Published in final edited form as:
Remote Sens Environ. 2005 May 15; 96(1): 87–97. doi:10.1016/j.rse.2005.02.001.

Assessing the accuracy of satellite derived global and national urban maps in Kenya

A. J. Tatem[1], **A. M. Noor**[2], and **S. I. Hay**[1,2]

[1]TALA Research Group, Tinbergen Building, Department of Zoology, University of Oxford, South Parks Road, Oxford, OX1 3PS, UK

[2]Public Health Group, KEMRI/Wellcome Trust Research Laboratories, P.O.Box 43640, 00100 GPO, Nairobi, Kenya

Abstract

Ninety percent of projected global urbanization will be concentrated in low income countries (United-Nations, 2004). This will have considerable environmental, economic and public health implications for those populations. Objective and efficient methods of delineating urban extent are a cross-sectoral need complicated by a diversity of urban definition rubrics world-wide. Large-area maps of urban extents are becoming increasingly available in the public domain, as are a wide-range of medium spatial resolution satellite imagery. Here we describe the extension of a methodology based on Landsat ETM and Radarsat imagery to the production of a human settlement map of Kenya. This map was then compared with five satellite imagery-derived, global maps of urban extent at Kenya national-level, against an expert opinion coverage for accuracy assessment. The results showed the map produced using medium spatial resolution satellite imagery was of comparable accuracy to the expert opinion coverage. The five global urban maps exhibited a range of inaccuracies, emphasising that care should be taken with use of these maps at national and sub-national scale.

Keywords

Urban area mapping; urbanization; accuracy assessment

Introduction

Approximately half of the World population live in urban areas (United Nations 2004). This is anticipated to exceed 60% by 2030, with 90% of projected urbanization occurring in low-income countries (United Nations 2004). These population shifts will affect large social, environmental, economic and public health impacts (Tatem and Hay 2004; Hay *et al.* 2005). There is therefore a need for timely information on human settlement distribution, location and size. While such information is more readily available for high-income countries, maps of settlements in low-income countries are often outdated, inaccurate or non-existent. Confounding these difficulties is the lack of a clear definition between governments and major international bodies of what constitutes an ‘urban area’(Tatem and Hay 2004; Hay *et al.* 2005).

The ambiguity over what constitutes an ‘urban’ area and consequently, ‘urbanization’ has lead to several attempts to generate global and continental-scale urban maps using consistent

Correspondence: Dr Andy Tatem, TALA Research Group, Tinbergen Building, Department of Zoology, University of Oxford, South Parks Road, Oxford, OX1 3PS, U.K. Tel/Fax: +44 (0)1865 271262; andy.tatem@zoo.ox.ac.uk..

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

methods. These range from utilising military mapping data (Digital Chart of the World (Danko 1992)), to population census data centred-approaches (Global Rural-Urban Mapping Project (Center for International Earth Science Information Network (CIESIN) *et al.* 2004)), through to a range of Earth observation satellite-based techniques. For example, large-scale maps of urban areas have been produced *via* expert interpretation of Landsat Thematic Mapper (TM) imagery (Africover, URL: http://www.africover.org, maps of 10 East-African countries), classification of Advanced Very High Resolution Radiometer (AVHRR) imagery (Hansen *et al.* 1998), interpretation of Defense Meteorological Satellite Program – Operational Linescan System (DMSP-OLS) nighttime lights imagery (Elvidge *et al.* 1996; Elvidge *et al.* 1997; Elvidge *et al.* 1999; Sutton 2003) and by fusing various imagery sources (Schneider *et al.* 2003). Table 1 provides details of the most widely used global urban coverages. For low-income countries, occasional contemporary maps of major settlements exist for use in city-scale analysis. For regional and national-level studies (Snow *et al.* 2005), however, these global maps are all that is available.

Here, automated approaches to national-scale urban mapping from medium scale satellite imagery are investigated. The application of such an approach to four districts in Kenya has been described, (Tatem *et al.* 2004) and its extension to national coverage of Kenya at 25 m spatial resolution also detailed. Important to the utility of the map produced is its accuracy compared to that of expert-opinion ground-validated data at a similar scale, as well as to widely-used global urban maps. Here, such an assessment is undertaken for Kenya in order to look to future research requirements and answer two questions, (i) can global urban maps be used at a national scale in a low-income country and (ii) does the processing involved in producing a national-level 25 m spatial resolution urban map produce accuracy increases significant enough to justify their adoption over existing global maps?

Methods

Kenya settlement map (KSM) production

Production of a 25 m spatial resolution settlement map for Kenya followed the methodology outlined in Tatem *et al* (2004), a brief description of the approach is provided here.

Landsat Enhanced Thematic Mapper (ETM) scenes of Kenya taken between January and March from the years 2000-2002 were acquired from the Global Land Cover Facility at the University of Maryland (URL: http://www.glcf.umiacs.umd.edu). Bands 1-5, 7 and 8 were initially atmospherically corrected using the approach outlined in Richter (Richter 1990), but a countrywide fine spatial resolution digital elevation model was unavailable for topographic correction. The scenes were then mosaiced together using Erdas Imagine™ version 8.6 (ERDAS 2002) to create a single countrywide coverage. The Landsat ETM Kenya mosaic was then georegistered using Erdas Imagine™ to a set of national vector coverages representing census sublocations, gazetted areas, roads, rivers and Africover country borders. By selecting the least cloud contaminated scenes from three different years, only a small section of the south-east coastal region was obscured by cloud in the final coverage. At the same time a Radarsat-1 synthetic aperture radar (SAR) 25m spatial resolution mosaic of Kenya was acquired from Radarsat International. This was also georegistered to the suite of vector coverages and processed to produce eight texture layers. These texture layers were calculated from the grey level cooccurrence matrix method (Haralick *et al.* 1973; Haralick 1979) and consisted of mean, variance, constrast, homogeneity, dissimilarity, correlation, entropy and angular second moment using a 7 by 7 pixel moving window (Haralick *et al.* 1973; Baraldi and Parmiggiani 1995; Tatem *et al.* 2004).

