Do we underestimate the global slum population?

Monika Kuffer
ITC, University of Twente
Enschede, NL
m.kuffer@utwente.nl

Claudio Persello
ITC, University of Twente
Enschede, NL
c.persello@utwente.nl

Karin Pfeffer (ITC)
ITC, University of Twente
Enschede, NL
k.pfeffer@utwente.nl

Richard Sliuzas
ITC, University of Twente
Enschede, NL
r.sliuzas@utwente.nl

Vinodkumar Rao
SPARC - (SDI affiliate),
Mumbai, India
vinodkumar.rao@gmail.com

***Abstract*—According to UN-Habitat, around one billion people live in slum conditions, this number is reported for the SDG indicator 11.1.1 (the proportion of urban population living in slums, informal settlements or inadequate housing). However, this number comes with many uncertainties. For several countries, estimates are not available, while for other countries reported data might not reflect the real population living in slum conditions. In this paper, we use Dar es Salaam in Tanzania as a showcase on how a combination of data extracted from remote sensing combined with locally available sample data and non-official data (e.g., from NGOs) could allow quantifying the degree of uncertainty about city-level slum population estimates. For the city of Dar es Salaam, the estimates based on the census data indicate that around 3 million of its inhabitants are living in slum-like conditions, while using a combination of household surveys, settlement level estimates from Shack/Slum Dwellers International combined with rooftop outlines extracted from Unmanned Aerial Vehicle (UAV) images, the estimated slum population is around 5 million. This raises the question of how much on a global level do we underestimate the number of people living in slum conditions and shows the potential of remote sensing to shed some light on this neglected issue.**

Keywords—SDG indicator, slums, informal settlement, deprived areas, population estimation, dasymetric modelling

I. Introduction

High urbanization rates in many global-south regions [1] and the insufficient provision of affordable housing to low-income groups contribute to the growth of slums, informal settlements and inadequate housing in urban areas (SDG indicator 11.1.1) [2]. According to UN-Habitat, a slum area is defined by the absence of improved water, adequate sanitation, tenure security, durable housing and insufficient living space of people [3]. To support pro-poor improvement strategies, accurate and up-to-date delineations of slum areas and related population data are required for local decision-making. However, available data are commonly inconsistent, or not available at a required aggregation level. Spatially disaggregated data (showing variations within cities) are usually not available and if available often not accessible to local communities and NGOs. Projects targeted to the urban poor commonly suffer from such missing base data. Census-based data available in many cities is rapidly outdated (by the time it is used), in particular in cities with high development dynamics. Thus such conventional approaches (e.g., based on household data collection) fail to provide a comprehensive account at city level over time. Generally, reported data by national governments for the indicator 11.1.1 has many uncertainties, e.g., they contain inconsistencies related to national and local policies of recognizing the size of the slum population (under-reporting). While at a municipal level also over-reporting has been observed [4]. Thus official slum statistics show large uncertainties, e.g., for the city of Ahmedabad, the Census of India (2011) reports 4% of slum inhabitants, while the municipality reports 18% slum population [3]. Many statistics do not account for officially unrecognized slum areas. In particular, the poorest section of the slum population living in new, temporary, or sub-rented shelters are commonly ignored. The few studies done using remote sensing based methods to estimate the slum population showed that official statistics possibly exclude a large number of inhabitants, sometimes 50% and more [5, 6].

Remote sensing (RS) studies [7-11] have shown the potential of satellite imagery analysis to provide consistent and timely information on locations and dynamics of urban slums. Such studies successfully showed the capability of Very High Resolution (VHR) satellite images to map and characterize urban poverty [12], largely drawing on locally adapted image features. The new group of advanced machine learning methods is overcoming the need to define locally adapted image features and is, therefore, more likely to be transferable across countries and cities. These studies [10, 13] demonstrated that advanced machine learning methods could capture the location of slums with an accuracy above 90%. The areas where these algorithms fail are commonly related to high- density formal areas where people have access to some basic services (e.g., resettlement colonies in India). However, inhabitants of such areas experience at least physical aspects of deprivation [14]. However, most remote sensing studies do not produce city-level information on the boundaries of slums (due to image and computational costs). Furthermore, to take the next step to provide population estimates in support of policy-relevant information is also commonly not done by remote sensing studies [15].

