

RESEARCH OUTPUTS / RÉSULTATS DE RECHERCHE

Improving aфриpop dataset with settlement extents extracted from rapideye for the border region comprising South-Africa, Swaziland and mozambique

Deleu, Julie; Franke, Jonas; Gebreslasie, Michael; Linard, Catherine

Published in:
Geospatial Health

DOI:
[10.4081/gh.2015.336](https://doi.org/10.4081/gh.2015.336)

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for pulished version (HARVARD):
Deleu, J, Franke, J, Gebreslasie, M & Linard, C 2015, 'Improving aфриpop dataset with settlement extents extracted from rapideye for the border region comprising South-Africa, Swaziland and mozambique', *Geospatial Health*, vol. 10, no. 2, 336, pp. 48-54. <https://doi.org/10.4081/gh.2015.336>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Improving AfriPop dataset with settlement extents extracted from RapidEye for the border region comprising South-Africa, Swaziland and Mozambique

Julie Deleu,¹ Jonas Franke,² Michael Gebreslasie,³ Catherine Linard⁴

¹*EUROSENSE, Wommel, Belgium;* ²*Remote Sensing Solutions GmbH, Baierbrunn, Germany;* ³*School of Agricultural, Earth and Environmental Sciences, University KwaZulu-Natal, Durban, South Africa;* ⁴*Biological Control and Spatial Ecology Lab, Université Libre de Bruxelles, Belgium*

Abstract

For modelling the spatial distribution of malaria incidence, accurate and detailed information on population size and distribution are of significant importance. Different, global, spatial, standard datasets of population distribution have been developed and are widely used. However, most of them are not up-to-date and the low spatial resolution of the input census data has limitations for contemporary, national-scale analyses. The AfriPop project, launched in July 2009, was initiated with the aim of producing detailed, contemporary and easily updatable population distribution datasets for the whole of Africa. High-resolution satellite sensors can help to further improve this dataset through the generation of high-resolution settlement layers at

greater spatial details. In the present study, the settlement extents included in the MALAREO land use classification were used to generate an enhanced and updated version of the AfriPop dataset for the study area covering southern Mozambique, eastern Swaziland and the malarious part of KwaZulu-Natal in South Africa. Results show that it is possible to easily produce a detailed and updated population distribution dataset applying the AfriPop modelling approach with the use of high-resolution settlement layers and population growth rates. The 2007 and 2011 population datasets are freely available as a product of the MALAREO project and can be downloaded from the project website.

Introduction

The population at risk of a vector-borne disease such as malaria is a function of the coincidence of human population and malaria transmission risk. The spatial and temporal distribution of malaria risk is closely related to a wide range of environmental parameters, such as the distance to water, land surface temperature, vegetation distribution, rainfall, elevation, *etc.* (Rogers *et al.*, 2002; Ceccato *et al.*, 2005). The spatial heterogeneities in malaria transmission and spread make risk maps and modelling methods valuable tools for understanding malaria dynamics and planning public health interventions (Elliott and Wartenberg, 2004; Ostfeld *et al.*, 2005; Riley, 2007). Yet, in deriving accurate estimates of population at risk of malaria, these spatial models must rely on accurate and detailed information on population distribution and density.

The most widely used global spatial datasets of population distribution in disease-related studies are the Gridded Population of the World (Deichmann *et al.*, 2001; Balk and Yetman, 2004), the Global Rural Urban Mapping Project (Balk *et al.*, 2006), LandScan (Dobson *et al.*, 2000), and the United Nations Environment Programme Population Databases (Deichmann, 1996). Characteristics of each dataset are outlined in Table 1.

Population distribution modelling methods over large spatial scales rely on redistributing populations within census units to obtain continuous population surfaces, *i.e.* gridded datasets with a number of inhabitants per cell. The four main global gridded population datasets described above are using four different interpolation methods for the spatial reallocation of census population. As a consequence, different sources of error and uncertainty are associated with these population datasets, which generally arise from i) the input data; ii) population growth forecast; and iii) the modelling procedure used (Linard and Tatem, 2012). Tatem *et al.* (2011) showed that, in the context of an

Correspondence: Julie Deleu, EUROSENSE, Nervierslaan 54, 1780 Wommel, Belgium. Tel: +32.2.4607000 - Fax: +322.4604958.
E-mail: julie.deleu@eurossense.com

Key words: Population mapping; Health; Malaria risk; Remote sensing; Southern Africa.