The sixteen different imagery layers were then combined for classifier training. Rather than risk valuable information loss and complications in texture measure derivation by

Europe PMC Funders Author Manuscripts

resampling the Radarsat imagery to the spatial resolution of the 30 m Landsat ETM imagery, nearest neighbour resampling was used to upscale the Landsat imagery to the 25 m spatial resolution of the SAR imagery. In a similar way, the Landsat ETM band 8 was resampled to 25 m spatial resolution to match the other imagery.

To overcome computational limitations, the imagery was split into 16 segments. In each segment, spatial-spectral segmentation was undertaken, as described in Tatem *et al* (2004), and separate training and classification was then carried out within each spectrally and spatially contiguous zone. A feed-forward neural network classifier was trained as described in Tatem *et al* (2004), but also using Kenya-wide georeferenced settlement points and visual interpretation of 15 m spatial resolution Landsat ETM band 8 to identify training pixels.

Validation data

Africover data at full spatial resolution (1:100,000) were acquired (URL: http://www.africover.org). The purpose of the Africover project is to establish a digital georeferenced database on land cover and a geographic referential for the whole of Africa (FAO 1997). It was initiated in response to numerous national requests across Africa for reliable land cover maps. The data production process involved initially the creation of an inventory of ‘possible’ Kenyan land cover classes from the FAO/UNEP international standard land cover classification system (Di Gregorio and Jansen 1996; Di Gregorio and Jansen 1998) through consultation of ancillary data, local knowledge and field survey. Local photo-interpreters then delineated visually land cover polygons as seen from Landsat TM imagery, and highlighted areas of uncertainty. Fieldwork was then undertaken to resolve such uncertainties associated with satellite imagery interpretation, and used to update the land cover database. Finally a full reinterpretation based on all available information was undertaken to result in a rigorously-defined land cover map. For this study, all land-cover polygons classified as an urban area, rural settlement or refugee camp were compared with census sub-location data and unionized to make them contiguous administrative groupings (244 from 327 polygons), as described in Hay *et al* (2004). For example, the cluster of 12 polygons of urban areas in and around Mombasa were unionized to one. The centroids of these unionized polygons were assigned names and population counts using a CBS database of municipalities, town councils and other urban centres from the 1999 population and housing census (CBS 2001) (133 of 244 polygons). Finally, we took the population and name of the parent sub-location for all remaining Africover polygons (n = 111).

Other urban maps

Five global urban coverages were obtained for testing and each is described below.

A global land cover classification at a spatial resolution of 1 km, using 14 years of imagery from the NASA/NOAA Pathfinder Land (PAL) Advanced Very High Resolution Radiometer (AVHRR) data set (Hansen *et al.* 1998) was downloaded from the Global Land Cover Facility, University of Maryland (URL: http://glcf.umiacs.umd.edu/data/landcover/), and the ‘urban’ land cover class extracted.

Another global land cover classification at 1 km spatial resolution was obtained, this time using one year of Moderate Resolution Imaging Spectrometer (MODIS) data (Strahler *et al.* 2003). This was downloaded from Boston University’s department of geography land cover and land cover change data sets (URL: http://duckwater.bu.edu/lc/mod12q1.html#2001001) and the ‘urban’ class extracted.

The Digital Chart of the World’s (DCW) Populated Places was originally developed for the US Defense Mapping Agency (DMA) using DMA data. DCW coverage depicts urbanized

areas of the world that can be represented as polygons at 1:1000000 scale. The areas represent the shape of an urbanized area as viewed by an air observer (Danko 1992).

A night-time lights map was obtained, produced using time series data from the Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS) for the year 2000, where pixel values are the average digital number values for the year. The human settlements product represents the stable lights minus identified gas flares (Elvidge *et al.* 1996; Elvidge *et al.* 1997; Elvidge *et al.* 1999).

Finally, a preliminary version of an urban-rural extent map for Kenya, created as part of the Global Rural-Urban Mapping Project (GRUMP) (Center for International Earth Science Information Network (CIESIN) *et al.* 2004) was obtained. This was produced by combining information from night-time lights, DCW, Tactical Pilotage Charts produced by the Australian Defense imagery and Geospatial Organization and some Landsat-derived polygons. Night-time lights were used as baseline, then any polygons identified by other sources not intersecting with the lights were added (Balk *et al.* 2004; Center for International Earth Science Information Network (CIESIN) *et al.* 2004). The project is on-going and limitations are documented in the accompanying paper(Balk *et al.* 2004)

Hereafter, each of the seven urban maps studied in this paper will be referred to as follows; AVHRR global land cover classification as AVHRR, MODIS global land cover classification as MODIS, Digital Chart of the World urban polygons as DCW, DMSP-OLS human settlements nighttime lights as DMSP-OLS, Global Rural-Urban Mapping Project Urban Extents as GRUMP-UE, Kenya settlement map produced using the methodology outlined in Tatem *et al* (2004) as KSM and Africover urban land cover layer as Africover. Table 1 details the main features of each.

