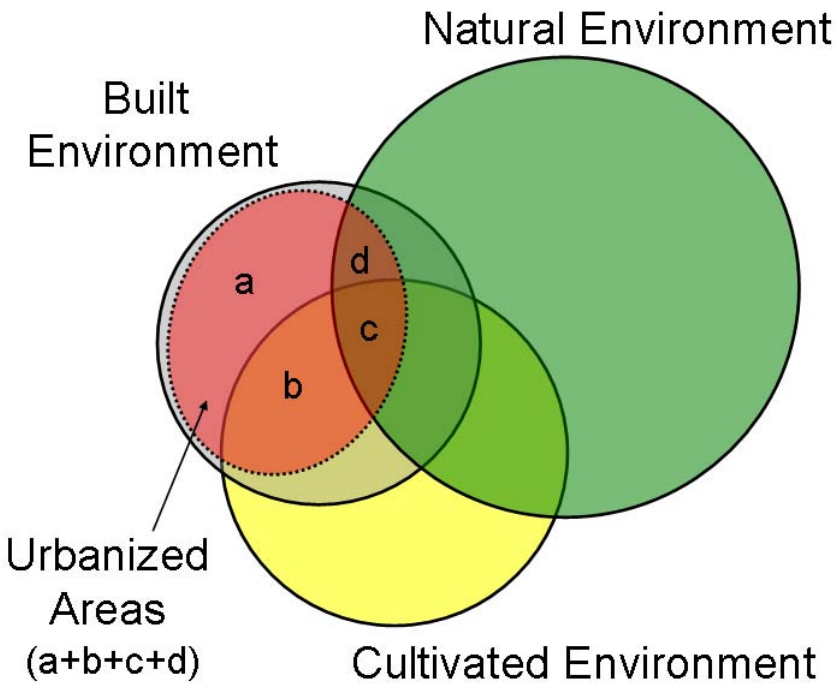


Figure 2.1 A Venn diagram depicting urban areas. The diagram shows urban areas in the context of the natural (green), built (gray), and cultivated (yellow) environments. Every coarse-resolution pixel belongs to one or more of these three environments. Urban areas (red) are those pixels that are dominated completely by the built environment (a), or contain mixtures of several environments but are majority-built (b, c, d). In this framework, major portions of the built environment can be considered non-urban (gray areas outside of the red oval). This diagram is designed only for visualizing thematic relationships and is not drawn to scale.

	Sample Framework	Assessment Data			
	Angel-Gazetteer	**Google Earth Validation Sites**	**Google Earth Boundaries**	**UCONN Maps**	**UWISC Maps**
Elements	3,949 cities	10,800 sites	107 cities	120 cities [a]	24 cities [a]
Source	tabular gazetteers	Google Earth (high-resolution)	Google Earth (high-resolution)	Landsat (medium-resolution)	Landsat (medium-resolution)
Author	Angel et al. (2005)	This study	This study	Angel et al. (2005)	Schneider & Woodcock (2008)
Section	Sect. 2.3.A	Section 2.4.A	Section 2.4.B	Sect. 2.4.B Sect. 2.4.C Sect. 2.4.D	Sect. 2.4.B Sect. 2.4.C Sect. 2.4.D

Note: (a) Four cities are duplicated in both the UCONN and UWISC samples.

Table 2.2 Assessment data sets. The sample framework is derived from the Angel et al. (2005) gazetteer (left-most column). Google Earth imagery is used to construct the validation sites and city boundaries (columns two and three), and the medium-resolution city maps were constructed in the course of two earlier studies (Angel et al. 2005; Schneider and Woodcock 2008).

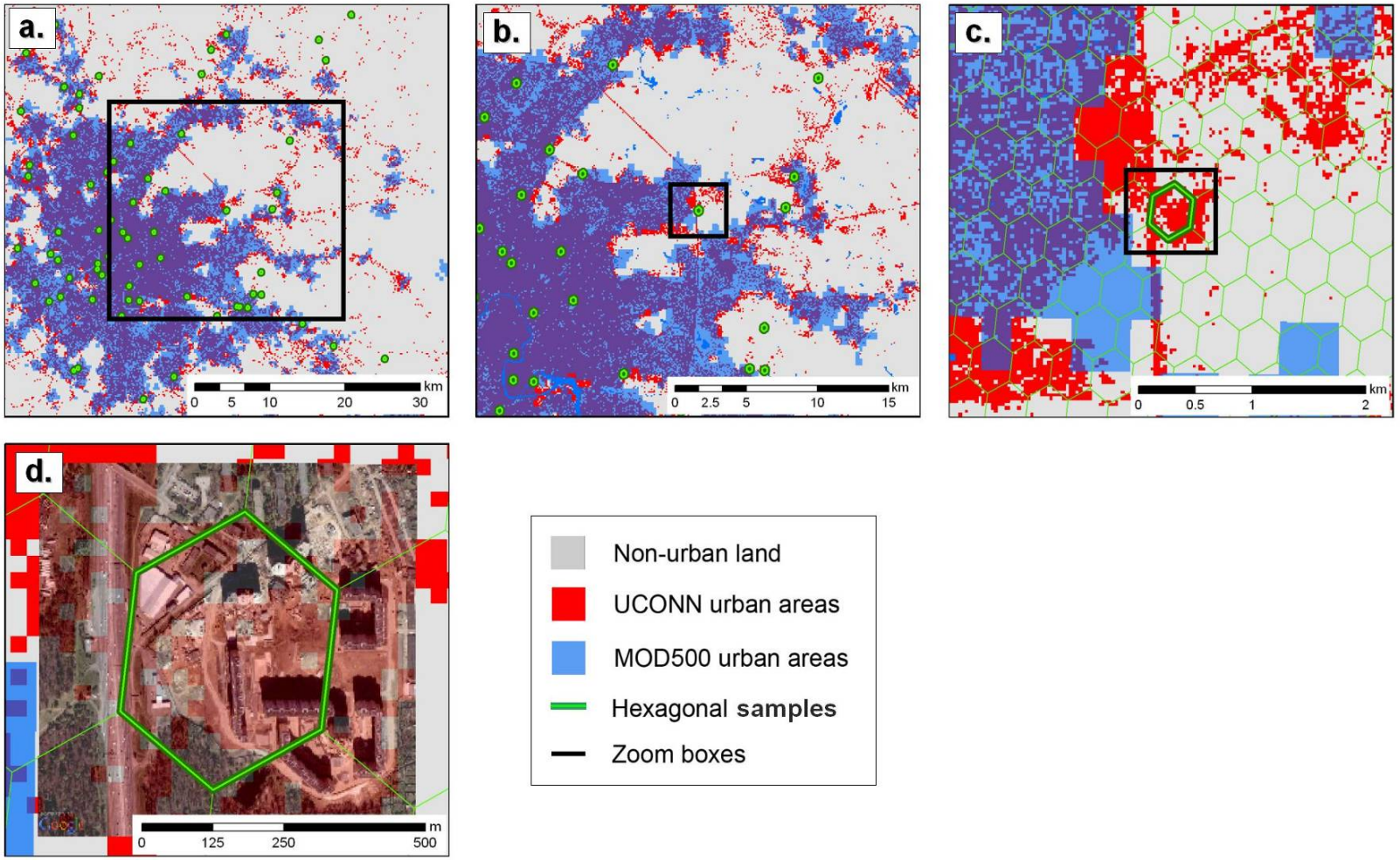
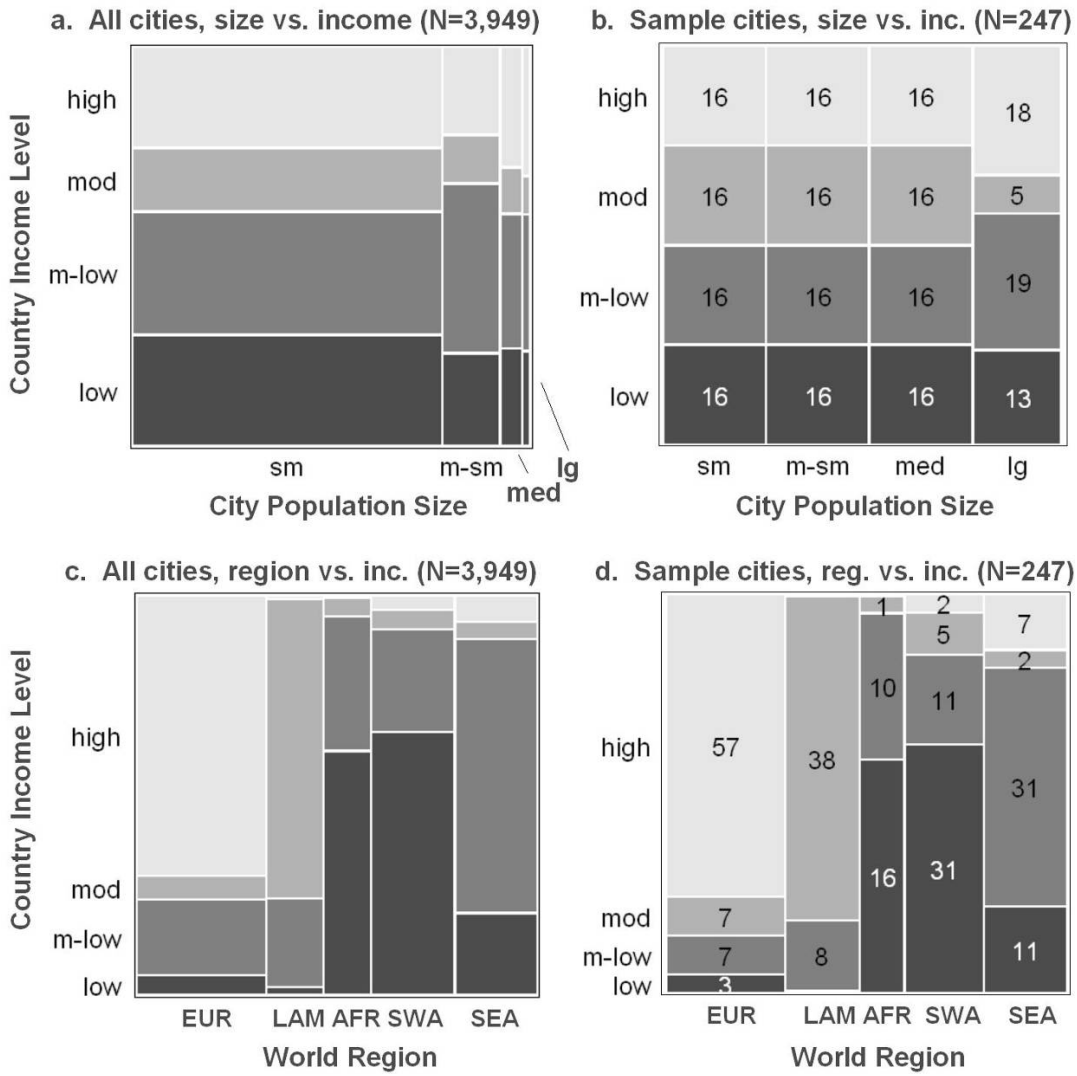


Figure 2.2 Assessment data sets for Moscow, Russia. Coarse-, medium-, and high-resolution assessment data sets for Moscow, Russia. MOD500 pixels (500 m) are depicted in pale blue (a-d), and UCONN pixels (28.5 m) are in red (a-d). The figures contain successive views (a-d), with zoom-boxes indicated in solid black. In Figures 2c and 2d the full hexagonal mesh is traced in green lines. Figure 2d includes the imagery from the Google Earth Land Cover Validation Tool used to assess the site.



Notes:

(1) City population size circa-2000 :
- Small (SM) : 0.1 to 0.5 M
- Medium-small (M-SM) : 0.5 to 1.5 M
- Medium (MED) : 1.5 to 4.2 M
- Large (LRG) : >4.2 M

(2) Country income circa-2000 :
- Low (LOW) : < $3K
- Moderate-low (M-LOW) : $3 to 5.2K
- Moderate (MOD) : $5.2 to 17K
- High (HIGH) : >$17K

(3) World Region : Europe and Other developed countries (EUR), Latin America and the Caribbean (LAM), Africa (AFR), South-Central and West Asia (SWA), SE and East Asia (SEA)

Figure 2.3 Sample and gazetteer cities stratified by city size, income and region. The distribution of cities with respect to country income level and world region for both the Angel et al. (2005) gazetteer (a,c) and the assessment sample (b,d). Country income is determined from *per-capita* gross domestic product (GDP) in purchasing power parity (World Bank 2008; UN 2008a). The abbreviations and breakpoints for income, city-size, and region are described in notes 1, 2, and 3 and in Angel et al. (2005).

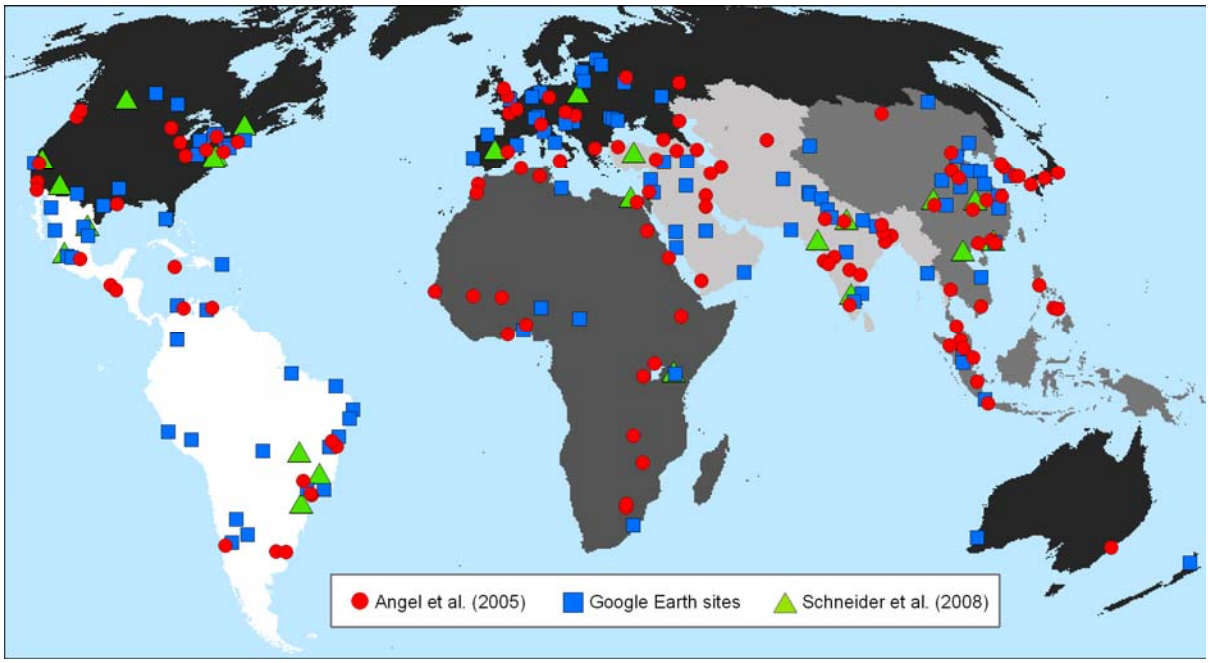


Figure 2.4 Map of assessment cities. This map shows the three assessment data sets: red circles for the 120 UCONN Landsat maps (Angel et al. 2005), green triangles for the 25 UWISC cities (Schneider and Woodcock 2008), and blue boxes for the 107 Google Earth urban boundaries (this study). The gray regional scheme corresponds to the World Bank continental divisions used to assemble the sample frame (Figures 3c and 3d).

		Reference data (Google Earth validation sites)					
		Non-urban	Urban	Total	Class prevalence (%)	User's accuracy (%)	
UWISC/ UCONN-maps	Non-urban	2,905	604	3,509	90.1	**82.8**	
	Urban	349	3,545	3,894	9.9	**91.0**	
	Total	3,254	4,149	7,403		***83.6***	**Overall accuracy (area-adjusted)**
	Producer's accuracy (%)	**89.3**	**85.4**			***87.1*** ± *0.4*	**Overall map accuracy**

Table 2.3 Tier one contingency table. Contingency table comparing the Landsat-based city map labels (UCONN / UWISC-map collections) to those derived from the Google Earth tool (GE-validation sites). The 7,403 sites in this table represent the cloud-free, high-resolution sites that were rated as majority urban or majority non-urban. Area-adjusted user's accuracy (83.6 percent) is based on the prevalence of non-urban and urban hexagons in the full UCONN/UWISC map collection (*class prevalence*, column four).

	VMAP0	GLC00	GLOBC	HYDE3	IMPSA	MOD500	MOD1K	GRUMP	8-map mean	class size (cities)
Total Omitted Cities	**8**	**30**	**15**	**2**	**0**	**0**	**1**	**7**	**7.9**	
Omission Rate (%)	**3.2**	**12.1**	**6.1**	**0.8**	**0**	**0**	**0.4**	**2.8**	**3.2**	
a. By Size (%)										
Small	9.4	28.1	17.2	1.6	0	0	1.6	3.1	**7.6**	64
Medium-Small	3.1	10.9	4.7	0	0	0	0	3.1	**2.7**	64
Medium	0	7.8	1.6	0	0	0	0	3.1	**1.6**	64
Large	0	0	0	1.8	0	0	0	1.8	**0.5**	55
b. By Income (%)										
Low	1.6	14.8	9.8	3.3	0	0	1.6	0	**3.9**	61
Moderate-Low	1.5	20.9	6.0	0	0	0	0	10.4	**4.9**	67
Moderate	9.4	7.5	5.7	0	0	0	0	0	**2.8**	53
High	1.5	4.5	3.0	0	0	0	0	0	**1.1**	66
c. By Region (%)										
Europe & ODC	0	1.3	0	0	0	0	0	0	**0.2**	75
Latin America and the Caribbean	11.1	6.7	6.7	0	0	0	0	0	**3.1**	45
Africa	0	14.8	3.7	3.7	0	0	0	0	**2.8**	27
South-Central and Western Asia	2.0	20.0	18.0	2.0	0	0	2.0	0	**5.5**	50
Southeast and East Asia	4.0	24.0	4.0	0	0	0	0	14.0	**5.8**	50

Table 2.4 City omission rates. Cities omitted by the global urban maps from the 247-city *UCONN/UWISC/GE-boundaries* dataset. Sections a, b, and c track omission rates by city-size, country-income, and world region. The right column is the mean omission rate for each division.

Omitted City	(pop., 1000s)	VMAP0	GLC00	GLOBC	HYDE3	IMPSA	MOD500	MOD1K	GRUMP
Lagos, Nigeria	(13,400)				●				
Dongguan, China	(6,400)	○							
Bogota, Colombia	(6,300)		○	○					
Wuhan, China	(5,200)								●
Kabul, Afghanistan	(2,600)		●	●				○	
Zhengzhou, China	(2,100)		●						
Baku, Azerbaijan	(1,900)		○	○					
Barranquilla, Colombia	(1,700)		○	○					
Sanaa, Yemen	(1,700)		○	○					
Nanyang, China	(1,600)		●						●
Yulin, China	(1,600)	○	●						
Taian, China	(1,500)		●						●
Yerevan, Armenia	(1,400)		●	●					
Yiyang, China	(1,300)		●						●
Fukuoka, Japan	(1,300)		○	○					
Huaian, China	(1,200)		●						
Shenzhen, China	(1,100)	●							
Leshan, China	(1,100)	○							
Omission Total (cities)		**1**	**8**	**2**	**1**	**0**	**0**	**0**	**4**
Omission Rate for Cities with Pop. > 1 million (%)		**1.1**	**8.4**	**2.1**	**1.1**	**0**	**0**	**0**	**4.2**

Note:
● – cities completely omitted
○ – cities with less than 5 km^2 urban land

Table 2.5 Omitted cities with populations greater than 1 million. The counts and rates in the bottom rows are for the 95 cities with populations greater than 1 million in the 247-city UCONN/UWISC/GE-boundaries sample.

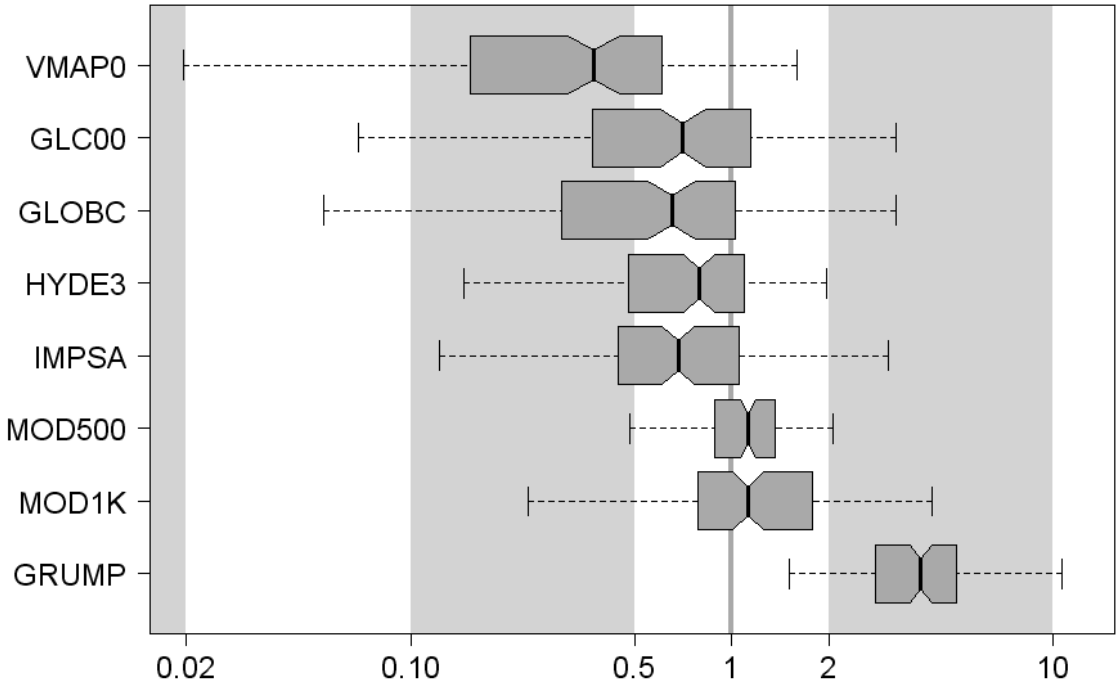


Figure 2.5 Errors in estimating city-size. Box-plots depict the log-ratios of city size; global map estimate over UCONN/UWISC map estimate (reference). The '1' line indicates perfect (1:1) agreement. Box-plots to the right of this line show those maps where city sizes are overestimated, and the box-plots to the left of this line indicate maps where city sizes are underestimated.

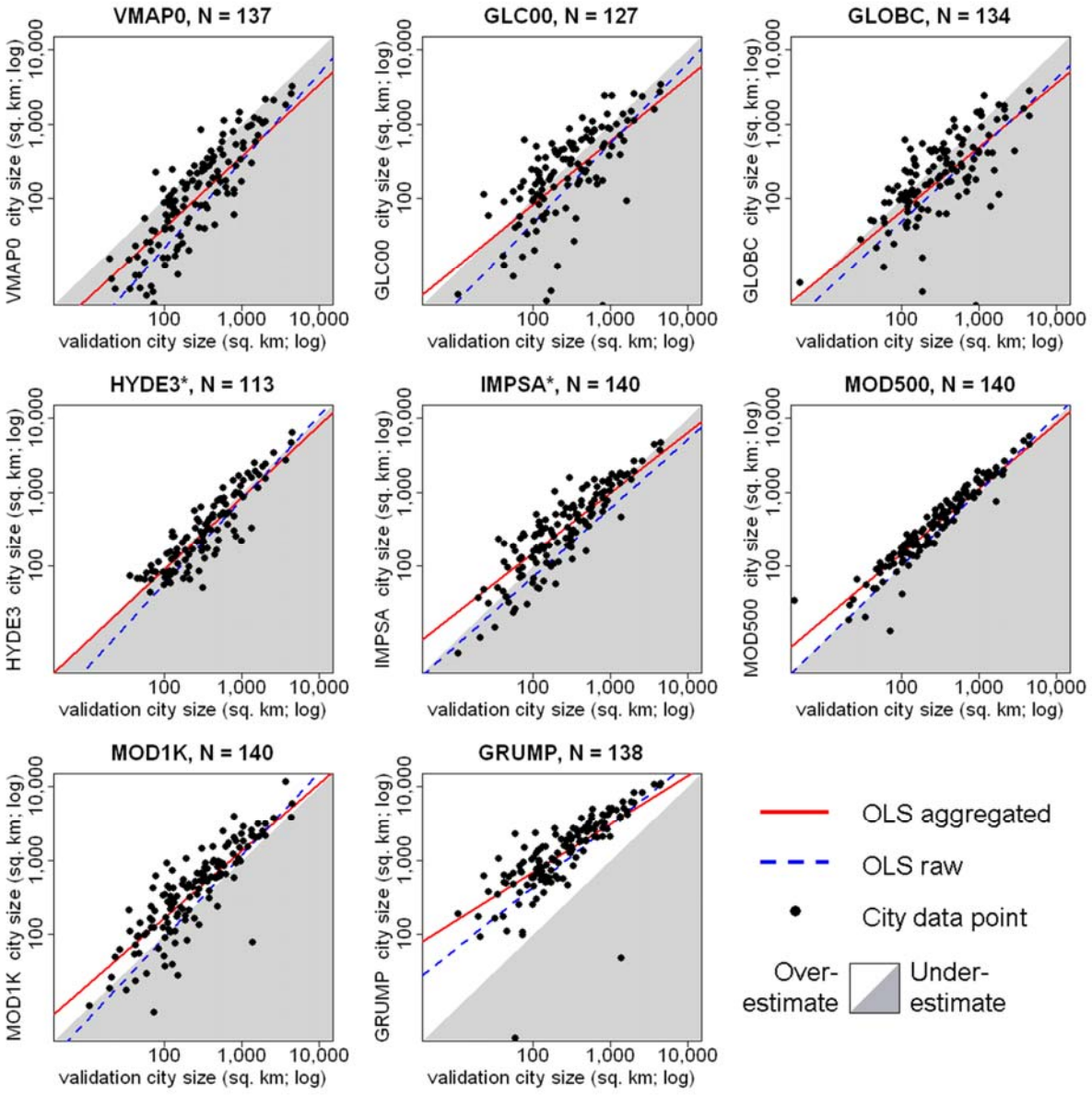


Figure 2.6 Scatter-plots of city size. UCONN/UWISC city size versus global urban map city size (log-log scale). Sample sizes (N) vary from map to map because only those cities that were not identified as omitted in **Section 2.4.B** were included. The x-axis is city size in square kilometers according to the UWISC / UCONN Landsat-based reference maps, where the medium-resolution pixels were first aggregated to the appropriate coarse-resolution using a simple majority rule and HYDE3 and IMPSA (asterisks) were thresholded at 50 and 20 percent, respectively. The gray triangle marks the 1:1 line; city sizes in the shaded region are underestimated and city sizes in the white region are overestimated. The red line marks the ordinary least squares (OLS) fit for aggregated data, while the dashed blue line marks the OLS fit based on city sizes from the raw medium-resolution UCONN/UWISC maps and the native, continuous HYDE3 and IMPSA maps.

	Sample size (N)	City Size Range for Global Map (sq. km)	Slope	Intercept (sq. km)	Standard Error of Slope	Standard Error of Intercept (sq. km)	Adjusted R^2
VMAP0	110	(5 - 1,443)	1.091	238.0	0.093	7.1	0.56
GLC00	101	(6 - 2,420)	0.602	258.6	0.080	5.6	0.36
GLOBC	105	(5 - 2,428)	0.605	278.2	0.090	6.0	0.30
HYDE3	111	(36 - 2,085)	0.735	161.9	0.043	5.9	0.73
IMPSA	112	(23 - 1,441)	0.993	122.8	0.076	3.4	0.60
MOD500	112	(13 - 2,137)	0.803	42.0	0.025	2.3	0.90
MOD1K	112	(9 - 3,971)	0.464	158.4	0.031	5.1	0.67
GRUMP	110	(97 - 5,127)	0.274	-14.8	0.015	-0.4	0.75

Table 2.6 City size correction coefficients. Correction factors for estimating reference (Landsat-based) city size using a linear univariate regression based on the extents from each global urban map. These regressions only include cities which have a non-zero extent in the global urban maps and an extent between 100 and 2000 sq. km in the reference (Landsat-based) map. All of the slopes are highly significant ($p < 0.01$) and all of the intercepts are significant ($p < 0.03$) expect for the GRUMP intercept ($p = 0.66$).

