 doi:10.20944/preprints201910.0242.v1

Critical Commentary. Need for an integrated deprived area "slum" mapping system (IDeAMapS) in LMICs

Dana R Thomson,[1,2]† Monika Kuffer,[3] Gianluca Boo,[2] Beatrice Hati,[4] Tais Grippa,[5] Helen Elsey,[6] Catherine Linard,[5,7] Ron Mahabir,[8] Catherine Kyobutungi,[9] Joshua Mulandi,[10] Dennis Mwaniki,[11] Robert Ndugwa,[11] Jack Makau,[4] Richard Sliuzas,[3] Salome Cheruiyot,[11] Kilion Nyambuga,[4] Nicholus Mboga,[5] Nicera Wanjiru,[4] Joao Porto De Albuquerque,[12]* Caroline Kabaria[9]*

1. Flowminder Foundation, UK
2. WorldPop Team, University of Southampton, UK
3. Monika Kuffer - Department of Urban and Regional Planning and Geo-Information Management, University of Twente, The Netherlands
4. Slum Dwellers International, Kenya
5. Institute for Environmental Management and Land-use Planning, Université Libre de Bruxelles, Belgium
6. Department of Global Health, University of York, UK
7. Department of Geography, Université de Namur, Belgium
8. Department of Computational and Data Sciences, George Mason University, USA
9. African Population & Health Research Center, Kenya
10. Slum Upgrading Team, UN-Habitat, Kenya
11. Global Urban Observatory, UN-Habitat, Kenya
12. Institute for Global Sustainable Development, University of Warwick, UK

† Corresponding author
* Co-senior authors

 doi:10.20944/preprints201910.0242.v1

Abstract

Ninety percent of the people added to the planet over the next 30 years will live in African and Asian cities, and a large portion of these populations will reside in deprived neighborhoods defined by slum conditions, informal settlement, or inadequate housing. The four current approaches to neighborhood deprivation mapping are largely silo-ed, and each fall short of producing accurate, timely, comparable maps that reflect local contexts. The first approach, classifying "slum households" in census and survey data and aggregating to administrative areas, reflects household-level rather than neighborhood-level deprivation. The second approach, field-based mapping, can produce the most accurate and context-relevant maps for a given neighborhood, however it requires substantial resources, preventing up-scaling. The third and fourth approaches, human interpretation and machine classification of satellite, aerial, or drone imagery, both overemphasize informal settlements, and fail to represent key social characteristics of deprived areas such as lack of tenure, exposure to pollution, and lack of basic public services. The latter, machine classification of imagery, can be automated and extended to incorporate new and multiple sources of data. This diverse collection of authors represent experts from these four approaches to neighborhood deprivation mapping. We summarize common areas of understanding, and present a set of requirements to produce maps of deprived urban areas that can be used by local-to-international stakeholders for advocacy, planning, and decision-making.

Keywords
satellite imagery, social indicator, urban, poverty, SDG

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0242.v1

Introduction

Most low- and middle-income countries (LMICs)[1] are in the midst of urban transitions, or will be soon, and are facing rapid growth of slum-like communities. Although urbanization has been associated with some of the greatest achievements in human history, including reduced mortality and the production of material wealth and wellbeing, it is also closely linked with socioeconomic inequalities that trap generations of families in perpetual cycles of poverty and insecurity (Diaz et al., 2017).

The United Nations (UN) expects that between 2018 and 2030, megacities such as Kinshasa (D.R. Congo), Delhi (India), and Dhaka (Bangladesh) will each add more than 700,000 people per year on average through 2030 (UN-DESA, 2018a). By this time the majority of the population in LMICs will be living in urban rather than rural areas, and by 2050 two-thirds of the population in LMICs is expected to be urban (UN-DESA, 2017). An estimated 2.5 billion people will be added to the planet by 2050, with 90% of that population increase concentrated in Asian and African cities alone (UN-DESA, 2018a). This is cause for concern given that many of the LMICs within these regions are currently facing various social, economic and development challenges, which impede their ability to adequately accommodate this future population growth (Mahabir et al., 2016).

To help cities better plan for future population growth, Sustainable Development Goal (SDG) 11 aims to "make cities and human settlements inclusive, safe, resilient and sustainable." Progress towards SDG 11 is measured, in part, by identifying the "proportion of urban population living in slums, informal settlements or inadequate housing" (UN-DESA, 2018b). Decision-makers at different levels (e.g. local, national, and international) and governments, non-governmental organizations, and private institutions alike require this information. Decision-makers use neighborhood deprivation maps to estimate numbers of people living in these areas (Angeles et al., 2009), allocate public services (Gruebner et al., 2014), plan and evaluate health policies and campaigns (Weeks et al., 2012); respond to humanitarian disasters (Bramante and Raju, 2013), and make long-term development decisions (Abbott, 2003; Chitekwe-Biti et al., 2012).