Funding: AfriPop project was partially funded by the European Commission (EC) under the Seventh Framework Programme (FP7). Earth observation data were provided by the European Space Agency (ESA) managed the GMES Space Component Data Access (GSC-DA), funded under the ESA-EC Agreement on the Implementation of the Space Component of GMES.

Note: the present paper includes results from MALAREO (2011-2013; http://cordis.europa.eu/project/rcn/97892_en.html), a project in the Seventh Framework Programme of the EU, aiming to stimulate and facilitate the use of earth observation in malaria control and management in South-Africa, Swaziland and Mozambique.

Received for publication: 25 February 2015.
Accepted for publication: 25 February 2015.

©Copyright J. Deleu *et al.*, 2015
Licensee PAGEPress, Italy
Geospatial Health 2015; 10:336
doi:10.4081/gh.2015.336

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (by-nc 3.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

endemic, vector-borne disease, the choice of spatial population dataset can have substantial effects on estimates of populations at risk of disease. This is especially true for low-income countries where estimates of national population totals are uncertain, *e.g.*, due to outdated census data and provided at low spatial resolutions.

The AfriPop project (<http://www.worldpop.org.uk/>), initiated in July 2009, aimed to produce detailed and open access gridded population distribution data for all African countries. The project was recently extended to Asia and Central and South America and renamed WorldPop (www.worldpop.org.uk). The AfriPop approach is designed to produce datasets based on freely available data and methods that can easily incorporate new data as it becomes available. AfriPop has therefore constructed a geographic information system (GIS)-linked database with census and official population data using the most recent and spatially detailed datasets available. Population count data have now been combined with land cover (LC) data to map population distributions across Africa at a finer spatial resolution than ever before (Linard *et al.*, 2012). Landsat-derived settlement extents were used to refine existing LC data sets and the results show that the use of the GlobCover dataset (2009; <http://due.esrin.esa.int/globcover/>), refined with detailed settlement extents, provides significantly more accurate results than previously for the production of gridded population datasets across large areas (Linard *et al.*, 2010b). AfriPop mainly uses 30 m Landsat enhanced thematic mapper satellite imagery for the mapping of settlements.

The AfriPop modelling method distinguishes between urban and rural population when redistributing populations. Major cities have already validated population numbers derived from census enumerations. The remaining rural population is redistributed using LC-based weightings. The refined LC data and fine resolution population data have been used to define population densities per LC class (*i.e.* the average number of people per pixel), following approaches previously outlined (Tatem *et al.*, 2007; Linard *et al.*, 2010b). These LC specific population densities were then used as weights to redistribute the rural populations within administrative units in the African countries.

Assessment results have shown that the AfriPop maps are more accurate than existing population map products as well as simple gridding of census data. Moreover, the WorldPop 100-m spatial resolution represents a finer mapping detail than has ever before been produced over large areas. Recent work, forming part of the AfriPop project, has shown that detailed satellite imagery-based mapping of settlements combined with LC information can be used to increase population model accuracies across large areas (Tatem *et al.*, 2007). Particularly in rural Africa, settlements are often characterized by very small spatial scales with small structures and non-dense development, which cannot be accurately detected with mid-resolution satellite data such as Landsat. New high-resolution satellite sensors that are able to cover large areas such as RapidEye (2011 version, provided under European Commission/European Space Agency GMES Space Component Data Access), SPOT or the upcoming Sentinels, create new opportunities for



Figure 1. Example of settlement extraction. (a) RapidEye true colour composite (R:3, G:2, B:1), subset shows the city of Naamacha in Mozambique; (b) same image as (a) but with the outlines of the RapidEye derived settlements overlaid; (c) same image as (a) but with the outlines of the AfriPop settlements overlaid.

Table 1. Global population datasets and their main characteristics.

