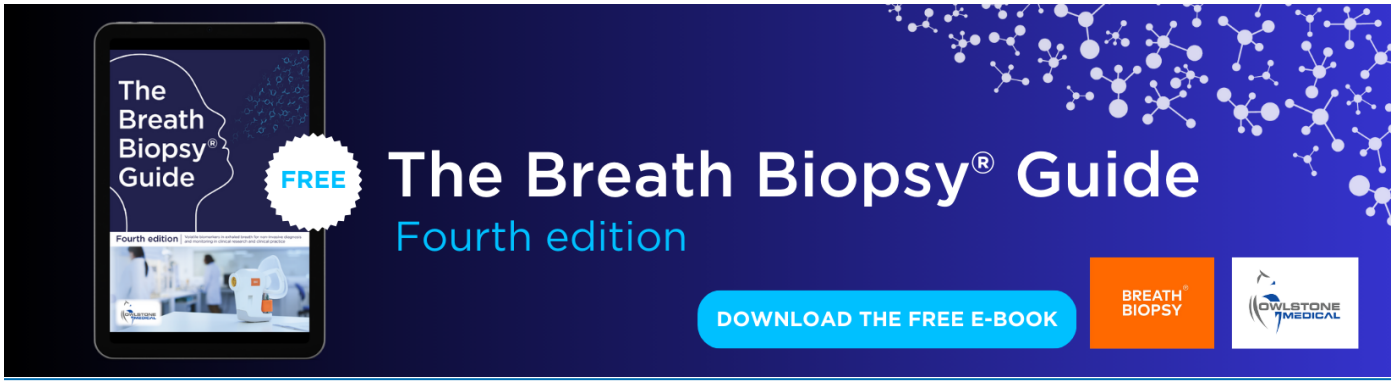
ENVIRONMENTAL RESEARCH
LETTERS

A new map of global urban extent from MODIS satellite data

To cite this article: A Schneider *et al* 2009 *Environ. Res. Lett.* **4** 044003



Environ. Res. Lett. **4** (2009) 044003 (11pp) doi:10.1088/1748-9326/4/4/044003

A new map of global urban extent from MODIS satellite data

A Schneider[1], **M A Friedl**[2] **and D Potere**[3]

[1] Center for Sustainability and the Global Environment, University of Wisconsin-Madison, 1710 University Avenue, Madison, WI 53726, USA
[2] Department of Geography and Environment, Boston University, 675 Commonwealth Avenue, Boston, MA 02215, USA
[3] Office of Population Research, Princeton University, 207 Wallace Hall, Princeton, NJ 08544, USA

E-mail: aschneider4@wisc.edu, friedl@bu.edu and dpotere@princeton.edu

Received 14 February 2009
Accepted for publication 21 September 2009
Published 12 October 2009
Online at stacks.iop.org/ERL/4/044003

Abstract
Although only a small percentage of global land cover, urban areas significantly alter climate, biogeochemistry, and hydrology at local, regional, and global scales. To understand the impact of urban areas on these processes, high quality, regularly updated information on the urban environment—including maps that monitor location and extent—is essential. Here we present results from efforts to map the global distribution of urban land use at 500 m spatial resolution using remotely sensed data from the Moderate Resolution Imaging Spectroradiometer (MODIS). Our approach uses a supervised decision tree classification algorithm that we process using region-specific parameters. An accuracy assessment based on sites from a stratified random sample of 140 cities shows that the new map has an overall accuracy of 93% ($k = 0.65$) at the pixel level and a high level of agreement at the city scale ($R^2 = 0.90$). Our results (available at http://sage.wisc.edu/urbanenvironment.html) also reveal that the land footprint of cities occupies less than 0.5% of the Earth's total land area.

Keywords: urban systems, cities, urbanization, land cover, ecoregions, global monitoring, remote sensing, decision trees, machine learning, data sets

1. Introduction

For the first time in history, more than 50% of the Earth's population now live in cities, towns and settlements (UN 2008). From an environmental standpoint, cities are viewed as an efficient way to concentrate intensive human impacts, while on the other hand, as a nexus of negative environmental impacts that cross scales and city boundaries (Mills 2007). It is clear that cities consume enormous amounts of resources, the by-products of urban activity and land use are numerous (Foley *et al* 2005), and recent studies demonstrate that the ecological footprint of many cities is significant and not sustainable (Kareiva *et al* 2007). Cities are also emerging as an important source of uncertainty in regional to global scale biogeophysical processes. For example, the impact of urban areas on atmospheric chemistry and aerosols is both pronounced and well-documented (Atkinson 2000). Urban land use influences local to regional climates through urban heat islands (Oke 1982, Quattrochi and Ridd 1994), impervious surfaces alter sensible and latent heat fluxes (Offerle *et al* 2006), and recent evidence has suggested that cities may also significantly affect precipitation regimes (Shepherd 2005). At larger scales, the debate continues on whether urban areas significantly impact global environmental processes; recent studies have demonstrated that accurate representation of urban land use is both important and poorly captured in current models (Oleson *et al* 2008). Accurate and timely information on the global distribution and nature of urban areas is therefore critical to a wide array of research questions related to the effect of humans on the regional and global environment (Kaye *et al* 2006).

Despite the growing importance of urban land in regional to global scale environmental studies, it remains extremely difficult to map urban areas at coarse scales due to the heterogeneous mix of land cover types in urban environments, the small area of urban land relative to the total land surface

Table 1. Ten global urban or urban-related maps listed in order of increasing global urban extent. (Abbreviations: NOAA, National Oceanographic and Atmospheric Administration; NASA, National Aeronautics and Space Administration; DOE, Department of Energy; MODIS, Moderate Resolution Imaging Spectroradiometer.)

Abbreviation[a]	Map (citation)	Definition of urban or urban-related feature	Resolution[b]	Extent (km^2)
VMAP	Vector Map Level Zero (Danko 1992)	Populated places	1:1 mil	276 000
GLC2000[c]	Global Land Cover 2000 (Bartholome and Belward 2005)	Artificial surfaces and associated areas	988 m	308 000
GlobCover[c]	GlobCover v2.2 (ESA 2008)	Artificial surfaces and associated areas (urban areas >50%)	309 m	313 000
HYDE[c]	History Database of the Global Environment v3 (Goldewijk 2005)	Urban area (built-up, cities)	9000 m	532 000
IMPSA	Global Impervious Surface Area (Elvidge *et al* 2007)	Density of impervious surface area	927 m	572 000
MODIS 500 m[c]	MODIS Urban Land Cover 500 m (Schneider *et al* 2009)	Areas dominated by built environment (>50%), including non-vegetated, human-constructed elements, with minimum mapping unit > 1 km ^{2}	463 m	657 000
MODIS 1 km[c]	MODIS Urban Land Cover 1 km (Schneider *et al* 2003)	Urban and built-up areas	927 m	727 000
GRUMP	Global Rural–Urban Mapping Project (CIESIN 2004)	Urban extent	927 m	3 524 000
Lights[d]	Nighttime Lights v2 (Imhoff *et al* 1997, Elvidge *et al* 2001)	Nighttime illumination intensity	927 m	NA
LandScan[d]	LandScan 2005 (Bhaduri *et al* 2002)	Ambient (average over 24 h) global population distribution	927 m	NA