In this paper, we present a showcase of how remote sensing data (i.e., rooftops extracted and land use/cover from VHR resolution imagery) can be combined with household samples and local non-official data to map and explore population data of slums. We employ dasymetric modelling [16, 17] to estimate the slum population and compare results with existing census-based estimates of the slum population. Generally, dasymetric modelling allows disaggregating spatial data (e.g., aggregated census ward population) to smaller spatial units employing ancillary data (e.g., building outlines or land use data) [18, 19], resulting in a detailed spatial representation of the population distribution.

II. Methodology

A. The case study area and data set

Dar es Salaam, the prime city and economic hub of Tanzania, had a population of almost 4.4 inhabitants, according to the last census in 2012 and is one of the fastest

978-1-7281-0009-8/19/$31.00 ©2019 IEEE

growing cities in Africa. An estimated 70% of its residents lives in deprived areas (locally referred to as informal settlements) [20]. Such areas are commonly under-serviced and in particularly lacking access to improved water and sanitation, as well as show other aspects of deprivation (e.g., street lighting [21]). Houses are commonly single-storey structures, where often part of one house is sub-rented. According to a national level UN-Habitat database, the amount of slum population in Tanzania reduced from 77% (in 1990) to 51% (in 2014). However, this data is a national aggregate with high uncertainty.

Recently, several initiatives created a rich repository of data on the city. For example, the community-based mapping project Ramani Huria created an extensive data set including the mapping of all rooftops from Unmanned Aerial Vehicle (UAV) images (Fig. 1). Derived data (e.g., rooftop outlines) for the entire city are available via OpenStreetMap (OSM). Another source of information on slum areas is the database built by Shack/Slum Dwellers International (SDI) available via the ‘Know your city’ portal, which provides rich information on the population, their available services and the history of individual settlements (Fig. 2). For some settlements the mapped settlement outlines and population estimates are also available. SDI advocates a city wide data

Fig.1 Example of UAV image to map rooftops, Dar es Salaam.

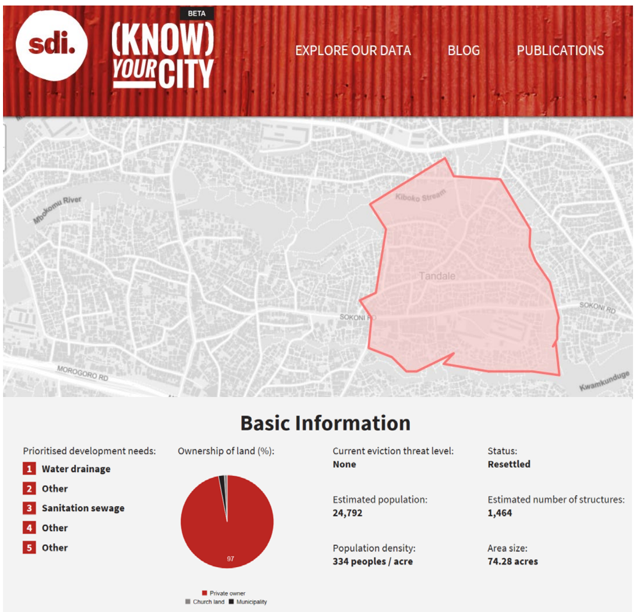

Fig. 2 Example of settlement based information by SDI.

collection and uses the "leave no one behind" principle, meaning every single slum needs to be counted, profiled and mapped.

Census population counts of 2012 for each sub-ward were also used (Fig. 4). Furthermore, a land use map showing the formal and informal areas, besides other general land use categories was available (reference date 2010/11). Moreover, a 2015 household survey conducted in five settlements distributed across the city [22] provided data on the number of the persons living under one roof and the rooftop areas.