	Cohen's Kappa		Overall Accuracy		True Skill Statistic		Sensitivity (1-Omission)	Specificity (1-Comission)
		(95% conf.)		(95% conf.)		(95% conf.)		
VMAP0	0.49	(0.41 -0.58)	0.977	(0.97 -0.98)	0.38	(0.30 -0.46)	0.38	0.996
GLC00	0.45	(0.37 -0.53)	0.970	(0.96 -0.98)	0.41	(0.33 -0.49)	0.42	0.987
GLOBC	0.46	(0.39 -0.54)	0.968	(0.96 -0.97)	0.38	(0.30 -0.46)	0.39	0.991
HYDE3	0.44	(0.36 -0.52)	0.969	(0.96 -0.97)	0.49	(0.39 -0.58)	0.51	0.982
IMPSA	0.61	(0.54 -0.68)	0.975	(0.97 -0.98)	0.59	(0.51 -0.67)	0.60	0.989
MOD500	0.63	(0.57 -0.69)	0.972	(0.97 -0.98)	0.73	(0.66 -0.80)	0.75	0.980
MOD1K	0.50	(0.44 -0.57)	0.960	(0.95 -0.97)	0.67	(0.59 -0.75)	0.70	0.969
GRUMP	0.22	(0.18 -0.25)	0.839	(0.83 -0.85)	0.75	(0.70 -0.80)	0.91	0.837
Mean	*0.48*		*0.954*		*0.55*		*0.58*	*0.966*

Table 2.7 Five measures of map agreement for the global urban maps. Each of the eight global urban m against the full UCONN/UWISC assessment data set in order to estimate five measures of overall map accuracy (incl confidence bounds).

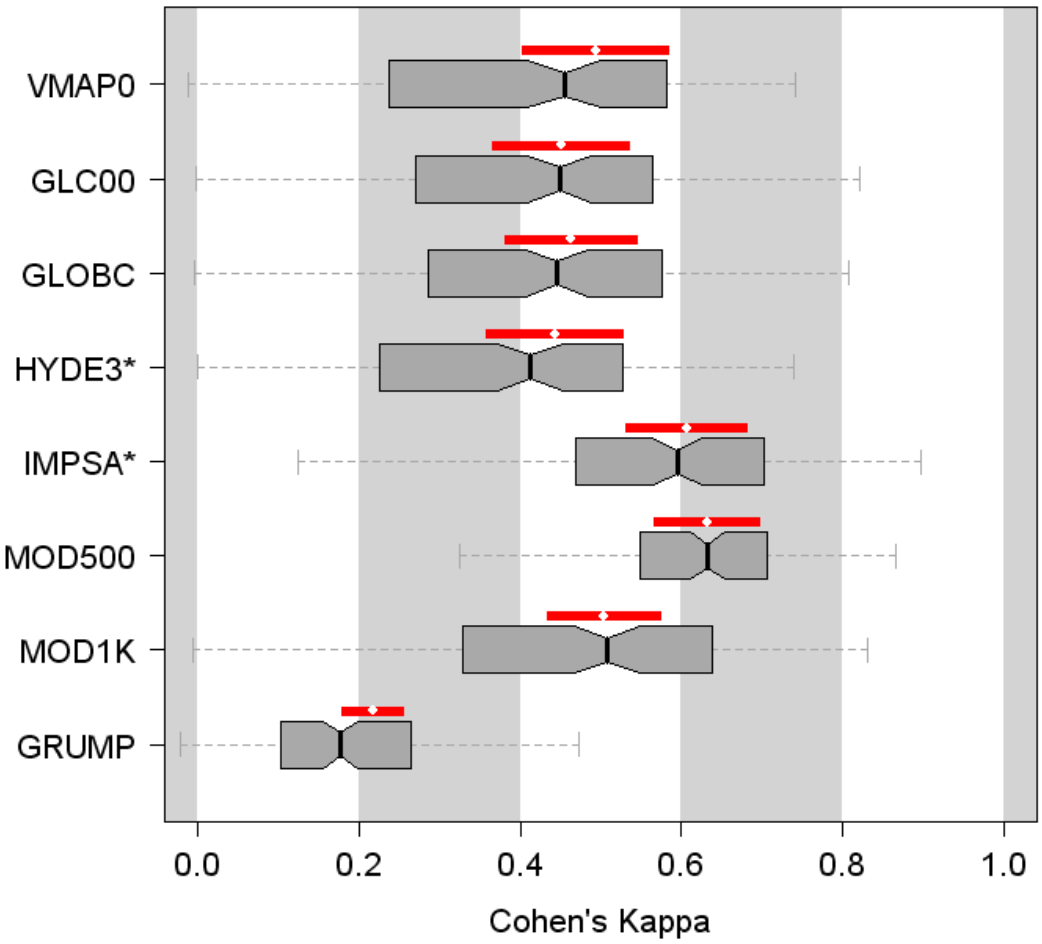


Figure 2.7 Cohen's Kappa estimates for each of the eight global urban maps. The UCONN / UWISC data sets were aggregated to the appropriate coarse-resolution and HYDE3 and IMPSA (asterisks) were thresholded at 50 and 20 percent, respectively. The box-plots depict the distribution of Kappa values across all 140 cities (a contingency table was constructed for each city). The red bars and white dots are Kappa values based on a pooled, weighted global contingency table for each map, with 95 percent confidence bounds (Fleiss et al. 1969).

	Europe and Other Developed Countries		Latin America and Caribbean		Africa		South-Central and Western Asia		Southeast and East Asia	
	Kappa	**TSS**	**Kappa**	**TSS**	**Kappa**	**TSS**	**Kappa**	**TSS**	**Kappa**	**TSS**
VMAP0	0.58	0.50	0.61	0.51	0.37	0.27	0.51	0.45	0.29	**0.19**
GLC00	0.53	0.48	0.57	**0.45**	0.52	0.71	0.48	0.59	0.37	0.32
GLOBC	0.50	**0.39**	0.59	0.48	0.53	0.64	0.46	0.50	0.43	0.38
HYDE3	0.57	0.54	0.62	0.63	0.35	**0.47**	0.50	0.49	0.36	0.50
IMPSA	**0.75**	0.75	**0.78**	0.71	0.57	0.50	0.52	**0.38**	0.56	0.49
MOD500	0.73	**0.78**	0.75	**0.86**	**0.65**	**0.76**	**0.62**	**0.80**	**0.58**	**0.66**
MOD1K	0.60	0.74	0.70	0.84	0.47	0.74	0.43	0.76	0.41	0.57
GRUMP	**0.28**	0.60	**0.23**	0.57	**0.25**	0.76	**0.17**	0.74	**0.26**	0.54
Global Mean	**0.57**	**0.60**	**0.61**	**0.63**	**0.46**	**0.61**	**0.46**	**0.59**	**0.41**	**0.46**

Table 2.8 Regional map agreement measures. Un-weighted Kappa and TSS values for each of the eight global urban maps for five world regions. The red and green values mark the regional minima and maxima, respectively.

	Section 2.4.B	Section 2.4.C	Section 2.4.D	
	1 - City Omission Rate (rank)	City Size R^2 (rank)	Cohen's Kappa (rank)	Mean Rank
VMAP0	5	6	4	**5.0**
GLC00	7	7	6	**6.7**
GLOBC	6	8	5	**6.3**
HYDE3	3	3	7	**4.3**
IMPSA	1*	5	2	**2.7**
MOD500	1*	1	1	**1.0**
MOD1K	2	4	3	**3.0**
GRUMP	4	2	8	**4.7**

Table 2.9 A summary of the accuracy results for each global urban map. The ranked global-scale results based on omitted cities, city size estimates, and overall map agreement. Ties are indicated with an asterisk.

Map Quality	Description	Recommended map
Highest map accuracy	Strongest agreement with medium- and high-resolution validation data	• **MOD500** has the highest accuracy • **IMPSA** follows closely for pixel-based accuracy measures and performs equally-well in terms of omitted cities
Highest spatial resolution	Greatest level of spatial detail	• **MOD500** is the best compromise between pixel size and accuracy • **GLOBC** has the highest spatial resolution but a very low mean accuracy
Most recent map	Underlying source data collected most recently	• **GLOBC** relies on the most current imagery (2005 MERIS data), but defaults to circa-2000 maps when confronted with cloud cover • **IMPSA** is based partly on 2004 LandScan data, and is a better compromise between current source data and accuracy • A **MOD500 update** will apply the methods behind **MOD500** to 2005 MODIS data
Regional focus	Highest within-region accuracy rates	• **MOD500** and **IMPSA** are the most accurate across nearly all regions • **VMAP0** and **GLC00** should be avoided in Southeast and East Asia because of low accuracies
Low urban commission errors	Few non-urban areas mistakenly labeled as urban	• **VMAP0** is the most conservative map of urban extent, although it has a high city omission rate • **GRUMP** should be avoided based on low specificity • **Alternatives:** (1) create a low urban commission error composite of all pixels which are labeled urban in more than one high-accuracy urban map, (2) threshold **IMPSA** at very high percent urban levels
Low urban omission errors	Few urban areas mistakenly labeled as non-urban	• **GRUMP** has the highest sensitivity, but also has a very high city omission rate • **MOD500** is a better compromise with a relatively high sensitivity and no urban omissions • **Alternatives**: (1) create a low-omission composite of any pixels labeled urban in a high-accuracy global map, or (2) threshold **IMPSA** at very low percent urban levels
Continuous-value urban field	Within-pixel percent of urban cover (as opposed to binary urban / non-urban)	• **IMPSA** has a mean accuracy rank that is higher than the other continuous-value map (HYDE3) • **Alternative:** aggregate MOD500 from 15” to 30” arc-seconds, and create a new global continuous value map (0, 25, 50, 75, and 100 percent urban)

Table 2.10 (previous page) Wise-use recommendations for global urban maps. The table lists those maps that are best-suited to several aspects of map quality.

CHAPTER THREE

URBAN EXPANSION AND THE GLOBAL NETWORK OF PROTECTED AREAS

3.1 INTRODUCTION AND LITERATURE REVIEW

Due in large part to rapid human population growth and associated land cover transitions, the biosphere now faces extinction rates of at least 100 times the pre-human background rate (Pimm et al. 1995; 2006). Providing food, fiber, and shelter to 6.5 billion humans demands a large share of ecosystem productivity (Imhoff et al. 2004), and is positively correlated with measures of both species loss and extinction threat (McKinney 2001; Hoekstra et al. 2005; Davies et al. 2006; Burgess et al. 2007). Urban areas and high density human settlements are of particular concern because they alter local ecosystems along a wide range of axes (McKinney 2002; Grimm et al. 2008), including: species diversity (Pauchard et al. 2006; Luck 2007), nutrient flows (Luck et al. 2001), phenology and net primary production (Zhang et al. 2004), and climate and biogeochemical systems (Calbo et al. 1998; Unger 2001; Oleson et al. 2008).

Establishing protected areas (PAs)[10] has been an important strategy in combating the human-driven loss of both natural habitat and species (Nagendra 2008; Jackson and Gaston 2008). Yet despite a rapid expansion of PAs over the past thirty years to include roughly 12 percent of the Earth's land surface (Lockwood et al. 2006), fundamental unknowns persist regarding our ability to achieve conservation goals in the context of a human population that

[10] A protected area as defined by the Convention on Biological Diversity is "a geographically defined area which is designated or regulated and managed to achieve specific conservation objectives."

will likely increase in size by roughly 50 percent through 2050 (UN 2008a). In this research, we use new global-scale physical and human geography datasets combined with a simple model of urban expansion to address three unknowns at the intersection of conservation and human settlement dynamics: in **Section 3.3.A** we assess how successful PAs have been at preserving the integrity of natural habitat, in **Section 3.3.B** we explore the degree to which the physical inaccessibility of PAs may safeguard them from future human disturbance, and in **Section 3.3.C** we model the potential regional and national-level impacts of urban expansion for PA vulnerability through 2050.

Prior to 2000, the questions at the core of our research could only be addressed at local scales and on a site-specific basis. If one simply asked, "How much of the Earth is designated for conservation, how much is undisturbed natural habitat, and what is their spatial relationship?" data quality, consistency, and availability would not have been up to the task. Those constraints have changed. Over the past five years, the conservation community has undertaken a surge in global mapping, constructing detailed maps of biodiversity hotspots, eco-regions, and species prevalence (Myers et al. 2000; Olson et al. 2001; Ceballos and Ehrlich 2006; Jetz et al. 2007, 2008) and significantly improving the spatial and temporal properties of the World Database of Protected Areas (WDPA; Chape et al. 2005). During the same timeframe, international research groups from both government and academia have employed new remote sensing and geographical information system (GIS) data sets to produce several global-scale maps of land cover (Herold et al. 2008) and urban areas (Potere and Schneider 2007, forthcoming; **Chapter 1**).

The initial line of global scale research opened up by these new datasets is focused on explaining the extent and spatial distribution of PAs and assessing their adequacy relative to conservation goals (Brooks et al. 2004; Rodrigues et al. 2004; Chape et al. 2005; Ceballos and Ehrlich 2006; Loucks et al. 2008; Soutullo et al. 2008). These gap analyses aim to identify misalignments between PAs and the world's threatened species. The underlying

assumption of many of these studies is that PAs are reasonably successful at safeguarding the species within their boundaries. A second line of analysis has recently emerged which questions these assumptions, focusing on characterizing the effectiveness of existing PAs at preserving undisturbed natural habitat (Nagendra 2008). Hoekstra et al. (2005) combined a circa-2000 global land cover map at 1 km spatial resolution (GLC 2000; Bartholome and Belward 2005) with a 2004 version of the WDPA to describe the ratio of protected habitat to human-appropriated habitat by Olson biome (Olson et al. 2001). Hoekstra et al. declared a 'biome crisis' based on the high ratio of habitat conversion to habitat protection across a variety of ecosystems. Lee and Jetz (2008) also use GLC 2000 to assess past incursions onto PAs in order to establish a benchmark for modeling future potential land cover transitions. In **Section 3.3.A**, we take a similar methodological approach to assessing global PA effectiveness, but with a new circa-2005 land cover map based on data from the Moderate Resolution Imaging Spectroradiometer (MODIS) at twice the spatial resolution of GLC 2000 (Friedl et al. 2002, forthcoming) and an updated 2007 version of the WDPA (WDPA 2007).

Joppa et al. (2008) use the GLC 2000 land cover map for a more detailed regional study of habitat loss in tropical moist forests. By focusing on the proximity of human-driven forest losses to park boundaries, Joppa et al. discriminate between regions where PA forests are protected *de facto* by inaccessibility and those protected *de jure* by legal enforcement of park boundaries. We follow this line of inquiry in **Section 3.3.B**, estimating the global and regional fraction of PAs that are effective primarily because of their physical inaccessibility. We measure accessibility using a combination of the MODIS circa-2000 500 m global urban map (Schneider et al. forthcoming) and a global slope and elevation map based on gap-filled Shuttle Radar Topography Mission data (SRTM; Reuter et al. 2007).

The most recent area of global PA analysis examines the vulnerability of PAs to future human encroachment. We estimate that more than 2.07 billion people (nearly one-

third of the human population) live within the biodiversity hotspots of Myers et al. (2000).[11] Others have observed strong regional-scale correlations between human population and species diversity (Balmford et al. 2001; Fjeldsa and Burgess 2008) and boundaries of PAs (Wittemyer et al. 2008). World population is expected to grow by nearly two billion through 2030, and almost all of that growth will likely occur within the cities of the developing world (UN 2007; Montgomery 2008). Combined with declining urban densities in many cities, this rapid global urbanization implies an equally rapid global urban expansion[12] (Angel et al. 2005; Mcdonald et al. 2008). Certainly these demographic forces will impact the world's PAs and undisturbed ecosystems (McGranahan and Satterthwaite 2002), yet thus far few studies have approached this problem in a global, quantitative manner. Early work on urban expansion and conservation was conducted prior to the advent of high-accuracy global urban maps (Morris and Kingston 2002) or was regional in scope (Ricketts and Imhoff 2003). Subsequent analyses have been conducted at extremely coarse spatial resolution (Lee and Jetz 2008), or using first-generation maps of urban areas based on remote sensing data from the 1990s (Mcdonald et al. 2008). In **Section 3.3.C**, we address this gap by exploring the degree to which future urban expansion may create new vulnerabilities in today's PA system with a simple, spatially-explicit, moderate-resolution model of urban growth driven by the most accurate contemporary global map of urban areas (MODIS 500 m; **Chapter 2**) and a set of probabilistic country-level demographic and economic time series derived from the Intergovernmental Panel on Climate Change scenarios (IPCC-SRES; Grubler et al. 2007; IIASA 2007).

The one precedent for this third portion of our analysis is the global urban growth model of Mcdonald et al. (2008). The Mcdonald model provides an important first global look at urban expansion in the context of global conservation efforts, assessing the impact of

[11] Based on our analysis of Conservation International's 34 biodiversity hotspots (Myers et al. 2000), and the 2006 release of the LandScan population map (Bhaduri et al. 2002).

[12] Urbanization refers to increases in the fraction of a population living in urban areas, and urban expansion refers to physical growth in the areal extent of cities.

increases in urbanized areas through 2030 on threatened species and PAs. There are a number of ways in which our model in **Section 3.3.C** represents a significant step forward in this new line of analysis, including: (1) the inclusion of a more accurate, spatially detailed, and current map of circa-2000 urban areas,[13] (2) the use of a model of urban population density that is country-specific and consistent with the circa-2000 global urban map,[14] (3) demographic change parameters driven by a probabilistic set of three separate integrated socio-economic scenarios,[15] (4) an underlying model framework that is at moderate resolution (500 m) and treats each pixel of land as an independent entity, and (5) a time span that is almost 50 percent longer, covering 2000-2050 in decadal increments.[16]

3.2 DATASETS

Because proximity and area are important aspects of our analysis, we began by rasterizing, co-registering, and projecting all datasets into a common Goode's Homolosine pseudo-cylindrical, equal-area, composite projection at 500 m-resolution (equivalent to 15" arc-seconds in a geographic or plate carré projection). For the PA, land cover, elevation, biome, and country-boundary datasets (described below), we checked alignment at all stages of the processing and differences between the original and projected maps were minimal.

---

[13] Our model relies on the MODIS 500 m urban map (Schneider et al. forthcoming), and Mcdonald et al. use the Global Rural Urban Mapping Project (GRUMP) dataset (CIESIN 2004). Compared with GRUMP, MODIS is more accurate (**Chapter 2**), relies on imagery that is six years more recent, and is conducted at twice the spatial resolution. GRUMP is ranked fourth in terms of its agreement with a global collection of high- and medium-resolution validation datasets, and because of the challenges inherent in interpreting the 1994-1995 Nighttime Lights data that underlies GRUMP, its overall extent is 5-10 times larger than any of the other seven global urban maps (Potere and Schneider 2007; **Chapter 1**).

[14] We construct our model of country-level urban population density to be consistent with the MODIS 500 m urban map. By contrast, the three global density constants used in the Mcdonald model (and derived from Angel et al. 2005) result in an estimate of circa-2000 urban land which is 7.23 times smaller than the model's underlying circa-2000 global urban map, GRUMP.

[15] Our model relies on three probabilistic, integrated scenarios for socio-economic change through 2050 (Grubler et al. 2007), while the Mcdonald model selects the median UN scenario through 2030. This difference is important because economic and demographic uncertainties are at least as large as those surrounding urban density.

[16] We rely on a more recent 2007 version of the WDPA, and then rasterize this vector map in order to treat each 500 m resolution pixel as an independent entity. The Mcdonald model operates only at the level of whole parks.

A. Protected area dataset

The 2007 World Database of Protected Areas (WDPA) is a global inventory of more than 108,000 conservation areas that is continually updated by the World Conservation Monitoring Center (WCMC), in partnership with the United Nations Environment Programme (Chape et al. 2005). In the most recent 2007 version, 80 percent of the terrestrial PA in the WDPA inventory is classified using the International Union for Conservation Nature (IUCN) categories or international treaties, and the remaining 20 percent are nationally designated PAs that have not yet been classified (Table 3.1). In Table 3.2, we construct three groups of protection status: (1) *Strict*: areas protected for conservation objectives (IUCN classes I-IV and international PA treaties), (2) *Managed extraction*: areas maintained with a goal of sustainable resource use and extraction (IUCN classes V-VI), and (3) *Unknown*: national PAs not yet classified into an IUCN category.

Although the WDPA has made significant strides in improving the spatial qualities of their park inventory, there are significant overlap areas that have competing IUCN labels. An improved version of the WDPA is in the final stages of development, but was not released in time to be included in this analysis (anticipated release December 2008, pers. corr.). For this study, we converted all PAs into a 500 m-resolution global raster, removed water pixels by applying the land-water mask from the MODIS 500 m land cover map, and resolved overlaps between PAs at the pixel level by prioritizing national sites over international treaty sites. We did not consider undesignated 'candidate' PA sites (714,000 sq. km of the WDPA) or Antarctica (3,500 sq. km of PA). Using this framework, the total coverage of terrestrial PAs within the 2007 WDPA is 16.6 million sq. km (12.5 percent of all land outside Antarctica).[17] Figure 3.1 is a global-scale view of the full WDPA.

[17] Our total differs from Chape et al.'s (2005) estimate of 18.4 million sq. km because we account for overlapping designations, apply a consistent land-water mask, use the updated 2007 WDPA, and exclude Antarctica. There may also be some small bias due to our rasterizing of the WDPA polygons, but this difference was estimated at less than 0.06 percent of the full WDPA.

B. Global land cover map

There are two major groups which provide satellite-derived, sub-1 km (moderate) resolution, circa-2005 estimates of global land cover: NASA's MODIS land cover project (Friedl et al. 2002, forthcoming), and the European Space Agency's (ESA) GlobCover map (ESA 2008; Arino et al. 2007). The MODIS 500 m land cover map is constructed using a supervised classification of one full year of eight-day-composited multispectral data from the MODIS sensors aboard NASA's *Aqua* and *Terra* satellites, while the GlobCover map relies on similar data from the Medium Resolution Imaging Spectrometer (MERIS) sensor aboard ESA's *ENVISAT* satellite. Although GlobCover is a higher resolution map (309 m versus 463 m), we choose to use the MODIS land cover map because: (1) the MODIS dataset is part of a six-year lineage of global maps that have been constructed using a consistent set of nadir-adjusted bi-directional reflectance MODIS imagery and supervised classification algorithms (Friedl et al. 2002; Schaaf et al. 2002) while GlobCover is the first MERIS-based effort; (2) when there are cloud cover problems for a given pixel, the MODIS map simply defaults to the most recent MODIS land cover map (produced annually), whereas when 2005 data is missing for GlobCover, it defaults to thirteen ancillary 'reference' maps of varying resolution, accuracy, and currency (ESA 2008);[18] and (3) the MODIS 500 m land cover map is aligned with the circa-2001 MODIS 500 m urban map (Schneider et al. forthcoming), which is critical to our modeling efforts in **Section 3.3.C** and is of higher accuracy than the urban areas in GlobCover or any other global urban map (**Chapter 2**).

The MODIS land cover map employs the sixteen-category International Geosphere-Biosphere Programme (IGBP) legend (plus water). For our analysis, we aggregate this legend into the five land cover groups of Table 3.2: (1) forest, (2) vegetated non-forest, (3)

[18] Globally, 2.6 percent of GlobCover is fill (3.82 million sq. km), but because cloud cover in the tropics accounts for many of these missing values, 25.9 percent of these fill values occur on protected areas within the WDPA. This 989,000 sq. km of filled PA represents 6.0 percent of the full WDPA.

barren, (4) urban areas, and (5) croplands.[19] The vegetated non-forest group is composed primarily of shrublands (33 percent), savannas (31 percent), and grasslands (22 percent). In order to subdivide the MODIS land cover map into individual countries, territories and regions, we modified the international boundaries map associated with LandScan 2006, a global 30" arc-second (~ 1 km) resolution map of human population density (Bhaduri et al. 2002; **Section 1.2.A**). We expanded the LandScan land-water mask to coincide with MODIS, assigned standard United Nations International Organization for Standardization (ISO) codes for 223 countries and territories (UN 2008b), and employed a modified version of the World Bank's major area scheme to define six world regions (Figure 3.1). We used forested PAs that have park boundaries that are visible in the MODIS land cover map to verify alignment of the WDPA and MODIS maps (three examples in Figure 3.2).

C. Physical geography data

To delineate biomes, we employ the sixteen-biome system of Olson et al. (2001). During processing and alignment checks, we corrected inconsistencies in the land-water mask for the vector version of this map by assigning uncategorized pixels to their nearest Olson biome-categorized neighbor. We then aggregated the sixteen Olson biomes into the four biome groups of Table 3.2: (1) tropical forest, (2) temperate forest, (3) vegetated non-forest, and (4) barren.[20] As in the case of the vegetated non-forest land cover group, the vegetated non-forest biome group is composed primarily of shrubland, savanna and grassland. For topographic data, we rely on a new 15" arc-second (500 m) global mosaic of void-filled 90 m Shuttle Radar Topography Mission elevation data (SRTM; Reuter et al. 2007). We assigned zero elevation to coastal areas outside the SRTM land-water boundary but within the MODIS 500 land-water mask. The SRTM mission did not extend above 60

[19] Although the IGBP 'barren or sparsely vegetated' class contains some vegetation, we include it in the 'barren' land cover group based on the mean characteristics of the class.

[20] Although, the Olson 'tundra' biome contains some sparse vegetation, we include it in the 'barren' biome group because of the mean characteristics of that biome.

degrees north, so for these areas our elevation mosaic relies on GTOPO30, a 30" arc-second topography dataset produced by the United States Geological Survey (Gesch and Larson 1996; USGS 1996). We generated a new global percent slope dataset at 500 m-resolution based on a Goode's Homolosine reprojection of the combined SRTM-GTOPO30 mosaic.

3.3 METHODS AND RESULTS

The analysis proceeds in three stages: we begin with an examination of the land cover composition of the WDPA using the new MODIS 500 m land cover map (**Section 3.3.A**), next we estimate the fraction of PAs which are protected from future encroachment by their physical inaccessibility (**Section 3.3.B**), and finally, we employ a simple model of urban expansion to explore several scenarios for potential urban-driven increases in vulnerable PA through 2050 (**Section 3.3.C**).

A. The land cover composition of protected areas

By overlaying the MODIS 500 m land cover map atop the WDPA (Figure 3.2), it is possible to characterize the land cover composition of the PAs. In Figures 3.3 and 3.4 and Table 3.3, we report PA land cover by protection status group, biome group, and world region. The bars on the left side of Figure 3.3 portray the land cover composition of PAs according to the four-group land cover scheme from Table 3.2, and the bars to the left of Figure 3.4 do the same by protection status (strict, managed extraction, or unknown). Overall, the vegetated non-forest land cover group (column totals in Table 3.3) makes up the largest share of PAs globally (7.3 million sq. km, or 44 percent of all PA), followed by forests (5.8 million sq. km, or 35 percent of all PA), and barren areas (3.1 million sq. km, or 19 percent of all PA). Undoubtedly much of the barren land cover group is protected *de facto* from future encroachment because of its poor suitability for human use, yet it constitutes the biome with the largest total area of strict-level protection, including a massive region of

strict protection on the Greenland icecap. Europe is the region with the largest amount of barren PA in absolute terms (940,000 sq. km) because Greenland is home to the world's largest glaciated PA. The two Asian regions have PAs containing 45 percent of the world's barren PA (gray bars, Figure 3.2) largely due to a huge desert PA in the Rub' al Khali of Saudi Arabia (541,000 sq. km) and similar-sized PAs in Tibet (780,000 sq. km). In contrast to the barren land cover group, the forest land cover group includes a very high fraction of the world's species diversity and biological productivity, making this group essential for meeting conservation objectives. Strict PAs contain the most forested land in absolute terms (2.3 million sq. km, or 40 percent of all forested PA), as does the tropical forest biome (3.7 million sq. km, or 63 percent of all forested PA) and the Latin America and Caribbean region (2.9 million sq. km, or 51 percent of all PA forest).