Dataset	Year(s) represented	Spatial resolution	Input data used	Source
LandScan	1998, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011	30 arcsec (~1 km)	Census, LC, elevation, slope, roads, populated areas / points	http://www.ornl.gov/sci/landscan
Gridded Population of the World	1990, 1995, 2000, 2005, 2010, 2015	2,5 arcmin (~5 km)	Census, water bodies (for masking)	http://sedac.ciesin.columbia.edu/gpw/global.jsp
Global Rural Urban Mapping Project	1990, 1995, 2000	30 arcsec (~1 km)	Census, populated areas, water bodies (for masking)	http://sedac.ciesin.columbia.edu/gpw/global.jsp
United Nations Environment Programme Global Population Database	2000	2,5 arcmin (~5 km)	Census, populated points, roads	http://na.unep.net/siouxfalls/datasets/datalist.php

LC, land cover.

improving spatial population values.

The present study aimed to further improve the AfriPop dataset by using settlement extents extracted from 5 m spatial resolution RapidEye data. A spatially enhanced and updated AfriPop dataset at a spatial resolution of 10 m was generated for the MALAREO study area covering southern Mozambique, eastern Swaziland and the malarious part of KwaZulu-Natal in South Africa and compared to existing datasets (http://cordis.europa.eu/project/rcn/97892_en.html).

Materials and Methods

Improving the AfriPop dataset with settlement extents derived from RapidEye is undertaken through three steps: i) settlement extraction from RapidEye images; ii) LC data refinement; and iii) population distribution modelling.

Settlement extraction

High-resolution Level 3, orthorectified RapidEye archive data were used to map the LC in the MALAREO study area. The RapidEye satellite system (<http://blackbridge.com/rapideye/index.html>) is a constellation of five identical satellites that provide data in five spectral bands with a spatial resolution of 5 m. Data were acquired between 18/07/2011 and 10/11/2011. The LC mapping approach was based on a classification scheme including 11 classes relevant to characterize the malaria vector habitat and the pattern of human activities (Franke *et al.*, in press). For this purpose, a hybrid classification scheme between LC and land use was used. The LC/use classification of the high-resolution data was conducted by applying an object-based image analysis with a predefined hierarchical rule-set in the software eCognition (Trimble Geospatial Imaging, Munich, Germany). The image objects were assigned to predefined classes according to decision rules based on spectral, spatial, shape, thematic or topologic criteria. The LC mapping approach has been outlined in more detail elsewhere (Franke *et al.*, in press). The MALAREO land use class *Settlement/Infrastructure* identifies built-up areas where the majority of people are located. Transportation infrastructure like roads and tracks were classified separately. The settlement delineations were used to generate an enhanced and updated version of the AfriPop dataset for the study area. The high spatial resolution of the RapidEye data allowed a detailed identification of human settlements, particularly of small homesteads in the rural area. Figure 1 compares the MALAREO settlement extents at 5 m with the AfriPop settlements at 30 m.

Land cover data refinement

The GlobCover dataset was refined with the more detailed settlement outlines derived from the RapidEye images following methods described elsewhere (Tatem *et al.*, 2005, 2007; Linard *et al.*, 2010b, 2012). GlobCover in combination with detailed settlement extents was found to be the most useful resource for the production of gridded population datasets across large areas (Linard *et al.*, 2010b). The GlobCover map was first resampled into 10-m spatial resolution by nearest-neighbour interpolation. The urban areas, which typically overestimate the settlement extent size, were removed and the surrounding classes correspondingly expanded to fill the remaining space. The MALAREO settlement map was also degraded to 10-m spatial resolution. These settlement extents were overlaid onto the GlobCover and the LCs beneath were replaced to produce a refined LC dataset.