All surfaces were georeferenced according to the accompanying projection information and checked against a single Kenya national boundary vector file. The nighttime lights layer used here was a continuous surface and therefore thresholding was undertaken to discriminate urban from non-urban and lessen “blooming effects” (Elvidge *et al.* 1996). The blooming effect is dependent upon the characteristics of the DMSP-OLS and results in an overestimation of the extent of most urban areas. Conversely, the extents of small low illumination settlements in low-income countries are often underestimated, leaving the map producer the quandary of whether to threshold the data or not. For the GRUMP data, the decision was made not to threshold to enable detection of such low illumination settlements (Balk *et al.* 2004), so as to provide a comparison, a threshold was applied here. In similar studies the choice of threshold has been based on ‘visual inspection’ (e.g. (Schneider *et al.* 2003)). Here, following Henderson *et al* (2003), a more quantitatively-derived choice of threshold was used, aimed at giving the lights layer the best possible chance by choosing the optimal threshold to match the Africover data. The Africover urban layer was rasterized to the same 1 km grid as the lights surface, then each integer threshold in the 0-63 range of the lights values was applied to produce 64 different binary night-time light-derived urban maps. These were linearly-regressed against the 1 km Africover layer and it was found that a threshold of 22 produced the largest correlation coefficient (0.34), and so was adopted throughout the remainder of the study.

Each raster 1 km urban map (AVHRR, DMSP-OLS, GRUMP-UE, MODIS) was resampled to 25 m spatial resolution and each vector urban map (DCW, Africover) was rasterized to the same 25 m grid. Figure 1 shows a subset of each of the five maps under test for Nairobi, with the Africover Nairobi vector boundary overlain. Rather than degrading maps to 1 km for comparison, this choice of spatial resolution fitted the aim of this paper to test whether global urban coverages could be used with accuracy at the national-scale, and whether the

effort involved in producing a national-scale medium spatial resolution urban map produced a justifiable improvement in accuracy. The 25 m spatial resolution also minimised any information loss from all layers under test that may have occurred at coarser resolutions. Accuracy assessment was then carried out by comparing each urban map to that of the Africover urban land cover layer via traditional per-pixel based approaches (Congalton 1991) complemented by maplet-based methods (Stoms 1996; Cihlar *et al.* 2000; Schneider *et al.* 2003).

Accuracy assessment

Per-pixel comparison—An independent set of 20,000 test pixels were used for per-pixel comparison. Of the 703,291,922 25m pixels in the Africover validation map, just 587,581 were urban (0.084%). Formal sampling approaches (Stehman 2001) involving random sampling across Kenya to derive a test set of pixels were therefore inappropriate. Instead, 10,000 of the 20,000 test pixels were randomly sampled from Africover's urban class, with the other 10,000 randomly sampled from the non-urban class. Producer's, User's and overall accuracies were calculated and displayed in a table (Congalton 1991).

Settlement size—While per-pixel approaches were useful in assessing overall map accuracy, the addition of comparisons of individual urban area sizes proved a useful extra assessment of map quality. The assessment compared the area mapped as urban for each of the 244 Africover-defined settlements against the same areas in the other urban maps.

Here, assumptions were required to define what constituted an individual settlement in each map. For DMSP-OLS, GRUMP-UE, DCW, MODIS and AVHRR this was simply any contiguous group of pixels mapped as urban. For KSM where each individual 25 m pixel had been independently mapped as urban or non-urban and consequently settlements consisted of scattered groups of pixels, an individual settlement was delineated where urban pixel density was greater than or equal to 50% of a 100 m moving window. Scatterplots of estimated settlement size against Africover-defined settlement size were produced to highlight any bias, and overall root mean square errors (RMSEs) were calculated for each urban map.

Settlement location—Per-pixel and settlement size assessments provide useful information regarding agreement between each urban map and the Africover data. However, neither provides information on any geographic pattern errors may display. To explore this factor, a series of 25 m buffers were produced around the Africover representation of Kenya's four most populous settlements; Nairobi, Mombasa, Nakuru and Kisumu (CBS 2001). Each buffer was overlaid on each of the urban maps, and the percentage of pixels classified as urban (i.e. incorrect classification) falling within the buffer were calculated. Buffers further and further from the Africover-defined settlement boundary were used in this manner until no urban pixels classified as belonging to the settlement in question in each urban map fell within the buffer. Percentages from each buffer for each settlement were plotted to examine the pattern of errors within the five urban maps being tested.

Results

Per-pixel comparison

Table 2 shows the results of per-pixel comparisons between Africover and the six settlement maps tested. The largest user's, producer's and overall accuracies for both urban and non-urban classes were produced by KSM with values above 87% in all cases. This indicates accurate classification of both urban and non-urban sample pixels. Similar levels of accuracy were exhibited by DMSP-OLS, DCW and AVHRR in terms of non-urban class user's

accuracy and urban class producer's accuracy. This indicates that where pixels were classified as urban, the majority were urban in Africover, and the non-urban Africover pixels mostly coincided with non-urban classified pixels in the three maps. However, DMSP-OLS, DCW and AVHRR each showed significantly lower urban user's accuracy and producer's non-urban accuracies, highlighting how each layer failed to identify a large number of settlements mapped by Africover. GRUMP-UE shows the lowest overall accuracy despite identifying more settlements than the other three global maps. The reason for this low accuracy is highlighted by the producer's and user's accuracies. The high urban user's accuracy indicates that 86.6% of the sample of Africover's urban-assigned pixels were mapped correctly, but the non-urban user's accuracy of just 25.8% shows that only a quarter of the non-urban pixels were mapped correctly. This suggests that GRUMP-UE overestimates consistently urban extent in Kenya. With very similar user's and producer's accuracies, MODIS fell between the two extremes, suggesting settlement overestimation in some cases and underestimation in others.