The bars on the right side of Figure 3.2 describe PA land that is mapped as human-dominated by MODIS land cover (right-side bars are in thousands of sq. km). These yellow cropland and red urban segments are a magnification of the black 'urban and crops' segments to the left of Figure 3.3. A 500 m-resolution pixel that is mapped as both protected by the WDPA and urban or cropland by MODIS can be explained in one of four ways: (1) it is a mislabeled pixel in the MODIS land cover map and is not actually human-dominated, (2) it is a mislabeled pixel in the WDPA and is not actually protected, (3) it is an *inholding*, an urban area or cropland pixel that existed prior to the establishment of the PA (a common occurrence in IUCN category V or VI lands), or (4) it is an *incursion,* a new urban or cropland pixel that extended into the PA subsequent to its establishment. Ultimately, only high-resolution imagery or a site visit can resolve the status of each of these overlapping protection-disturbance areas. Until a significant sample of these overlaps can be assessed in detail, our findings should be regarded as a preliminary indication of the overall pattern of both inholdings and incursions on PA lands. For brevity we will refer to

these potential inholdings and incursions simply as 'incursions' or 'overlaps' from here forward.

Overall, urban area and cropland represent a small fraction of the total PA system (2.3 percent), but a very large amount of land in absolute terms; the 384,000 sq. km of inholdings and incursions in the PA system is 184 times the size of Yosemite National Park in the United States (US). The managed extraction protection class has the largest share of incursions and inholdings (167,000 sq. km or 43 percent of all incursions), as does the temperate forest biome group (188,000 sq. km or 49 percent), and Europe (169,000 sq. km or 44 percent). The exceptionally large share of urban and cropland PA in Europe (5.5 percent of all European PAs) may be explained by the large number of European IUCN class V and VI areas, parks that were likely established with significant inholdings due in part to Europe's relatively long history of agriculture-based human settlement. As of 2005, the PAs within the biologically important Latin America and Caribbean and Africa regions had the least amount of incursions or inholdings in both absolute and relative terms (0.5 percent and 1.0 percent of within-region PAs, respectively).

Figure and Table 3.4 shift the focus exclusively onto urban incursions or inholdings. Although urban areas represent a much smaller fraction of incursions and inholdings than croplands, they are much more permanently lost in terms of conservation objectives. From the bars to the right of Figure 3.4, it is clear that the overall extent of urban incursions and inholdings follows that of croplands: most urban PA is in the temperate forest biome group (12,270 sq. km) and Europe (8,640 sq. km), and most of these urban overlaps occur within the managed extraction protection group (12,080 sq. km). By comparison, there are very few urban incursions or inholdings in Africa or South-central and West Asia (690 and 900 sq. km, respectively). Altogether, developing countries account for just 33 percent of the 19,600 sq. km of observed urban PA incursions and inholdings circa-2000.

In assigning MODIS land cover classes to our five land cover groups (Table 3.2), we chose not to associate the 8.5 million sq. km of 'cropland and natural vegetation mosaic' to a disturbance category, instead assigning it to the vegetated non-forest group. In making this conservative assignment, our methods depart from those of both Hoekstra et al. (2005) and Joppa et al. (2008). Our decision reflects a relatively low degree of confidence in mixed mosaic classes because of the difficulty of distinguishing small-scale agriculture and pasture from undisturbed heterogeneous grassland, shrub, or savanna landscapes. Because there has not yet been a global-scale map accuracy assessment aimed at agriculture, it is difficult to make this decision with a high degree of confidence.[21] If, instead, we include the MODIS mosaic pixels in the disturbance category, total PA incursions increase by 140 percent (an additional 537,000 sq. km of incursions; 35 percent *strict* PAs, 41 percent *managed extraction*, and 24 percent *unknown*). By biome, these additional mosaic incursions align well with urban and cropland incursions: the temperate forest biome group accounts for the largest share of mosaic incursions (206,000 sq. km or 38 percent). By region, however, mosaic incursions are far more prevalent in Latin America (203,000 sq. km or 38 percent) and Africa (166,000 sq. km or 31 percent), increasing the total incursions in those biologically critical regions by 1,015 and 519 percent, respectively. Our estimate of 384,000 sq. km of PA inholdings and incursions is therefore probably a conservative lower limit for human disturbance on protected areas, particularly for Africa and Latin America.

B. Geographic 'security' of protected areas

In this section, we turn our attention to those portions of PAs which are not already disturbed by human action, but which still face risks due to their physical geography. Within

[21] Although there are no agriculture-specific global map accuracy assessments, Ramankutty et al. (2008) offer an independent estimate of the total extent of the world's croplands: 12.2 – 17.1 million sq. km (90 percent confidence bounds). If we assume that the MODIS 500 mosaic class is 50 percent cropland and 50 percent undisturbed, including the cropland portion of this class would increase the MODIS global cropland estimate to 16.4 million sq. km—well within the Ramankutty et al. total.

PAs, Joppa et al. (2008) differentiate between *de facto* protection resulting from the physical isolation of a PA and *de jure* protection resulting from the legal status of the PA. Because they are unavoidably linked to the effectiveness of park management, measures of *de jure* protection are complex, poorly understood, and not currently gathered with sufficient detail or consistency for global analysis. By contrast, the degree of *de facto* protection afforded to a PA can be conceived of as a function of two geographic variables: access and productivity. Access can be characterized by physical attributes such as slope and elevation, or the distance separating a PA from transportation infrastructure and moderate-density human settlements. Productivity refers to the potential of a PA to function as a useful human-dominated landscape and can be measured directly with metrics such as net primary productivity and subsurface mineral potential, or indirectly as a function of water availability, soil composition, and climate. In our analysis, we construct a measure of *de facto* protection status by focusing on access. We use mean slope and distance to the nearest urban area as proxies for PA accessibility. Our assumption is that areas of high slope, high elevation, and extreme distance from urban areas are less accessible to human appropriation that those on relatively low slopes, at low elevations, and in close proximity to the built environment. Of course, the distance separating PAs and the nearest urban area is also an important factor in its own right because urban proximate PAs are subject to changes in ecosystem function (**Section 3.1**), potential increases in non-land cover-related extractive processes like the bush meat trade (East et al. 2005), and elevated costs of conservation management (Bruner et al. 2004). Examining the global urban-proximity of PA is a continuation of earlier work (Potere and Schneider 2008); similar and independent proximity analysis is also present in McDonald et al. (2008).

The challenge is to define formal relationships between the probability of PA encroachment and slope and urban proximity. Ideally we would model risk of encroachment using a time series of land cover change and PAs. In our analysis, we have an urban area

map from just one period (circa-2001), so we pursue a more modest goal of constructing a pair of simple linear thresholds that constrain the degree to which a PA pixel is susceptible to human appropriation based on accessibility. We begin by examining the global distribution of slope for urban areas and croplands (the two land cover classes considered human disturbance in **Section 3.3.A**). In the examples of Figure 3.5, the constraining role played by slope is clear for (a) Salt Lake City, USA, (b) Tehran, Iran, and (c) central and northern Italy. In all of these locations, urban borders extend right to the edge of the moderate-slope (6-10 percent) and high-slope (11+ percent) mountainous areas. The black dotted and dashed curves in Figures 3.6a-c generalize these observations using the full global slope and urban maps, demonstrating the importance of slope in limiting urban and agricultural expansion. Globally, 96.7 percent of all urban areas and 95.3 percent of all croplands lie on land that is at less than 6 percent slope. We also examined elevation as a factor in limiting urban growth. Although there was a strong association between urban areas and low-lying regions, the cumulative frequency of elevation dropped off much less sharply than for slope, and elevation did not significantly alter the global urban suitability surface based on slope alone. Based on these results, we consider slopes of 0-5 percent suitable for urban or agricultural expansion and slopes greater than 6 percent unsuitable.[22]

To determine an appropriate vulnerability threshold for the proximity of PA to urban areas, we examined the global distribution of the distance between croplands and the nearest urban area (dotted lines, Figures 3.6d-f). In defining urban areas, we only considered those contiguous patches of urban land that were four sq. km or larger, in this way ignoring small and isolated pieces of built environment that are of little ecological importance and far less likely to form the nucleus of future urban expansion.[23] Although the

[22] We also constructed country-level thresholds for slope suitability. Although there was considerable heterogeneity in the country-level 95-percent slope quantiles, the overall distribution of suitability based on these thresholds deviated only slightly from those derived from simple global thresholds.

[23] When determining patch-size, we defined patch membership as any pixel within two km of its nearest urban neighbor. Our threshold of four sq. km is somewhat less conservative than the 10 sq.

rise in cumulative frequency is less sharp than that of slope (Figure 3.6a-c), there is a strong spatial relationship between urban areas and croplands: much of the world's croplands lie in close proximity to urban areas. This is expected given the important role of agricultural areas in providing the food, fiber, and wood surpluses that make urban settlement possible. We choose 30 km as a proximity to urban areas below which a PA is exposed to increased chances of encroachment; 82.9 percent of all croplands are within this range of urban areas. We acknowledge that both our slope and urban proximity thresholds are somewhat arbitrary, but fortunately the overall rank order of slope and urban proximity remains fairly constant across a wide range of slope and proximity thresholds (order of the colored group/region lines in Figure 3.6). This is important because the primary purpose of this analysis is to compare the *relative* degrees of *de facto* protection across protection groups, biome groups, and world regions. These relative relationships are robust to small perturbations of the slope and urban proximity thresholds.

To construct the colored group/region curves in Figure 3.6, we included only non-urban PA pixels. The closer these curves are to the black dotted urban and cropland curves, the more susceptible that PA group/region is to human appropriation. Overall, it is clear that a much higher proportion of PAs are suitable for encroachment based on slope than based on distance to urban areas: note the scale change for the x-axes of (a) through (c) versus (d) through (f). For slope (upper three plots), although all regions/groups have at least 60 percent exposure to flat slopes (0-5 percent), the unknown protection status PAs, barren PAs, and African PAs have the lowest slopes (and are therefore more exposed to potential encroachment). Temperate forest PAs and the PAs of Southeast and East Asia and Other Developed Countries (ODC) have the highest slope distribution of any group/region and are therefore less exposed to potential encroachment. For urban proximity (lower three

km threshold of Zhang et al. (2004) in their analysis of the impact of urban proximity on local vegetation phenology, and is based in part in the fact that our urban map is at twice the spatial resolution of that used by Zhang et al.

plots), the managed extraction PAs are more urban-proximate (more than 20 percent of these PAs are within 30 km or urban areas) and are therefore more exposed to potential encroachment than other protection groups, likewise for the temperate forest biome and the developed countries of Europe and ODC. Europe is the only region in the distance plot (Figure 3.6f) to have a concave down curve; all other curves are convex. This may be explained by the large amount of cropland in European PAs (Figure 3.3). It is not possible to directly compare these proximity findings with those of Mcdonald et al. (2008) because that analysis relied on whole park-level statistics as opposed to our pixel-level observations. We chose not to report mean and median slope or urban proximity statistics because the distributions are highly skewed, and most of the focus is more properly placed on PAs with slopes from 0-20 percent and within 60 km of urban areas.

Because both slope and urban proximity are important proxies for access (and *de facto* protection status), it is important to estimate the union of all undisturbed (non-cropland and non-urban) PA pixels that are jointly safe from encroachment (slope steeper than five percent and more than 30 km from urban area). Likewise, it is useful to estimate the area of undisturbed PAs that are jointly exposed to the threat of encroachment (slope of 0-5 percent and within 30 km of an urban area). Globally, 2.7 million sq. km of undisturbed PA (17 percent) is safe from encroachment (*de facto* PA) based on our slope and urban proximity thresholds, while 1.3 million sq. km (8 percent) is exposed to encroachment (*de jure* PA) based on the same factors (Table 3.5). There are some interesting paradoxes in the relative amounts of safe and exposed PA by group/region: (1) by virtue of its high slope, the temperate forest biome group has the largest area of 'safe' PA (1 million sq. km or 27 percent of all undisturbed temperate forest PA), but because of its urban proximity it also has the most exposed PA (630,000 sq. km or 17 percent), and (2) Africa has both the smallest fraction of safe PAs (5.6 percent) and the smallest fraction of exposed PAs (4.6 percent). Overall, the ODC and Southeast and East Asia regions have the largest area of safe PAs in

both absolute and relative terms (25.8 and 34.0 percent of all with-region PA, respectively), and ODC and Europe have the most exposed protected area (11.1 and 12.4 percent, respectively). Globally, tropical forest PAs have 39 percent less 'safe' area than temperate forests, yet these tropical forests harbor a disproportionately large fraction of the Earth's biodiversity.

C. A simple model of urban expansion and protected areas

One of the principal risks faced by urban-proximate protected areas is the potential for future urban incursion. In this section, we introduce a methodology for converting existing country-level models of urban demographic change into estimates of future urban extent on a 500 m-resolution global grid (**Section 3.3.C.1**). This simple model allows us to generate urban expansion envelopes on decadal time scales from 2000-2050. We intersect these envelopes with the WDPA, and explore the implications of urban expansion for the world's PAs at global, regional, and national scales (**Section 3.3.C.2**). Finally, in **Section 3.3.C.3** we introduce two scenarios for changes in the density of cities to describe the implications of compact versus extensive[24] development for the world's protected areas.

(1) Model description. For country-level estimates of urban population and economic output through 2050, we turn to the International Institute for Applied Systems Analysis (IIASA) national scenarios of demographic change (Grubler et al. 2007). The IIASA research is one of four institutional-level downscaling projects (Gaffin et al. 2004; Bengtsson et al. 2007; van Vuuren et al. 2007) aimed at increasing the spatial resolution of the regional Intergovernmental Panel on Climate Change-Special Report on Emissions Scenarios (IPCC-SRES; Nakicenovic and Swart 2000; Van Vuuren et al. 2006). We choose Grubler et al. because they use a probabilistic population projection methodology (Lutz 1996; Lutz et al.

[24] The term 'sprawl' is sometimes used to describe low-density urban settlement, but we prefer 'extensive' because of its less ambiguous nature.

2001), they model urban and rural populations separately, and they update all of the SRES scenarios based on the results of circa-2000 censuses. Grubler et al. explore three of the four SRES demographic and economic storylines (Table 3.6 and Figures 3.7a-b). Their modifications to these scenarios are discussed more completely in Riahi et al. (2007). Collectively, these three scenarios are designed to encompass a wide range of alternate futures regarding demographic transition and changes in the world economy. Because of differences in bulk changes to global population, we label the A2r scenario as 'high-growth,' the B2 scenario as 'medium-growth,' and the B1 scenario as 'low-growth.' Although these labels apply well to the total global population estimates (Figure 3.7), over certain time intervals and in certain regions there are deviations from these patterns.

In order to estimate urban expansion, we begin by assuming that the MODIS 500 m-resolution urban map (Schneider et al. forthcoming), which was produced with circa-2001 imagery, represents the best estimate of circa-2000 urban land cover (**Chapter 2**).[25] Like all of the SRES downscaling efforts, our algorithm makes a number of simplifying assumptions with respect to urban population density and the spatial form of urban expansion. Figure 3.8 is a schematic of the four main steps in our model. In the first step, we estimate country-level urban population density using each country's urban population from the IIASA model and the extent of urban area within each country from the MODIS urban map.[26] The number of urban residents in a given country divided by the urban area within that country yields the mean urban population density for the country circa-2000. Although there are indications that the majority of cities in both developing and developed countries are experiencing a decline in urban density (Angel et al. 2005; Schneider and

[25] The four SRES downscaling projects also conduct a second stage of sub-national modeling, downscaling from their country-level results to global grids of between 30 km and 60 km in spatial resolution (measured at the equator). These grids are at too coarse a resolution for our analysis.

[26] There are 38 small island nations and territories that are not tracked in the IIASA method, but are mapped by MODIS and described in our LandScan-based international boundary file. For these countries, we rely on UN median urban population projections through 2030. Beyond 2030, we make the conservative assumption that these populations levels remain constant. The total urban area of these islands is only 0.06 percent of the MODIS 500 m urban map.

Woodcock 2008; Mcdonald et al. 2008), we make the strong simplifying assumption that the country-level urban densities estimated for 2000 remain constant over time. Our model is conservative; even small decreases in mean urban density would significantly increase total global urban expansion. Later, in **Section 3.3.C.3**, we relax this assumption of constant densities and report model projections based on compact and extensive density trajectories.

In step two (Figure 3.8), for each decade of the IIASA projection, we estimate the decadal increase (or decrease) in each country's urban residents. We then divide this increment by the country's population density from step one, yielding an estimate of the increase (or decrease) in urban area during that decade. For the small number of countries and decades where the urban population decreases (Europe, Japan, etc.), we do not model urban area contraction, instead we keep the amount of urban land constant at the level of the previous decade (zero increase). This is a reasonable assumption considering that urban land is rarely converted back into natural vegetation, even when that urban land becomes uninhabited.

Steps one and two produce a decadal estimate of the amount of new urban land within each country. In step three, we construct an estimate of urban suitability in order to spatially allocate this new urban land. We begin by masking out water pixels based on the MODIS 500 m-resolution land cover map, and eliminating high slope pixels (> 6 percent) based on our global topographic map (Reuter et al. 2007; **Section 3.3.B**). Next, we assume that urban expansion is most likely to occur near existing urban patches greater than four sq. km in size. We rank all of the low-slope, non-water, non-urban pixels within each country based on their urban proximity. This proximity is a rough proxy for urban suitability. In the fourth and final step, we select enough urban-proximate pixels in the circa-2000 map to meet the decadal urban expansion requirements from step two (starting with the most urban-proximate pixels and working outwards). The same process is repeated for each

country that has urban area in the MODIS map (200 out of 223 countries and territories), forming an estimate of urban land cover circa-2010. This map is then substituted for the circa-2000 map in Figure 3.8, and steps two through four are iterated for the remaining decadal intervals. There is a considerable literature surrounding the modeling of urban expansion at the grid-level, and most of these models rely on some form of proximity to the built environment (Herold et al. 2003). Our decision to use an untransformed measure of urban proximity (as opposed to an integrative, gravity-based measure) as our primary predictor of urban growth considerably improves the efficiency of our processes, allowing us to estimate a series of global-scale models. This overall approach aligns well with that of Mcdonald et al. (2008).

(2) Model results. Figure 3.9 (left axis, upper three curves) plots cumulative urban expansion from 2000-2050 under the three IIASA-SRES storylines. All three scenarios are essentially in agreement with respect to total urban expansion through 2020. Urban areas cross the one million sq. km point by 2030, reaching 1.03 million sq. km, an increase of 57 percent over the circa-2000 MODIS urban extent of 657,000 sq. km.[27] By 2050, the medium-growth storyline projects a global urban extent of 1.22 million sq. km (1.16 – 1.38 million sq. km for low- and high-growth storylines, respectively). Under the high-growth storyline, the 2050 urban extent is more than twice the circa-2000 urban extent observed using MODIS, representing a doubling of the world's urban land area.

When we turn to the question of PA incursions, it is important to stress that an underlying assumption of our model is that PAs completely lack enforcement. In allocating urban expansion (step four, Figure 3.8), the protection status of a pixel has no impact on whether that pixel is reclassified as new urban land. Our estimates of PA 'losses' should be more conservatively thought of as proxies for the vulnerability of the PA network to human-

[27] Because Mcdonald et al. (2008) do not report areal estimates of urban expansion or PA incursions, it is not possible to compare our model outputs.

induced pressures stemming from urban expansion. The overall pattern for changes in PA vulnerability follows a similar trajectory as that of overall urban expansion, albeit on a much smaller scale (Figure 3.9, right axis). The three PA vulnerability storylines are in rough agreement through 2020, indicating growth in vulnerable PA of 5,600 sq. km over the 19,600 sq. km of PA incursions observed circa-2000 (medium-growth scenario). By 2050, the three storylines are more distinct with the low- and high-growth scenarios separated by 34 percent and the medium-growth storyline projecting new vulnerabilities of 12,000 sq. km. The high-growth 2050 estimate of 15,700 sq. km represents an increase of 80 percent over the total urban incursions and inholdings observed in **Section 3.3.A** (Table 3.4). Overall, global urban extent increases more sharply than PA losses; the fraction of urban areas that intersect the WDPA declines over the period 2000-2050 from 2.98 to 2.58 percent.

There is a high degree of geographic variation in the distribution of these protected area losses. Figures 3.10a-b describe the cumulative regional and global total of vulnerable PAs for 2000-2050 based on the medium-growth storyline, and Table 3.7 does the same based on all three IIASA-SRES urban population storylines. The vulnerability bar plot segments in Figure 3.10 are colored according to protection group (left axes), and the black lines represent the total urban population according to the IIASA medium-growth model (right axes). Europe, despite having the largest share of 'safe' PAs (**Section 3.3.B**), faces the least amount of potential PA incursions due in large part to its anticipated population decline (800 sq. km through 2050, an increase over observed circa-2000 incursions of just 9 percent). Conversely, despite the relative safety of PAs in Africa, that region faces the largest increase in vulnerable PAs in absolute and relative terms (2,700 sq. km, an increase of 392 percent over observed year 2000 incursions and inholdings). This is driven primarily by a projected African urban population increase of 326 percent through 2050. The equally important (and largely tropical) PAs of Latin America and the Caribbean and Asia (east and

west combined) also face a sharp increase in potential incursions through 2050: 2,600 and 3,800 sq. km, respectively. Primarily because of their rapid population growth rates, the developing regions of Africa, Latin America and the Caribbean, and Asia account for 76 percent of all of the new vulnerable PA through 2050, and virtually all of the vulnerable PA on strictly protected land (red bars, Figure 3.10). This is a striking reversal when one considers that these same developing countries account for only 33 percent of all the observed urban PA incursions in **Section 3.3.A**.

Table 3.8 lists the 15 countries with the largest PA vulnerabilities due to potential urban expansion from 2000-2050, under the medium-growth scenario. Collectively, these 15 countries represent 80.5 percent of all the projected increases in vulnerable PA. It is not surprising that several of the world's most populous countries are on top of this list, including the US, China, and Brazil (ranks 1-3 in Table 3.8). Altogether, only three countries from Europe and ODC are on the list, with the rest constituting countries of the developing world. Some of the percentage increases in urban PAs are striking; Uganda and Ethiopia are both predicted to experience percent increases in vulnerable PA of 1,712 and 762 percent through 2050, respectively. Also tracked is the fraction of these vulnerable areas that are in strict PAs; these vary widely, with countries such as Saudi Arabia, China, and Nigeria experiencing essentially no strict PA losses, and India and Uzbekistan experiencing all of their potential incursions in strictly managed PA. These strict percentage statistics are complicated by the considerable international variation in the composition of PAs by protection level. Table 3.8 also includes two factors that are potential drivers of the 2000-2050 vulnerability estimates: the fraction of a country's PA considered 'exposed' based on accessibility (**Section 3.3.B)**, and the percentage increase in urban area (or population) from 2000. What is clear from this small sample is that the demographic forces behind urbanization are linked far more clearly to potential PA incursions than to measures of accessibility.

In Figure 3.11, we plot the percent increase in vulnerable PA through 2050 and the absolute extent of all potential PA incursions over the same time period for each of the 121 countries that experienced a potential contraction of PA under the medium-growth scenario. The countries above the dotted horizontal line account for 80.5 percent of all vulnerable PA (Table 3.8). Overall, there is a clear relationship between both the magnitude of potential PA incursions through 2050 and the amount of observed incursions in 2000; this is expected considering the basic structure of our urban proximity-based growth model. The distribution of developing countries (red dots) and developed countries (black dots) are clearly separable. In relative terms (x-axis), most European and ODC countries are predicted to experience percent increases in urban PA incursions of less than 100 percent (the developed country median is 20 percent), while most developing countries have increases in vulnerable PA of more than 100 percent (the developing country median is 161 percent). In absolute terms (y-axis), the two groups of countries and territories are less easily separable, in part because of very large potential incursions in the United States, Germany, and France.

(3) Density scenarios. There is a negative relationship between per capita gross domestic product in the year 2000 and the population density estimates that are part of our urban model (Figure 3.12a). The direction of this relationship coincides well with the assumption that more affluent and mobile societies currently have a preference for more spacious, sprawling urban environments. In addition to tracking urban and rural populations, the IIASA-SRES scenarios also model urban and rural gross domestic product (1990-2100), presenting a potential opportunity to model GDP-linked changes to urban density over time. Because we are limited to one time-period for our global MODIS urban map, it is not yet possible to exploit this GDP-density relationship in a temporal model. To estimate such a model, we would need a new version of the circa-2001 MODIS map based on

another year of data, or a global urban change-map with characteristics similar to the city-scale maps of Schneider and Woodcock (2008).

Given that we do not have adequate data to model density changes using economic indicators, we chose to examine past city-scale density changes in two global urbanization studies (Angel et al. 2005; Schneider and Woodcock 2008). Figure 3.12b is a histogram of the percent change in density from 1990 to 2000 for 113 cities in the combined sample. The density changes are normally distributed with a mean annual density change of -1.48 percent (standard deviation of 1.81 percent). The area-weighted density change from 1990 to 2000 is -1.02 percent per year. Based on these findings we estimate two new urban modules, an *extensive urban expansion* module where country-level urban population densities decline by 1.0 percent per year as observed in the Angel et al. (2005) and Schneider and Woodcock (2008) data, and an optimistic *compact urban expansion* module where country-level urban densities increase by 1.0 percent per year. In keeping with the range of densities observed in our country-level estimates, we restrict densities to between 1,000 and 50,000 people per sq. km.[28] These density change modules are roughly equivalent in overall magnitude to those of Mcdonald et al. (2008), but our decadal density changes are phased in gradually and are conducted at the country-level.[29]

Using the medium-growth demographic scenario, through 2050 the extensive and compact urban expansion modules generate 683,000 and 476,200 sq. km of new urban area, respectively. Likewise, these two disparate density modules generate 14,400 and 10,100 sq. km of new vulnerable PA, respectively. For both urban expansion and potential PA incursions, the differences between model runs with the compact versus the extensive modules are comparable to the differences between the low-growth and high-growth

[28] We hold density constant from 2000-2050 for those cities that begin in 2000 with densities outside the range of 1-50 thousand people / sq. km. When the 1.0 percent density decline/increase places a new density outside this range, we hold the new density constant at the limit of the range.

[29] Mcdonald et al. (2008) alter their three global density constants by 50 percent, while our ± 1 percent annual changes accumulate to a roughly 60 percent change in density through 2050.

demographic scenarios. When the low-growth demographic scenario is combined with the compact urban expansion module, total urban expansion through 2050 is just 387,400 sq. km with 9,000 sq. km of potential urban incursions. In contrast, when the high-growth demographic scenario is combined with the extensive urban expansion module, total urban expansion through 2050 is 946,400 sq. km with 20,800 sq. km of potential new PA incursions. Certainly, these two combinations of demographic and urban planning decisions represent very distinct futures; the more optimistic dense and low-growth scenarios generate roughly one-third of the urban growth and one-half the increase in vulnerable PA.

3.4 CONCLUSIONS AND FUTURE DIRECTIONS

Based on an analysis of a new satellite-derived moderate-resolution (500 m) land cover map and the 2007 World Database of Protected Areas, we estimate the land cover composition of protected areas (PAs) globally (**Section 3.3.A**). In total, 384,000 sq. km of PA is classified as cropland or urban areas; these incursions or inholdings are 2.3 percent of the full PA system. By biome group, temperate and tropical forests represent the largest share of these overlaps (60 percent). By region, European PAs have the most incursions and inholdings (5.5 percent of all within-region PA), and the regions of Latin America and the Caribbean and Africa have the least (0.5 and 1.0 percent, respectively). Focusing on urbanized PA, which is the most permanent loss of natural habitat, developing countries account for 33 percent of the 19,600 sq. km total. In order to address the resilience of PAs to future incursions, we use global topographic and land cover data to estimate the portion of PAs that are protected *de facto*, through their inaccessibility (**Section 3.3.B**). We find that 82.9 percent of all croplands are within 30 km of urban areas, and flat slopes of 0-5 percent encompass 96.7 percent of all urban areas and 95.3 percent of all croplands. Based on these observations, we estimate that 2.7 million sq. km of PA (17 percent) is relatively safe from

urban encroachment due to its poor accessibility, while 1.3 million sq. km (8 percent) is exposed to potential urban or agricultural encroachment.