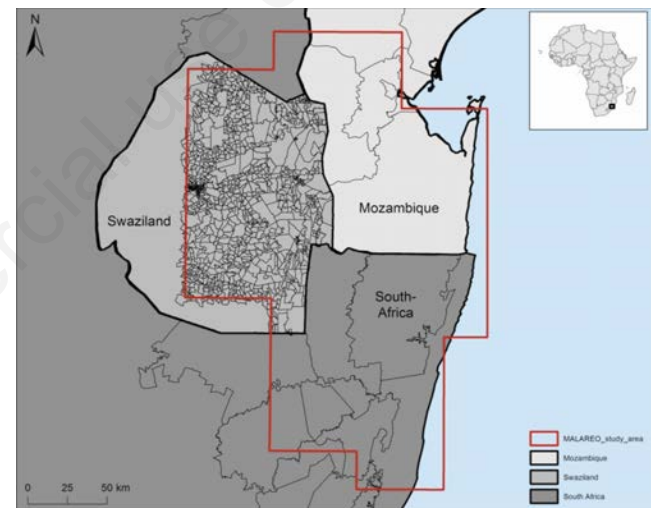


Figure 2. The highest levels of administrative boundaries for which national census data is available for the MALAREO study area.

Table 2. Average population densities for each land cover class in the MALAREO study area.

GlobCover aggregated LC classes	Population density (people/km ²)
Cultivated terrestrial areas and managed lands	129.34
Natural and semi-natural terrestrial vegetation-woody/trees	72.55
Natural and seminatural terrestrial vegetation-shrubs	64.97
Natural and seminatural terrestrial vegetation-herbaceous	18.22
Natural and semi-natural sparse terrestrial vegetation	36.36
Natural and seminatural aquatic vegetation	53.44
Bare areas	16.02
Settlements (extracted from RapidEye)	2486.65
Waterbodies	0

LC, land cover.

Population distribution modelling

Human population census data and corresponding administrative unit boundaries were obtained at the highest level available from the most recent census datasets in Mozambique (2007, administrative level 2 - district), South-Africa (2007, administrative level 3 - municipality), and Swaziland (2007, administrative level 4 – enumeration area) (Figure 2). The aggregated census counts were redistributed at a finer spatial scale using the AfriPop modelling method (Linard *et al.*, 2012). Since the RapidEye 5-m resolution data do not include every small hut, population density weightings were assigned to the LC classes. The relative population densities per LC class were defined for each class of the refined LC dataset based on the majority method (Linard *et al.*, 2010b, 2012) using the whole study area as training data, *i.e.* the average population density of each specific LC class was calculated based on the administrative units that had recorded this LC class for the majority of their pixels. Zero population was assigned to classes with no human habitation (mainly water bodies) and to all pixels that were not settlements and located in protected areas (IUCN and UNEP, 2012: <http://www.protectedplanet.net>). These LC-specific population densities were then used as weights to redistribute populations within the administrative units (Table 2).

A population distribution map for 2007 was processed based on census 2007 data. Additionally, population density estimations for 2011 were calculated based on 2011 census data (at the municipality level) for South Africa. For Swaziland and Mozambique, UN rural and urban annual growth rates for the periods 2005-2010 and 2010-2015 were used to predict populations density values for 2011 (Linard *et al.*, 2010a). These growth rates were taken from the UN World Urbanization Prospects Database, 2011 version (<http://esa.un.org/unup/>).

Comparison with existing datasets

Two comparison tests were undertaken between the original AfriPop dataset and the enhanced MALAREO population density map, *i.e.* per-pixel comparison between settlement extents and per-pixel absolute differences between population datasets.

With respect to the former, the MALAREO settlement extents derived from RapidEye 2011 imagery were compared with the AfriPop settlements based on 2005 Landsat 30-m data. The accuracy assessment was based on per-pixel comparison as described by Tatem *et al.* (2005). An independent set of 2000 test pixels were used for the per-pixel comparison. A set of 1000 of the 2000 test pixels was randomly sampled from the AfriPop urban class, while the other 1000 test pixels were randomly sampled from the AfriPop non-urban class and producer's, user's and overall accuracies calculated (Congalton, 1991).

To make the latter comparison possible, population datasets of the same year (2007) were used and resampled to the 10-m spatial resolution. The resulting grid-based difference measures led to interesting insights on where the differences between datasets are the most significant.