[a] Bold type indicates the maps assessed in this paper. HYDE was not included due to its coarse spatial resolution.
[b] For maps in a native geographic projection, the resolution describes pixel size at the equator.
[c] These maps are multi-class land cover maps with an urban class.
[d] Maps depicting urban-related features.

area, and the significant differences in how different groups and disciplines define the term 'urban'. Each of the datasets that have emerged during the last decade (table 1) suffers from limitations related to these scale and definitional issues (Potere *et al* 2009). Moreover, the maps differ by an order of magnitude in how they depict urban areas (0.3 million km^2 for Vector Map, Danko 1992, to 3.5 million km^2 for the Global Rural–Urban Mapping Project, CIESIN 2004). The extreme variability in these estimates calls into question the accuracy of each map's depiction of urban and built-up land, and yet past efforts to validate the maps have been minimal.

In this letter, we present results from recent efforts to produce a new global map of urban land based on a new approach that uses remotely sensed data in association with a global stratification of 'urban ecoregions'. This work builds on our past work using Moderate Resolution Imaging Spectroradiometer (MODIS) data at 1 km spatial resolution (Schneider *et al* 2003, 2005) in coordination with the MODIS Collection 4 Global Land Cover Product (Friedl *et al* 2002, 2009). The goal of producing this new map is to address several key deficiencies in the Collection 4 (C4) map arising from confusion between built-up areas, bare ground and shrublands, as well as begin development of a database of urban land surface characteristics for multiple time periods (2001–2010). To this end, the new dataset is produced using newly released Collection 5 (C5) MODIS data circa 2001–2002 with increased spatial resolution (500 m). In the following sections, we describe our methods and results, and briefly highlight key findings from our accuracy assessment of the new global urban map.

2. Defining urban extent

We begin our analysis with a conceptual framework for characterizing urbanized areas in regional and global mapping studies. We define urban areas based on the physical attributes and composition of the Earth's land cover: urban areas are places dominated by the built environment. The 'built environment' includes all non-vegetative, human-constructed elements, such as roads, buildings, runways, etc (i.e. human-made surfaces), and 'dominated' implies coverage greater than 50% of a given landscape unit (the pixel). When

vegetation (e.g. a park) dominates a pixel, these areas are not considered urban, even though in terms of land use, they may function as urban space. The term 'impervious surface' is often used interchangeably with 'built environment' (Ridd 1995), but we prefer the more direct term 'built environment' because of uncertainty and known scaling issues surrounding the impervious surface concept (Small and Lu 2006, Stow *et al* 2007). Finally, our definition also includes a minimum mapping unit: urban areas are contiguous patches of built-up land greater than 1 km^2.

In general, global urban mapping efforts have been beset by the lack of a consistent, unambiguous definition of 'urban areas'. Each group in table 1 approaches the task from a different perspective, employing methodologies that draw on a combination of satellite imagery, census information, and other maps. However, the definitions utilized in each map (table 1) do not necessarily reflect each group's methodological approach to mapping urban areas. Rather, previous work has shown that the representation of urban land is often tied closely to the input data: maps from census data correspond to population distribution, those utilizing Nighttime Lights data correlate with national income levels, while maps from multispectral data align most closely with 'built-up areas' (Potere and Schneider 2007, Potere *et al* 2009). To provide context for our accuracy assessment of the new MODIS 500 m urban map, we therefore include four maps that define urban areas similar to our approach—GLC2000, GlobCover, Global Impervious Surface Area dataset (IMPSA), and GRUMP—as well as the C4 MODIS 1 km map.

3. Methods

3.1. Supervised classification of MODIS data

The classification methodology for this research relies on a supervised decision tree algorithm (C4.5), a non-parametric classifier shown to be particularly effective for coarse resolution datasets with complex, non-linear relationships between features and classes, and noisy or missing data (Friedl and Brodley 1997, Friedl *et al* 2002). Decision tree construction involves the recursive partitioning of a set of training data, which is split into increasingly homogeneous subsets based on statistical tests applied to the feature values (the satellite data). Once the decision tree has been estimated, these rules are then applied to the entire image to produce a classified map.

To improve classification accuracy, the decision tree algorithm is used in conjunction with boosting, an ensemble classification technique that improves class discrimination by estimating multiple classifiers while systematically varying the training sample (Quinlan 1996). The final classification is produced by an accuracy-weighted vote across all classifications. Boosting has been shown to be equivalent a form of additive logistic regression, and as a result, probabilities of class membership can be assigned for each class at every pixel (McIver and Friedl 2001).

Our classification approach employs a one-year time series of MODIS data to exploit spectral and temporal properties in land cover types. Specifically, we utilize the differences in temporal signatures for urban and rural areas that result from phenological differences between vegetation inside and outside the city. For example, the spectral signatures of an urbanized plot and a fallow agricultural plot may appear similar at any given time in medium-coarse resolution data, and are therefore easily confused during classification. However, over the course of one year, the signatures for the urban and agricultural plots will vary due to differences in bud-burst, vegetation abundance, etc. Accordingly, the MODIS data inputs include one year (18 February 2001 to 17 February 2002) of 8 day composites of the seven land bands and the enhanced vegetation index (EVI) at 463.3 m spatial resolution. All input data are adjusted to a nadir-viewing angle to reduce the effect of varying illumination and viewing geometries (Schaaf *et al* 2002). The 8 day values are aggregated to 32 day averages to reduce the frequency of missing values from cloud cover, and monthly and yearly minima, maxima and means for each band are included as inputs to increase classification accuracy. Finally, the training data include 1860 sites selected across 17 land cover classes by manual interpretation of Landsat and Google Earth imagery, as well as a set of urban training sites chosen from 182 cities located across the globe.

While the previous approach for C4 MODIS data also utilized the C4.5 algorithm, the new methods differ in several key ways. Rather than rely on external datasets to constrain the classification as in the C4 methodology, the C5 methodology does not include this step. The increased resolution of the MODIS data and improvements to the training site database are typically sufficient to generate the final urban map for most regions. For areas where class confusion does result—typically semi-arid/arid regions without significant settlement—we exploit the ability of the boosted decision trees to produce class membership probabilities. We run the classification algorithm twice: classification 1 utilizes the full set of land cover exemplars that includes urban sites, and classification 2 excludes the urban training sites. The first classification classifies both the urban core and mixed urban spaces correctly, with the caveat that some non-urban areas are erroneously labeled urban land (e.g. expanses of semi-arid shrubland). The urban class probabilities are extracted from classification 1, and areas of confusion are determined based on low membership to the urban class. For these pixels, we then take advantage of the information in the second classification (without urban sites) to modify the urban class probabilities using Bayes rule. To complete this step we rely on the urban ecoregion stratification described below.