To place some bounds on the extent and spatial distribution of potential future incursions onto this 1.3 million sq. km of exposed PA, we construct a simple global model for urban expansion. By employing three distinct demographic scenarios and three modules for changes in urban density, we predict a near doubling of urban areas through 2050, from 0.66 to 1.22 million sq. km (**Section 3.3.C**). Assuming that PAs have no ability to deter urban encroachment, this massive urban expansion would directly consume an additional 12,000 sq. km of PA, a 61 percent increase over observed circa-2000 urban incursions. Despite the relative inaccessibility of PAs in Africa, by 2050 that region faces the largest increase in vulnerable PA in absolute and relative terms (2,700 sq. km, an increase of 392 percent over circa-2000 extents). Overall, the developing regions of Africa, Latin America and the Caribbean, and Asia will account for 76 percent of all potential urban PA incursions through 2050, and virtually all of the potential incursions on strictly protected land. This is a significant reversal given that the same developing countries account for only 33 percent of all the urban incursions on PA observed in **Section 3.3.A**. Finally, by combining high-growth demographic scenarios with sprawling urban density modules, we predict global increases in vulnerable PA twice the size of those predictions derived from low-growth and compact alternative futures.

In each of the three sections of this analysis, there are significant opportunities for further study. In **Section 3.3.A**, we highlighted the fact that those pixels that are labeled as croplands or urban areas in MOD500 and labeled as PA in the WDPA may in fact be attributed to map errors. Only a field survey or recent high-resolution imagery can resolve this uncertainty. A web-based Google Earth land cover validation tool used in **Chapter 2** and **Appendix A** could be helpful in reducing this uncertainty. In applying that tool to assess more than 10,000 validation sites, expert analysts required only 11 seconds per site to

detect the presence or absence of built environment. The concept would be to build a stratified random sample of the potential incursions and inholdings from **Section 3.3.A**, and to 'visit' those sites using the high-resolution circa-2005 imagery of Google Earth. Because the tool is web-based, it is relatively easy to assemble a large team of collaborators from multiple locations. With an adequate number of sample sites, this assessment work would allow us to place confidence limits on our estimates of current urban incursions and inholdings, and to identify important instances of illicit land cover change within PA.

In **Section 3.3.B**, we identified two aspects of *de facto* protection status, access and productivity, but we chose to explore only access. There are a number of global datasets of relevance to constraining productivity, including the global MODIS-based net primary production map (Zhao et al. 2005). After assembling both access and productivity attributes, ideally one would formally model the likelihood of PA incursion using regional land cover change datasets. In the absence of regional-to-global scale land cover change data, it may be possible to use the new annual MODIS 500 m land cover maps (Friedl et al. 2002; forthcoming) as a rough indication of aggregate land cover change within PAs.[30]

The urban expansion model of **Section 3.3.C** is a recent development, and there are a wide range of ongoing improvements which extend to all aspects of the algorithm design (Figure 3.8). As mentioned earlier in step one of our growth algorithm, it may be possible to more formally model changes in urban density using the economic modules of the Grubler et al. (2007) time series. Instead of relying exclusively on urban proximity to spatially allocate new urban areas (step four), it may be possible to build a more complex, gravity-based measure of urban suitability (step three). The 140 medium-resolution (28.5 m), two-date city maps in the combined Schneider and Woodcock (2008) and Angel et al. (2005) datasets offer a potential means for fitting such a model. In tabulating the impacts of urban

30 Although the MODIS land cover maps are not explicitly intended for land cover change analysis (pers. corr.), it may be possible to assess the temporal pattern in PA incursions using spatially and thematically aggregated versions of these land cover maps.

expansion (**Section 3.3.C.2**), we did not consider the biological implications of urban expansion by drawing upon new global species diversity and threat datasets (as in Mcdonald et al. 2008); a similar analysis of the 500 m urban expansion projections is planned.

The most significant potential improvement to the urban growth model in **Section 3.3.C** is closely related to the core conservation questions that motivate this research. We make the assumption in our urban expansion model that PAs are completely without enforcement. It is possible to modify this assumption in step four, allowing PAs to resist urban encroachment. This resistance could be set in a binary fashion or continuously. By inserting the management effectiveness of PAs into the model, it will be possible to test a wide range of global-scale conservation strategies. For example, should conservation managers focus most on those parks that contain the most threatened species, on those parks that face the most pressure from human-related activities, or on those that are the most pristine? A parameter for PA effectiveness would also allow modification of the spatial structure of the PA network over time, allowing managers to consider the effects of adding new PAs to high-priority regions. There is little doubt that urbanization will alter the distribution of land suitable for future PAs; modeling those alterations is the most reasonable way to improve our ability to identify and acquire high-priority reserve sites before they urbanize.

Finally, from the standpoint of its utility to conservation planners, there is a clear need to integrate croplands and pasture into our model of urban expansion. Fully 94.7 percent of all incursions and inholdings observed in the circa-2005 MODIS map were croplands. In a non-spatial modeling framework, Tilman et al. (2001) estimate that an additional 10 million sq. km of agricultural land will come into production by 2050 (3.5 million sq. km for cropland and 5.4 million sq. km for pasture). This expansion of agricultural land is 3.7 times the total urban expansion predicted under the most pessimistic combination of model parameters from **Section 3.3.C.3**. Our observation that 81.9 percent

of all cropland now lies within 30 km of urban areas is a likely first step in finding a reasonable strategy for spatially allocating cropland expansion in an integrated agro-urban expansion model. Designing such a strategy is the next large hurdle to delivering a land cover change model that meets the needs of the conservation planning community. The benefits of an integrated model are highlighted by our findings, which point to a clear need to account for both demographic pressure and urban and rural planning decisions in designing sustainable conservation strategies within the context of a rapidly urbanizing world.

	Purpose / Size in millions of sq. km (Percent of all Declared Protected Areas)
IUCN Category Ia	**Strict nature reserve**: protected area managed mainly for science. 0.74 (4%) *Definition:* Area of land and/or sea possessing some outstanding or representative ecosystems, geological or physiological features and/or species, available primarily for scientific research and/or environmental monitoring.
IUCN Category Ib	**Wilderness area**: protected area mainly for wilderness protection. 0.35 (2%) *Definition:* Large area of unmodified or slightly modified land, and/or sea, retaining its natural character and influence, without permanent or significant habitation, which is protected and managed so as to preserve its natural condition.
IUCN Category II	**National park**: protected area managed mainly for ecosystem protection and recreation. 3.83 (23%) *Definition:* Natural area of land and/or sea, designated to (a) protect the ecological integrity of one or more ecosystems for present and future generations, (b) exclude exploitation or occupation inimical to the purposes of designation of the area and (c) provide a foundation for spiritual, scientific, educational, recreational and visitor opportunities, all of which must be environmentally and culturally compatible.
IUCN Category III	**Natural Monument**: protected area managed mainly for conservation of specific natural features. 0.20 (1%) *Definition:* Area containing one, or more, specific natural or natural/cultural feature which is of outstanding or unique value because of its inherent rarity, representative or aesthetic qualities or cultural significance.
IUCN Category IV	**Habitat/species management area**: protected area managed mainly for conservation through management intervention. 2.43 (15%) *Definition:* Area of land and/or sea subject to active intervention for management purposes so as to ensure the maintenance of habitats and/or to meet the requirements of specific species.
IUCN Category V	**Protected landscape/seascape**: protected area managed mainly for landscape/seascape conservation and recreation. 2.13 (13%) *Definition:* Area of land, with coast and sea as appropriate, where the interaction of people and nature over time has produced an area of distinct character with significant aesthetic, ecological and/or cultural value, and often with high biological diversity. Safeguarding the integrity of this traditional interaction is vital to the protection, maintenance and evolution of such an area.
IUCN Category VI	**Managed resource protected areas**: protected areas managed mainly for the sustainable use of natural ecosystems. 3.19 (19%) *Definition:* Area containing predominantly unmodified natural systems, managed to ensure long term protection and maintenance of biological diversity, while providing at the same time a sustainable flow of natural products and services to meet community needs.
National Sites	3.33 (20%) *Definition:* Protected areas designated by national governments that have not yet been classified according to the IUCN classification scheme.
World Heritage Areas	0.23 (1%) *Definition:* Sites significant to 'world cultural and natural heritage.' These sites are listed in the World Heritage Convention, which is administered by the United Nations Educational, Scientific and Cultural Organization (UNESCO) World Heritage Committee.
Ramsar Wetlands	0.18 (1%) *Definition:* Sites which belong to an international treaty for the protection of wetlands first signed in Ramsar, Iran, in 1972.
UNESCO Man and the Biosphere	0.02 (0.1%) *Definition:* Sites protected for research and training activities under the UNESCO Man and the Biosphere program.

Table 3.1 (previous page). Protected area classification system. Definitions from the IUCN (1994) and Lockwood et al. (2006). The global sizes reported for each class are from this analysis and are free of overlaps or inconsistencies in land-water boundaries; each parcel of protected land can only belong to one category. Conflicts were resolved by assigning priority to national sites (in descending order of IUCN protection), followed by international sites (World Heritage, Ramsar, or Man and the Biosphere). Antarctic protected areas are excluded from this analysis.

Protection Groups		millions sq. km *(% PAs)*	World Database of Protected Areas (percent of group total)
1	Strict	8.0 *(48%)*	National Park (IUCN II; 48%) • Habitat/species management area (IUCN IV; 31%) • Strict nature reserve (IUCN Ia; 9%) • Wilderness area (IUCN Ib; 4%) • Ramsar Wetlands (3%) • National monument (IUCN III; 2%) • UNESCO World Heritage Areas (2%) • UNESCO Man and Biosphere (0.2%)
2	Managed extraction	5.3 *(32%)*	Managed resource protected areas (IUCN VI; 60%) • Managed landscape (IUCN V; 40%)
3	Unknown	3.3 *(20%)*	Unclassified protection (100%)

Biome Groups		millions sq. km *(% land)*	Olson Biome (percent of group total)
1	Tropical Forest	24.7 (*19%)*	Tropical and subtropical moist broadleaf forests (81%) • Tropical and subtropical dry broadleaf forests (15%) • Tropical and subtropical coniferous forests (3%) • Mangroves (1%)
2	Temperate Forest	35.8 *(27%)*	Boreal forests and taiga (44%) • Temperate broadleaf and mixed forests (35%) • Temperate conifer forests (12%) • Mediterranean forests, woodlands, and scrub (9%)
3	Vegetated non-forest	35.3 *(26%)*	Tropical and subtropical grasslands, savannas, and shrublands (55%) • Temperate grasslands, savannas, and shrublands (27%) • Montane grasslands and shrublands (15%) • Flooded grasslands and savannas (3%)
4	Barren	50.1 *(28%)*	Deserts and xeric shrublands (74%) • Tundra (21%) • Rock and ice (5%)

Land Cover Groups		millions sq. km *(% land)*	Land Cover Class (percent of group total)
1	Forest	29.0 *(22%)*	Evergreen broadleaf forest (46%) • Mixed forest (21%) • Evergreen needleleaf forest (17%) • Deciduous needleleaf forest (9%) • Deciduous broadleaf forest (7%)
2	Vegetated non-forest	70.0 *(52%)*	Open shrubland (29%) • Grasslands (22%) • Woody savannas (19%) • Savannas (12%) • Cropland/natural vegetation mosaic (12%) • Closed shrubland (4%) • Permanent wetlands (2%)
3	Barren	34.1 *(16%)*	Barren or sparsely vegetated (84%) • Snow and ice (16%)
4	Urban	0.7 *(0.5%)*	Urban and built-up (100%)
5	Croplands	12.2 *(9%)*	Croplands (100%)

Table 3.2 Thematic aggregations. Protected area classes from the 2007 World Database of Protected Areas, Olson biomes from Olson et al. (2001), and IGBP land cover classes from the 2005 MODIS 500 meter land cover map. Group components are listed in order of their proportion of the total group area. All percentages exclude Antarctica.

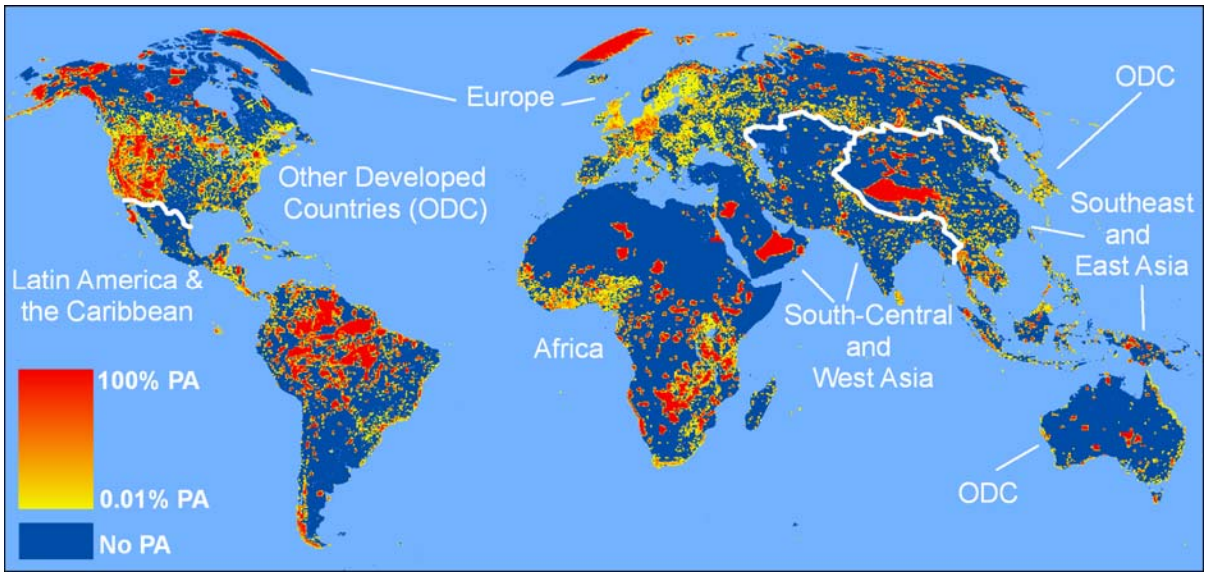


Figure 3.1 Protected areas and six world regions. The map is in the Goode's Homolosine projection, the protected areas are from the World Database of Protected Areas (WDPA 2007), and the world regions are modified from a World Bank five-region scheme: *Europe* including Greenland, Russia, and Ukraine; *Other developed countries* including the USA, Canada, Japan, and Australia and New Zealand; *Latin America and the Caribbean* including all of South and Central America and Mexico; *Africa* including North Africa; *South-central and West Asia* including Turkey, Georgia, and India; and *Southeast and East Asia* including China and Indonesia.

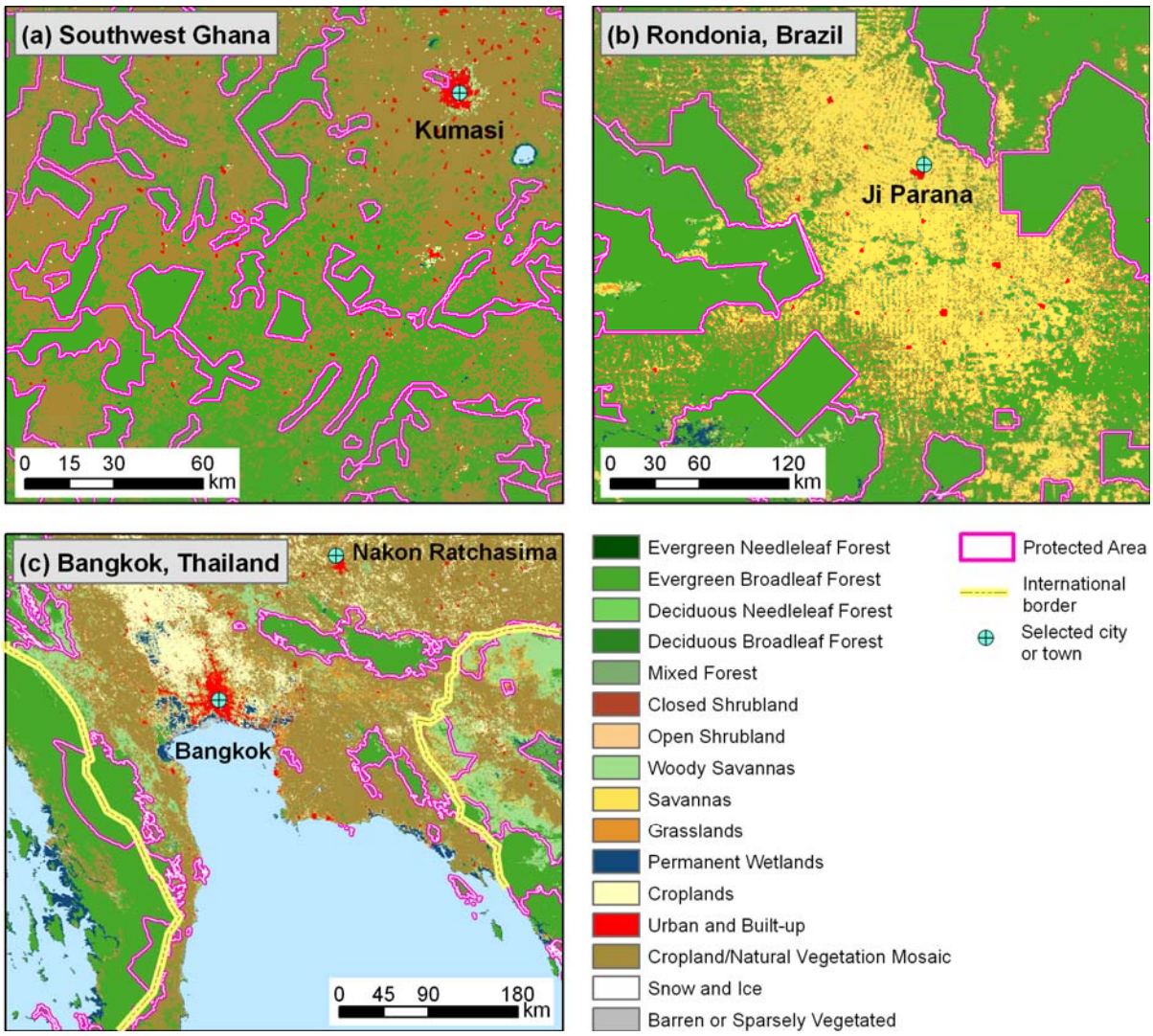


Figure 3.2 Protected areas and the MODIS 500 m land cover map. World Database of Protected Areas (WDPA 2007; red and white lines) and the 2005 MODIS 500 meter land cover map: (a) in the vicinity of Kumasi, Ghana, (b) in the state of Rondônia, Brazil, and (c) and in the region surrounding Bangkok, Thailand. The scale of these plots decreases from 1:2 million (a), to 1:4 million (b), and 1:6 million (c). All of the WDPA sites are either nationally or internationally designated (part of our strict, managed extraction, or unknown classes). Note the excellent alignment between the 2007 version of WDPA and the MODIS land cover map, evident in those regions where logging has encroached right to the border of the protected areas. The border between Thailand and Myanmar/Burma is also evident due to differences in forest clearance rates.

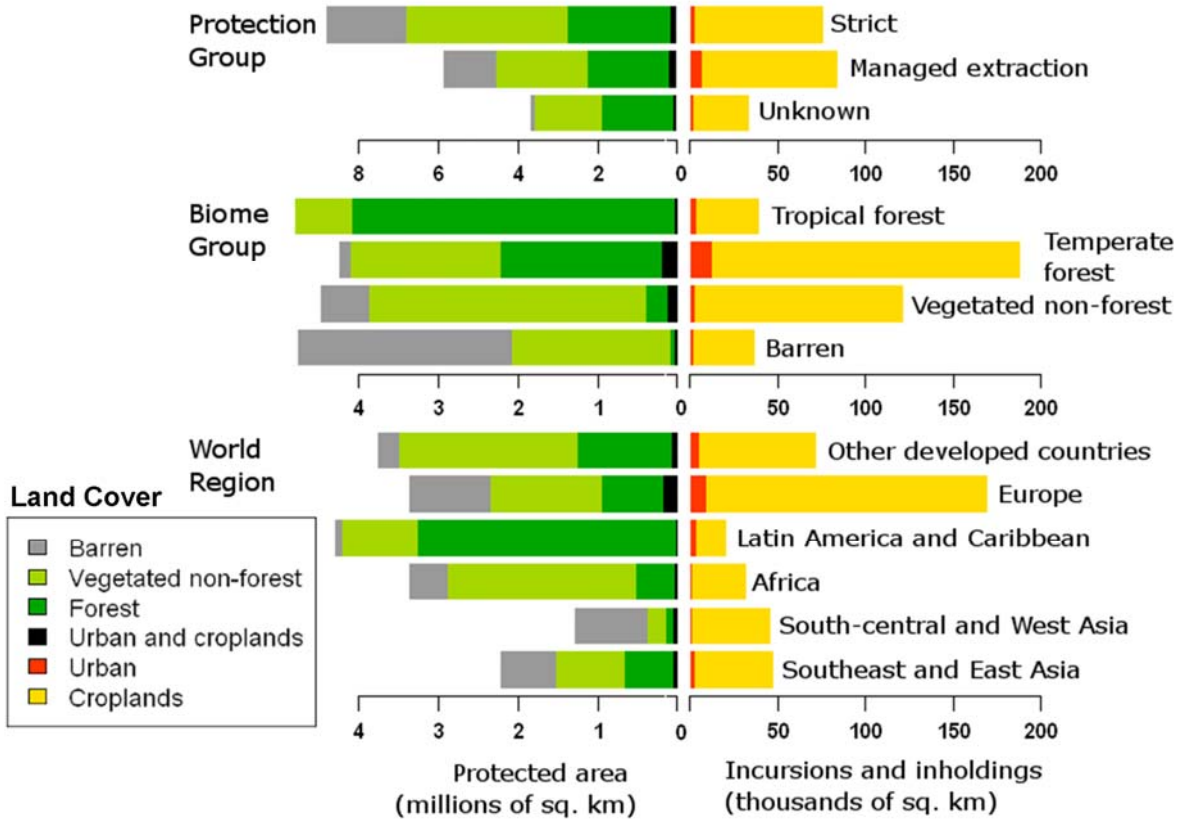


Figure 3.3 Land cover composition of protected areas. The left bars describe the land cover composition of protected areas (in millions of sq. km), and the bars on the right describe inholdings and incursions within these protected areas (in thousands of sq. km). Note the scale change for the protected class groups (upper-left three bars; 0-8 versus 0-4 million sq. km). The land cover categories are based on the new MODIS 500 meter 2005 land cover map (Friedl et al. 2002) and protection status is based on the 2007 World Database of Protected Areas (WDPA 2007).

	Land Cover (thousands of sq. km)						
	Forest Group	Vegetat-ed Non-forest Group	Barren Group	Urban Area or Crop Group	Total PA (x1000 sq. km)	Pct. of PA in Urban Area or Crops	Pct. of Group / Region in PA
Total	**5,790**	**7,310**	**3,140**	**384**	**16,630**	**2.3%**	**12.5%**
Percent of all PA	***35%***	***44%***	***19%***	***2.3%***	NA	NA	NA
Protection Group							
Strict	2,320	3,680	1,840	151	7,990	1.9%	NA
Managed Extraction	1,860	2,090	1,210	167	5,320	3.1%	NA
Unknown	1,620	1,550	100	66	3,330	2.0%	NA
Biome Group							
Tropical Forest	3,670	650	4	39	4,370	0.9%	17.7%
Temperate Forest	1,840	1,700	140	188	3,860	4.9%	10.8%
Vegetated Non-Forest	240	3,150	550	121	4,060	3.0%	11.5%
Barren	40	1,820	2,440	36	4,340	0.8%	11.5%
World Region							
Other Developed Countries	1,070	2,030	240	71	3,410	2.1%	12.5%
Europe	700	1,270	940	169	3,070	5.5%	12.4%
Latin America and Caribbean	2,940	860	82	20	3,910	0.5%	19.1%
Africa	450	2,140	430	32	3,050	1.0%	10.2%
South-central and West Asia	84	220	820	45	1,170	3.9%	7.4%
SE and East Asia	560	790	620	47	2,020	2.3%	13.0%

Table 3.3 Land cover composition of protected areas. Land cover is based on the new MODIS 500 meter 2005 land cover map (Friedl et al. 2002) and protection status is based on the 2007 World Database of Protected Areas (WDPA 2007). The percent protected column (rightmost) does not include Antarctica.

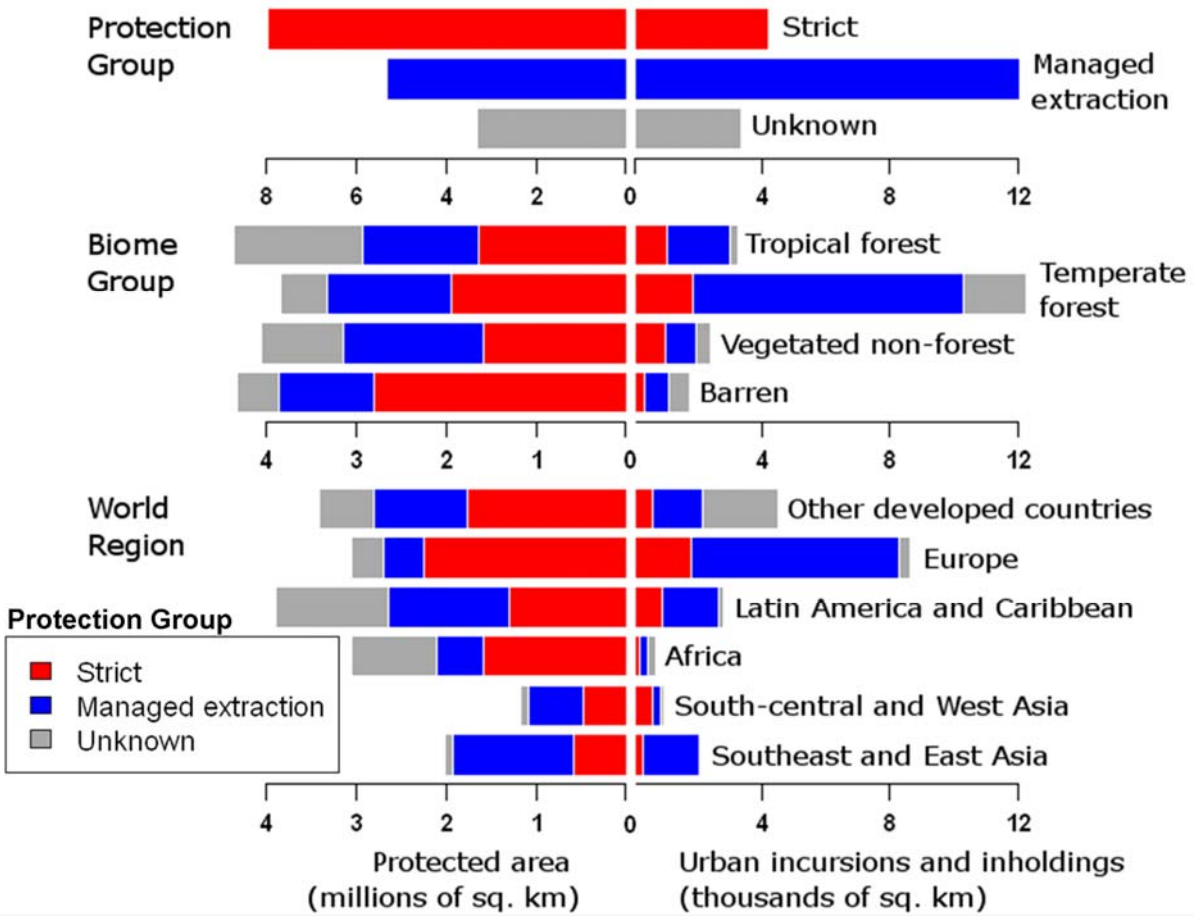


Figure 3.4 Protection composition of the global protected area network (millions of sq. km; left), and total urban incursions as of 2005 (thousands of sq. km; right). Note the scale change for the protected class groups (upper-left bars; 0-8 versus 0-4 million sq. km). Land cover is based on the new MODIS 500 meter 2005 land cover map (Friedl et al. 2002) and protection status is based on the 2007 World Database of Protected Areas (WDPA 2007).