Settlement size

Figures 2a-f show log-log scale scatterplots of Africover-derived settlement size against estimated settlement size from the five maps under test, with one-to-one lines superimposed. The plots show varying numbers of points dependent upon how many settlements each map shows, e.g. AVHRR only identifies nine settlements in Kenya, so only nine points are plotted. The RMSE values however are calculated by comparing the 244 Africover-identified urban areas with each layer, e.g. for AVHRR the estimated size of 235 of the settlements was zero.

RMSE values show large variations between urban maps. KSM shows the lowest error of 5.01 km^2, while AVHRR and DCW reveal similar results with RMSEs below 7 km^2. These low errors are not reflected for DMSP-OLS and GRUMP-UE with RMSEs of 43 km^2 and 340 km^2 respectively. MODIS has an RMSE of just 13.4 km^2, but interestingly, only 22 of the 150+ settlements mapped actually match those defined by Africover.

The location of each of the points above or below the one-to-one line on the scatterplots indicates whether an urban area has been over- or underestimated, and by how much compared to the Africover-derived areas. Figures 2a-e show that DCW and AVHRR map settlement size without bias, but DMSP-OLS, MODIS and GRUMP-UE consistently overestimate urban area, and in the case of GRUMP-UE, by large amounts. In contrast, while urban area is estimated accurately for many areas, there is a tendency towards underestimation of area by KSM.

Settlement location

Figures 3a-d show comparisons of locational accuracy for Kenya's four most populous settlements in terms of percentage of incorrectly classified pixels in each 25 m buffer ring surrounding the Africover-defined extent. Figures 3a, c and d, representing Nairobi, Nakuru and Kisumu reveal similar findings. In each case GPW-UR clearly shows locational error and overestimation of urban extent, and in the case of Nairobi still classifies some pixels incorrectly as urban almost 23 km from the Africover-defined urban boundary. Locational error and extent overestimation is also shown to a lesser extent in all figures by DMSP-OLS and MODIS. The other three urban maps consistently show more accurate location and extent estimation, each with no pixels incorrectly classified within 4 km of the Africover-defined boundaries for all four settlements.

Europe PMC Funders Author Manuscripts

Discussion

This study has revealed that semi-automated settlement mapping from a combination of medium spatial resolution optical, SAR and derived texture layers can produce a map of similar accuracy to one derived through time-consuming and costly visual interpretation of imagery and ground surveys. For Kenya, this mapping approach has also proved to be significantly more accurate in terms of mapping urban extent, size and location than any of the more widely available global urban maps. Results suggest that care should be taken in use of any of the five global urban maps tested here at a sub-national level. This must be balanced by the fact that in many instances there will be no alternative and is further elaborated below.

Limits and Transferability

While the results here do highlight inaccuracies associated with five widely-used global urban maps, it is of course important to emphasise that these findings are within the bounds of Kenya. However, this in no way means the methods, findings and implications are limited solely to within its national borders. Kenya represents a diverse country both ecologically and in terms of its urban landscape, ranging from the large international city of Nairobi, to the densely-populated highlands and the small, lowly-lit, isolated settlements of the arid north. In this way, it is representative of the vast majority of the worlds' low-income countries and suggests that the implications arising from this paper may be extrapolated with caution.

KSM

Focussing on the comparisons between Africover and KSM, it is clear that at per-pixel level strong agreement exists. Figure 2(f) however reveals a bias in KSM towards underestimating urban area extent. While this may at first appear a worrying feature, figure 1(f) demonstrates two causal factors.

Firstly, the Africover production process involved an object-based approach whereby the interpreter simply drew a boundary around what appeared visually to be 'urban'. The process used to produce KSM relied on an automated pixel-based approach to examine whether or not each individual pixel was similar enough to the 'urban' training signatures to be classified itself as urban. The heterogeneous nature of urban areas when viewed at 25 m spatial resolution, therefore produced a very speckled classification of an individual settlement. Whereas in Africover, everything within the interpreter-defined urban boundary was classed as 'urban', a pixel containing e.g. grass or water within supposed settlement boundaries would not be classified as 'urban' by the Tatem *et al* (2004) procedure. Hence, even after the application of 100 m majority filtering consistent underestimation occurred. Whether the correct 'urban extent' should include areas of water or parkland is open to interpretation, and KSM could be further processed to accommodate these.

Secondly, fig 1(f) also highlights how discrepancies occur when defining a settlement boundary. In fig 1(f), KSM has identified clearly many small settlements connected to Nairobi as well as apparent extensions to the city outside the Africover Nairobi boundary. Defining exactly where large urban areas end and separate surrounding smaller ones begin is a major problem.