B. *Overview of the methodology*

Present methods of the SDG indicator 11.1.1 rely on nationally reported statistics, which may neglect unrecognized, newly developing, low-density peri-urban, temporary and small clusters of slums. The few reported experiments to estimate slum populations combining ground-based data with VHR imagery show that national statistics might dramatically underestimate the number of slum inhabitants [5, 6, 23]. Furthermore, several countries (e.g., Chile, Cuba, Malaysia, Paraguay) do not report any slum population [24]. Such data inconsistencies constrain the monitoring of global and local efforts to eradicate poverty and to provide basic services and a safe living environment to all urban inhabitants. However, location-based information on slum areas can be provided by VHR satellite imagery or UAV images [25], giving all deprived areas a location on maps. Therefore, the available rooftop outlines were downloaded from OSM. Combining such information within dasymetric modelling allows to make an estimate of the population distribution [26]. Generally, dasymetric modelling can be split in top-down and bottom-up models. Top-down models commonly start with population census data and use additional spatial data to disaggregate the census data, while bottom-up models build on local micro-scale data (here rooftops) to create spatial delineations, thereby avoiding data limitations of census statistics.

To estimate the number of people living in slum-like conditions, we employ a bottom-up approach using available data of all OSM rooftops, covering the city of Dar es Salaam, data from the 2015 household survey and manually extracted settlement outlines from the SDI website to derive an estimate of the roof area (in m^2 per person) in slums across the city. One estimate is based on the household survey and the second uses the SDI data. To exclude non-residential areas, the available land use map is used, and OSM buildings located in industrial, commercial and institutional zones are excluded. To obtain all slum buildings, all buildings within planned areas are excluded. Using this ratio, based on the OSM building layer, we estimate how many people live in each slum building for the entire city. As the target unit is not individual buildings and to balance for variations, aggregation to larger units is necessary to provide an area based estimate.

III. Resutls

Using the 2015 household survey an estimated ratio of around 9.8 m^2 roof area per person is obtained. The estimated roof area per person using the SDI database, which provides some slums boundaries and the estimated population per settlement, is 10.6 m^2. These two roof area estimates are combined with all OSM buildings (excluding commercial, industrial and institutional zones and excluding buildings with non-residential uses). The results show (Tab. I), that both slum population estimates are much higher than the Census estimate of around 3 million slum inhabitants and arrive at around 5 million people that are forced to live in slum conditions. When assuming that the inhabitants of formal areas are well captured, this would indicate that the total population of Dar es Salaam is beyond 6 million with around 80% living in slum-like conditions. These figures should not be understood as the final truth, but an indication that present estimates of slum inhabitants are highly uncertain and most likely overlook a large number of people whose needs should be considered for the planning and provision of basic infrastructure and services.

TABLE I. ESTIMATES OF THE SLUM POPULATION IN DAR ES SALAAM.

	Roof Area per person in m^2			Slum population estimates (persons)			
	Census	HH survey	SDI samples	Census[a]	HH survey	SDI samples	Combined estimate
Total area	19.07			4,364,541			
Slum areas	16.85	9.84	10.57	3,055,179	5,232,405	4,868,623	5,043,903
Formal areas	24.25			1,309,362			

[a.] Based on an estimate of 70% slum population.

The main advantage of a dasymetric population estimation model that builds on building-level information is its capability to aggregate and provide data at any desired spatial unit beyond buildings (e.g., a slum population density map – Fig. 3), in support of planning and decision-making processes. The data also allows to quantify the density of uncounted slum population (not counted persons per ha) for all sub-ward (Fig. 4). This shows that there is a spatial concentration of a high density of possibly uncounted slum inhabitants in a relatively central zone outside the city centre and along the major roads (there might be also a chance to overestimate the population in buildings with mixed uses). Planning of infrastructure improvements in such areas should receive particular attention.