	Urban Incursions and Inholdings by Protection Group				
Biome Group	Strict (sq. km)	Managed Extraction (sq. km)	Unknown (sq. km)	Total urban area (sq. km)	Percent of PA classified as urban area
Tropical Forest	1,030	1,940	250	3,220	(0.07%)
Temperate Forest	1,840	8,480	1,950	12,270	(0.32%)
Vegetated Non-Forest	990	930	460	2,380	(0.06%)
Barren	340	730	660	1,730	(0.04%)
World Region					
Other Developed Countries	550	1,560	2,400	4,500	(0.13%)
Europe	1,780	6,500	360	8,640	(0.28%)
Latin America and Caribbean	880	1,780	150	2,810	(0.07%)
Africa	160	230	300	690	(0.02%)
South-central and West Asia	590	220	90	900	(0.08%)
SE and East Asia	240	1,800	30	2,070	(0.10%)
Total	**4,200**	**12,080**	**3,320**	**19,600**	**(0.12%)**

Table 3.4 Urban incursions and inholdings. Urban area is based on the new MODIS 500 meter urban map (Schneider et al. forthcoming) and protection status is based on the 2007 World Database of Protected Areas (WDPA 2007).

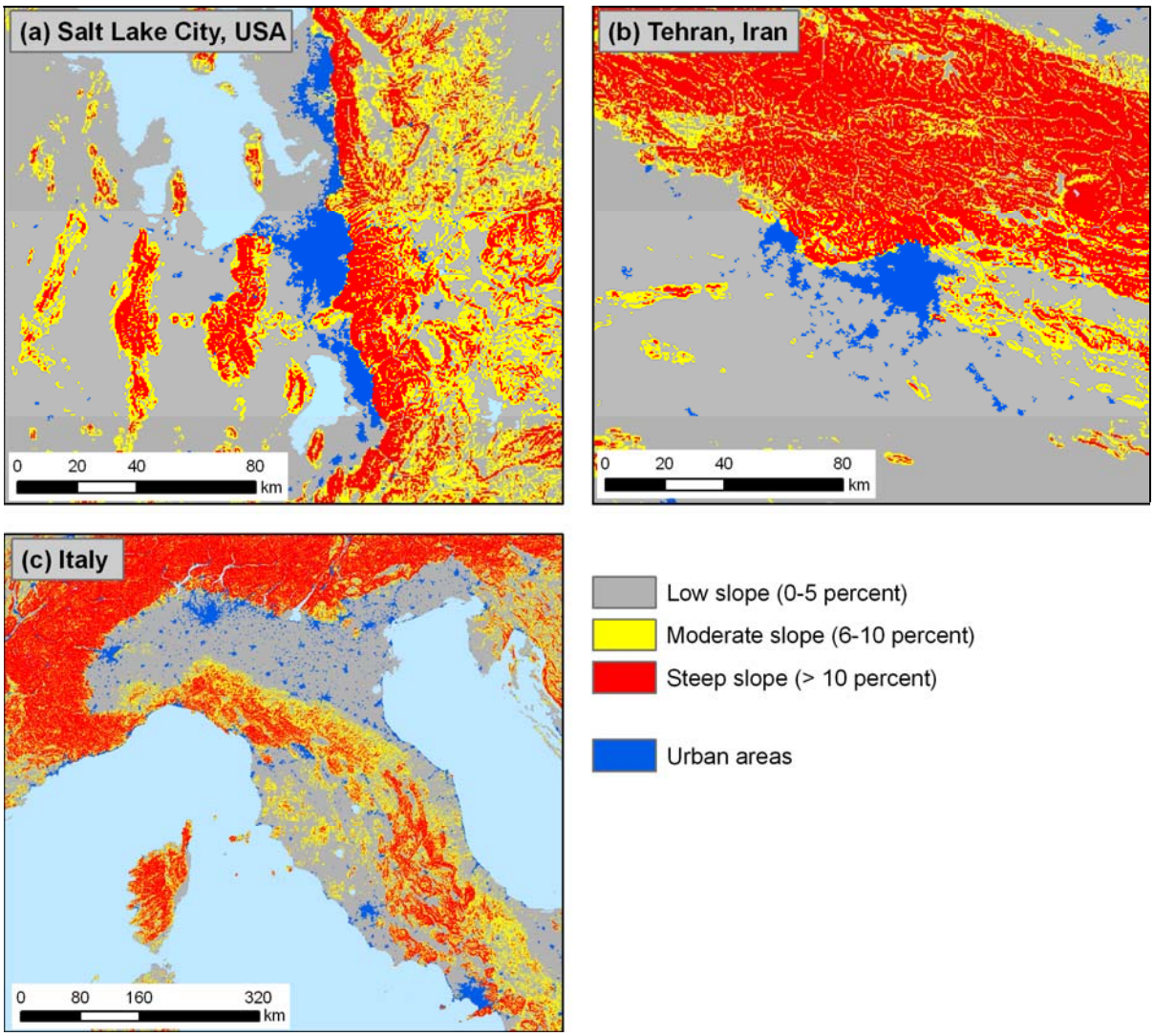


Figure 3.5 Slope and urban areas examples. Three categories of slope derived from gap-filled Shuttle Radar Topography Mission data (SRTM; Reuter et al. 2007), and the MODIS 500 meter urban map (blue) in: (a) Salt Lake City, USA, (b) Tehran, Iran, and (c) Italy. Red indicates high slope (11-71 percent), yellow moderate slope (6-10 percent), and gray low slope (0-5 percent). The three examples are illustrative subsets of the global slope and urban maps used to establish urban suitability thresholds.

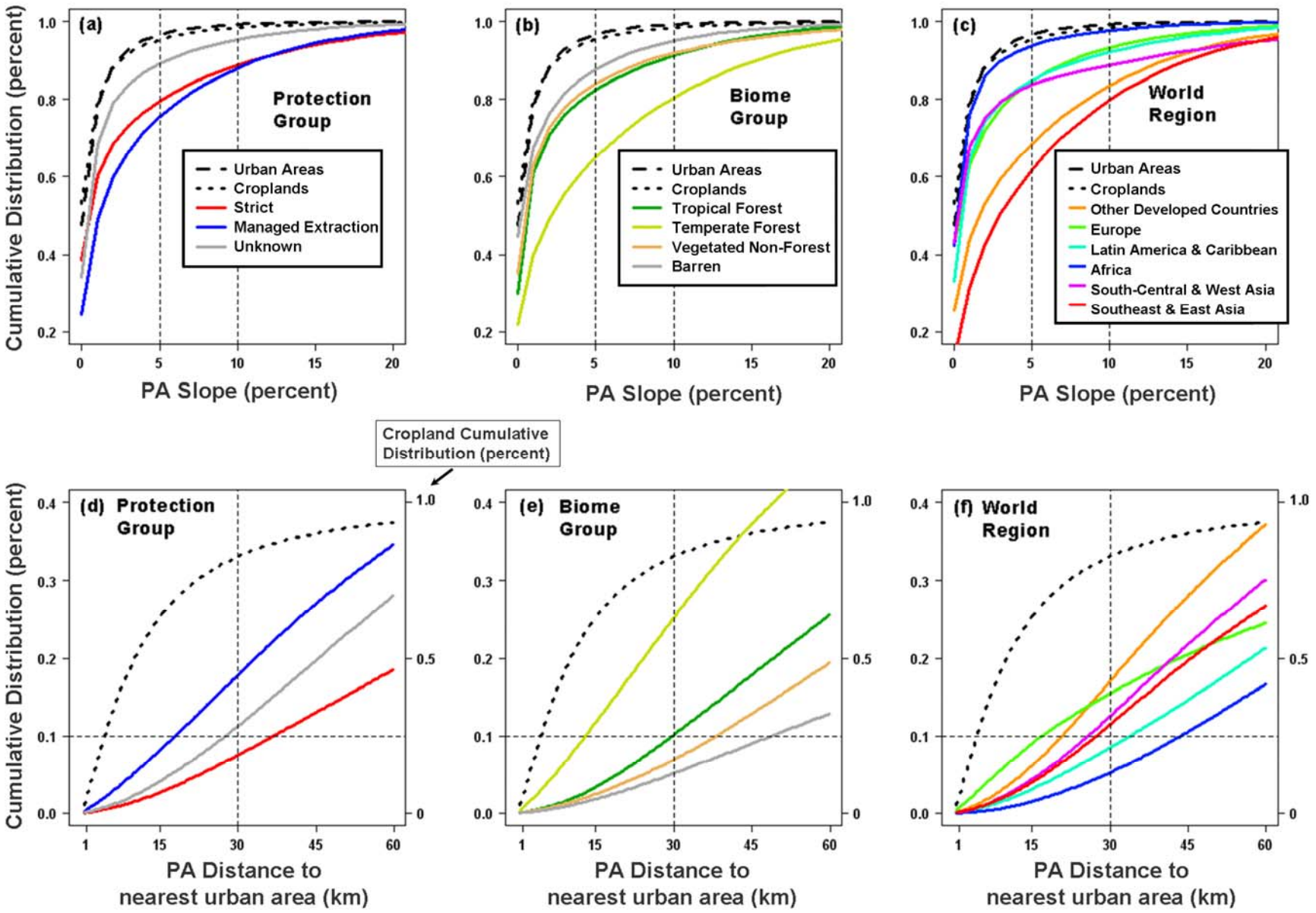

(a)
Protection Group
Urban Areas
Croplands
Strict
Managed Extraction
Unknown
(b)
Biome Group
Urban Areas
Croplands
Tropical Forest
Temperate Forest
Vegetated Non-Forest
Barren
(c)
World Region
Urban Areas
Croplands
Other Developed Countries
Europe
Latin America & Caribbean
Africa
South-Central & West Asia
Southeast & East Asia
Cumulative Distribution (percent)
PA Slope (percent)
Cropland Cumulative Distribution (percent)
(d) Protection Group
(e) Biome Group
(f) World Region
PA Distance to nearest urban area (km)

Figure 3.6 (previous page) Protected areas by slope and urban proximity. The cumulative frequency of protected area pixels by slope (a-c; upper three plots) and distance to nearest urban land (d-e; lower three plots). Protected areas are grouped by protection class (a,d; left plots), biome group (b,e; middle plots), and world region (c,f; right plots). For the slope plots, the full global distribution of urban area and cropland classes are included for comparison (black dashed and dotted lines, respectively). For the distance to urban plots, the right-axis is the cumulative frequency for the full global cropland class (black dotted line). Slope is based on the gap-filled Shuttle Radar Topography Mission data (SRTM; Reuter et al. 2007) and urban areas are from the MODIS 500 meter urban map (Schneider et al. forthcoming).

	Protected Area Slope Steeper than 5%		Protected Area greater than 30 km from Urban Area		Union of 'Safe' Protected Area		Union of 'Exposed' Protected Area	
Protection Group	thousands of sq. km (percent of total undisturbed protected area in group/region)							
Strict	1,630	(20.8%)	7,380	(94.2%)	1,410	(18.0%)	390	(5.0%)
Managed Extraction	1,300	(25.3%)	4,360	(84.7%)	960	(18.6%)	620	(12.0%)
Unknown	360	(11.0%)	2,950	(90.4%)	300	(9.2%)	320	(9.8%)
Biome Group								
Tropical Forest	770	(17.8%)	3,920	(90.5%)	610	(14.1%)	290	(6.7%)
Temperate Forest	1,340	(36.5%)	2,870	(78.2%)	1,000	(27.3%)	630	(17.2%)
Vegetated Non-Forest	650	(16.5%)	3,780	(96.0%)	590	(15.0%)	230	(5.8%)
Barren	530	(12.3%)	4,110	(95.5%)	490	(11.4%)	180	(4.2%)
World Region								
Other Developed Countries	1,080	(32.4%)	2,830	(84.8%)	860	(25.8%)	370	(11.1%)
Europe	470	(16.2%)	2,590	(89.3%)	360	(12.4%)	360	(12.4%)
Latin America and Caribbean	590	(15.2%)	3,570	(91.8%)	480	(12.3%)	220	(5.7%)
Africa	190	(6.3%)	2,890	(95.8%)	170	(5.6%)	140	(4.6%)
South-central and West Asia	190	(16.9%)	1,020	(90.7%)	150	(13.3%)	110	(9.8%)
Southeast and East Asia	770	(39.0%)	1,780	(90.2%)	670	(34.0%)	130	(6.6%)
Total	**3,290**	**(20.3%)**	**14,680**	**(90.4%)**	**2,680**	**(16.5%)**	**1,320**	**(8.1%)**

Table 3.5 The slope of protected areas and their distance to urban areas. Slope is based on the gap-filled Shuttle Radar Topography Mission data (SRTM; Reuter et al. 2007) and urban areas from the MODIS 500 meter urban map (Schneider et al. forthcoming).

Population	**High-growth (A2r)**	**Medium-growth (B2)**	**Low-growth (B1)**
Population size	High	Medium-high	Low
Demographic transition	Delayed and slow	Medium	Rapid
Long-term fertility levels	Near or below replacement	Converging to replacement	Well below replacement
Urbanization	**(A2r)**	**(B2)**	**(B1)**
Urbanization rates	High	Medium	Low
Mega-city growth	High	Localized (Asia)	Low (constrained)
Urban-rural gradient	Medium-high	Medium	Converging to zero
Income	**(A2r)**	**(B2)**	**(B1)**
Income growth	Medium-low	Medium	High
Income convergence	Very low (initially diverging)	Medium-low	Very rapid
Domestic/International price differences	persist initially, slow convergence post-2040	Medium convergence	Rapid convergence

Table 3.6 Three socio-economic change scenarios. Three Intergovernmental Panel on Climate Change - Special Report on Emission Scenarios (IPCC-SRES) scenarios as downscaled from regional to national scale by Grubler et al. (2007). This table is modified directly from Grubler et al. (2007). From here forward, we use the label 'high-growth' for the A2r scenario, 'medium-growth' for B2, and 'low-growth' for B1.

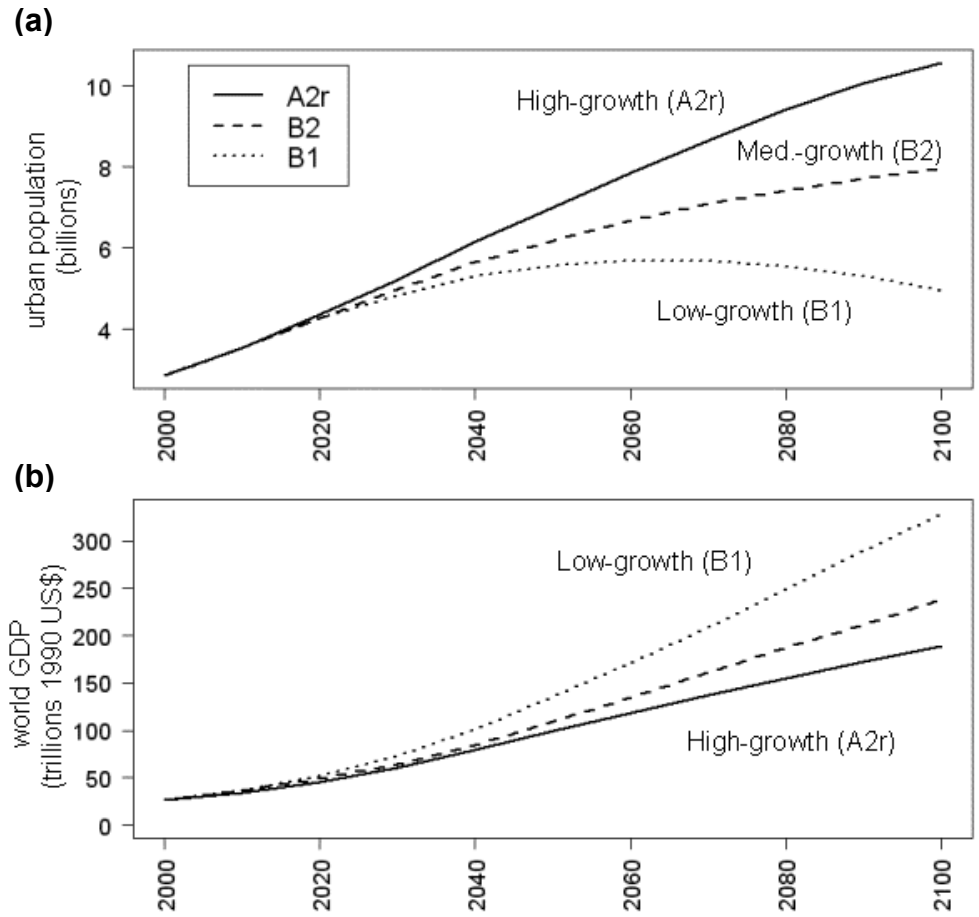


Figure 3.7 Urban population size and gross domestic product, three scenarios. Urban population size (a) and world gross domestic product (b) from three Intergovernmental Panel on Climate Change - Special Report on Emission Scenarios (IPCC-SRES) scenarios as downscaled from regional to national scale by Grubler et al. (2007).

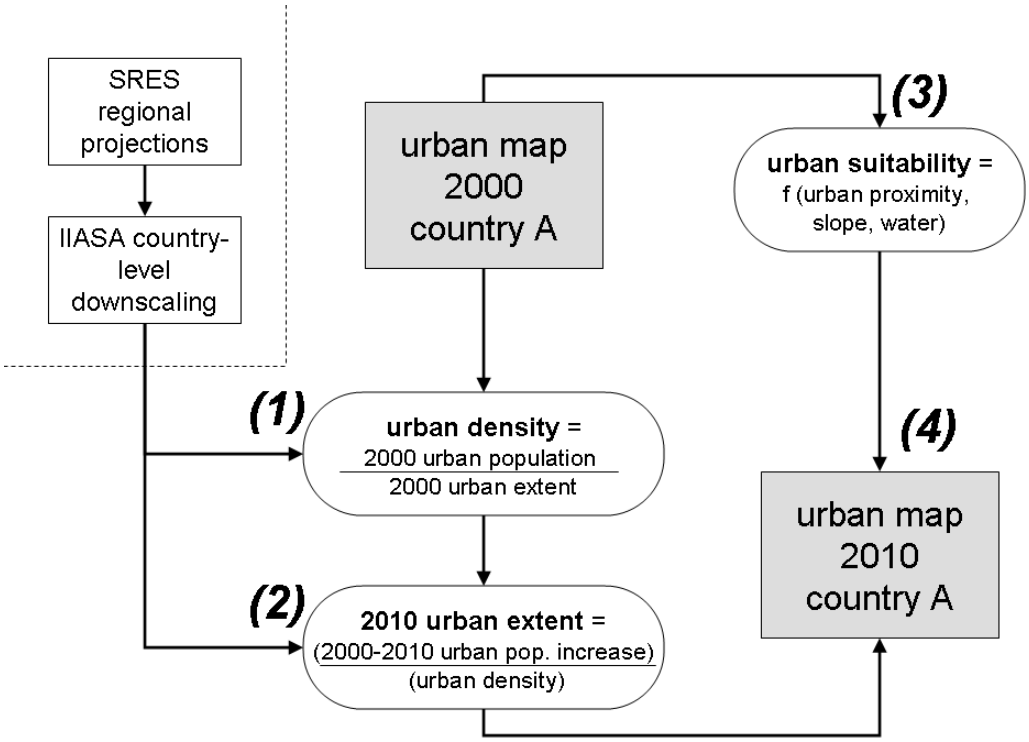


Figure 3.8 A simple model for global urban expansion. The principal inputs are the Special report on Emissions Scenarios storylines as downscaled by Grubler et al. (2007; upper left box), and the MODIS 500 meter urban map (Schneider et al. forthcoming; lower right box).

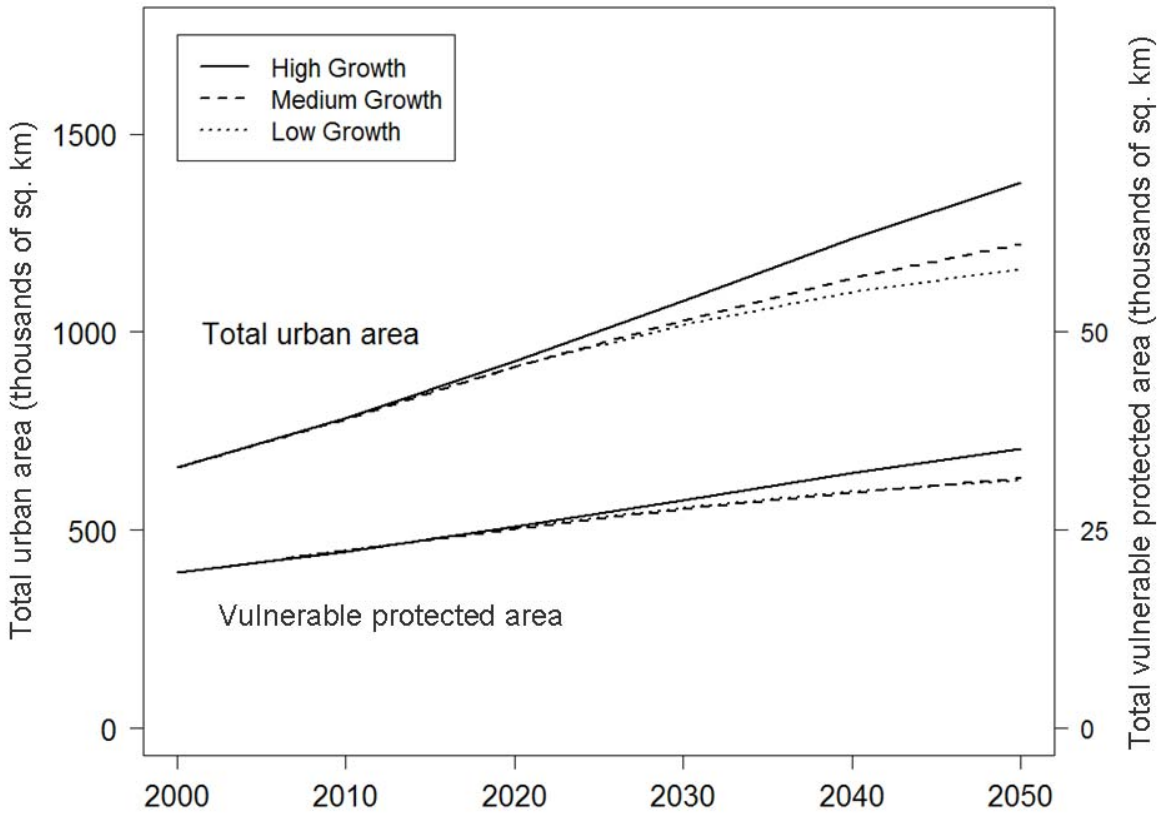


Figure 3.9 Global model estimates of urban expansion and increases in vulnerable protected area. Model estimates of total urban expansion (top curves; left axis) and total extent of vulnerable protected area due to potential urban encroachment (bottom curves; right axis) under the three Grubler et al. (2007) population growth scenarios. The year 2000 baseline is observed urban area (657,000 sq. km) and observed urban incursions or inholdings on protected areas (19,600 sq. km).

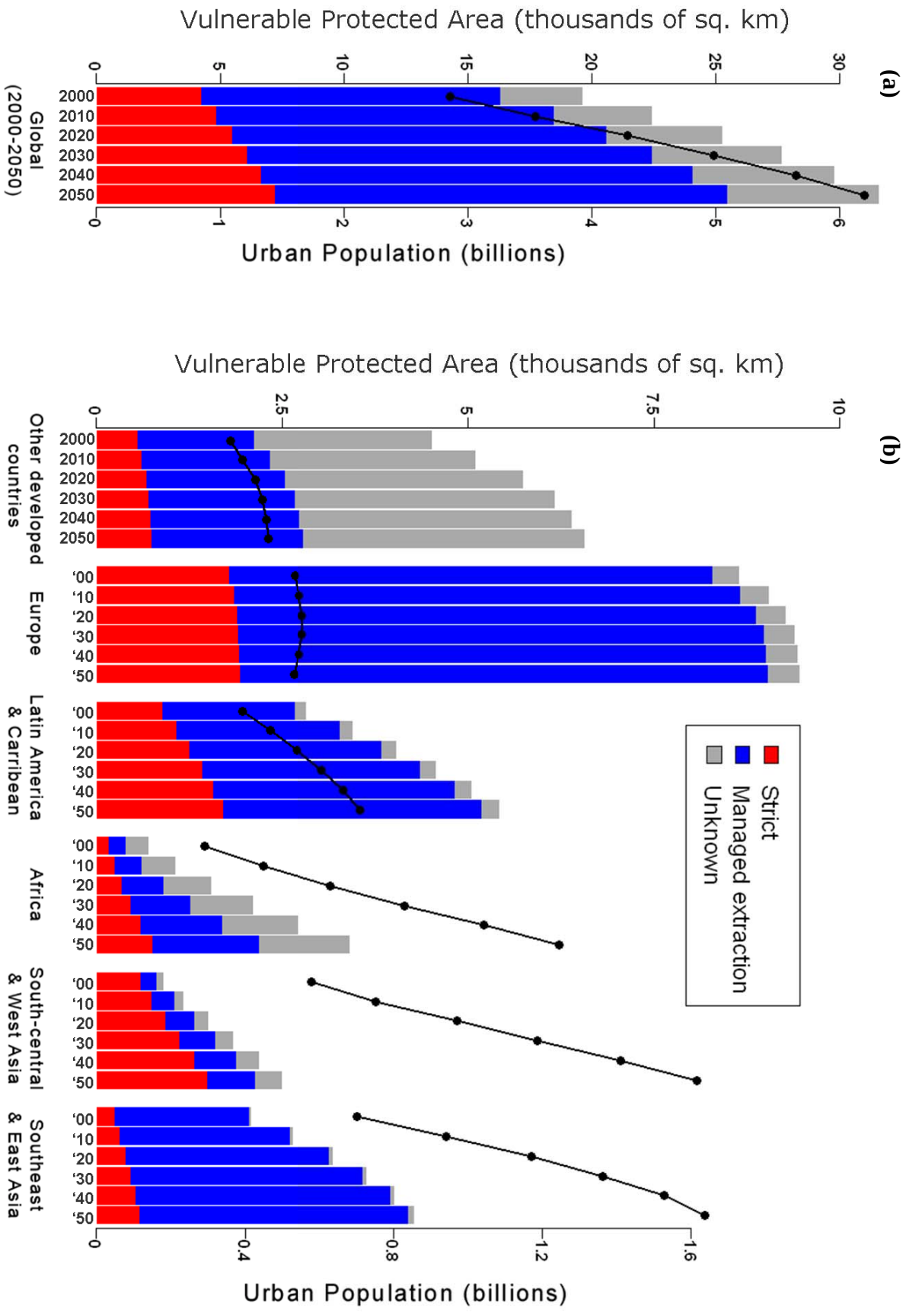

(a)
Vulnerable Protected Area (thousands of sq. km)
0
5
10
15
20
25
30
Global (2000-2050)
2000
2010
2020
2030
2040
2050
0
1
2
3
4
5
6
Urban Population (billions)
(b)
Vulnerable Protected Area (thousands of sq. km)
0
2.5
5
7.5
10
Other developed countries
Europe
Latin America & Carribean
Africa
South-central & West Asia
Southeast & East Asia
'00
'10
'20
'30
'40
'50
Strict
Managed extraction
Unknown
0
0.4
0.8
1.2
1.6
Urban Population (billions)

Figure 3.10 (previous page) Regional model estimates of vulnerable protected areas 2000-2050. Cumulative model estimates of potential urban incursions from 2000-2050 using the medium-growth scenario (B2), globally (a) and for six world regions (b). The leftmost bar within each six-bar-group is the baseline area of observed circa-2000 urban incursions (19,600 sq. km globally). The subsequent five bars of each six-bar group portray the cumulative total of vulnerable protected area (areas reclassified as urban in the medium-growth model) over five decades: 2000-10, 2010-20, 2020-30, 2030-40, and 2040-50. The colors of the bars correspond to protection status (same color scheme as in Figure 3.3 and 3.4), and are measured on the left axes of (a) and (b) in thousands of sq. km. The black bars are the total urban population according to the medium growth model, and are measured along the right axes of both (a) and (b) in millions of urban residents.