Finally, while Africover's object-based and spatial resolution of analysis resulted in the identification of a total of 244 separate urban areas, KSM identified a massive 8621 distinct groups of 10 or more contiguous 'urban' pixels. Whether all these groups of pixels identified as settlements really are settlements is difficult to validate given the lack of available data. In a limited test, 87% of the pixel groups were within 500 m of a major road, suggesting

correct classification in most cases. There are far more than 244 individual settlements in Kenya (CBS 2001) and just visual examination reveals that even the most rigorous of classification strategies in the form of Africover have missed even occasional significant settlements such as the municipal capital of Wajir district a town of some 45,000 inhabitants.

The global maps

Given the aims of each global urban map producer are to create a global urban coverage at 1 km spatial resolution for the point in time of production, it is not surprising that each faired poorly in a comparison with 2002 data at national-level and at 25 m spatial resolution. However, quite how inaccurate each was at national-level is cause for concern, particularly given that for many of the low-income regions of the World there simply is no contemporary alternative for regional to national-level analysis. While some maps and imagery sources are in continual development, provide feedback mechanisms for errors found and readily document limitations (e.g. (Elvidge *et al.* 1999; Balk *et al.* 2004), others provide no such flexibility and warnings to users.

The Africover map reveals that there are at least 72 settlements in Kenya of over 1 km^2 in area, yet only GRUMP-UE and MODIS come anywhere close to identifying this many. Figure 2 shows, however, that only a small proportion of the settlements identified by MODIS coincide with those mapped by Africover. Worryingly, further visual comparison of MODIS with the other urban maps and Landsat TM imagery suggests that many of the 150+ individual groups of ‘urban’ pixels have been incorrectly classified. The identification of just 12 and 15 settlements in DCW and AVHRR respectively, represents a very worrying statistic should these maps be utilised at a sub-national scale. Interestingly, for the settlements that were mapped, DCW and AVHRR showed little bias in area estimation, reflected in their relatively low RMSEs. However, the other three global maps tested showed a bias towards overestimation of urban extent, especially GRUMP-UE, exhibiting a massive RMSE of 340.45 km^2. In each of these urban maps there were settlements estimated to be over an order of magnitude larger in area than that estimated by Africover. In some cases this can be explained either by artefacts of comparing 1 km spatial resolution maps at 25 m spatial resolution, or the fact that a group of contiguous pixels may represent more than one settlement. In the case of GRUMP-UE, the area overestimation is due to the use of unthresholded nighttime lights data, incorporating the blooming effect associated with this imagery. Compensation for this effect is vital if the principal usage is to obtain urban area estimates, otherwise overestimation of urban coverage will occur (e.g. (Keiser *et al.* 2004)) potentially invalidating inferences (Bremen *et al.* 2004) based on the estimates (Hay and Tatem 2005). The results of this study for DMSP-OLS and GRUMP-UE are evidence of the quandary facing those producing global settlement maps using currently available nighttime lights data; either leave the data as it is and risk overestimating settlement size or threshold and risk missing multiple settlements. A new generation of annual global OLS nighttime lights currently in production should help take away this difficult choice with the ability to detect small, low illumination settlements while thresholding large, well-electrified cities that tend to bloom (Balk *et al.* 2004). Another approach currently at pilot study stage aims to combine MODIS and DMSP-OLS with other data for more effective urban detection (Schneider *et al.* 2003).The provision of improved imagery sources and development of such multi-source approaches suggests many of the problems highlighted in this paper may be reduced or eliminated in the near future.

Conclusions

With projected urbanization concentrated in low-income countries, improved knowledge on settlement distribution, location and size is important for the assessment of environmental,

Europe PMC Funders Author Manuscripts

socioeconomic and demographic change. The urban maps examined here represent the first steps in filling this knowledge gap.

Two principal research questions were defined at the outset of this paper: (i) can global urban maps be used accurately at a national scale for a low-income country?, and (ii) does the processing involved in producing a national-level 25 m spatial resolution urban map produce significant enough accuracy increases to justify its adoption over global-level maps at a national scale?

In many low-income regions of the World, the only up-to-date maps of urban extents are those produced on a global-scale. However, in response to research question (i), the various inaccuracies highlighted show that to conduct research or form policy decisions at a national-level based solely on urban data from the AVHRR, DCW, DMSP-OLS, GRUMP-UE and MODIS global maps in their current incarnations would require very careful interpretation. Results here emphasise clearly that such maps were produced on a global scale with the aim of being used in global studies, and thus should be used exclusively for such purposes.

This study has also shown that a semi-automated approach to settlement mapping based on medium spatial resolution satellite imagery can produce a map of comparable accuracy to one produced through visual interpretation and ground survey. In response to research question (ii), results show that such a map exhibits much greater accuracy in urban delineation justifying adoption over any global map tested here. In low-income regions where the speed of urbanisation is creating the greatest need for contemporary maps and where resources are not available for map update *via* traditional survey-based methodologies, this may represent a cost-effective viable alternative.

At a time when urbanization in low-income countries is receiving much attention (UN Habitat 2003; UN Habitat 2004), satellite imagery has an important role to play in aiding the production of urban maps for city-, regional-, national- and global-scale analysis. Users of urban maps need to be made aware of limitations, and providers should remain flexible in their ability to update, adapt and incorporate new information, imagery and approaches. If this can be achieved, then satellite imagery-derived urban maps will prove vital tools in extending the study of, among others, population (Hay *et al.* 2005), childhood mortality (Balk *et al.* 2004) and malaria burden (Hay *et al.* 2004; Hay *et al.* 2005) across the low-income regions of the World. Continued successful usage will in turn provide feedback on optimal approaches for mapping improvement.