Accuracy assessment

Validation and accuracy assessment of high-resolution population data is challenging as few independent reference data are generally available. Map accuracies can be tested in target regions where reference data are available at a higher administrative unit level than used in the population estimation modelling. The National Statistical Office of South Africa obtained such census counts which were carried out in 2011 for a part of the study area in KwaZulu-Natal. The goal of the accuracy assessment was twofold. Firstly, the predicted population totals per

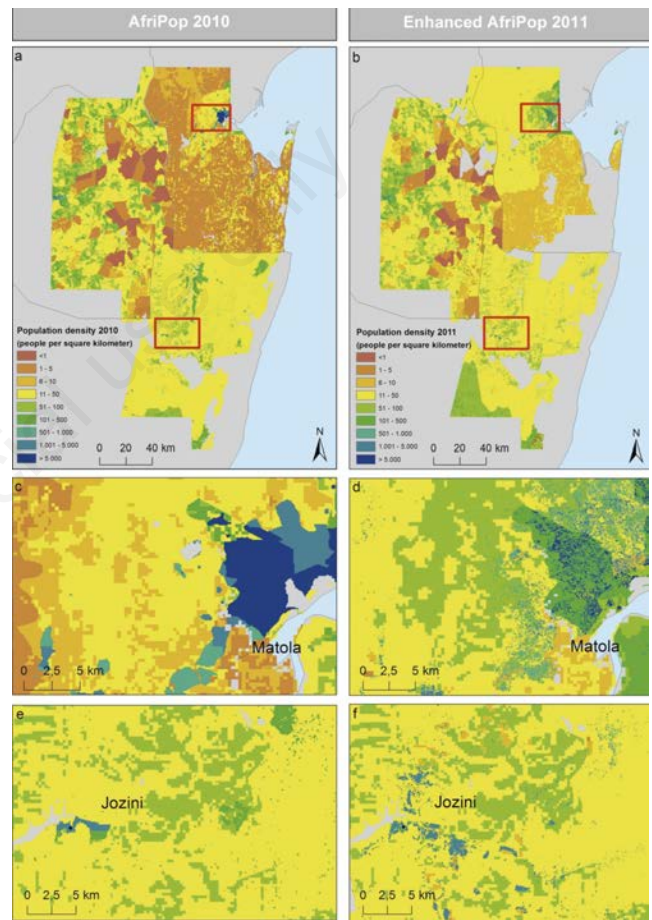


Figure 3. Comparison of gridded population datasets for MALAREO study area between AfriPop 2010 and the enhanced AfriPop 2011. (a) AfriPop dataset showing population numbers in each 100×100 meter grid square; (b) newly created population dataset showing population numbers in each 10×10 meter grid square, subset of the AfriPop dataset showing detail around; (c) Matola; and (e) Jozini. Subset of the enhanced AfriPop dataset showing detail around (d) Matola and (f) Jozini.

Table 3. Result of the per-pixel comparison between the AfriPop settlements (2005) and the MALAREO settlement extents derived from RapidEye 2011 imagery.

	Urban class		Non-urban class		Overall accuracy
	Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy	
AfriPop	95.1	19.5	55.2	99.0	59.3

Values are expressed as percentage.

municipality were compared to the population census data used in the modelling process. The second goal was to assess the correlation between the number of people predicted at validation data level (ward level) with the actual population received from the National Statistical Office of South Africa. To support the error analysis and comparison of the original and the enhanced AfriPop 2011 map, the difference between the estimated and actual population for the KwaZulu-Natal wards completely covered by the study area was calculated. Accuracy statistics including mean absolute error (MAE), root mean square error (RMSE) and correlation coefficient (CC) were computed.

Results

The final results of the AfriPop improvement is freely available as a product of the MALAREO project and can be downloaded via http://cordis.europa.eu/project/rcn/97892_en.html. A brief description follows below.

Enhanced AfriPop population density map

The AfriPop modelling method was designed to produce datasets based on freely available data and methods that can easily incorporate new data as in the future. In this paper a spatially enhanced and updated AfriPop dataset for the MALAREO study area was generated based on the RapidEye 2011 settlement map and census data of 2007, whereby UN urban/rural growth rates for 2011 were additionally considered in order to account for recent population trends.