World Region	New Urban Area 2000-50 (sq. km)	New Vulnerable Protected Area 2000-50 (sq. km)	Percent Increase in Vulnerable Protected Area 2000-50	Percent Increase in Urban Area 2000-50	Percent Increase in Urban Population 2000-2050
Other Developed Countries	56,800 *(77100-87600)*	2,100 *(2800-3200)*	46% *(62-70)*	36% *(48-55)*	28% *(40-46)*
Europe	11,300 *(18100-17300)*	800 *(1600-1400)*	9% *(18-16)*	8% *(12-12)*	0% *(10-11)*
Latin America & Caribbean	72,600 *(62200-95300)*	2,600 *(2200-3400)*	93% *(80-122)*	77% *(66-101)*	81% *(69-107)*
Africa	117,200 *(92900-135100)*	2,700 *(2100-3100)*	392% *(300-447)*	283% *(224-326)*	326% *(254-374)*
South-central and West Asia	146,400 *(140000-196000)*	1,600 *(1500-2100)*	177% *(172-235)*	156% *(149-209)*	179% *(154-229)*
Southeast and East Asia	162,000 *(110500-189400)*	2,200 *(1500-2500)*	106% *(72-121)*	137% *(94-160)*	133% *(89-157)*
Total	566,300 *(500900-720700)*	12,000 *(11700-15700)*	61% *(60-80)*	86% *(76-110)*	117% *(95-146)*

Note: Asterisks indicate regions where the 'medium-growth' estimate is not bracketed by the 'low-' and 'high-growth' scenarios; under IPCC-SRES Europe and ODC have their lowest population growth rates in the 'medium-growth' (B2) scenario (Grubler et al. 2007).

Table 3.7 Modeled increases in urban areas and vulnerable protected areas through 2050. The table shows regional and global projections from the medium-growth scenarios, with the low- and high-growth scenario projections in parentheses.

	New Vulnerable PA 2000-50 (sq. km)	Percent of New Vulnerable PA on Strict PA	Percent Increase in Vulnerable PA (2000-50)	Percent Increase in Urban Area (2000-50)	Percent of PA 'Exposed' in 2000 (Table 3.5)	New Urban Population 2000-50 (millions)
USA	1,950	7%	55%	41%	14.7%	87.3
China	1,723	1%	101%	136%	6.1%	598.8
Brazil	1,316	24%	96%	68%	5.6%	94.2
Ethiopia	788	4%	762%	719%	5.2%	71.9
Venezuela	777	24%	89%	98%	6.6%	20.6
Uganda	497	26%	1712%	1380%	4.9%	45.0
Germany	421	5%	8%	5%	83.7%	3.4
Saudi Arabia	399	0%	229%	192%	4.3%	36.3
Pakistan	398	34%	364%	333%	18.7%	157.4
India	393	100%	138%	148%	22%	418.1
Nigeria	244	1%	399%	325%	24.5%	161.5
DR Congo	219	59%	275%	586%	3.3%	84.7
France	191	19%	32%	23%	43.1%	10.3
Indonesia	167	68%	123%	160%	7.1%	138.5
Uzbekistan	153	100%	107%	116%	17.1%	10.5

Table 3.8 The top fifteen country-level increases in vulnerable protected area through 2050. All results are for the medium-growth scenario under the assumption of constant urban population density.

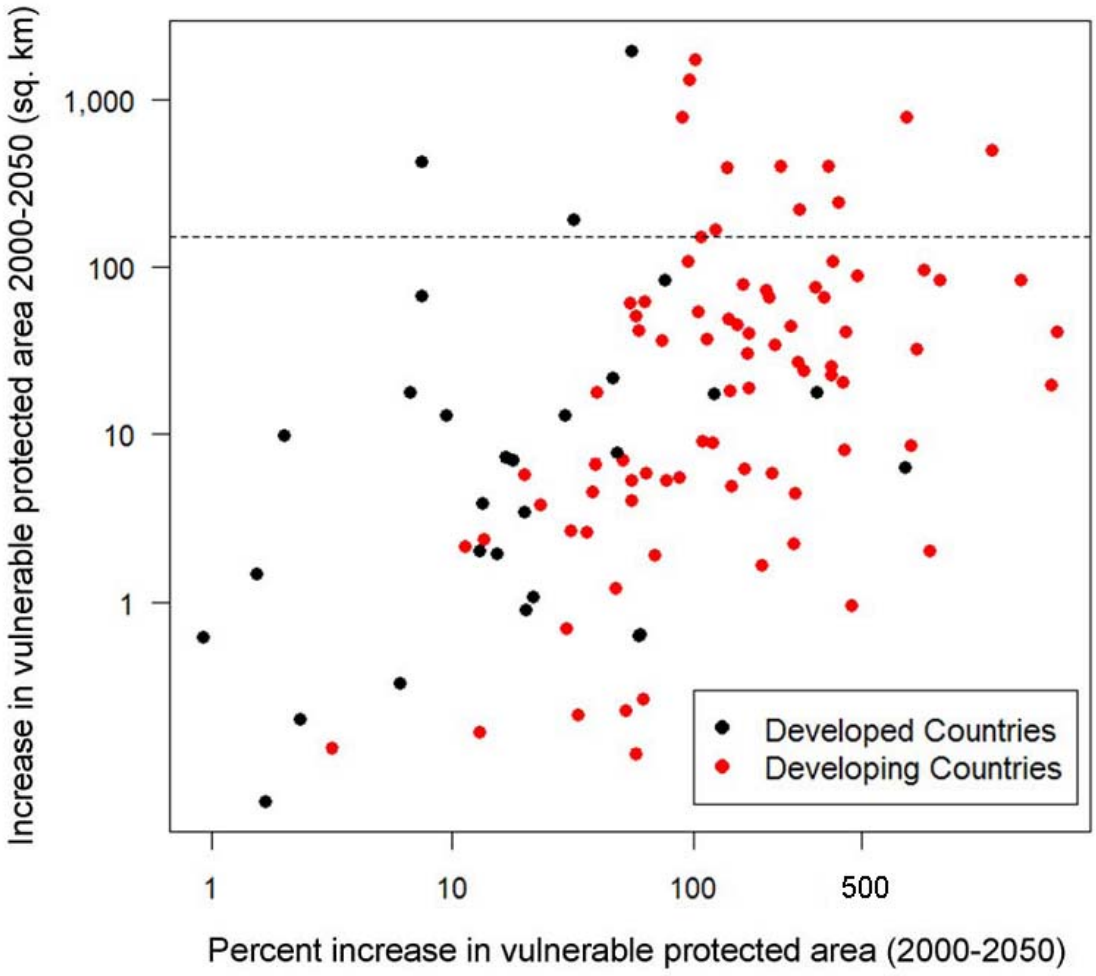


Figure 3.11 Country-level projections of vulnerable protected areas 2000-2050. Includes 121 countries and territories, under the medium-growth scenario. Black dots are for Europe and Other developed countries and red dots are for the rest of the world. The 15 countries above the dotted line account for 80.5 percent of all new vulnerable protected area from 2000-2050; all of these countries are listed in Table 3.8.

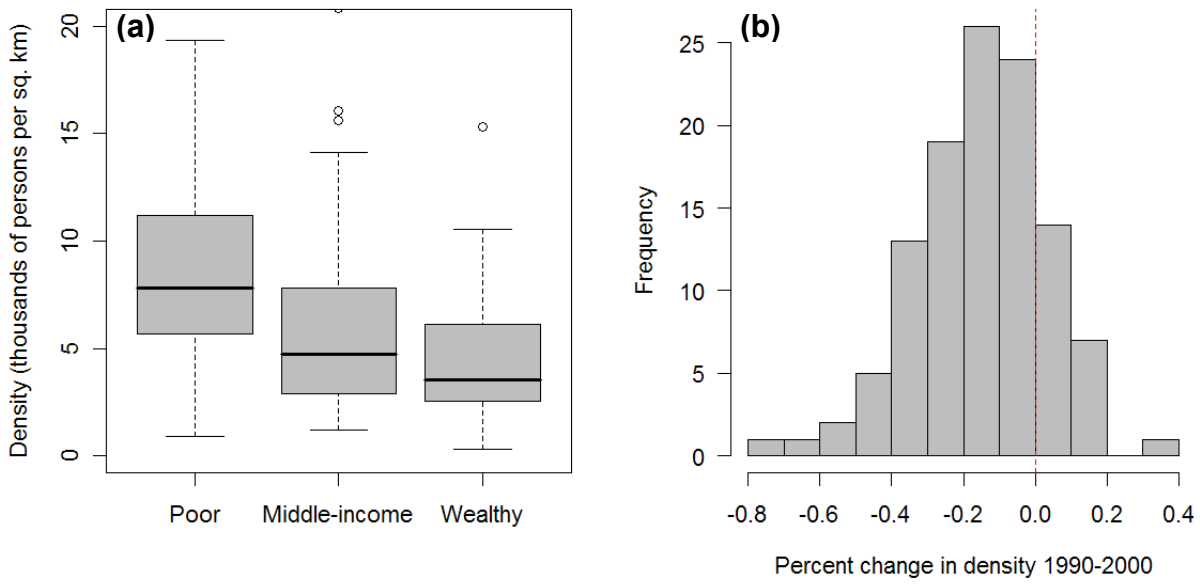


Figure 3.12 Distribution of country- and city-level urban population density. Figure 3.12a is a box-plot of country-level urban population density based on IIASA circa-2000 national population (Grubler et al. 2007) and MODIS 500 meter urban (Schneider et al. forthcoming). The density values are split into three quintiles based on circa-2000 per capita gross domestic product (GDP) from the World Bank (World Bank 2008; UN 2008a). There are 191 countries included in the figure (9 outliers are not visible in the plot extent). Figure 3.12b is a histogram of percent change in urban population density from 1990-2000 in a pooled sample of Landsat-based densities from 113 cities in the Angel et al. (2005) and Schneider and Woodcock (2008) analyses of global urbanization.

CONCLUSION

Human activities are an important driver of global change processes, and the settlements we call home place an anchor on the spatial distribution of these human-induced changes. A global map of human settlements is essential for (a) illuminating the regional to global scale connections between settlement characteristics and the health of both human populations and the natural ecosystems on which they depend, (b) planning for and responding to natural disasters, and (c) charting the course of contemporary urban expansion. Although settlements occur in both rural and urban areas, because more than half of the population dwells in cities and nearly two billion new urbanites are on track to arrive by 2030, a global map of urban areas represents a significant first step in constructing a global human settlement map. To be of use, such a global urban map should be accuracy-assessed, of moderate resolution, and regularly updated. Today, eight groups from around the world have created global urban maps that meet at least one of these three basic requirements. Our challenge has been to systematically compare the maps in a uniform analysis environment (**Chapter 1**), to select the map that most accurately portrays urban areas (**Chapter 2**), and to use the most accurate map to understand the impact that urban expansion may have on areas set aside for conservation (**Chapter 3**). In the course of meeting the three research objectives, we have introduced a number of new methods, including: the use of hexagonal equal-area Discrete Global Grids (DGGs) for global map comparison, the web-based implementation of a Land Cover Validation Tool (LCVT), and the design of a spatially-explicit model for global urban expansion.

Despite the considerable resources allocated to the task of creating global urban maps, **Chapter 1** revealed that the eight maps differ by as much as an order of magnitude in their estimates of the total areal extent of the Earth's urban land (0.27 – 3.52 million sq. km). Inter-map differences persist at the scale of urban patches, countries, and world

regions. To better visualize the maps and to conduct a quantitative map comparison, we employed a hexagonal system of equal-area Discrete Global Grids (DGGs). An analysis of the spatial distribution of urban land based on these DGGs across a wide range of spatial resolutions (29-264 km) has revealed that inter-map correlations are highest in North America ($\overline{r}$ = 0.83), lowest in Asia ($\overline{r}$ = 0.66), and intermediate in Europe, South and Central America, and Sub-Saharan Africa ($\overline{r}$ = 0.79). The large observed inter-map differences are likely driven by a combination of several factors, including the timing of map construction, underlying source data and methods, map resolution and class enumeration, and fundamental differences in each group's approach to defining urban areas.

When confronted with such a high degree of heterogeneity in a collection of maps that claims to describe the same feature, the question naturally arises: which maps correspond most closely to built-up patterns observed on the ground? We addressed this long-overlooked issue in the two-tiered accuracy assessment of **Chapter 2.** We began by first establishing an operational definition of what we mean by 'urban areas'. This was an essential task because of the lack of consensus surrounding the urban concept within both the social and physical science communities. For the purposes of our assessment, we define 'urban' based on physical attributes: urban areas are landscapes that are dominated by the built environment, where 'built environment' encompasses all non-vegetative human-constructed elements, including roads, buildings, runways, industrial facilities, etc., and 'dominated' implies coverage greater than or equal to fifty-percent of a given landscape unit (in this case, moderate-resolution pixels). This physically based definition of urban areas aligns well with the definitions that are implicit in the methods of the eight global urban maps under review.

With an operational definition of urban areas, we next turned to our two-tiered accuracy assessment. In this approach, we began by using high-resolution imagery to assess the accuracy of a collection of medium-resolution city-scale maps (tier one), and next we

used this medium-resolution map collection to assess our eight coarse-resolution global maps (tier two). By comparing the 140 medium-resolution maps to more than 10,000 high-resolution validation sites constructed with imagery from our new web-based Google Earth Validation Tool, we found map accuracy to be both high (upper-80 to lower-90 percent overall map agreement) and consistent with previous assessments. In the second-tier of our accuracy assessment, we leveraged these medium-resolution city maps for three separate sets of global map accuracy measures: (a) presence / absence of major cities, (b) estimation of city size, and (c) pixel-based measures of map agreement. The new MODIS 500 m-resolution urban map had the best accuracy rank for all three measures (mean rank of 1.0), followed by the nighttime lights-based Global Impervious Surface Area map (mean rank of 2.7). Based on these accuracy estimates and the characteristics of each map, we concluded **Chapter 2** by offering recommendations for wise-use aimed at a diverse group of potential map users.

Having established the MODIS 500 m map as the most accurate global representation of urban areas, in **Chapter 3**, we turned to the task of applying that global urban map to the problem of global-scale biodiversity conservation. Our analysis of the most recent global inventory of protected areas indicates that worldwide there are 16.6 million sq. km of terrestrial parks. In total, 384,000 sq. km of these protected areas (2.3 percent) is mapped as cropland or urban incursions by the MODIS land cover map. Focusing on urbanized protected areas, which are the most permanent losses of natural habitat, developing countries account for 33.0 percent of the 19,600 sq. km total. Of the 16.3 million sq. km of protected areas that remain free of human disturbance, we estimate that 2.7 million sq. km (16.5 percent) is relatively safe from urban encroachment because of high slope or poor proximity to urban areas, while 1.3 million sq. km (8.1 percent) is exposed to potential urban or agricultural encroachment.

In the final section of **Chapter 3**, we examine which of the 1.3 million sq. km of exposed parklands are most vulnerable to urban expansion through 2050. We constructed a simple model of urban growth based on the MODIS 500 m urban map, three storylines for demographic change, and a range of assumptions regarding future trends in urban population density. Our medium-growth model predicts a near doubling of urban areas through 2050, from 0.66 to 1.22 million sq. km. Assuming that protected areas have no ability to deter urban encroachment, this massive urban expansion will directly consume an additional 12,000 sq. km of parks, a 61 percent increase over circa-2000 observations. Despite the relative inaccessibility of parks in Africa, by 2050 that region faces the largest losses of protected areas (2,700 sq. km, an increase of 392 percent). Overall, the developing regions of Africa, Latin America and the Caribbean, and Asia will account for 75.8 percent of all incursions through 2050. An optimistic low population growth, compact urban future (increasing urban densities) implies roughly half the urban incursions projected assuming a more pessimistic high population growth, sprawling urban future (declining urban densities). The differences between these two scenarios highlight the need to account for demographic pressure and urban planning when designing conservation strategies in a rapidly urbanizing world.

Overall, this research has (a) illuminated a surprising degree of disagreement in our understanding of the spatial distribution of urban areas, (b) identified the MODIS 500 m urban map as the global urban map that most accurately describes areas dominated by the built environment, and (c) established the importance of urban expansion for global-scale conservation planning. The tools developed in applying hexagonal discrete global grids to the problem of map comparison (**Chapter 1**) and sample design (**Chapter 2**) are of value to an expanding community of global-scale researchers. Likewise, the Google Earth LCVT (**Chapter 2; Appendix A**) is a concept that has the potential for use outside of the urban mapping community. By providing all of the web-based components necessary to assemble

a geographically-distributed team of land cover observers, LCVT makes it possible to perform a wide range of land cover analyses on any region covered by the global GE archive.

Because this is the first global, systematic analysis of urban maps, there are a number of opportunities for future research, including:

(1) Expanded assessment data. Our assessment data does not include places with fewer than 100,000 inhabitants, in effect omitting up to a quarter of all urban residents and all rural components of the built environment. Fortunately, our assessment methodology is designed for scalability, opening the door for wider assessments that draw on new validation data from high-resolution imagery archives and fieldwork in order to reduce the uncertainty surrounding the location and extent of small cities, towns, and villages. This scalability of our assessment methods is important for a second reason: because contemporary global urban maps are dated as soon as they arrive, there will be a steady stream of new maps in need of independent accuracy assessments.

(2) Map fusion. The overall focus of our assessment was to select the single 'best' map of urban areas. Instead, one could imagine using the assessment data to build a series of composite global urban maps that meet the needs of a wide range of user communities and achieve higher accuracies than any single urban map. Cities are complex and heterogeneous, and no one binary definition of urban areas is likely to satisfy all research interests or application areas. The advantage of a set of composite or 'fusion' maps would be the ability to combine the strengths of several independent global maps; the disadvantage would be the potential confounding of multiple error structures.

(3) Improved models of urban expansion. There are opportunities for refining all aspects of the urban expansion algorithm, including: introducing more formal models of urban density based on economic parameters, creating more sophisticated measures of urban suitability, and using multi-temporal land cover maps to fit model parameters. In addition, the model framework could be expanded to include the future growth in cropland.

A combined agro-urban expansion model would much more fully encompass the full range of future human-induced land cover transitions through 2050. Finally, it is possible to insert a parameter into the expansion model that describes the effectiveness of protected areas at preserving natural habitat and species (on a site-specific or regional basis). Such a model would generate more realistic projections of future urban incursions and would facilitate new research aimed at comparing a suite of global-scale conservation strategies.

We are in the early days of global urban observation and modeling, and there remain numerous opportunities to improve our basic understanding of urban areas. Given current and projected rates of population growth and urbanization, there is no doubt that urban expansion will play a critical role in altering human and physical geographies through the next fifty years. Investing in an improved understanding of the extent and spatial distribution of our world's cities is an important early step in building a more sustainable urban future.

APPENDIX A

THE HORIZONTAL POSITIONAL ACCURACY OF GOOGLE EARTH'S HIGH-RESOLUTION IMAGERY ARCHIVE [31]

A.1 INTRODUCTION

With more than 200 million users since its release in June 2005 (Google 2007), Google Earth (GE) has recently been recognized for its potential to significantly improve the visualization and dissemination of scientific data (Butler 2006; Compieta et al. 2007; Guralnick et al. 2007). Yet the imagery which underlies GE has potential applications that extend beyond visualization; the archive could contribute directly to land-cover and land-use change science (LCLUC). GE now hosts high-resolution (< 2.5 m) imagery from 2000-2008 that spans more than twenty percent of the Earth's land surface, and more than a third of the human population (Google 2006). Imagery at these resolutions allows human observers to readily discriminate between major natural land cover classes and to discern components of the human built environment, including: individual houses, industrial facilities, and roads (IRARS 1996; Leachtenauer 1997). Some scientists have recently begun using this rapidly expanding, cost-free imagery source (Cha and Park 2007; Thenkabail et al. 2007; Zhang et al. 2008; Constantine and Dunne 2008), but the GE high-resolution imagery archive remains a largely unexploited resource for the scientific analysis and description of the Earth's land surface.

Although high-resolution imagery has long played a role in scientific inquiry, the 1999 and 2001 launches (respectively) of the commercial imaging satellites IKONOS and

[31] **Appendix A** is based on Potere (forthcoming).

QuickBird, have generated increased interest in methods that facilitate the efficient extraction of scientifically relevant information from high-resolution imagery. Improvements in algorithm design and computational power have steadily reduced the analytic obstacles for leveraging this imagery, yet the cost of commercial imagery remains prohibitive for many science applications. This cost landscape changed in 2005, when Google began hosting high-resolution commercial imagery at reduced spectral and spatial resolution on its cost-free Google Earth and Google Maps applications.

GE high-resolution imagery does not contain an infrared band and sometimes has a slightly coarser spatial resolution than the original images provided directly from the sensor operators, yet a user of the GE environment is often able to readily discern land cover type, disturbance events, and other relevant attributes based solely on the GE imagery. Users also have a number of additional resources to rely upon, including: detailed digital elevation models which allow three-dimensional viewing of the imagery, more than five million geo-referenced photos from services such as Panoramio (Panoramio 2007), and a rapidly expanding set of vector and image-based overlays from a wide range of commercial geospatial services companies, scientific and government organizations, and millions of individual members of the GE community.

To make Google's high-resolution imagery as useful as possible, it is necessary to more fully characterize the temporal, spectral, and spatial properties of the archive. Up to this point, Google has been unwilling to release detailed information regarding any of these aspects of their holdings. Of all the desired attributes, geo-registration is the most readily tested. Errors in image alignment are apparent in the form of disjoint linear features such as roads and coastlines (Figure A.1). In the face of these errors, the question arises: how trustworthy is the horizontal positional information in this archive? Large errors in geo-registration would limit the scientific utility of the GE archive. In this analysis, we address this important question of horizontal positional accuracy.

A.2 DATA AND METHODS

The dataset under review in this analysis is the GE high-resolution imagery archive—a global collection of images at roughly 2.5 m-resolution. The bulk of the high-resolution images in GE are from DigitalGlobe's QuickBird satellite, a polar orbiting sensor that produces sub-meter resolution imagery with a horizontal accuracy of 23 m (90 percent confidence interval; DigitalGlobe 2008). Where no high-resolution imagery is available, GE defaults to medium-resolution Landsat imagery. There is no detailed documentation publicly available on the process by which GE ingests high-resolution imagery; however, it is apparent that the overall strategy has been to reproject all images into a geographic (platte carré) projection using the WGS-84 datum. The high-resolution GE imagery is true color, with less dynamic range than the original underlying QuickBird, SPOT 5, or aerial photography images. GE's acquisition of high-resolution imagery is a continuous process, with sporadic reports of updates from both Google and the GE user community. There is no meta-data readily available regarding image acquisition dates, spectral transformations, or spatial interpolations. When a GE user zooms into a sufficiently fine level of detail, the name of the imagery provider is sometimes displayed together with a year, but because there is no global vector layer describing scene boundaries, it remains impossible to ascertain acquisition dates with confidence.

To characterize the horizontal positional accuracy of the high-resolution GE archive, we compare the locations of 436 control points in the GE imagery to their equivalent positions in the Landsat GeoCover data set, which has positional accuracy of 50 m root-mean-squared error (RMSE; Masek et al. 2001; Tucker et al. 2004). In an ideal assessment of spatial accuracy, we would determine the position of these control points through a global ground-based campaign using global positioning satellites (GPS). However, GeoCover images allow us to more rapidly and inexpensively assess global horizontal positional

accuracy. In our clustered sampling scheme, we select an average of four points in each of 109 cities. By clustering around cities, we took advantages of GE's city-focused image acquisition strategy. The 109 cities we selected are a sub-set of a pre-existing random 120-city sample that was stratified based on world-region, national income, and city-size (Angel et al. 2005). The Landsat imagery for these cities was either provided as part of holdings related to the Angel et al. urban study, or was downloaded directly from the Global Land Cover Facility (GLCF 2007). The GE imagery was downloaded in November 2007.

The 120 cities that we examined are depicted in Figure A.2, where it is clear that 11 cities (9 percent) had only medium-resolution Landsat GeoCover imagery available on GE, and the remaining 109 had high-resolution imagery available for at least half of the city area. Only the 109 cities with high-resolution GE imagery were included in our study. For these cities, we began by identifying control points that could be readily located within the medium-resolution reference images. We generated a GE keyhole markup language (kml) file with the centroids of the 109 cities, and added approximately four control points per city to this kml file using standard GE tools.

The following attributes guided our process for selecting control points: (a) strong vegetative/non-vegetative contrast, (b) large enough to be apparent in 28.5 m Landsat images, (c) long and linear in form, preferably cross-shaped so that location can be constrained on more than one axis, (d) thin enough to involve only one or two medium-resolution pixels, (e) located within a relatively stable region that would be unlikely to undergo land cover change in the period separating the Landsat and GE image acquisitions. Some examples of useful features include: airport taxiways, road intersections, canal crossings, bridges, and monuments. Features that satisfied most of the criteria, but were avoided because of the potential for land cover change included: agricultural field boundaries, forest logging operations, river features, and new construction events. Once all

of the points were added to the kml file, we converted the kml file to an ArcGIS point shapefile.

In order to locate these control points within the Landsat GeoCover imagery, we used ArcGIS to view clips of the full Landsat scenes in a 4-5-3 band combination (RGB; Landsat bands 453; near-infrared, short-wave infrared, red) and standard linear stretches (Figure A.3). An Arc script linked the view-extent of the GE and ArcGIS windows. Control points were identified independently in the Landsat image, added to a shapefile for each city, and re-projected into a geographic projection. These Landsat-derived points were then compared to the equivalent control points from the GE imagery. For each point-pair, an error vector was estimated, with the Landsat control point as the origin and the GE control point as the terminus. For each of these vectors, a direction and magnitude was estimated.

A.3 RESULTS AND DISCUSSION

The overall accuracy of the full sample (436 control points) is 39.7 m RMSE (Table A.1; Figure A.4a), with a range of 0.4 to 171.6 m. The city with the lowest mean offset is Pittsburgh, USA, and the highest is Anqing, China (5.4 and 163.3 m, respectively). Only five cities contain control points with errors greater than 100 m. The 48 control points derived from aerial photography have an accuracy of 41.3 m RMSE, which is significantly poorer than the accuracy of the satellite-derived control points (22.8 m RMSE; t-test p-value < 0.01). Developed country control points (Europe, United States, Canada, Japan, Australia, and New Zealand) have an accuracy of 24.1 m RMSE, which is significantly better than the developing country accuracy (44.4 m RMSE; t-test p-value < 0.01). There are no apparent patterns in the direction of the error vectors (Figure A.4b), and no significant correlations between the magnitude or direction of the error vectors and the latitude of the control points. Because details of the GE geo-registration process are not publicly available, it is not possible to speculate as to the underlying cause of trends in the magnitude of error vectors.

However, these findings indicate that the uncertainty surrounding the position of GE imagery (39.7 m RMSE) is less than that of the Landsat GeoCover dataset (50 m RMSE). Conservatively, GE horizontal accuracy is 89.7 m RMSE (the sum of the Landsat and GE errors), adequate for GE high-resolution imagery to be of use in analyzing moderate-resolution remote sensing imagery and imagery-derived products.

This geospatial analysis removes one important obstacle to using GE imagery for science applications, yet the spatial resolutions, spectral attributes, and temporal qualities of the GE archive remain uncharacterized. The open nature of the GE archive means that it is possible to build data mining tools that would allow one to map the spatial and spectral properties of the full archive, however there is no similar opportunity for reducing the temporal unknowns. These unknowns in the dates of image acquisitions are of particular concern to the land-cover and land-use change (LCLUC) community. Because the vast majority of Google's high-resolution imagery comes from the QuickBird satellite (launched October 2001), we can place some bounds on image acquisitions, but as time passes this window continues to widen. Unless Google provides vector-based meta-data on their high-resolution GE imagery archive, these unknowns will continue to limit many promising scientific applications.