Acknowledgments

We thank Andy Nelson, Alastair Graham, Marvin Bauer and three anonymous reviewers for comments on earlier drafts of the manuscript. We are grateful to the Africover team for their efficiency and advice. We are similarly grateful to Deborah Balk and Greg Yetman for supplying alpha versions of GPW-UR for testing and scrutiny. SIH and AJT are funded by a Research Career Development Fellowship from the Wellcome Trust (No.06904) to SIH. We acknowledge the support of the Kenyan Medical Research Institute (KEMRI) and this paper is published with the permission of its director.

References

Balk, D.; Pozzi, F.; Yetman, G.; Deichmann, U.; Nelson, A. The distribution of people and the dimension of place: Methodologies to improve global population estimates in urban and rural areas. CIESIN, Columbia University; New York: 2004.

Balk D, Pullum T, Storeyguard A, Greenwell F, Neuman M. A spatial analysis of childhood mortality in West Africa. Population, Space and Place. 2004; 10:175–216.

Baraldi A, Parmiggiani F. An investigation of the textural characteristics associated with the gray level cooccurrence matrix statistical parameters. IEEE Transactions on Geoscience and Remote Sensing. 1995; 33:653–666.

Bremen JG, Alilio MS, Mills A. Conquering the intolerable burden of malaria: what's new, what's needed: a summary. American Journal of Tropical Medicine and Hygiene. 2004; 71:1–15.

CBS. 1999 population and housing census: counting our people for development. Volume 1: Population distribution by administrative areas and urban centres. Central Bureau of Statistics (CBS), Ministry of Finance and Planning, Government of Kenya; Nairobi: 2001.

Center for International Earth Science Information Network (CIESIN). Columbia University. International Food Policy Research Institute (IFPRI). The World Bank. Centro International de Agricultura Tropical (CIAT). Global Rural-Urban Mapping Project (GRUMP): Urban Extents. CIESIN, Columbia University; Palisades, New York: 2004.

Cihlar JR, Latifovic J, Chen J, Beaubien Z, Magnussen S. Selecting representative high resolution sample images of land cover studies, Part 2: Application to estimating land cover composition. Remote Sensing of Environment. 2000; 72:127–138.

Congalton RG. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment. 1991; 37:35–46.

Danko DM. The digital chart of the world project. Photogrammetric Engineering and Remote Sensing. 1992; 58:1125–1128.

Di Gregorio, A.; Jansen, LJM. Part I - Technical document on the Africover land cover classification scheme. FAO Africover Land Cover Classification; 1996. p. 4-33.

Di Gregorio, A.; Jansen, LJM. Land Cover Classification System (LCCS): classification concepts and user manual. Environment and Natural Resources Service (SDRN); Rome: 1998. p. 157

Elvidge C, Baugh K, Kihn EA, Davis ER. Mapping city lights with nighttime data from the DMSP operational linescan system. Photogrammetric Engineering and Remote Sensing. 1996; 63:727–734.

Elvidge C, Baugh KE, Hobson VR, Kihn EA, Kroehl HW, Davis ER, Cocero D. Satellite inventory of human settlements using nocturnal radiation emissions: a contribution to the global toolchest. Global Change Biology. 1997; 3:387–395.

Elvidge CD, Baugh KE, Dietz JB, Bland T, Sutton PC, Kroehl HW. Radiance calibration of DMSP-OLS low-light imaging data of human settlements. Remote Sensing of Environment. 1999; 68:77–88.

ERDAS. ERDAS Field Guide. 6th Edition. ERDAS Inc.; Atlanta: 2002.

FAO. Africover Land Cover Classification. FAO; Rome: 1997.

Hansen, M.; DeFries, R.; Townshend, JRG.; Sohlberg, R. 1 km land cover classification derived from AVHRR. The Global Land Cover Facility; College Park, Maryland: 1998.

Haralick RM. Statistical and structural approaches to texture. Proceeding of the IEEE. 1979; 67:786–804.

Haralick RM, Shanmugam K, Dinstein I. Texture features for image classification. IEEE Transactions on Systems, Man, and Cybernetics. 1973; 3:610–621.

Hay SI, Guerra CA, Tatem AJ, Atkinson PM, Snow RW. Urbanization, malaria transmission and disease burden in Africa. Nature Reviews Microbiology. 2005; 3:81–90.

Hay SI, Guerra CA, Tatem AJ, Noor AM, Snow RW. The global distribution and population at risk of malaria: past, present and future. Lancet Infectious Diseases. 2004:4.

Hay SI, Noor AM, Nelson A, Tatem AJ. Demography for epidemiology: the precision of large-area human population maps. International Journal of Epidemiology. 2005 Submitted.

Hay SI, Tatem AJ. Remote sensing of malaria in urban areas: two scales, two problems. American Journal of Tropical Medicine and Hygiene. 2005 in press.

Keiser J, Utzinger J, Caldas de Castro M, Smith TA, Tanner M, Singer BH. Urbanization in sub-Saharan Africa and implication for malaria control. American Journal of Tropical Medicine and Hygiene. 2004; 71:118–127. [PubMed: 15331827]

Richter R. A fast atmospheric correction algorithm applied to Landsat TM images. International Journal of Remote Sensing. 1990; 11:159–166.