Comparison with existing datasets

Figure 3 shows a visual comparison between the original AfriPop dataset and the enhanced MALAREO population density map. The comparison between the settlement extent used by AfriPop and the MALAREO settlements demonstrates the much more detailed representation of the urban areas using high-resolution RapidEye data. The delineation of the settlements is much more accurate which has a significant effect on the population distribution model. Both modelling approaches concentrate the population in settlements and weigh populations outside urban areas based on different LC types.

Per-pixel comparison between settlement extents

Table 3 shows the result of the per-pixel comparison between the AfriPop settlements and the MALAREO settlement extents derived from RapidEye imagery. The MALAREO LC map has an overall accuracy of 80.7% assessed by a stratified random sampling approach (Franke *et al.*, in press). The higher resolution and the high overall accuracy seen in the table suggest that the urban class is more accurately represented in the MALAREO LC map. In this comparison, the level of accuracy produced by AfriPop in terms of non-urban user's accuracy and urban class producer's accuracy is above 95%. This indicates that the majority of areas is classified as urban according to AfriPop; this was also the case for RapidEye, while the non-urban AfriPop pixels mostly coincided with non-urban classified pixels. The low user's accuracy of the urban class, on the other hand, indicates that only 19.5% of the urban-assigned pixels were mapped correctly. This suggests that AfriPop overestimates urban extent in this study area. The non-urban producer's accuracy indicates that only 55.2% of the pixels classified as non-urban were mapped as non-urban in AfriPop. The reason for the low overall accuracy of 59.3% is highlighted by the producer's and user's accuracies. It is important to emphasise that these findings are only valid for this

study area. Obviously the differences in spatial resolution and acquisition time are the main factors for the wide difference in accuracy. The use of the recent and more accurate RapidEye data is an improvement and an update of the AfriPop database.

Per-pixel absolute differences between population datasets

Figure 4 shows that for the large majority of pixels, the absolute difference between the original AfriPop 2007 dataset and the enhanced MALAREO 2007 population density map is lower than 1 person per

Table 4. Error statistics for AfriPop 2011 and Enhanced AfriPop 2011 compared with detailed census 2011 data.

	MAE	RMSE	CC
AfriPop 2011	1063.36	1729.59	0.483
Enhanced AfriPop 2011	607.99	974.16	0.756

MAE, mean absolute error; RMSE, root mean square error; CC, correlation coefficient.

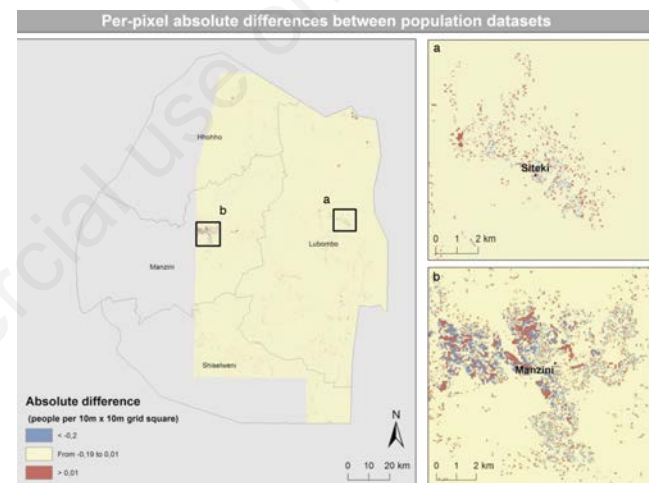


Figure 4. Per-pixel absolute differences between original AfriPop 2007 dataset and the enhanced AfriPop 2007 population density map for the study area located in Swaziland. Subsets show detail around the cities (a) Siteki and (b) Manzini.

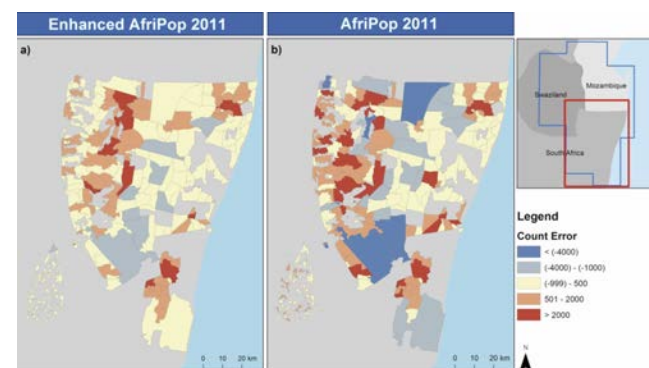


Figure 5. Quantitative error assessment by ward. a) Enhanced AfriPop 2011, and b) original AfriPop 2011. Populations underestimated are highlighted in blue and overestimated values are indicated in red.