IV. DISCUSSION

The developed methodology allows localization of the SDG indicator 11.1.1, and highlights uncertainties in available slum population estimates. The possibility that the slum population of Dar es Salaam is underestimated by up to 2 million people, raises the question of how much do we possibly underestimate the global slum population? This has to be asked in the context, knowing that many governments do not cover the entire population in slum conditions (e.g., excluding temporary and small settlements) [27, 28]. Furthermore, some countries (e.g., India) use minimum size conditions to define slums. The uncertainty affects the reporting on the SDG indicator 11.1.1, but also other SDG indicators, e.g., 1.4.1 or 11.3.1, which require robust information on population estimates, also in slums. However, when dramatically underestimating the slum population in cities, which can be more than 50% of the urban population, we do not have a good basis for other SDG indicators. Localizing information on slums and their inhabitants, is essential for local authorities to develop specific improvement strategies for their inhabitants in need. Mapped locations of informal settlements and having robust estimates of their populations are critical for disaster mitigation and adaptation towards the planning for inclusive and resilient cities. Remote sensing can contribute to address these knowledge gaps. In this study, we used delineated rooftop outlines, but such data can, in principle, be extracted in a semi-automatic manner, as well as boundaries of slums can be updated with the latest generation of Convolutional Neural Networks (CNNs) that show an outstanding capability to solve complex classification problems [10, 13, 29]. However, this methodology neglects differences within

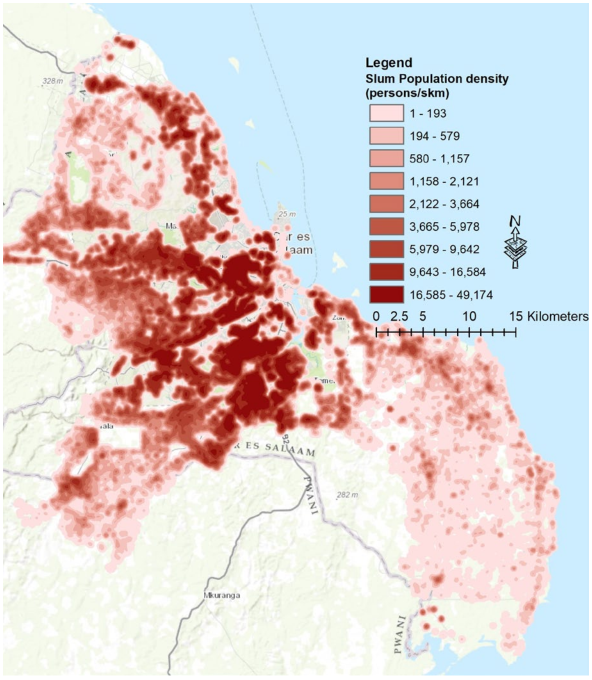

Fig. 3 Kernel density estimate of slum population across the city of Dar es Salaam, Tanzania (shown on top of a topographic base map).

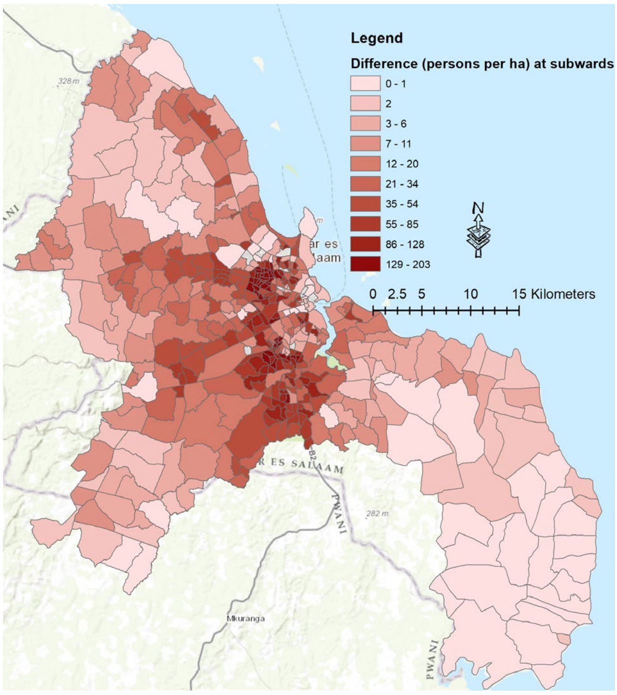

Fig. 4 Difference of local estimates and census based estimates of slum popualtion per ha, Dar es Salaam, Tanzania (shown on top of a topographic base map).

the city, e.g., central slums might have a different ratio compared to peripheral slums, which should be improved in a further study taking variations into account.