3.2. Urban ecoregionalization

A key element of the new methodology is a global stratification of land based on the natural, physical and structural components of urban areas. While urbanized areas are some of the most complex landscapes in the world, there is surprising regularity in city structure, configuration, composition, and vegetation types within geographic regions and by level of economic development. Our approach exploits these similarities to define 16 quasi-homogeneous areas that we term *urban ecoregions* (table 2). We base the stratification on: (1) a

Table 2. The 16 *urban ecoregion* stratification designed for map processing, analysis and validation.

Ecoregion	No	Geographic region	Example areas
Temperate broadleaf, mixed forest	1	North America, Japan, Australia	Eastern US, Canada
	2	Europe, Japan	Germany, France, Japan
	3	Eastern Asia	Eastern China
Temperate grassland, shrubland	4	Americas, Central Asia, Australia	US, Argentina, Australia
	5	Middle East, Central Asia	Turkey, Georgia
Tropical broadleaf forest	6	South America	Brazil, Colombia, Guatemala
	7	Sub-Saharan Africa	Dem. Republic of the Congo
Tropical–subtropical mixed forest	8	South-Central Asia, South-East Asia	China, India
Tropical–subtropical savannah, grasslands	9	South America	Southeastern Brazil, Paraguay
	10	Sub-Saharan Africa	Ghana, Kenya, Tanzania
Tropical–subtropical grasslands	11	South America, Southern Africa	Chile, Peru, South Africa
Mediterranean	12	North America, Southern Europe, North Africa	California, Italy, Spain
Arid, semi-arid desert	13	Africa, Middle East, Central Asia, Australia	Sahara Desert
Arid, semi-arid steppe, shrubland	14	Central Asia	Western China
Boreal forest, tundra	15	North America, Northern Eurasia	Canada, Russia
Permanent ice, snow	16	North, south pole	Antarctica

biome designation to summarize climate and vegetation trends (Olson *et al* 2001), (2) level of economic development defined by per-capita gross domestic product (GDP) (UN 2008); and (3) regional differences in city structure, organization and historic development (Bairoch 1988).

Using this stratification, we identify the regions that require further refinement to the training sites or input data, or post-processing. In addition to estimation of posterior probabilities using prior information from the land cover probabilities, post-processing includes masking of problem areas, application of the MODIS 500 m water mask, and hand-editing.

4. Results

Figure 1 illustrates the results of the new MODIS Collection 5 map of urban extent for two urbanized regions, Washington DC and Guangzhou, China. For comparison, the same region is shown for a Landsat-based 30 m classification (figures 1(d) and (j)) and the previously released MODIS-based map at 1 km (figures 1(f) and (l)). The pattern of urban land is similar across the C4 and C5 maps, except the MODIS 500 m map provides greater detail on the edge of the city as well as within the urban fabric compared to the 1 km map. To illustrate how these maps might be utilized at coarsened resolutions, we have included views of each region where the MODIS 500 m map has been aggregated to 2 km (figures 1(a) and (g)) and 8 km spatial resolution (figures 1(b) and (h)). This step effectively converts the map legends from binary (urban/non-urban) to continuous (percentage urban). Continental views of the new MODIS map are included in the appendix (figures A.1(a1)–(a6)).

Regionally, our results reveal that previous estimates of urban extent (2–3%, CIESIN 2004) drawn from global urban maps may over-estimate the true extent of built-up areas. The MODIS 500 m map shows that urban land area varies from only 0.17% of total continental land area in Africa to 0.67 in North America, with most regions near the continental average of 0.5% urbanized (e.g. South America, 0.47%; Asia, 0.53%). The exception is the European land mass (1.78%), a result that is to be expected given the extensive urban morphology in this region (table 3).

To assess the accuracy of the MODIS 500 m map, we compiled a geographically comprehensive set of Landsat-based maps for 140 cities (30 m resolution, Angel *et al* 2005, Schneider and Woodcock 2008). The cities were selected using a random-stratified sampling design based on population, geographic region and income, and are independent of the training exemplars used during classification of the MODIS 500 m map. We conducted an independent assessment of the 140 Landsat-based maps to ensure that these data provide a statistically defensible basis for characterizing the accuracy of the MODIS 500 m map (Potere *et al* 2009). Using an independent set of 10 000 random samples labeled from very high resolution imagery (4 m) in Google Earth, the pooled confusion matrix results showed that the maps range in accuracy from 82.8 to 91.0%. While this accuracy assessment is based on subjective labeling of sites as urban or non-urban using photo-interpretation of very high resolution data, we employed multiple analysts in a double-blind procedure to reduce uncertainty and bias during analysis. Thus, we feel confident that these data are suitable as reference data in this study.

The 140 reference maps were resampled to match the spatial resolution of the MODIS 500 m map and each global urban map (0.3–1 km^2), and binary urban/non-urban contingency matrices were estimated to calculate the level of agreement between the reference city map and the global map of interest. The results reveal that overall accuracy is generally high across all map sources (figure 2(a)), with mean accuracy rates ranging from 73% (GRUMP) to over 93%