10×10 m grid square. These pixels are mainly located outside urban areas where the human population density is generally close to zero. However, the absolute differences can be much higher in more densely populated places. The urban extent of major cities (e.g. Siteki or Manzini) strongly influenced the results (Figure 4a and 4b). The urban extents derived from the RapidEye dataset are generally smaller, resulting in significantly higher population densities in city centres. In general, the main difference between the datasets is that more villages, small towns and homesteads are represented in the MALAREO settlement dataset.

Accuracy assessment

The ward-level count error maps for the original and the enhanced AfriPop 2011 map are shown in Figure 5. Blue areas indicate underestimation of population, while red areas indicate overestimation. The Enhanced AfriPop method produces more accurate results, with many more administrative units showing low error values compared to the original AfriPop. The results of the error analysis in Table 4 suggest that the variance in individual errors was less for the Enhanced AfriPop than the AfriPop as the Enhanced AfriPop approach produced lower RMSE and MAE values. The Enhanced AfriPop model also shows the highest correlation between estimated and observed values at 0.75 compared to the AfriPop (0.48) method.

Discussion

Population distribution modelling is of major importance for estimating the population at risk of infectious diseases, and ultimately, disease burden. The primary aim of this work was to explore the impact of the use of detailed settlement layers on population density modelling. The results presented here show that using accurate and spatially detailed settlement extents is significantly improving population distribution modelling. The comparison between both settlement extent datasets demonstrates the much more detailed representation of the urban areas using high-resolution RapidEye data. The delineation of the settlements is more accurate, which has a significant effect on the population distribution model. Moreover, the settlement layer derived from the high-resolution RapidEye data incorporates data of small settlements and homesteads that were not represented in the AfriPop dataset. This study also proved that the population distribution modelling approach developed by AfriPop is appropriate to easily incorporate new data. Recent and upcoming new satellite systems will provide higher resolution data during the next years, which can constantly be used to improve population figures at higher spatial scales.

Detailed spatial population data is very valuable to support the daily work of the national malaria control programs in the project area including the planning of indoor residual spraying and monitoring coverage of insecticide-treated bed nets, and to identify focal areas of high malaria transmission risk. The integration of detailed population figures can substantially improve the planning of the distribution of insecticide treated nets, the estimation of the required amount of insecticides for indoor residual spraying in an area and can thus result in more effective malaria control measures. Moreover, this population data can be spatially linked to malaria transmission risk maps to quantify the number of people residing in specific risk zones and thus derive population at risk numbers. Such maps highlighting transmission areas (due to high contact rates between people and mosquitoes) can be produced by combining population data with vector density maps or malaria incidence or prevalence maps.

The main limitation of modelling population distribution is the fact that census data is collected at different administrative unit levels. This level varies across countries which is inherited to the final AfriPop data and thus results in different sizes of the spatial units of similar population density, which is also apparent in Figure 3. The smaller the administrative unit of the census data, the more detailed the redistribution of population by the AfriPop model. This issue could be minimized by the use of higher resolution LC classifications also for non-urban areas. In the present study, GlobCover data was used for non-urban areas, which does not allow the weighted redistribution of population figures at the same spatial detail as in urban areas in which RapidEye data was used. With new satellite systems such as the Sentinels, it is expected that high-resolution LC data for large areas will be available which can substantially improve the population modelling in low density areas. Also, other ancillary data such as road and railway networks or topography could improve population distribution modelling methods.