V. Conclusions

This paper opens a debate about whether we might underestimate the slum population in cities and by how much? Recent developments in remote sensing allow slum settlements and rooftop outlines (e.g., via UAV or VHR satellite images) to be rapidly mapped. Such data, combined with local data on building occupancy, permits building level estimates of slum inhabitants that can be aggregated to any desired spatial unit beyond the building. The initial experiment done in Dar es Salaam supports the hypothesis that official statistics on the SDG indicator 11.1.1 might dramatically underestimate the number of people living in slums. For further research, such experiments need to incorporate variations across different parts of the city (e.g., differences in roof area per person depending on locational characteristics) and the inclusion of additional cities to produce more evidence for the hypothesis that the available global statistics significantly underestimate the number of slum dwellers.

References

[1] UN-Habitat, State of the world's cities 2010- 2011: Bridging the urban divide. Nairobi, Kenya, 2010.

[2] UN-Habitat, Slums Almanac 2015-16. Tracking Improvement in the Lives of Slum Dwellers. Nairobi, 2016.

[3] UN-Habitat, "Metadata: Indicator 11.1.1: Proportion of urban population living in slums, informal settlements or inadequate housing," Accessed on: 22-10-2017Available: https://unstats.un.org/sdgs/metadata/files/Metadata-11-01-01.pdf

[4] G. Leonita, M. Kuffer, R. Sliuzas, and C. Persello, "Machine Learning-Based Slum Mapping in Support of Slum Upgrading Programs: The Case of Bandung City, Indonesia," Remote Sens., vol. 10, no. 10, p. 1522, 2018.

[5] H. Taubenböck and M. Wurm, "Ich weiß, dass ich nichts weiß—Bevölkerungsschätzung in der Megacity Mumbai," in Globale Urbanisierung: Perspektive aus dem All, H. Taubenböck, M. Wurm, T. Esch, and S. Dech, Eds. Berlin/Heidelberg, Germany: Springer, 2015, pp. 171–178.

[6] T. Veljanovski, U. Kanjir, P. Pehani, K. Oštir, and P. Kovačič, "Object-based image analysis of VHR satellite imagery for population estimation in informal settlement Kibera-Nairobi, Kenya," in Remote Sensing - Applications, B. Escalante, Ed. Rijeka, Croatia: InTech, 2012, pp. 407-434.

[7] J. C. Duque, J. E. Patino, and A. Betancourt, "Exploring the Potential of Machine Learning for Automatic Slum Identification from VHR Imagery," Remote Sens., vol. 9, no. 9, p. 895, 2017.

[8] M. Wurm and H. Taubenböck, "Detecting social groups from space – Assessment of remote sensing-based mapped morphological slums using income data," Remote Sens. Lett., vol. 9, no. 1, pp. 41-50, 2018.

[9] H. Taubenböck, N. J. Kraff, and M. Wurm, "The morphology of the arrival city - A global categorization based on literature surveys and remotely sensed data," Appl. Geogr., vol. 92, pp. 150-167, 2018.

[10] C. Persello and A. Stein, "Deep Fully Convolutional Networks for the Detection of Informal Settlements in VHR Images," IEEE Geosci. Remote Sens. Lett., vol. 14, no. 12, pp. 2325-2329, 2017.

[11] M. Kuffer, K. Pfeffer, R. Sliuzas, and I. Baud, "Extraction of slum areas from VHR imagery using GLCM variance," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 9, no. 5, pp. 1830–1840, 2016.

[12] M. Kuffer, K. Pfeffer, R. Sliuzas, I. Baud, and M. van Maarseveen, "Capturing the Diversity of Deprived Areas with Image-Based Features: The Case of Mumbai," Remote Sens., vol. 9, no. 4, p. 384, 2017.