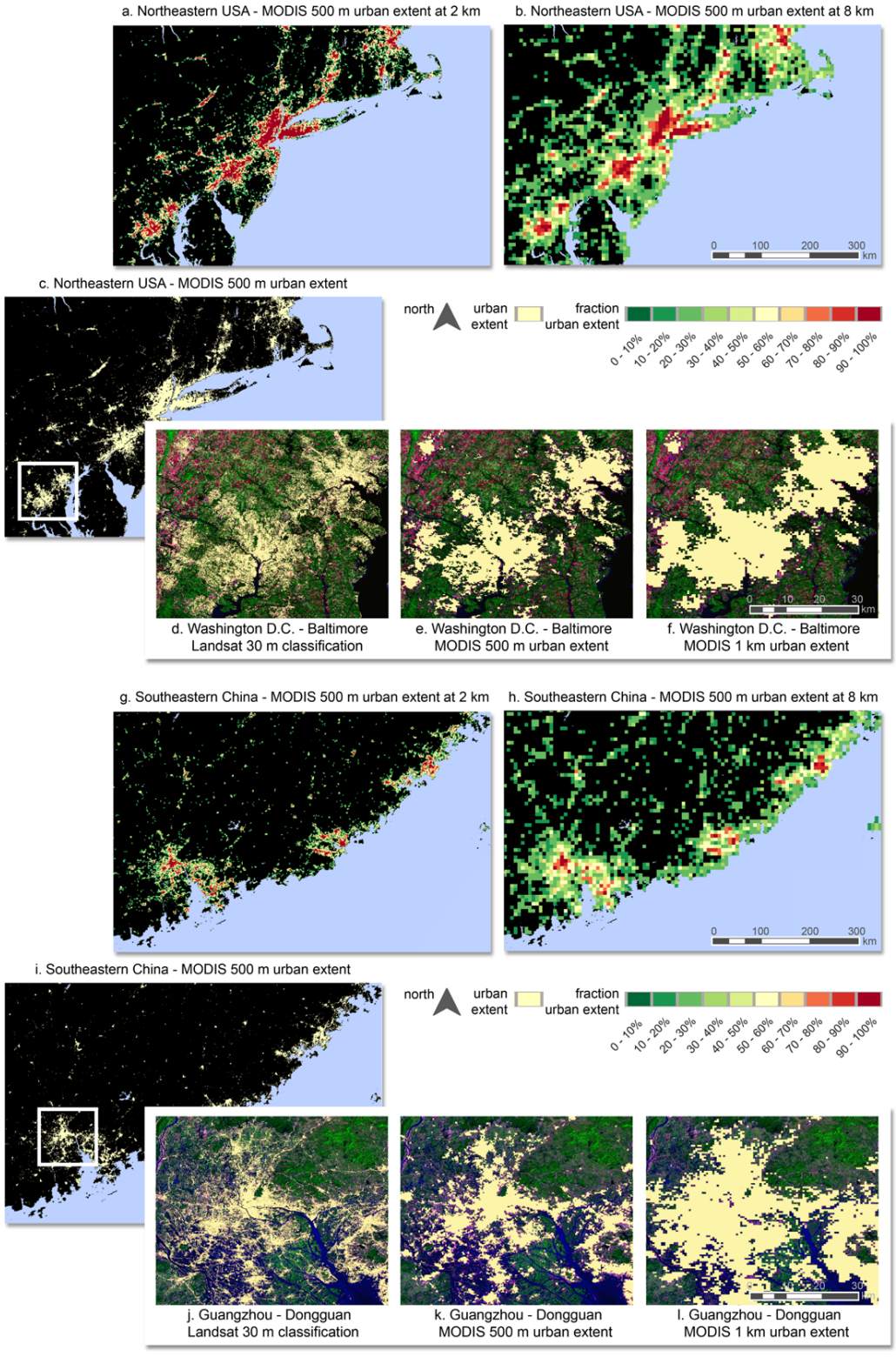


Figure 1. An illustration of the new MODIS 500 m global urban map for two urbanized regions: the Northeastern United States (a)–(f) and Southeastern China (g)–(l). In addition to the binary map (c), (i), the map has been aggregated to 2 and 8 km resolution for display (a), (b), (g) and (h). The inset maps (shown in rows) include a Landsat-based classification (30 m resolution, Schneider and Woodcock 2008), the new map of urban extent from MODIS 500 m data, and the previous version of the MODIS-based map (1 km resolution).

(MODIS 500 m map). The MODIS 500 m map has the highest accuracy, the lowest standard deviation (±4.5%, as compared to >±7.2% in other maps), and the fewest outliers below 75%. Differences among the maps are summarized in the kappa statistic: the MODIS 500 m map has the highest mean kappa values (0.65), while IMPSA, GlobCover, GLC2000, and MODIS 1 km have kappa values ranging from 0.38 to 0.50, and GRUMP has the lowest mean kappa of 0.28.

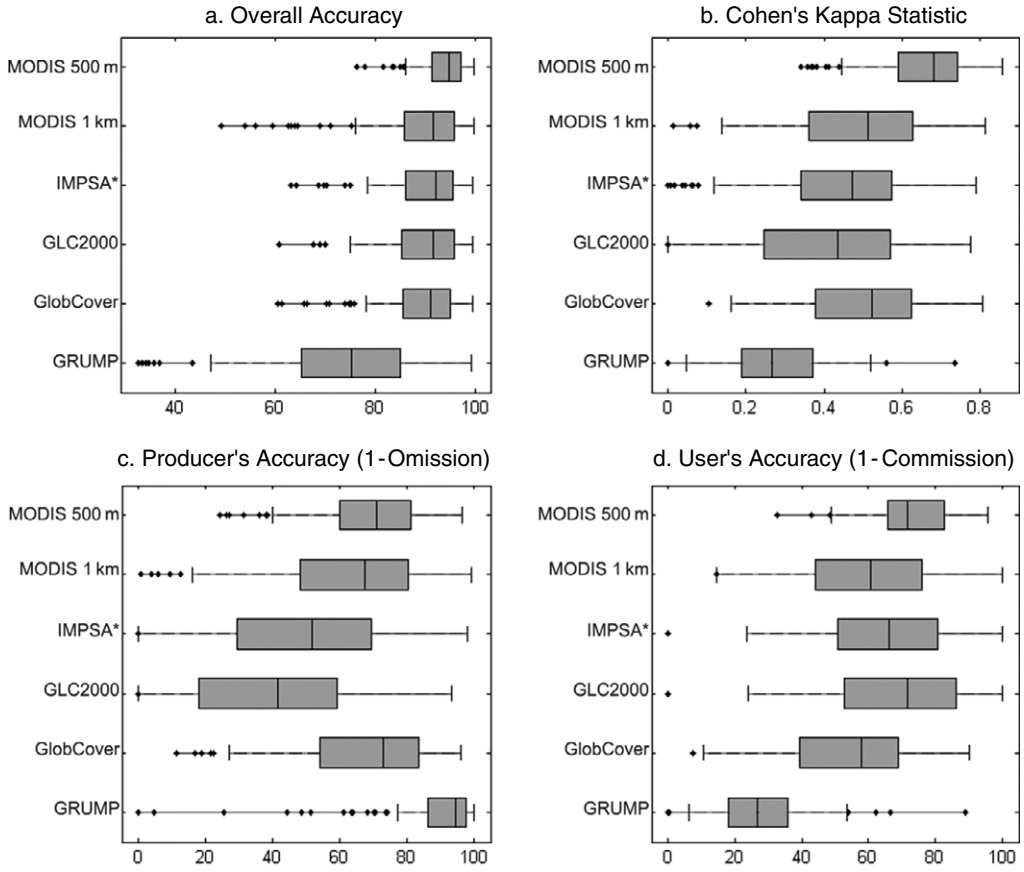


Figure 2. Box plots of accuracy statistics for the MODIS 500 m map and five additional global urban maps using the 140 city validation sample. The figure shows four measures to assess the accuracy of the global urban maps at the scale of the city: (a) overall accuracy, (b) the kappa statistic, (c) producer's accuracy (or sensitivity, 1-commission error), and (d) user's accuracy (or specificity, 1-commission error). The six global urban maps shown along the *y*-axis are the MODIS-based maps of urban extent at 500 m and 1 km spatial resolution, NOAA's Impervious Surface Area map (IMPSA), Global Land Cover 2000 (GLC 2000), the recently released GlobCover map, and CIESIN's Global Rural–Urban Mapping Project (GRUMP).