Schneider A, Friedl MA, McIver DK, Woodcock CE. Mapping urban areas by fusing multiple sources of coarse resolution remotely sensed data. Photogrammetric Engineering and Remote Sensing. 2003; 69:1377–1386.

Snow RW, Guerra CA, Noor AM, Myint HY, Hay SI. The global distribution of clinical episodes of *Plasmodium falciparum malaria*. Nature. 2005 in press.

Stehman SV. Statistical rigor and practical utility in thematic map accuracy assessment. Photogrammetric Engineering and Remote Sensing. 2001:67.

Stoms DM. Validating large area land cover databases with Maplets. Geocarto International. 1996; 11:87–95.

Strahler, A.; Friedl, MA.; Schaaf, CLB.; Zhang, X.; Hodges, JCF. Validation of the consistent - year v003 MODIS land cover product. Boston University; Boston: 2003.

Sutton P. A scale-adjusted measure of “urban sprawl” using nighttime satellite imagery. Remote Sensing of Environment. 2003; 86:353–369.

Tatem AJ, Hay SI. Measuring urbanization pattern and extent for malaria research: a review of remote sensing approaches. Journal of Urban Health. 2004; 81:363–376. [PubMed: 15273262]

Tatem AJ, Noor AM, Hay SI. Defining approaches to settlement mapping for public health management in Kenya using medium spatial resolution satellite imagery. Remote Sensing of Environment. 2004; 93:42–52. [PubMed: 22581984]

UN Habitat. The challenge of slums - global report on human settlements 2003. 2003.

UN Habitat. Report of the second session of the World Urban Forum. 2004.

United Nations. World urbanisation prospects, 2004 revision, United Nations. 2004.

Europe PMC Funders Author Manuscripts

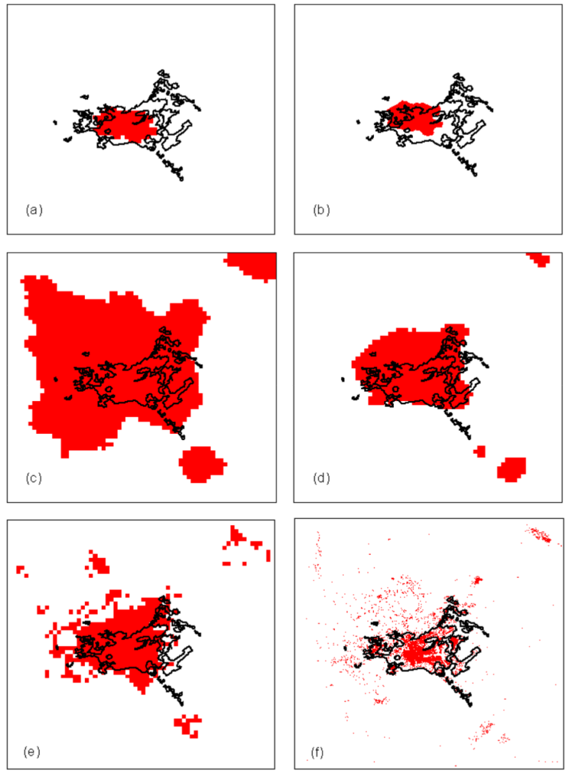
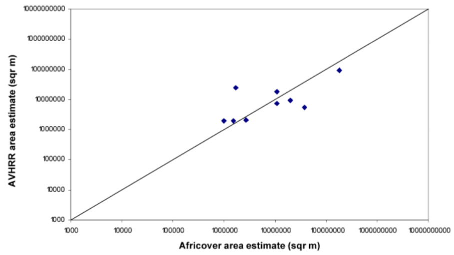
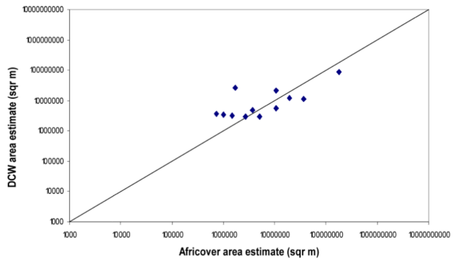


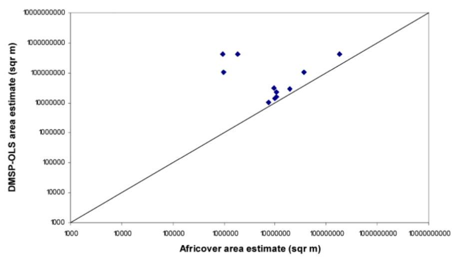
Figure 1.
Nairobi mapped by the six settlement maps under test, with the Africover Nairobi vector boundary overlain: (a) AVHRR; (b) DCW; (c) GRUMP-UE, (d) DMSP-OLS, (e) MODIS, (f) KSM.