Conclusions

The outcomes of this study have a high potential for transferability to other countries and replicability in regard to other environmental related diseases such as helminth infections (e.g. schistosomiasis and soil-transmitted helminthiasis). Schistosomiasis is considered the *second leading parasitic killer in the world after malaria* (Global Network on Neglected Tropical Diseases, 2014: <http://www.globalnetwork.org/>). The World Health Organization strategy for schistosomiasis control focuses on the reduction of disease through periodic, targeted treatment with the anti-schistosomal drug praziquantel. This involves regular large-scale treatment of at-risk population groups. Patterns of schistosome infection and disease vary strongly by demographic variables such as age, sex and socioeconomic status (Brooker, 2007). Spatial population data can be used for the planning of the mass drug administration including the calculation of at-risk age groups and the estimation of the number of tablets required. The combination of demographic and epidemiological data can support and improve the creation of reliable maps on the spatial distribution of schistosomiasis, so that interventions can target communities at highest risk.

Furthermore, high-resolution data from the European Space Agency sentinel missions will further stimulate the use of earth observation techniques for malaria control and epidemiology also because of the related free and open data policy.

References

- Balk D, Yetman G, 2004. The global distribution of population: evaluating the gains in resolution refinement. Center for International Earth Science Information Network, New York, NY, USA.
- Balk DL, Deichmann U, Yetman G, Pozzi F, Hay SI, Nelson A, 2006. Determining global population distribution: methods, applications and data. *Adv Parasit* 62:119-56.
- Brooker S, 2007. Spatial epidemiology of human schistosomiasis in Africa: risk models, transmission dynamics and control. *T Roy Soc Trop Med H* 101:1-8.
- Ceccato P, Connor PJ, Jeanne I, Thomson MC, 2005. Application of geographical information systems and remote sensing technologies for assessing and monitoring malaria risk. *Parasitologia* 47:81-96.



- Congalton RG, 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens Environ* 37:35-46.
- Deichmann U, 1996. A review of spatial population database design and modelling. National Center for Geographic Information and Analysis-University of California, Santa Barbara, CA, USA.
- Deichmann U, Balk D, Yetman G, 2001. Transforming population data for interdisciplinary usages: from census to grid. Center for International Earth Science Information Network, Washington, DC, USA.
- Dobson JE, Bright EA, Coleman PR, Durfee RC, Worley BA, 2000. LandScan: a global population database for estimating populations at risk. *Photogramm Eng Rem S* 66:849-57.
- Elliott P, Wartenberg D, 2004. Spatial epidemiology: current approaches and future challenges. *Environ Health Persp* 112:998-1006.
- Franke J, Gebreslasie M, Bauwens I, Deleu J, Siegert F, 2015. Earth observation in the support of malaria control and epidemiology. MALAREO monitoring approaches. *Geosp Health* 10:335.
- Linard C, Alegana VA, Noor AM, Snow RW, Tatem AJ, 2010a. A high resolution spatial population database of Somalia for disease risk mapping. *Int J Health Geogr* 9:45.
- Linard C, Gilbert M, Snow RW, Noor AM, Tatem AJ, 2012. Settlement patterns, population distribution and accessibility measures at the national level. *PLoS One* 7:e31743.
- Linard C, Gilbert M, Tatem AJ, 2010b. Assessing the use of global land cover data for guiding large area population distribution modelling. *Geoj Lib* 76:525-38.
- Linard C, Tatem AJ, 2012. Large-scale spatial population databases in infectious disease research. *Int J Heal Geogr* 11:7.
- Ostfeld RS, Glass GE, Keesing F, 2005. Spatial epidemiology: an emerging (or reemerging) discipline. *Trends Ecol Evol* 20:328-36.
- Riley S, 2007. Large-scale spatial-transmission models of infectious disease. *Science* 316:1298.
- Rogers DJ, Randolph SE, Snow RW, Hay SI, 2002. Satellite imagery in the study and forecast of malaria. *Nature* 415:710-5.
- Tatem AJ, Campiz N, Gething PW, Snow RW, Linard C, 2011. The effects of spatial population dataset choice on population at risk of disease estimates. *Popul Health Metr* 9:4.
- Tatem AJ, Noor AM, Hay SI, 2005. Assessing the accuracy of satellite derived global and national urban maps in Kenya. *Remote Sens Environ* 96:87-97.