Figure 2(c) shows the distribution of producer's accuracies across the 140 city sample for the coarse resolution urban maps; this measure conveys the error of omission. Two of the maps (IMPSA, GLC2000) have mean/median values below 50%, which indicates that these datasets are missing urban land in their classifications. GRUMP, however, has a high producer's accuracy, at nearly 90%. Although GRUMP does not *miss* many true urban pixels, the map has a large number of erroneously labeled urban pixels evident from its low user's accuracy. User's accuracy (figure 2(d)) reflects errors of commission, and the distribution of user's accuracies mirrors the results of the overall accuracy measure (figure 2(a)): the MODIS 500 m map has the highest user's accuracy (72.9%), with GLC2000 and IMPSA close behind (65.6 and 64.8%).

In addition to the pixel-based assessment, we also assessed the MODIS 500 m at the scale of the city. The city size estimates were derived from the 140 Landsat-based classifications (native resolution), which range in size from 20 to 8000 km^2. Figure 3 illustrates how the estimates of urban size (*x*-axis) compare to estimates obtained from the MODIS 500 m map and the five global urban map sources (*y*-axis). The results show that the MODIS 500 m map has good agreement with the reference data when compared against the other sources, a low root mean square error (RMSE), 142.6 m, and a high R^2 of 0.90.

5. Conclusions

Urban areas are an increasingly important component of the global environment, yet they remain one of the most challenging areas for conducting research. Model parameterizations (e.g. climate, meteorological, biogeochemical, hydrological models) are particularly difficult for urban areas given the complex three-dimensional structure of cities and the mixture of surface types with contrasting radiative, thermal and moisture characteristics. It is therefore essential that regional–global maps of urban land use not only display the point location of cities or the spatial distribution of population, but also provide up-to-date information regarding the extent, growth, and physical characteristics of urban land.

This letter presents a new, global moderate resolution map of urban extent circa 2001–2002, with several improvements over currently available map sources. The increased spatial resolution and radiometric quality of the MODIS C5 data (500 m) has allowed a four fold increase in spatial detail. Our accuracy assessment shows that the MODIS 500 m map provides the most realistic depiction of global urban land use

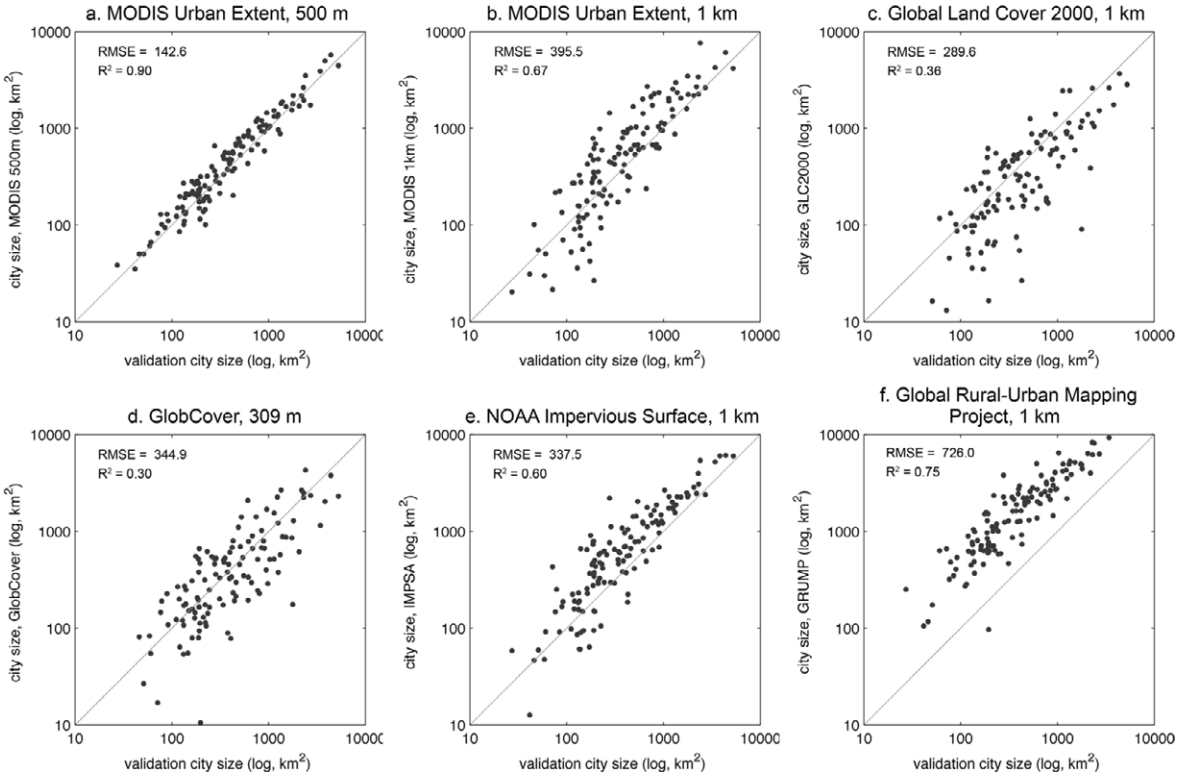


Figure 3. Scatter plots of the 140 cities in the validation sample, where each plot shows the city size from the high resolution Landsat-based reference maps (*x*-axis, assumed to be 'truth') and the global urban maps (*y*-axis). Note the log scale on both axes.

Table 3. The areal extent of each of the six global urban maps (in square kilometers) for eleven world regions.