(a)



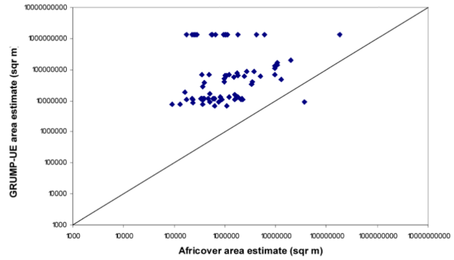
(b)



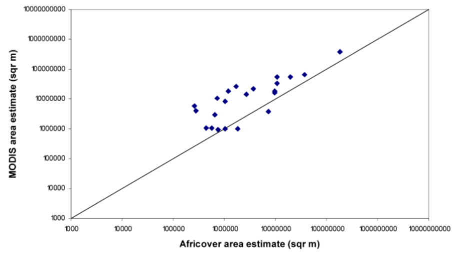
(c)

Europe PMC Funders Author Manuscripts

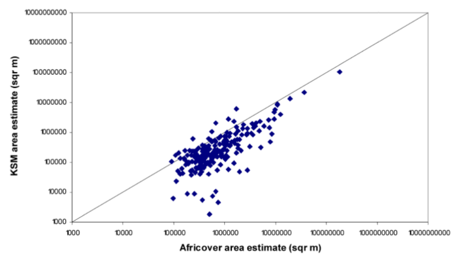
Europe PMC Funders Author Manuscripts



(d)



(e)



(f)

Figure 2.

Scatterplots of Africover-estimated settlement size against estimated settlement size by (a) AVHRR, RMSE = 6.6 km^2; (b) DCW, RMSE = 6.6 km^2; (c) DMSP-OLS, RMSE = 43.0 km^2; (d) GRUMP-UE, RMSE = 340.5 km^2; (e) MODIS, RMSE = 13.4 km^2; (f) KSM, RMSE = 5.0 km^2.

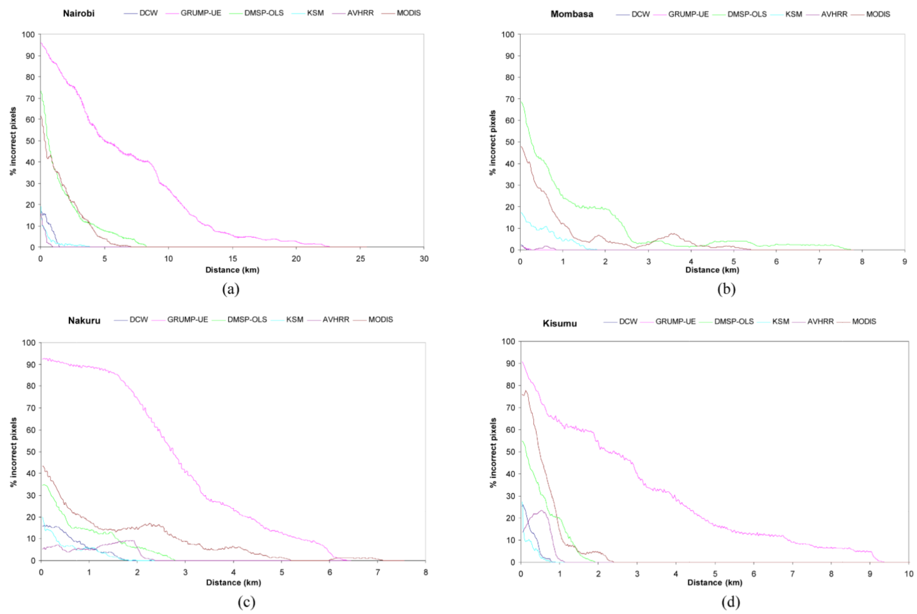


Figure 3.
A comparison of locational accuracy in the 6 urban maps tested for (a) Nairobi; (b) Mombasa; (c) Nakuru; (d) Kisumu.

Europe PMC Funders Author Manuscripts

Table 1

Features of each of the seven global urban coverages compared in this study. Africover = FAO Africover: Kenya urban land cover; DCW = Digital Chart of the World; Urbanized area polygons; GRUMP-UE = Global Rural-Urban Mapping Project Urban Extents; DMSP-OLS = Defense Meteorological Satellite Program-Operational Linescan System: Nighttime Lights of the World; Human Settlements 2000; AVHRR = AVHRR Global Land Cover Classification urban land cover class; MODIS = MODIS Land Cover Product Binary Data from Boston University; KSM = Medium spatial resolution satellite imagery derived Kenya settlement map.

Urban Map	Spatial Resolution	Production Year	Number of individual settlements mapped for Kenya	Reference/Source
Africover	Vector polygons: 1:100000	2002	244	FAO 1997; http://www.africover.org/
DCW	Vector polygons: 1:1000000	1993	12	Danko 1992
GRUMP-UE	1 km	2004	84	CIESIN 2004, Balk 2004
DMSP-OLS	1 km	2000	12 (using threshold of 22)	http://dmsp.ngdc.noaa.gov/html/download_world_change_pair.html
AVHRR	1 km	1998	15	Hansen *et al* 1998
MODIS	1 km	2002	Approximately 150	Strahler *et al* 2003; http://duckwater.bu.edu/lc/mod12q1.html
KSM	25 m	2004	Dependent upon filter applied	Tatem *et al* 2004 for methodology

Europe PMC Funders Author Manuscripts

Table 2

Results of per-pixel comparison of each urban map under test with Africover.

	Urban Class		Non-Urban Class		
Urban Map	**Producer's Accuracy (%)**	**User's Accuracy (%)**	**Producer's Accuracy (%)**	**User's Accuracy (%)**	**Overall Accuracy (%)**
AVHRR	90.1	29.9	58.1	97	63.5
DCW	89.1	30.5	58.1	96.3	63.4
DMSP-OLS	85.6	53.2	66.1	91.1	72.1
GRUMP-UE	53.9	56.6	65.8	25.8	56.2
MODIS	62	64.9	63.1	60.2	62.5
KSM	98	87.6	88.8	98.2	92.9