[13] N. O. Mboga, C. Persello, J. Bergado, and A. Stein, "Detection of informal settlements from VHR images using Convolutional Neural Networks," Remote Sens., vol. 9, no. 11, p. 1106, 2017.

[14] S. Patel, R. Sliuzas, and N. Mathur, "The risk of impoverishment in urban development-induced displacement and resettlement in Ahmedabad," Environ. Urban., vol. 27 no. 1, pp. 231-256, 2015.

[15] M. Kuffer et al., "The Scope of Earth-Observation to Improve the Consistency of the SDG Slum Indicator," ISPRS Int. J. Geo-Inf., vol. 7, no. 11, p. 428, 2018.

[16] K. Cockx and F. Canters, "Incorporating spatial non-stationarity to improve dasymetric mapping of population," Appl. Geogr., vol. 63, pp. 220-230, 2015.

[17] N. N. Nagle, B. P. Buttenfield, S. Leyk, and S. Speilman, "Dasymetric Modeling and Uncertainty," Ann. Assoc. Am. Geogr., vol. 104, no. 1, pp. 80-95, 2014.

[18] M. Langford, G. Higgs, J. Radcliffe, and S. White, "Urban population distribution models and service accessibility estimation," Comput. Environ. Urban Syst., vol. 32, no. 1, pp. 66-80, 2008.

[19] D. J. Briggs, J. Gulliver, D. Fecht, and D. M. Vienneau, "Dasymetric modelling of small-area population distribution using land cover and light emissions data," Remote Sens. Environ., vol. 108, no. 4, pp. 451-466, 2007.

[20] UN-Habitat, "Informal settlements and Finance in Dar es Salaam, Tanzania," Nairobi, Kenya2010.

[21] M. Kuffer, R. Sliuzas, M. F. A. M. van Maarseveen, K. Pfeffer, and I. Baud, "City nighttime light variations using ISS images," presented at the Joint urban remote sensing event (JURSE) 2017, Dubai, United Arab Emirates, 6-8 March 2017.

[22] D. A. Kaaya, "Explaining variations in informal neighborhoods' consolidation levels in Dar es Salaam, Tanzania," University of Twente Faculty of Geo-Information and Earth Observation (ITC), Enschede, 2017.

[23] O. Kit and M. Lüdeke, "Automated detection of slum area change in Hyderabad, India using multitemporal satellite imagery," ISPRS J. Photogramm. Remote Sens., vol. 83, pp. 130-137, 2013.

[24] World Bank. (2017, 22-10). Population living in slums (% of urban population). Available: https://data.worldbank.org/indicator/EN.POP.SLUM.UR.ZS

[25] C. M. Gevaert, C. Persello, R. Sliuzas, and G. Vosselman, "Informal settlement classification using point-cloud and image-based features from UAV data," ISPRS J. Photogramm. Remote Sens., vol. 125, pp. 225–236, 2017.

[26] M. Kuffer and R. Sliuzas, "Modelling day and night time population using a 3D urban model," presented at the Real Corp 2014, Vienna, Austria 21-23 May, 2014. Available: http://www.corp.at/archive/CORP2014_46.pdf

[27] R. Sliuzas, M. Kuffer, K. Pfeffer, C. M. Gevaert, and C. Persello, "Slum mapping: From space to unmanned aerial vehicle based approaches," presented at the Joint Urban Remote Sensing Event (JURSE), Dubai, United Arab Emirates, 6-8 March, 2017.

[28] M. Kuffer, K. Pfeffer, and R. Sliuzas, "Slums from space—15 years of slum mapping using remote sensing," Remote Sens., vol. 8, no. 6, p. 455, 2016.

[29] C. M. Gevaert, R. Sliuzas, C. Persello, and G. Vosselman, "Opportunities for UAV mapping to support unplanned settlement upgrading," presented at the GeoTech Rwanda 2015, Kigali, Rwanda, 18-20 Nov., 2015.