Region	Urban land area (km^2) and percentage urban land (percentage of total global land area)					
	Global Land Cover 2000	GlobCover	Impervious Surface Areas[a]	MODIS 500 m Urban Extent	MODIS 1 km Urban Extent	Global Rural–Urban Mapping Project
North America	86 536 (0.48)	32 456 (0.18)	86 788 (0.48)	129 904 (0.72)	125 398 (0.70)	888 353 (4.93)
Central America, Caribbean	3 468 (0.13)	5 322 (0.20)	17 581 (0.67)	13 099 (0.50)	10 274 (0.39)	154 951 (5.95)
South America	10 731 (0.06)	10 801 (0.06)	35 382 (0.20)	82 242 (0.47)	42 876 (0.24)	374 942 (2.14)
Western Europe	70 919 (1.75)	55 764 (1.37)	53 262 (1.31)	85 900 (2.11)	125 342 (3.08)	536 770 (13.21)
Eastern Europe	35 937 (0.20)	28 232 (0.16)	34 540 (0.19)	63 494 (0.36)	68 487 (0.38)	301 596 (1.69)
Sub-Saharan Africa	17 937 (0.07)	20 458 (0.09)	49 788 (0.21)	31 052 (0.13)	39 621 (0.17)	144 996 (0.60)
Western Asia, North Africa	16 905 (0.17)	17 285 (0.17)	34 492 (0.34)	37 782 (0.37)	44 039 (0.43)	222 113 (2.17)
South-Central Asia	31 680 (0.30)	29 690 (0.28)	112 296 (1.07)	64 973 (0.62)	86 298 (0.82)	350 383 (3.33)
East Asia	12 630 (0.11)	69 401 (0.61)	103 266 (0.91)	110 514 (0.97)	162 645 (1.43)	402 530 (3.55)
South-East Asia, Pacific Islands	11 819 (0.26)	7 416 (0.16)	40 933 (0.89)	29 197 (0.63)	12 597 (0.27)	102 290 (2.22)
Australia, New Zealand	9 446 (0.12)	35 954 (0.46)	3 161 (0.04)	10 602 (0.14)	9 366 (0.12)	45 027 (0.58)
Total urban land area (km^2)[b]	308 007 (0.24)	312 779 (0.24)	571 504 (0.44)	658 760 (0.51)	726 943 (0.57)	3524 109 (2.74)

[a] Impervious Surface Area Map is thresholded at 20% to produce estimates of urban land.

[b] Total land area excludes Antarctica and Greenland.

among the available datasets. This analysis also presents the first global validation effort performed for any of the global urban maps, and provides important information to the user community on the quality and suitability of this map for a range of applications.

Moving forward, the MODIS 500 m urban map provides a foundation for refined representations of global urban land use. It is clear from recent research efforts that an extended database of land surface characteristics for urban areas is greatly needed. With these needs in mind, our ongoing efforts are focused on: (1) creating maps of urban extent circa 2005 and 2009; (2) providing sub-pixel estimates of urban land use and vegetation; and (3) differentiating core downtown areas from low density residential areas.

Acknowledgments

The authors wish to thank Solly Angel and Dan Civco for generous use of their datasets, Scott Macomber and Damien Sulla-Menashe for technical support, and Mutlu Ozdogan for comments on an earlier draft of this paper. This work was supported by NASA grant NNX08AE61A.

Appendix

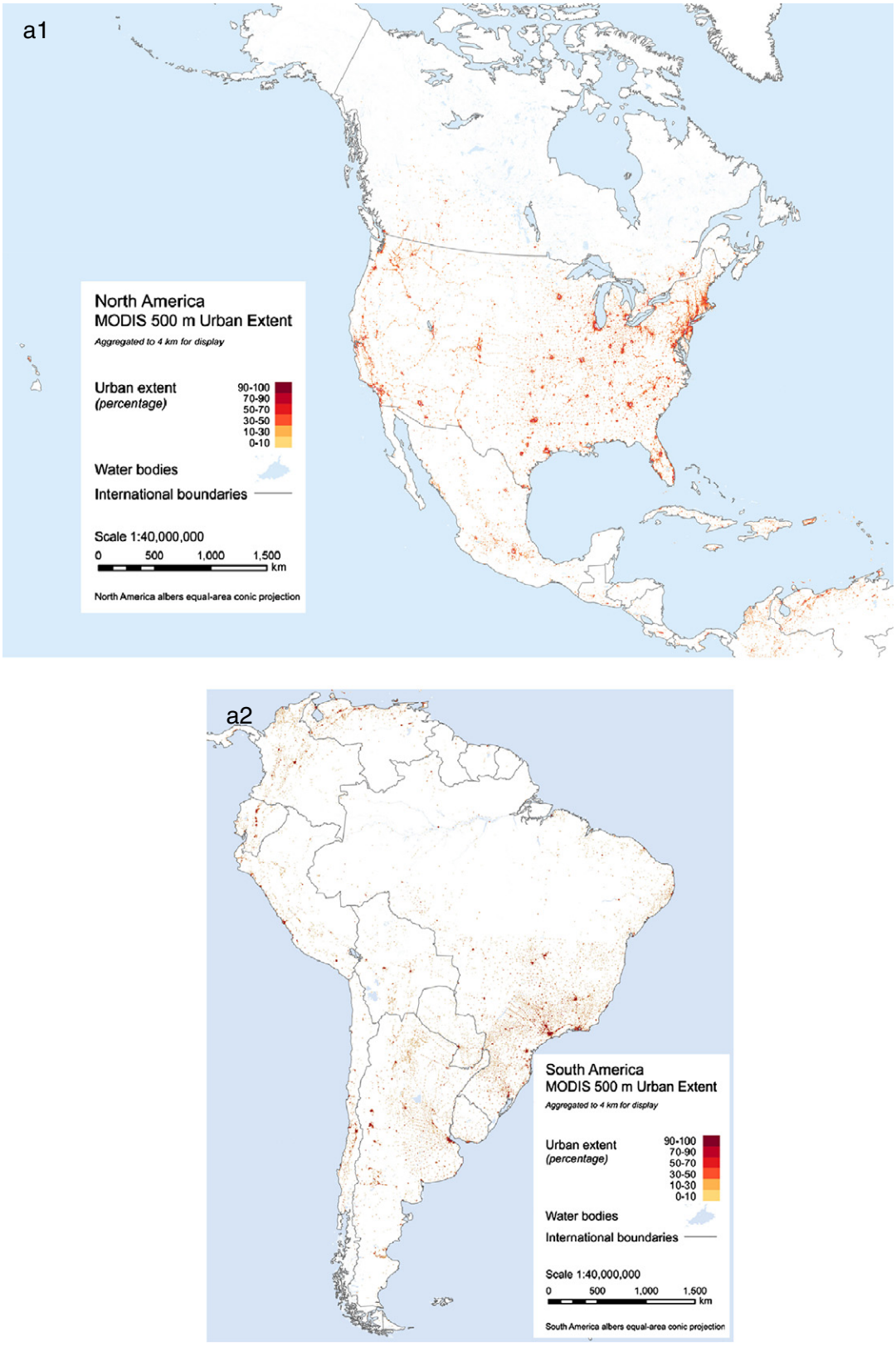


Figure A.1. (a1)–(a6) Continental views of the new MODIS 500 m global urban map. For viewing purposes, the 463.3 m resolution has been aggregated to 2 km resolution; this step yields a continuous value map where each 2 km pixel depicts the percentage of cells labeled as 'urban' in the native resolution map.