The Land Cover Validation Tool (LCVT) is one example of the kind of science applications that are made possible with the GE high-resolution archive (Figure A.5). LCVT was designed to facilitate the assessment of global land cover and human population maps, the training of supervised land cover classifiers, and the development of land cover change detection algorithms. This web-based application allows users to select a global, random, stratified sample of validation sites. These sites are portrayed with the Google Earth high-resolution imagery archive through the use of the Google Maps Application Programming Interface (API). In this example, the sites are 13-hectare hexagons (visible in red in Figure A.5) that were created using Discrete Global Grid software (Sahr et al. 2003). There are 100

sites for each of the 120 cities in the Angel et al. (2005) global sample, and these sites are stratified based on the fraction of urban area within each hexagon according to a library of Landsat-derived urban maps. Because the tool is web-based, multiple users can simultaneously assess sites using the web-form on the right of Figure A.5. User responses are collected in real-time as part of an online database. Although the Google Maps API is an excellent new resource for the global change community, the power of applications such as the LCVT would be greatly enhanced if Google opted to share temporal, spatial, and spectral meta-data on their high-resolution imagery holdings.

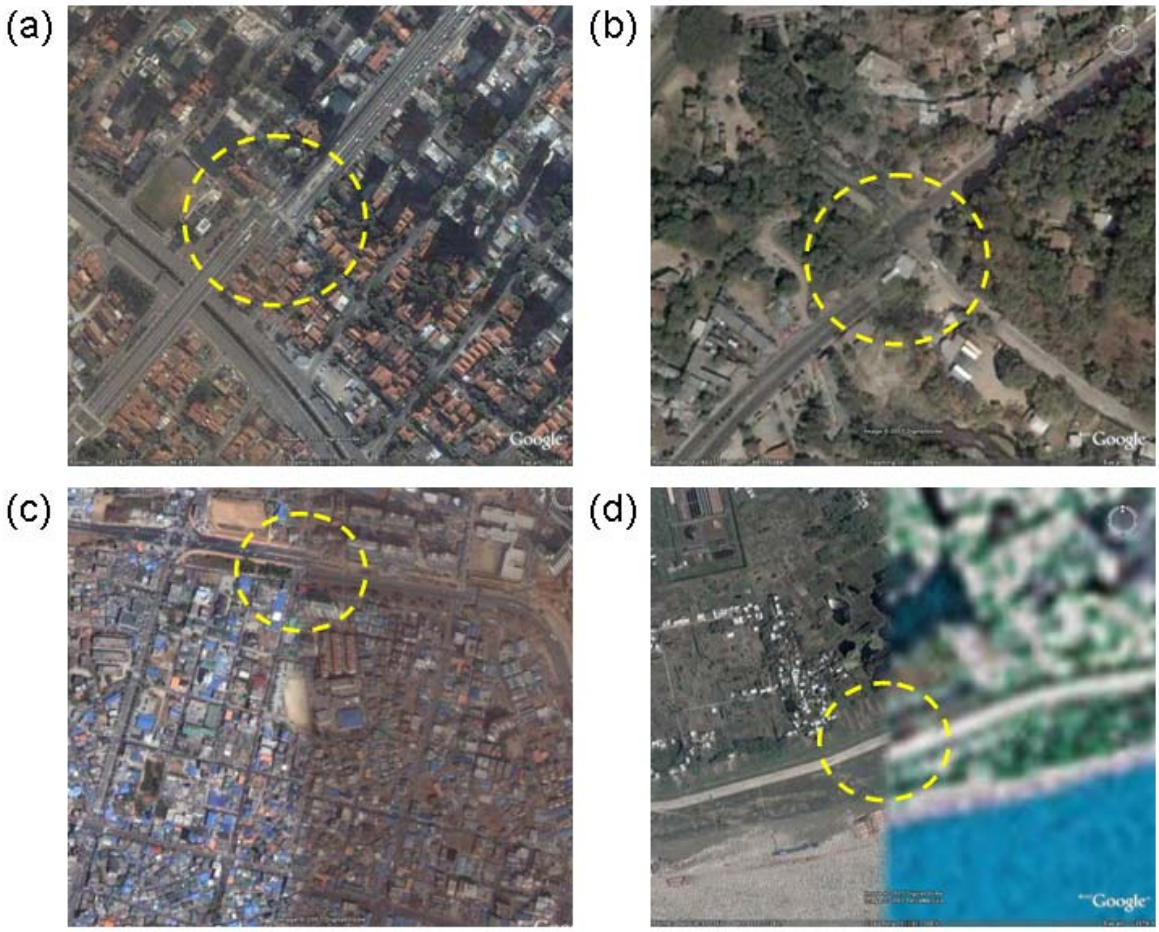


Figure A.1 Apparent geo-registration problems in Google Earth imagery. Problem areas are highlighted with dashed yellow circles. Plots are for **(a)** Sao Paolo, Brazil, **(b)** San Salvador, El Salvador, **(c)** Chonan, South Korea, and **(d)** Anqing, China. The Anqing example includes both Landsat (right side) and Quickbird imagery (left side).

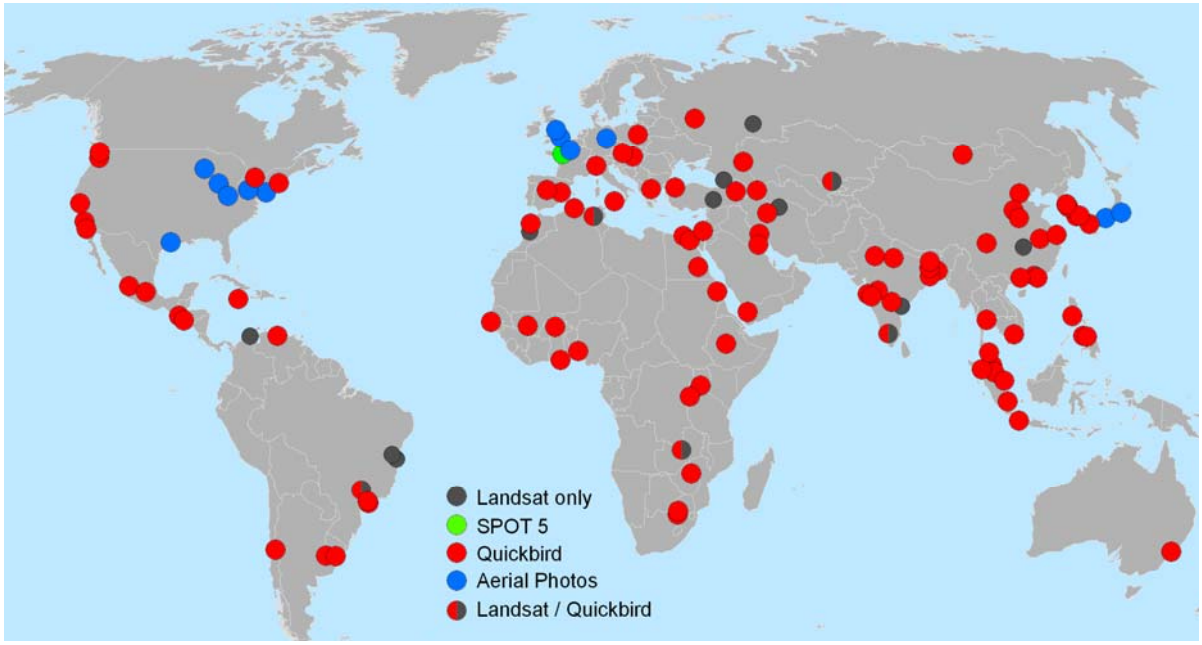


Figure A.2 Map of imagery sources for 120 cities. This figure depicts the 120-city global urban sample from Angel et al. (2005). Colored circles mark cities where high-resolution Google Earth imagery was available (109 cities) and gray circles mark cities where only Landsat GeoCover base imagery was available (11 cities). Only the 109 cities with high-resolution imagery were included in this study.

	Root Mean Squared Error (RMSE) (meters)	Mean Error (meters)	Standard Deviation (σ) (meters)	Range (meters)	Control Points (N)	Sample Cities
Other Developed Countries	22.6	19.0	12.3	(0.9 - 53.0)	64	16
Europe	25.7	23.1	11.4	(1.6 - 46.0)	60	15
Latin America and Caribbean	41.4	30.9	27.9	(0.4 – 111.7)	54	14
Western and SC Asia	42.3	31.5	28.5	(1.4 – 115.0)	78	19
SE and East Asia	45.9	32.7	32.4	(1.0 – 171.6)	108	27
Africa	46.2	35.7	29.6	(2.7 – 160.9)	72	18
Total	***39.7***	***29.4***	***26.6***	***(0.4 – 171.6)***	***436***	***109***

Table A.1 Error vector magnitudes by world region and globally. 'Other developed countries' refers to the United States, Canada, Australia, New Zealand, and Japan.

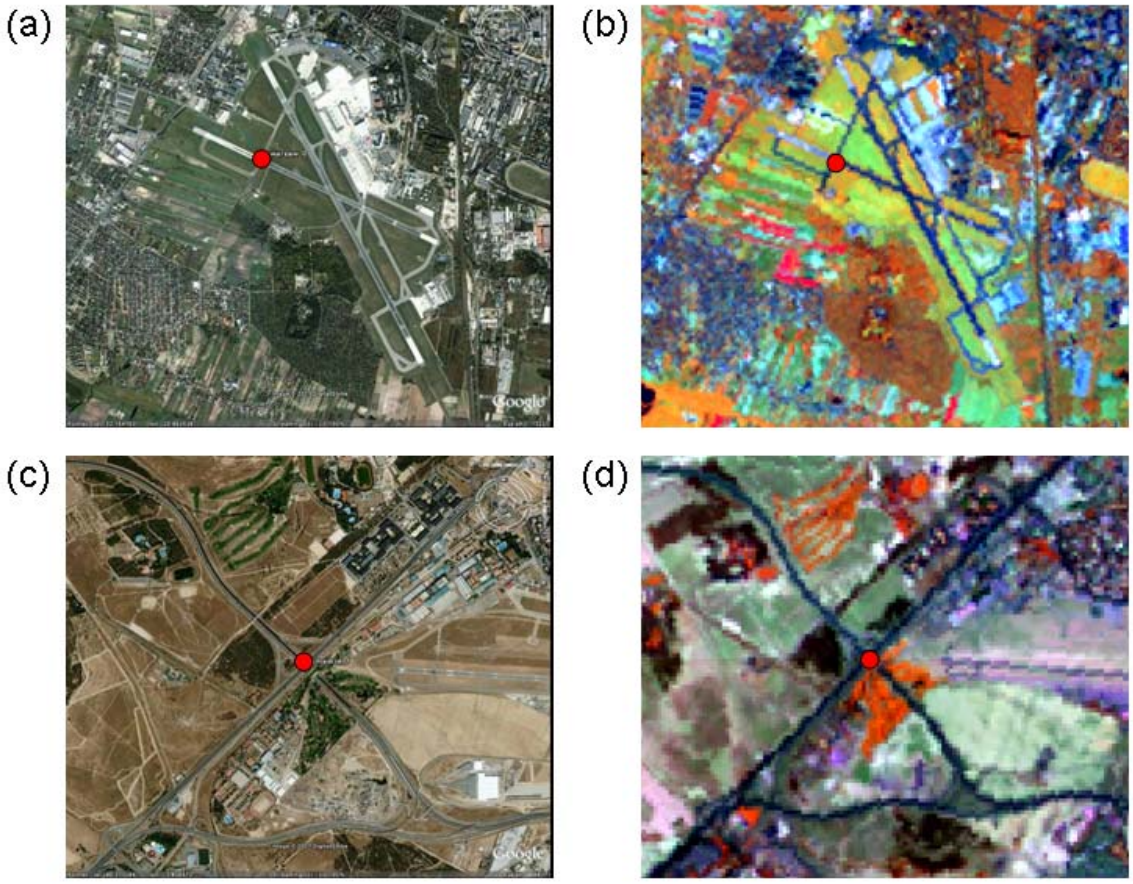


Figure A.3 Google Earth and Landsat control points. Google Earth **(a, c)** and Landsat **(b, d)** control points for an airport in Warsaw, Poland **(a, b)**, and a highway intersection in Madrid, Spain **(c, d)**. Control point pairs appear as red circles; the Warsaw and Madrid windows are 5 km and 3km in width, respectively.

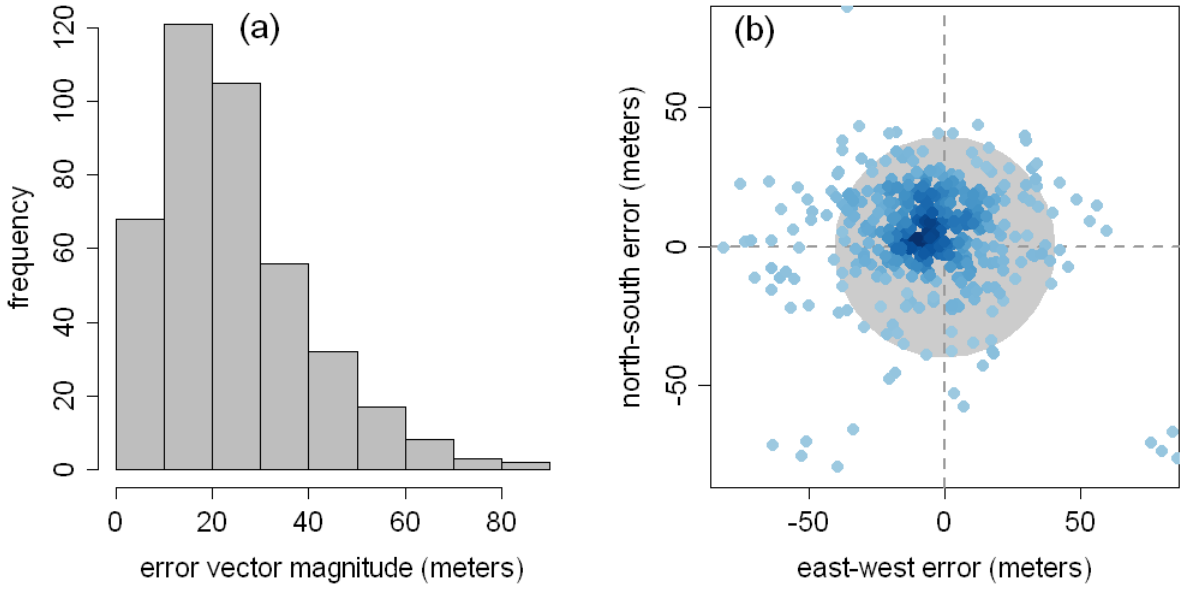


Figure A.4 Error vectors for 412 control sites. **(a)** Magnitude of horizontal error vectors for 412 control points (points from six cities with mean errors greater than 1 standard deviation from the global mean are not plotted). **(b)** Error vectors for 413 control points (23 large-magnitude error vectors are not plotted). The gray crosshair marks the origin of the error vectors (Landsat GeoCover), the blue density-shaded dots are termini (GE), and the gray ring describes the overall accuracy (39.7 meters RMSE).

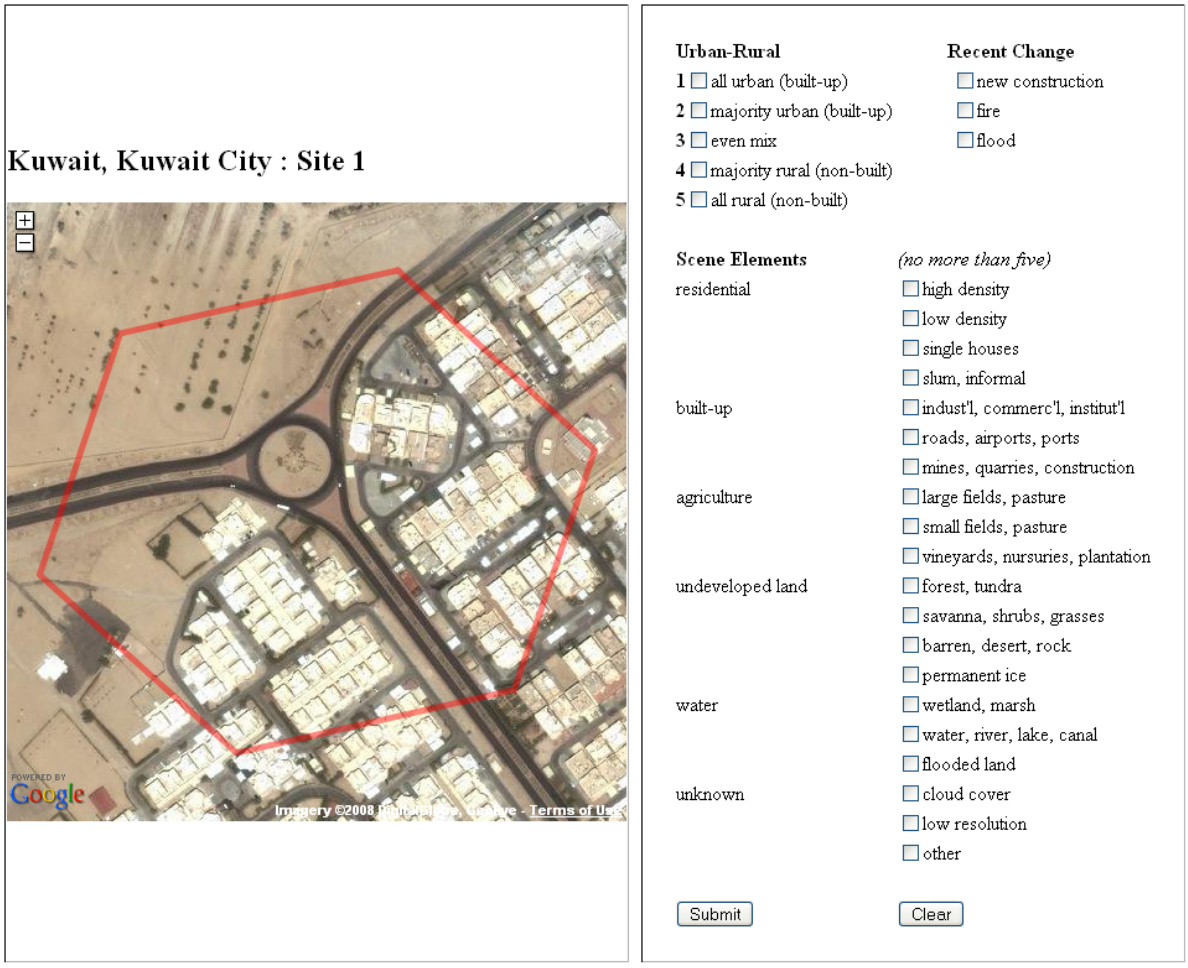


Figure A.5 A screen-shot from the Land Cover Validation Tool v 1.0. The site covers 13-hectares in Kuwait City, Kuwait. Entries from the right panels are captured in an online database. The imagery is served by the Google Maps application programming interface (API).

BIBLIOGRAPHY

Alberti, M., 2005. The effects of urban patterns on ecosystem function, *International Regional Science Review* 28:2, 168-192.

Allouche, O., A. Tsoar, and R. Kadmon, 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS), *Journal of Applied Ecology* 43, 1223-1232.

Angel, S., S.C. Sheppard, and D.L. Civco, 2005. *The Dynamics of Global Urban Expansion*, The World Bank, Washington, DC. (accessed April 2007, http://www.williams.edu/Economics/UrbanGrowth/WorkingPapers.htm).

Arino, O., D. Gross, F. Ranera, L. Bourg, M. Leroy, P. Bicheron, J. Latham, A. Di Gregorio, C. Brockman, R. Witt, P. Defourny, C. Vancutsem, M. Herold, J. Sambale, F. Achard, L. Durieux, S. Plummer, and J. L. Weber, 2007. GlobCover: ESA service for global land cover from MERIS, International Geoscience and Remote Sensing Symposium, 23-28 July 2007, Barcelona, Spain. doi: 10.1109/IGARSS.2007.4423328

Balmford, A., J. L. Moore, T. Brooks, N. Burgess, L. A. Hansen, P. Williams, and C. Rahbek, 2001. Conservation conflicts across Africa, *Science* 291, 2616-2619.

Bartholome, E., and A.S. Belward, 2005. GLC2000: a new approach to global land cover mapping from Earth observation data, *International Journal of Remote Sensing* 26, 1959-1977.

Bengtsson, M., Y. Shen, and T. Oki, 2007. A SRES-based gridded population dataset for 1990-2100, *Population and Environment* 28:113-131.Bhaduri, B., Bright, E., Coleman, P., and Dobson, J., 2002. LandScan: locating people is what matters, *Geoinfomatics*, 5: 34-37.

Berry, B.J.L. 1961. City size distributions and economic development, *Economic Development and Cultural Change* 9:4, 573-588.

Bhaduri, B., E. Bright, P. Coleman, and J. Dobson, 2002. LandScan: locating people is what matters, *Geoinfomatics* 5, 34-37.

Bourne, J., 2000. Louisiana's vanishing wetlands: going, going…, *Science* 289, 1860-1863.

Brockerhoff, M, and E. Brennan, 1998. The poverty of cities in developing regions, *Population and Development Review* 24:1, 75-114.

Brockerhoff, M., 2000. An urbanizing world, *Population Bulletin* 55:3, 3-5.

Brooks, T.M., M.I. Bakarr, T. Boucher, G. A. B. da Fonseca, C. Hilton-Taylor, J. M. Hoekstra, T. Moritz, S. Olivieri, J. Parrish, R. L. Pressey, A.S.L. Rodrigues, W. Sechrest, A. Stattersfield, W. Strahm, and S.N. Stuart, 2004. Coverage provided by the global protected-area system: is it enough? *BioScience*, 54:12, 1081-1091.

Bruner, A.G., R.E. Gullison, and A. Balmford, 2004. Financial costs and shortfalls of managing and expanding protected-area systems in developing countries, *BioScience*, 54, 1119-1126.

Burgess, N.D., A. Balmford, N.J. Cordeiro, J. Fjeldsa, W. Kuper, C. Rahbek, E.W. Sanderson, J.P.W. Scharlemann, J.H. Sommer, and P.H. Williams, 2007. Correlations among species distributions, human density and human infrastructure across the high biodiversity tropical mountains of Africa, *Biological Conservation* 2:164-177.

Burgess, R., 2000, "The compact city debate: A global perspective," in Jenks M. and Burgess, R., eds., *Compact Cities: Sustainable Urban Forms for Developing Countries*, Spon Press, New York, New York.

Butler, D, 2006. The web-wide world, *Nature* 439, 776-779.

Calbo, J., W. Pan, M. Webster, R.G. Prinn, and G.J. McRae, 1998. Parameterization of urban sub-grid scale processes in global atmospheric chemistry models, *Journal of Geophysical Research* 103, 3437-3467.

Castells, M., 2000. *The rise of the network society*, Blackwell Publishers, Malden, Massachusetts.

Ceballos, G. and P.R. Ehrlich, 2006. Global mammal distributions, biodiversity hotspots, and conservation. *Proc. Natl. Acad. Sci.* 103:19, 374–379. (doi:10.1073/ pnas.0609334103)

Cha, S., and C. Park, 2007. The utilization of Google Earth images as reference data for the multitemporal land cover classification with MODIS data of North Korea, *Korean Journal of Remote Sensing* 23:5, 483-491.

Chape, S., J. Harrison, M. Spalding, and I. Lysenko, 2005. Measuring the extent and effectiveness of protected areas as an indicator for meeting biodiversity targets, *Phil. Trans. R. Soc. B* 360:1454, 443-455.

CIESIN (Center for International Earth Science Information Network), 2004. Global Rural-Urban Mapping Project (GRUMP), Alpha Version: Urban Extents. http://sedac.ciesin.columbia.edu/gpw (accessed April 15, 2007).

Cohen, B., 2004. Urban growth in developing countries: A review of current trends and a caution regarding existing forecasts, *World Development* 32:1, 23-51.

Cohen, J., 1960. A coefficient of agreement for nominal scales, *Educational and Psychological Measurement* 20, 37-46.

Compieta, P., S. Di Martino, M. Bertolotto, F. Ferrucci, and T. Kechadi, 2007. Exploring spatio-temporal data mining and visualization, *Journal of Visual Languages and Computing* 18, 255-279.

Congalton, R. G. and K. Green, 1999. *Assessing the Accuracy of Remotely Sensed Data: Principals and Practices*, CRC Press, New York, New York.

Constantine, J.A., and T. Dunne, 2008. Meander cutoff and the controls on the production of oxbow lakes, *Geology* 36(1), 23-26.

Danko, D.M., 1992. The Digital Chart of the World Project, *Photogrammetric Engineering and Remote Sensing* 58, 1125-1128.

Davies, R., C. Orme, V. Olson, G. Thomas, S. Ross, T. Ding, P. Rasmussen, A. Strattersfield, P. Bennett, T. Blackburn, I. Owens, and K. Gaston, 2006. Human impacts and the global distribution of extinction risk, *Proceedings of the Royal Society B* 273, 2127-2133.

Dobson, J.E., E.A. Bright, P.R. Coleman, R.C. Durfee, and B.A. Worley, 2000. Landscan: A global population database for estimating populations at risk. *Photogrammetric Engineering and Remote Sensing* 66, 849-857.

Doll, C.N.H., J.P. Muller, and J.G. Morley, 2006. Mapping regional economic activity from night-time light satellite imagery, *Ecological Economics* 57, 75-92.

Douglas, I., 1994. "Human settlements," in W. B. Meyer and B. L. Turner II, eds., *Changes in land use and land cover: a global perspective*. Cambridge University Press, Cambridge, United Kingdom.

Dye, C., 2008. Health and urban living, *Science* 319:766-769.

East, T., N. F. Kumpel, E. J. Milner-Gulland, and J. M. Rowcliffe, 2005. Determinants of urban bushmeat consumption in Rio Muni, Equatorial Guinea, *Biological Conservation* 126, 206-215.

Ehrlich, P.R., 1991. Population diversity and the future of ecosystems, *Science* 254, 175.

El Araby, M., 2002. Urban growth and environmental degradation, *Cities* 19, 389-400.

Ellis, E.C. and N. Ramankutty, 2008. Putting people in the map: anthropogenic biomes of the world, *Frontiers in Ecology and the Environment* 6:8, 439-447.

Elvidge, C., B. T. Tuttle, P.C. Sutton, K. E. Baugh, A. T. Howard, C. Milesi, B. L. Bhaduri, and R. Nemani, 2007. Global Distribution and Density of Constructed Impervious Surfaces, *Sensors* 7, 1962-1979.

Elvidge, C., M.L. Imhoff, K.E. Baugh, V.R. Hobson, I. Nelson, J. Safran, J.B. Dietz, and B.T. Tuttle, 2001. Nighttime lights of the world: 1994-95, *ISPRS Journal of Photogrammetry and Remote Sensing* 56, 81-99.

ESA (European Space Agency), 2008. *GlobCover Product Description Manual.* ftp://uranus.esrin.esa.int/pub/globcover_v2/GLOBCOVER_PDM_I2.1.pdf

Fielding, A.H. and J. F. Bell, 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models, *Environmental Conservation* 24, 38-49.

Fjeldsa, J. and N. D. Burgess, 2008. The coincidence of biodiversity patterns and human settlement in Africa, *Afr. J. Ecol.* 46, 33-42.

Fleiss, J.L., J. Cohen, and B.S. Everitt, 1969. Large sample standard errors of kappa and weighted kappa, *Psychological Bulletin* 72:5, 323-327.

Foley, J.A., R. DeFries, G.P. Asner, C. Barford, G. Bonan, S.R. Carpenter, F.S. Chapin, M.T. Coe, G.C. Daily, H.K. Gibbs, J.H. Helkowski, T. Holloway, E.A. Howard, C.J. Kucharik, C. Monfreda, J.A. Patz, I.C. Prentice, N. Ramankutty, and P.K. Syder, 2005. Global Consequences of Land Use, *Science* 309, 570-574.

Folke, C., A. Jansson, J. Larsson, and R. Costanza, 1997. Ecosystem appropriation by cities, *Ambio* 26, 167-172.

Foody, G. M., 2006. What is the difference between two maps? A remote senser's view, *Journal of Geographical Systems* 8, 119-130.

Foody, G. M., 2007. Map comparison in GIS, *Progress in Physical Geography* 31, 439-445.

Forbes, A.D., 1995. Classification algorithm evaluation: five performance measures based on confusion matrices, *Journal of Clinical Monitoring* 11, 189-206.

Friedl, M., D. McIver, J. Hodges, X. Zhang, D. Muchoney, A. Strahler, C. E. Woodcock, S. Gopal, A. Schneider, A. Cooper, A. Baccini, F. Gao, and C. Schaaf, 2002. Global land cover mapping from MODIS: algorithms and early results, *Remote Sensing of Environment* 83, 287–302.

Gaffin, S.R., C. Rosenzweig, X. Xing, and G. Yetman, 2004. Downscaling and geospatial gridding of socio-economic projections from the IPCC Special Report on Emissions Scenarios (SRES), *Global Environmental Change* 14: 105-123.

Gesch, D.B., and K. S. Larson, 1996. Techniques for development of global 1-kilometer digital elevation models. In: Pecora Thirteen, Human Interactions with the Environment - Perspectives from Space, Sioux Falls, South Dakota, August 20-22, 1996.

Giri, C., Z.L. Zhu, and B. Reed, 2005. A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets, *Remote Sensing of the Environment* 94, 123-132.

GLCF (Global Land Cover Facility), 2007. http://glcf.umiacs.umd.edu (accessed February 2007).

Goldewijk, K. and N. Ramankutty, 2004. Land cover change over the last three centuries due to human activities: the availability of new global data sets, *GeoJournal* 61, 335-344.

Goldewijk, K., 2001. Estimating global land use change over the past 300 years: the HYDE database, *Global Biogeochemical Cycles* 15, 417-434.

Goldewijk, K., 2005. Three centuries of global population growth: A spatially referenced population density database for 1700 – 2000, *Population and Environment* 26, 343-367.

Google, Inc. "Press Release: Google Announces Major Imagery Update for Google Earth, New Tools and Innovations in Mapping," 12 June 2006, Mountain View, California, USA. (available at: www.google.copm/press/pressrel/geoday.html)

Google, Inc. "Press Release: Introducing Google Earth Outreach," 26 June 2007, Mountain View, California, USA. (available at: www.google.copm/press/pressrel/outreach_20070625.html)

Greyner, R., C. Orme, S. Jackson, G. Thomas, R. Davies, T. Davies, K. Jones, V. Olson, R. Ridgely, P. Rasmussen, T. Ding, P. Bennett, T, Blackburn, K. Gaston, J. Gittleman, and I. Owens, 2006. Global distribution and conservation of rare and threatened vertebrates, *Nature* 444:2, 93-96.

Grimm, N.B., S.H. Faeth, N.E. Golubiewski, C. L. Redman, J. Wu, X. Bai, and J.M. Briggs, 2008. Global change and the ecology of cities, *Science* 319: 756-760.

Grübler, A., 1990. "Technology," in W. B. Meyer and B. L. Turner II, eds., *Changes in land use and land cover: a global perspective*. Cambridge University Press, Cambridge, United Kingdom.

Grubler, A., B. O'Neill, K. Riahi, V. Chirkov, A. Goujon, P. Kolp, I. Prommer, S. Scherbov, and E. Slentoe, 2007. Regional, national, and spatially explicit scenarios of demographic and economic change based on SRES, *Technological Forecasting and Social Change* 74: 980-1029.

Guralnick, R.P., A.W. Hill, and M. Lane, 2007. Towards a collaborative, global infrastructure for biodiversity assessment, *Ecology Letters* 10, 663-672.

Hansen, M.C. and B. Reed, 2000. A comparison of the IGBP DISCover and University of Maryland 1 km global land cover products, *International Journal of Remote Sensing* 21, 1365–1373.

Hansen, M.C., R.S. Defries, R.G. Townsend, and R. Sohlberg, 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach, *International Journal of Remote Sensing* 21, 1331-1364.

Henderson, M., E.T. Yeh, P. Gong, C. Elvidge, and K. Baugh, 2003. Validation of urban boundaries derived from global night-time satellite imagery, *International Journal of Remote Sensing* 24, 595-609.

Herold, M., N.C. Goldstein, and K.C. Clarke, 2003. The spatiotemporal form of urban growth: measurement, analysis and modeling, *Remote Sensing of Environment* 86, 286-302.

Herold, M., P. Mayaux, C.E. Woodcock, A. Baccini, and C. Schmullius, 2008. Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets, *Remote Sensing of the Environment* 112:5, 2538-2556.

Hoekstra, J.M., T.M. Boucher, T.H. Ricketts, C. Roberts, 2005. Confronting a biome crisis: global disparities of habitat loss and protection, *Ecology Letters* 8:1, 23-29.

IIASA (International Institute for Applied System Analysis), 2007. Greenhouse Gas Initiative (GGI) Scenario Database (http://www.iiasa.ac.at/Research/GGI/DB/).

Imhoff, M.L., L. Bounoua, T. Ricketts, R. Harriss, and W.T. Lawrence, 2004. Global patterns in human consumption of net primary production, *Nature* 429, 870-873.

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 accurately map urban areas, *Remote Sensing of Environment* 61, 361-370.

IPCC (Intergovernmental Panel on Climate Change), 2007. *Working Group II Report: Climate Change 2007, Impacts, Adaptation and Vulnerability – Summary for Policymakers*, IPCC Secretariat, Geneva, Switzerland.

IRARS (Imagery Resolution Assessment and Reporting Standards Committee), 1996. *Civil National Imagery Interpretability Rating Scale (NIIRS) Reference Guide* (available at: http://www.fas.org/irp/imint/niirs_c/index.html).

IUCN (International Union for the Conservation of Nature), 1994. *Guidelines for Protected Areas Management Categories,* IUCN, Cambridge, United Kingdom.

Jackson, S. F. and K. J. Gaston, 2008. Land use change and the dependence of national priority species on protected areas, *Global Change Biology* 14, 2132-2138.

Jensen, J.R. and D.L. Toll, 1982. Detecting residential land use development at the urban fringe, *Photogrammetric Engineering and Remote Sensing* 48, 629–643.

Jensen, J. R. and D. C. Cowen, 1999. Remote Sensing of Urban/Suburban Infrastructure and Socio-economic Attributes, *Photogrammetric Engineering & Remote Sensing* 65:5, 611-622.

Jetz, W., D.S. Wilcove, and A.P. Dobson, 2007. Projected impacts of climate and land-use change on the global diversity of birds, *PLoS Biol*. 5, 1211–1219.

Jetz, W., H. Kreft, G. Ceballos, and J. Mutke, 2008. Global associations between terrestrial producer and vertebrate consumer diversity, *Proc. R. Soc. B*, doi:10.1098/rspb.2008.1005

Joppa, L.N., S.R. Loarie, and S.L. Pimm, 2008. On the protection of "protected areas," *Proc. Natl. Acad. Sci. USA* 105:18, 6673-6678.

Jung, M., K. Henkel, M. Herold, and G. Churkina, 2006. Exploiting synergies of global land cover products for carbon cycle modeling, *Remote Sensing of Environment* 101, 534-553.

Kaye, J., P.M. Groffman, N.B. Grumm, L.A. Baker, and R.V. Pouyat, 2006. A distinct urban biogeochemistry? *Trends in Ecology and Evolution* 21:4, 192-198.

Keiser, J., J. Utzinger, M.C. De Castro, T.A. Smith, M. Tanner, and B.H. Singer, 2004. Urbanization in sub-Saharan Africa and implication for malaria control, *American Journal of Tropical Medicine and Hygiene* 71:2, 118-127.

Kendall, M., 1938. A New Measure of Rank Correlation, *Biometrika* 30, 81-89.

Landis, J. R. and G. G. Koch, 1977. The measurement of observer agreement for categorical data, *Biometrics* 33, 169-174.

Latifovic, R. and I. Olthof, 2004. Accuracy assessment using sub-pixel fractional error matrices of global land cover products derived from satellite data, *Remote Sensing of Environment*, 90:153-165.

Leachtenauer, J.C., K. Daniel, T.P. Vogl, 1997, "Digitizing Corona imagery: Quality vs. cost," In R.A. McDonald ed., *Corona: Between the Sun & the Earth, The first NRO reconnaissance eye in space*, American Society Photogrammetry and Remote Sensing, Washington, D.C., USA, 189-203.

Lee, T. M. and W. Jetz, 2008. Future battlegrounds for conservation under global change, *Proc. R. Soc. B* 275, 1261-1270.

Liu, C., P. Frazier, and L. Kumar, 2007. Comparative assessment of the measures of thematic classification accuracy, *Remote Sensing of the Environment* 107, 606-616.

Lobo, J.M., A. Jimenez-Valverde, and R. Real, 2007. AUC: a misleading measure of the performance of predictive distribution models, *Global Ecology and Biogeography* 17, 145-151.

Lockwood, M., G.L. Worboys, and A. Kothari, 2006. *Managing protected areas: a global guide,* Earthscan, London.

Longley, P.A. and V. Mesev, 2000. On the measurement and generalization of urban form, *Environment and Planning A* 32, 473–488.

Loucks, C.T.H. Ricketts, R. Naidoo, J. Lamoreux and J. Hoekstra, 2008. Explaining the global pattern of protected area coverage: relative importance of vertebrate biodiversity, human activities and agricultural suitability, *J. Biogeogr*. 35, 1337-1348.

Loveland, T.R., B.C. Reed, J. F. Brown, D.O. Ohlen, J. Zhu, L. Yang, and J.W. Merchant, 2000. Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR data, *International Journal of Remote Sensing* 21, 1303-1330.

Luck, A.L., G.D. Jenerette, J. Wu, and N.B. Grimm, 2001. The urban funnel model and the spatially heterogeneous ecological footprint, *Ecosystems* 4:8, 782-796.

Luck, G.W. 2007. A review of the relationships between human population density and biodiversity, *Biological Reviews* 82, 607-645.

Lutz, W. ed, 1996. *The future population of the world: what can we assume today?* Earthscan, London, United Kingdom.

Lutz, W.W. Sanderson, and S. Scherbov, 2001. The end of world population growth, *Nature* 412: 543-545.

Manel, S., H.C. Williams, and S.J. Ormerod, 2001. Evaluating presence-absence models in ecology: the need to account for prevalence, *Journal of Applied Ecology* 38, 921-931.

Masek, J., M. Honzak, S.N. Goward, P. Liu, and E. Pak, 2001. Landsat-7 ETM+ as an observatory for land cover: Initial radiometric and geometric comparisons with Landsat-5 Thematic Mapper, *Remote Sensing of the Environment* 78, 118-130.

Massey, D., 1996. The age of extremes: concentrated affluence and poverty in the twenty-first century, *Demography* 33:4, 395-412.

Massey, D., 2005. *Strangers in a Strange Land: Humans in an Urbanizing World*, W.W. Norton & Company, New York, New York.

Mayaux, P., H. Eva, J. Gallego, A.H. Strahler, M. Herold, S. Agrawal, S. Naumov, E. Miranda, C. Bella, C. Ordoyne, Y. Kopin, and P. Roy, 2006. Validation of the Global Land Cover 2000 map, *IEEE Transactions on Geoscience and Remote Sensing* 44, 1728-1739.

McCallum, I., M. Obersteiner, S. Nilsson, and A. Shvidenko, 2006. A spatial comparison of four satellite derived 1 km global land cover datasets, *International Journal of Applied Earth Observation and Geoinformation* 8, 246-255.

Mcdonald, R. I., P. Kareiva, and R.T.T. Forman, 2008. The implications of current and future urbanization for global protected areas and biodiversity conservation, *Biological Conservation* 141, 1695-1703.

McGranahan, G. and D. Satterthwaite, 2002. The environmental dimensions of sustainable development for cities, *Geography* 87, 213-226.

McGranahan, G. and Satterthwaite, D., 2000. "Environmental health of ecological sustainability: Reconciling the brown and green agendas in urban development", in Pugh, C., ed., *Sustainable Cities in Developing Countries: Theory and Practice at the Millennium*, Earthscan Publishers, London, United Kingdom, 53-72.

Mckinney, M. L., 2001. Role of human population size in raising bird and mammal threat among nations, *Animal Conservation* 4, 45-57.

Mckinney, M. L., 2002. Urbanization, Biodiversity, and Conservation, *BioScience* 52:10, 883-890.

McPherson, J. M., W. Jetz, and D.J. Rogers, 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? J*ournal of Applied Ecology* 41, 811-823.

Milesi, C., C.D. Elvidge, R.R. Nemani, and S.W. Running, 2003. Assessing the impact of urban land development on net primary productivity in the southeastern United States, *Remote Sensing of the Environment* 86:3, 401-410.

MODIS (Moderate Resolution Imaging Spectroradiometer) Land Cover Team, 2003. *MODIS MOD12 Land Cover and Land Cover Dynamics Products User Guide*, V003 Validation: Validation of the Consistent-year V003 MODIS Land Cover Product (last accessed June 2008, http://www-modis.bu.edu/landcover/userguidelc/consistent.htm).

Monserud, R, and R. Leemans, 1992. Comparing global vegetation maps with the Kappa statistic, *Ecological Modeling* 62, 275-293.

Montgomery, M., 2008. The urban transformation of the developing world, *Science* 319, 761-764.

Montgomery, M., R. Stren, B. Cohen, and H. Reed, 2003. *Cities Transformed: Demographic Change and Its Implications in the Developing World*, National Academies Press, Washington, D.C.

Morris, D.W., S.R. Kingston, 2002. Predicting future threats to biodiversity from habitat selection by humans, *Evolutionary Research* 4, 787-810.

Myers, N., R. Mittermeier, C. Mittermeier, G. da Fonseca, and J. Kent, 2000. Biodiversity hotspots for conservation priorities, *Nature* 403:24, 853-858.

Nagendra, H., 2008. Do parks work? Impact of protected areas on land cover clearing, *Ambio* 37:5, 330-337.

Nakicenovic, N., and R. Stuart, eds., 2000. *Special report on emissions scenarios: A special report of working group III of the intergovernmental panel on climate change*, Cambridge University Press, Cambridge, United Kingdom.

NGDC (National Geophysical Data Center), 2007 (Nighttime Lights data available at: *http://www.ngdc.noaa.gov/dmsp/sensors/ols.html*).

Nishiyama, Y., S. Osada, and Y. Sato, 2008. OLS estimation and the t-test revisited in rank-size rule regression, *Journal of Regional Science* 48:4, 691-715.

Oke, T.R., 1982. The energetic basis of the urban heat island, *Quarterly Journal of the Royal Meteorological Society* 108, 1-24.

Oleson, K. W., G. B. Bonan, J. Feddema, and M. Vertenstein, 2008. An urban parameterization for a global climate model. Part I: Formulation and evaluation for two cities, *Journal of Applied Meteorology and Climatology* 47:4, 1038-1060.

Olson, D.M., E. Dinerstein, E.D. Wikremanayake, N.D. Burgess, G.V.N. Powell, E.C. Underwood, J.A. D'Amico, I. Itoua, H.E. Strand, J.C. Morrison, C.J. Loucks, T.F. Allnutt, T.H. Ricketts, T. Kura, J.F. Lamoreux, W.W. Wttengel, P. Hedao, and K.R. Kassem, 2001. Terrestrial ecoregions of the world: a new map of life on Earth, *Bioscience* 51, 933-938.

Ozdogan, M., and C.E. Woodcock, 2006. Resolution dependent errors in remote sensing of cultivated areas, *Remote Sensing of the Environment* 103, 203-217.

Panoramio, "Press Release: One million registered users and five million photos uploaded," 25 Oct 2007. (available at: http://www.panoramio.com/blog/1-million-registered-users-and-5-million-photos-uploaded/)

Pauchard, A, M. Aguayo, E Peria, and R. Urrutia, 2006. Multiple effects of urbanization on the biodiversity of developing countries: The case of a fast-growing metropolitan area (Concepcion, Chile), *Biological Conservation* 127, 272-281.

Pickett, S., W. Burch, S. Dalton, T. Foresman, M. Grove, and R. Rowntree, 1997. A conceptual framework for the study of human ecosystems in urban areas, *Urban Ecosystems* 1, 186-199.

Pickett, S.T.A., M.L Cadenasso, J.M. Grove, P.M. Groffman, L.E. Band, C.G. Boone, W.R. Burch, C.S.B. Grimmond, J. Hom, J.C. Jenkins, N.L. Law, C.H. Nilon, R.V. Pouyat, K. Szlavecz, P.S. Warren, and M.A. Wilson, 2008. Beyond urban legends: An emerging framework of urban ecology, as illustrated by the Baltimore Ecosystem Study, *Bioscience* 58:2, 139-150.

Pimm, S.L., G.J. Russel, J.L. Gittleman, T.M. Brooks, 1995. The future of biodiversity, *Science* 269, 347-350.

Pimm, S.L., P. Raven, A. Peterson, C.H. Sekercioglu, and P.R. Ehrlich, 2006. Forest losses predict bird extinctions in eastern North America, *Proc. Natl. Acad. Sci. USA* 103, 10941-10946.

Potere, D., forthcoming. The horizontal positional accuracy of Google Earth's high-resolution imagery archive, *Sensors*.

Potere, D., and A. Schneider, 2008. Urbanization and the global network of protected areas, Population Association of America Annual Meeting, New Orleans, Louisiana, April 2008.

Potere, D., and A. Schneider, 2007. A critical look at representations of urban areas in global maps, *GeoJournal* 69, 55-80.

Potere, D., and A. Schneider, forthcoming, "Comparison of global urban maps," Chapter 13 in *Global Mapping of Human Settlement: Experiences, Data Sets, and Prospects*, Taylor and Francis, New York, New York.

Quigley, J.M., 1998. Urban diversity and economic growth, *Journal of Economic Perspectives* 37, 426-434.

Ramankutty, N., A.T. Evan, C. Monfreda, and J.A. Foley, 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000, *Global Biogeochemical Cycles* 22, doi: 10.1029/2007GB002952.

Rees, W.E., 1992. Ecological footprint and appropriated carrying capacity: what urban economics leaves out, *Environment and Urbanization* 4, 121-130.

Reuter, H.I., A. Nelson, and A. Jarvis, 2007. An evaluation of void filling interpolation methods for SRTM data, *International Journal of Geographic Information Science* 21:9, 983-1008.

Riahi, K., A. Grubler, and N. Nakicenovic, 2007. Scenarios of long-term socio-economic and environmental development under climate stabilization, *Technological Forecasting and Social Change* 74, 887-935.

Ricketts, T., and M. Imhoff, 2003. Biodiversity, urban areas, and agriculture: locating priority ecoregions for conservation, *Conservation Ecology* 8, 1-15.

Ridd, M. K. 1995. Exploring a V-I-S (Vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing – comparative anatomy for cities, *International Journal of Remote Sensing* 16, 2165-2185.

Rodrigues, A., S.J. Andelman, M.I. Bakarr, L. Boitani, T.M. Brooks, R.M Cowling, L.D.C. Fishpool, G. A. B de Fonseca, K. J. Gaston, M. Hoffmann, J.S. Long, P.A. Marquet, J.D. Pilgrim, R.L. Pressey, J. Schipper, W. Sechrest, S.N. Sturart, L.G. Underhill, R.W. Waller, M. E. J. Watts, and X. Yan, 2004. Effectiveness of the global protected area network in representing species biodiversity, *Nature* 428:8, 640-643.

Sahr, K., D. White, and A. Kimerling, 2003. Geodesic Discrete Global Grid Systems, *Cartography and Geographic Information Science* 30, 121-134.

Sassen, S., 1994. *Cities in a World Economy*, Pine Forge-Sage Press, Thousand Oaks, California.

Schaaf, C.B., F. Gao, A.H. Strahler, W. Lucht, X. Li, T. Tsang, N. Strugnell, X. Zhang, Y. Jin, J. Muller, P. Lewis, M. Barnsley, P. Hobson, M. Disney, G. Roberts, M. Dunderdale, C. Doll, R. d'Entremont, B. Hu, S. Liang, J.L. Privette, and D. Roy, 2002. First operational BRDF, albedo nadir reflectance products from MODIS, *Remote Sensing of the Environment* 83, 135-148.

Schneider, A. and C.E. Woodcock, 2008. Compact, dispersed, fragmented, extensive? A comparison of urban expansion in twenty-five global cities using remotely sensed data, pattern metrics and census information, *Urban Studies* 45, 659-692.

Schneider, A., M. A. Friedl and D. Potere, forthcoming. A new map of global urban extent from MODIS data, *Geophysical Research Letters*.

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, 1377-1386.

See, L. M., and S. Fritz, 2006. A method to compare and improve land cover datasets: Application to the GLC-2000 and MODIS land cover products, *IEEE Transactions in Geoscience and Remote Sensing* 44:7, 1740-1746.

Small, C., 2002. Reflectance properties of pervious and impervious surfaces, *Proceedings of the ASPRS Pecora Land Remote Sensing Conference*, Denver CO.

Small, C., 2003. High spatial resolution spectral mixture analysis of urban reflectance, *Remote Sensing of the Environment* 88, 170-186.

Small, C., 2005. A global analysis of urban reflectance, *International Journal of Remote Sensing* 26:4, 661-681.

Small, C., F. Pozzi, and C.D. Elvidge, 2005. Spatial analysis of global urban extent from DMSP-OLS night lights, *Remote Sensing of the Environment* 96, 277-291.

Small, C., J.E. Cohen, 2004. Continental physiography, climate, and the global distribution of human population, *Current Anthropology* 45:2, 269-277.

Soutullo, A., M. De Castro, and V. Urios, 2008. Linking political and scientifically derived targets for global biodiversity conservation: implications for the expansion of the global network of protected areas, *Diversity and Distributions* 14, 604-613.

Stren, R., R. White, and J. Whitney, 1992. *Sustainable Cities: Urbanization and the Environment in International Perspective,* Westview Press, Boulder, Colorado.

Sutton, P., D. Roberts, C. Elvidge, and H. Meij, 1997. A comparison of night-time satellite imagery and population density for the continental United States. *Photogrammetric Engineering and Remote Sensing* 63, 1303-1313.

Sutton, P.C. and R. Costanza, 2002. Global estimates of market and non-market values derived from nighttime satellite imagery, land cover, and ecosystem service valuation, *Ecological Economics* 41, 509-527.

Tatem, A. J., A.M. Noor, and S.I. Hay, 2005. Assessing the accuracy of satellite derived global and national urban maps in Kenya, *Remote Sensing of the Environment* 96, 87-97.

Thenkabail, P.S., M. Biradar, P. Noojipady, X. Cai, V. Dheeravath, Y. Li, M. Velpuri, M. Gumma, and S. Pandey, 2007. Sub-pixel area calculation methods for estimating irrigated areas, *Sensors* 7, 2519-2538.

Tilman, D., J. Fargione, B. Wolff, C. D'Antonio, A. Dobson, R. Howarth, D. Schindler, W.H. Schlesinger, D. Simberloff, and D. Swackhamer, 2001. Forecasting agriculturally driven global environmental change, *Science* 292:5515, 281-284.

Travis, J., 2005. Scientists' fears come true as hurricane floods New Orleans, *Science* 309, 1656-1659.

Treitz, P.M., P.J. Howarth, and P. Gong, 1992. Application of satellite and GIS technologies for land cover and land use mapping at the rural urban fringe: a case study, *Photogrammetric Engineering and Remote Sensing* 58, 439–448

Tucker, C.J., D.M. Grant, and J.D. Dykstra, 2004. NASA's global orthorectified Landsat data set, *Photogrammetric Engineering and Remote Sensing* 70, 313-322.

Turner, B.L., E.F. Lambin, and A. Reenberg, 2007. The emergence of land change science for global environmental change and sustainability, *Proc. Natl. Acad. Sci. USA* 104:52, 20666-20671.

Unger, J., Z. Sumeghy, Z. Bottyan, and L. Musci, 2001. Land use and meteorological aspects of the urban heat island, *Meteorological Applications* 8, 189-194.

UN-HABITAT (United Nations Human Settlements Programme), 2003. *The Challenge of Slums: Global Report on Human Settlements, 2003,* Earthscan Publications, London, United Kingdom.

UN-HABITAT (United Nations Human Settlements Programme), 2007. "Chapter 1.1: Citizens of the World: Urban Trends in the 21st Century", in *The State of the World's Cities Report 2006/2007*, Earthscan, London, United Kingdom, 4-12.

United Nations (Department of Economic and Social Affairs, Population Division), 2005. *United Nations World Urbanization Prospects: The 2005 Revision*, online at http://esa.un.org/unup, last accessed 15 April 2007.

United Nations (Department of Economic and Social Affairs, Population Division), 2008a. *World Urbanization Prospects: The 2007 Revision (*accessed June 2008).

United Nations (Statistics Division), 2008b. Standard Area or Country Codes for Statistical Use. (http://unstats.un.org/unsd/methods/m49/m49.htm, accessed September 2008).

United Nations (Department of Economic and Social Affairs), 2008c. *World Urbanization Prospects: The 2007 Revision*, Unpublished data extracted from proUrban database (POP/DB/WUP/Rev.2007), (access date: 26 June 2008).

USGS (United States Geological Survey), 1996. GTOPO30 online product documentation, (http://edc.usgs.gov/products/elevation/gtopo30/README.html).

Utzinger, J. and J. Keiser, 2006. Urbanization and tropical health—then and now, *Annals of Tropical Medicine and Parasitology* 100, 517-533.

Van Oosterom, P., S. Zlatanova, E.M. Fendel, eds., 2005. *Geo-information for disaster management*, Springer, New York, New York.

Van Vliet, W., 2002. Cities in a globalizing world: from engines of growth to agents of change. *Environment and Urbanization* 14, 31-40.

Van Vuuren, D.P., and B.C. O'Neill, 2006. The consistency of IPCC's SRES scenarios to recent literature and recent projections, *Climatic Change* 75, 9-46.

Van Vuuren, D.P., P.L. Lucas, and H. Hilderink, 2007. Downscaling drivers of global environmental change: Enabling use of global SRES scenarios at the national and grid levels, *Global Environmental Change* 17, 114-130.

Vogelmann, J.E., S.M. Howard, L. Yang, C.R. Larson, B.K. Wylie, and N. Van Driel, 2001. Completion of the 1990s National Land Cover Data set for the coterminous United States from Landsat Thematic Mapper data and ancillary data sources, *Photogrammetric Engineering and Remote Sensing* 67, 650-662.

WDPA (World Database of Protected Areas), 2007. UNEP-WCMC and IUCN World Commission on Protected Areas December 2007. (http://www.unep-wcmc.org/wdpa, accessed March 2008).

Welch, R. 1980. Monitoring urban population and energy utilization patterns from satellite data, *Remote Sensing of the Environment* 9, 1-9.

Wittemyer, G., P. Elsen, W.T. Bean, A.C.O. Burton, and J. S. Brashares, 2008. Accelerated human population growth at protected area edges, *Science* 321:5885, 123-126.

World Bank, 2008. World Development Indicators, http://web.worldbank.org (accessed February 20, 2008).

Wratten, E., 1995. Conceptualizing urban poverty, *Environment and Urbanization* 7, 11-38.

Yang, L., C. Huange, C. Homer, B. Wylie, and M. Coan, 2003. An approach for mapping large-are impervious surfaces: Synergistic use of Landsat 7 ETM+ and high spatial resolution imagery, *Canadian Journal of Remote Sensing* 29, 230-240.

Zhang, B., A. NanShan, Z. Huang, C. Yi, and F. Qin, 2008. Meanders of the Jialing River in China: Morphology and formation, *Chinese Science Bulletin* 53:2, 267-281.

Zhang, X.Y., M.A. Friedl, C.B. Schaaf, A.H. Strahler, and A. Schneider, 2004. The footprint of urban climates on vegetation phenology, *Geophysical Research Letters* 31:12, L12209.

Zhao, M., F.A. Heinsch, R.R. Nemani, and S.W. Running, 2005. Improvements of the MODIS terrestrial gross and net primary production global data set, *Remote Sensing of the Environment* 95:2, 164-176.