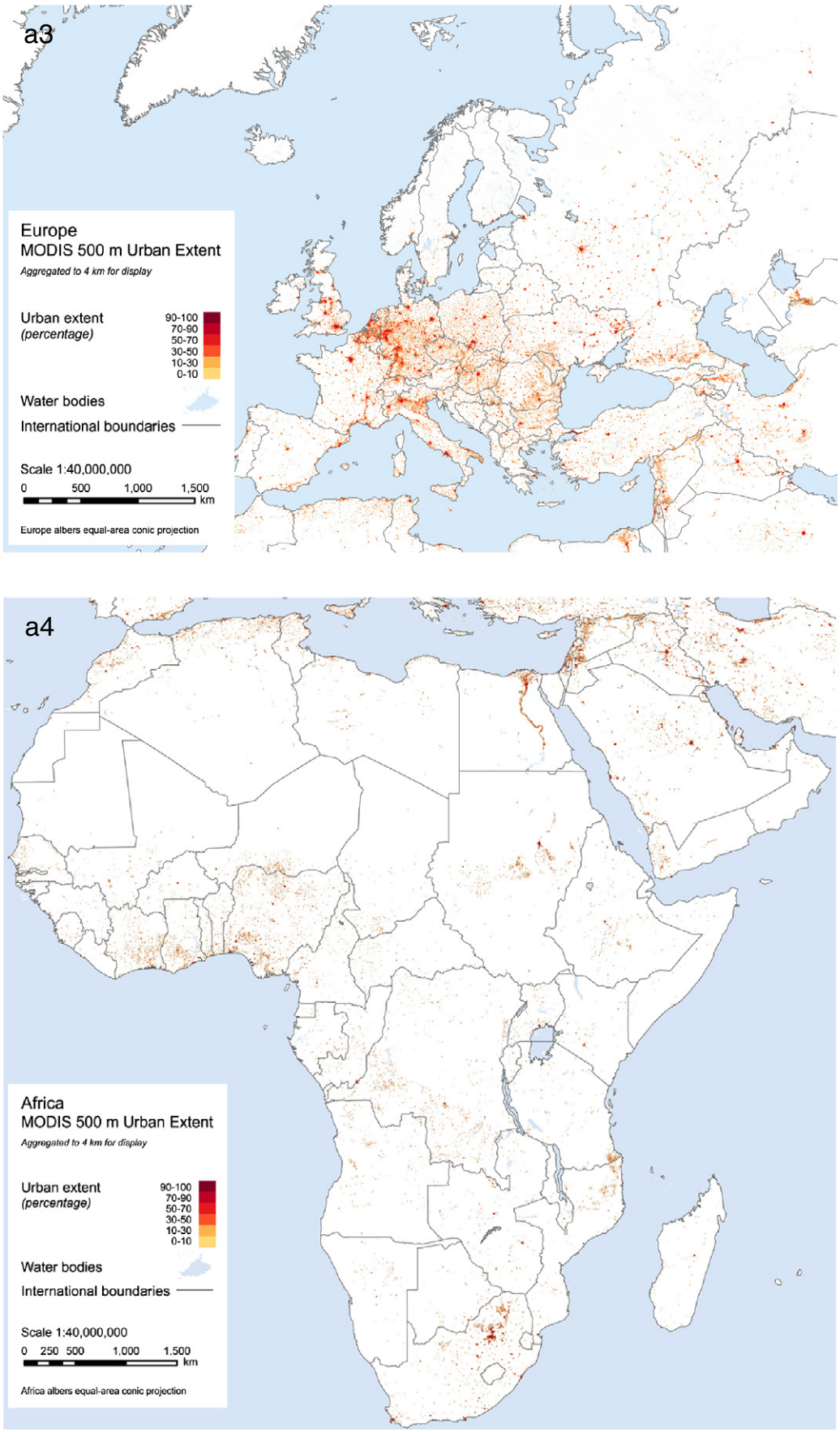


Figure A.1. (Continued.)

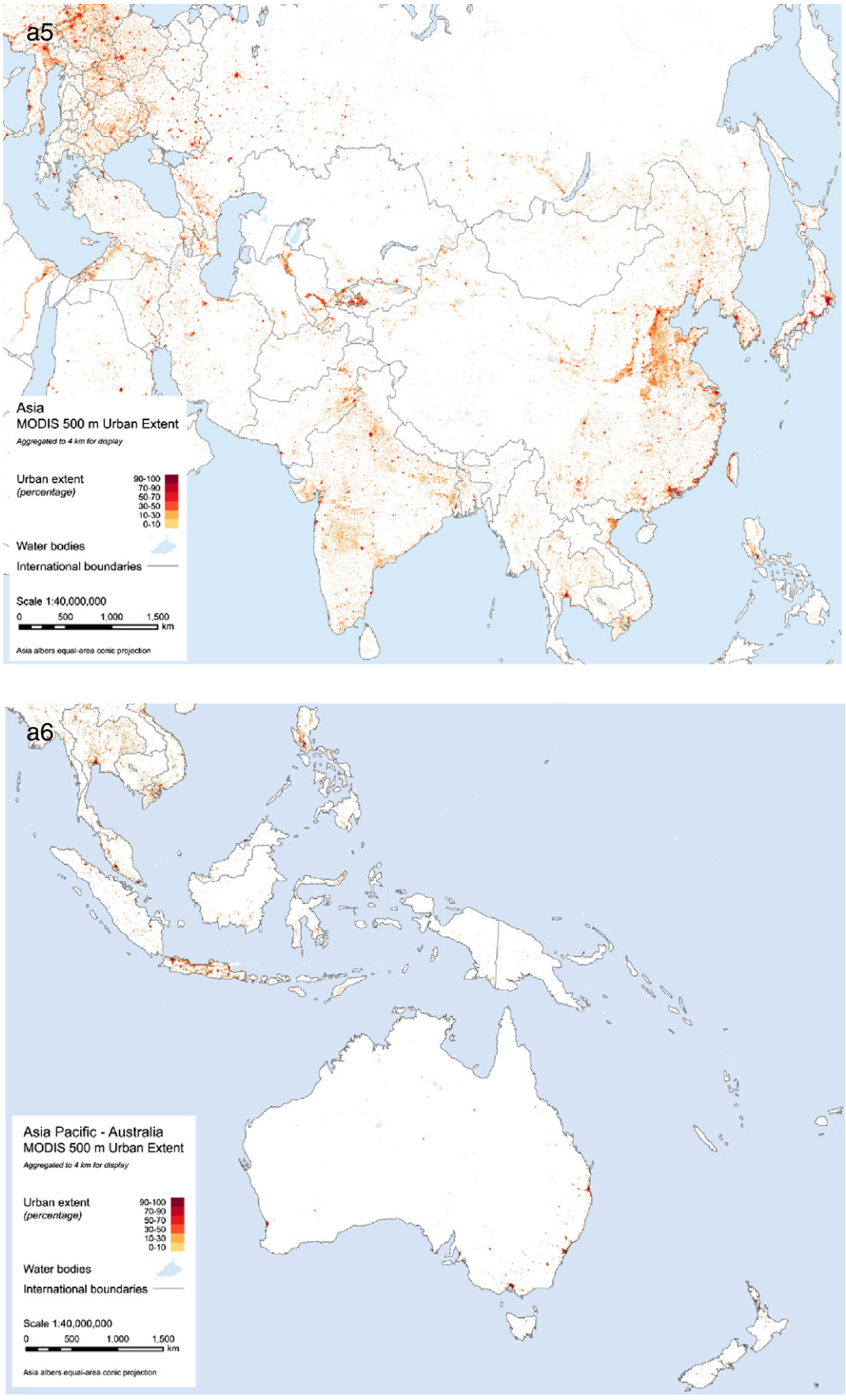


Figure A.1. (Continued.)

References

Angel S, Sheppard S and Civco D 2005 The dynamics of global urban expansion *World Bank Report* (Washington, DC: The World Bank)

Atkinson R 2000 Atmospheric chemistry of VOCs and NOx *Atmos. Environ.* **34** 2063–101

Bairoch P 1988 *Cities and Economic Development: From the Dawn of History to the Present* (Chicago, IL: The University of Chicago Press)

Bartholome E and Belward A 2005 GLC2000: a new approach to global land cover mapping from Earth observation data *Int. J. Remote Sens.* **26** 1959–77

Bhaduri B, Bright E, Coleman P and Dobson J 2002 LandScan: locating people is what matters *Geoinfomatics* **5** 34–7

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 last accessed 21 July 2009

Danko D 1992 The digital chart of the world project *Photogramm. Eng. Remote Sens.* **58** 1125–8

Elvidge C, Imhoff M, Baugh K, Hobson V, Nelson I, Safran J, Dietz J and Tuttle B 2001 Nighttime lights of the world: 1994–95 *ISPRS J. Photogramm. Remote Sens.* **56** 81–99

Elvidge C, Tuttle B, Sutton P, Baugh K, Howard A, Milesi C, Bhaduri B and Nemani R 2007 Global distribution and density of constructed impervious surfaces *Sensors* **7** 1962–79

ESA (European Space Agency) 2008 The Ionia GlobCover project *The GlobCover Portal* http://ionia1.esrin.esa.int last accessed 21 July 2009

Foley J *et al* 2005 Global consequences of land use *Science* **309** 570–4

Friedl M and Brodley C 1997 Decision tree classification of land cover from remotely sensed data *Remote Sens. Environ.* **61** 399–409

Friedl M, Sulla-Menashe D, Tan B, Schneider A, Ramankutty N, Sibley A and Huang X 2009 MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets *Remote Sens. Environ.* at press

Friedl M *et al* 2002 Global land cover mapping from MODIS: algorithms and early results *Remote Sens. Environ.* **83** 287–302

Goldewijk K 2005 Three centuries of global population growth: a spatially referenced population density database for 1700–2000 *Populat. Environ.* **26** 343–67

Imhoff M L, Lawrence W, Stutzer D and Elvidge C 1997 A technique for using composite DMSP/OLS 'city lights' satellite data to map urban areas *Remote Sens. Environ.* **61** 361–70

Kareiva P, Watts S, McDonald R and Boucher T 2007 Domesticated nature: shaping landscapes and ecosystems for human welfare *Science* **316** 1866–9

Kaye J, Groffman P, Grimm N, Baker L and Pouyat R 2006 A distinct urban biogeochemistry? *Trends Ecol. Evolut.* **21** 192–9

McIver D and Friedl M 2001 Estimating pixel-scale land cover classification confidence using non-parametric machine learning methods *IEEE Trans. Geosci. Remote Sens.* **39** 1959–68

Mills G 2007 Cities as agents of global change *Int. J. Climatol.* **27** 1849–57

Offerle B, Grimmond C, Fortuniak K and Pawlak W 2006 Intraurban differences of surface energy fluxes in a central European city *J. Appl. Meteorol. Climatol.* **45** 125–36

Oke T R 1982 The energetic basis of the urban heat island *Q. J. R. Meteorol. Soc.* **108** 1–24

Oleson K, Bonan G, Feddema J, Vertenstein M and Grimmond C S B 2008 An urban parameterization for a global climate model. Part 1: formulation and evaluation for two cities *J. Appl. Meteorol.* **47** 1038–60

Olson D M *et al* 2001 Terrestrial ecoregions of the world: a new map of life on Earth *BioScience* **51** 933–8

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

Potere D, Schneider A, Schlomo A and Civco D 2009 Mapping urban areas on a global scale: which of the eight maps now available is more accurate? *Int. J. Remote Sens.* at press

Quattrochi D A and Ridd M K 1994 Measurements and analysis of thermal energy responses from discrete urban surfaces using remote sensing data *Int. J. Remote Sens.* **15** 1991–2002

Quinlan J R 1996 Bagging, boosting and C4.5 *AAAI-96: Proc. 13th National Conf. on Artificial Intelligence (Portland, Oregon, Aug. 1996)* (Menlo Park, CA: AAAI Press) pp 725–30

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 *Int. J. Remote Sens.* **16** 2165–85

Schaaf C B *et al* 2002 First operational BRDF albedo nadir reflectance products from MODIS *Remote Sens. Environ.* **83** 135–48

Schneider A, Friedl M, McIver D and Woodcock C 2003 Mapping urban areas by fusing multiple sources of coarse resolution remotely sensed data *Photogramm. Eng. Remote Sens.* **69** 1377–86

Schneider A, Friedl M and Potere D 2009 Monitoring urban areas globally using MODIS 500 m data: new methods based on urban ecoregions *Remote Sens. Environ.* in review

Schneider A, Friedl M and Woodcock C 2005 Mapping urban areas by fusing multiple sources of coarse resolution remotely sensed data: global results *Proc. 5th Int. Symp. of Remote Sensing of Urban Areas (Tempe, AZ, March 2005)*

Schneider A and Woodcock C 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 Stud.* **45** 659–92

Shepherd M 2005 A review of current investigations of urban-induced rainfall and recommendations for the future *Earth Interact.* **9** 1–27

Small C and Lu J 2006 Estimation and vicarious validation of urban vegetation abundance by spectral mixture analysis *Remote Sens. Environ.* **100** 441–56

Stow D, Lopez A, Lippitt C, Hinton S and Weeks J 2007 Object-based classification of residential land use within Accra, Ghana based on QuickBird satellite data *Int. J. Remote Sens.* **28** 5167–73

UN (United Nations) 2008 *World Urbanization Prospects: The 2007 Revision* Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat http://esa.un.org/unup last accessed 21 July 2009