Despite more than two decades of effort, slums, informal settlements or areas of inadequate housing are still not mapped accurately and routinely across LMICs. The problem is twofold. First, there is no universal definition of what a deprived neighborhood is. Second, there are no established, universally applicable best practices for mapping such areas. As a result, there are no data repositories of consistent, up-to-date, publicly accessible data on deprived areas within cities. This paper, with contributions from a diverse group of international experts, outlines the need integrate and leverage the strengths of existing approaches to routinely, and accurately map deprived urban areas in LMIC cities to support SDG 11 and decision-making.

[1] According to the World Bank, a LMIC in 2018-19 had a GNI per capita between 996 - 3,895 US dollars. (https://blogs.worldbank.org/opendata/new-country-classifications-income-level-2018-2019)

Slum are versus slum household

The term “slum” has been used to belittle and marginalize groups in some contexts, and it is used as an identity-marker among residents in other contexts (Nuissl and Heinrichs, 2013). Further debate may be needed to identify terminology that reflects the assets and humanity of all individuals in cities, including those who live in slums, informal settlements, and inadequate housing (Corburn and Cohen, 2012). “Favela”, “ghetto”, “barrio”, or “shantytown” may be more commonly used terms in some cities; however, each of these labels also comes with a specific political and social history. Recognizing these limitations with the term “slum,” we instead use the term “deprived areas” to refer to urban residents of slums, informal settlements or inadequate housing in line with SDG 11.

A number of efforts have been made to define deprived urban neighborhoods including expert meetings in 2002 (UN-Habitat et al., 2002), 2008 (Sliuzas et al., 2008), and 2017 (UN-Habitat, 2017); published frameworks (Lilford et al., 2019; Mahabir et al., 2016); and operational definitions within Earth Observation (EO) research (Kohli et al., 2012; Kuffer et al., 2014; Mahabir, et al., 2018). Despite efforts over the last 20 years, no universal definition has been achieved to describe deprived urban areas. This is due, in large part, to the enormous diversity and dynamism of slums and informal settlements, and because perceptions of neighborhood deprivation is relative to other nearby communities (Nuissl and Heinrichs, 2013). Both high-rise blocks of social housing, and clusters of self-built shacks are commonly referred to as “slums” in LMIC cities (Nuissl and Heinrichs, 2013).

UN-Habitat provides a widely accepted definition to classify a household or group of individuals as a “slum household” if they lack any of the following: durable housing, sufficient living space, safe water, adequate sanitation, or security of tenure (UN-Habitat, 2007). Household tenure, however, is generally not measured in censuses and surveys, so it is routinely excluded from this definition in practice.

Although relatively easy to operationalize, a household-level definition of deprivation fails to account for some of the most important area-level risks and outcomes that result from living in slums, informal settlements, and inadequate housing (Table 1). Deprived areas reflect multiple social, environmental and ecological factors that together pose risks to health and well-being beyond household characteristics. Living in a deprived neighborhood can increase the incidence of disease via exposure to animal vectors and crowding of buildings, injuries such as fire, vulnerability to extreme weather events, higher incidence of crime, and physical and social barriers to services (Ezeh et al., 2017). The “slum household” definition reflects household- rather than area-level poverty, which poses risks such as increased incidence of disease due to crowding in the home and economic barriers to services. Furthermore, the household-based definition has been shown to overestimate deprived areas in some contexts, classifying neighborhoods as “slums” that are not considered as such locally (Engstrom et al., 2013) and labelling almost entire cities as “slums” such as in the case of Addis Ababa (Lemma et al., 2006).

Table 1. Definition of a deprived area (slum, informal settlement, area of inadequate housing) versus "slum household"

Deprived area	**Slum Household**
Reflects *social, environmental, and ecological risk factors* to health and wellbeing above and beyond household and individual characteristics	Reflects *household poverty* risk factors to individual health and wellbeing
Indicators include: ● Social risk - e.g. no social safety net, crime ● Environmental risk - e.g. flood zone, slopes ● Lack of facilities - e.g. schools, health facilities ● Lack of infrastructure - e.g. roads, bus service ● Unplanned urbanization - e.g. small, high-density, disorganized buildings ● Contamination - e.g. open sewer, trash piles ● Land use/rights - e.g. non-residential zoning	Indicators include: ● Non-durable walls, floor, or roof ● Too few sleeping rooms ● Lack of safe water source ● Lack of adequate toilet ● Lack of tenure of home (usually not measurable)

The risks of belonging to a "slum household" within a deprived neighborhood act simultaneously to exacerbate individual health and wellbeing (Figure 1). Within-household disparities, for example due to gender or age differences, can also shape individual health and wellbeing through restricting access to resources such as food or health care (Harris-Fry et al., 2017), though, this is beyond the scope of this paper.

Members of both "slum" and "non-slum households" located in a deprived neighborhood face multiple area-level risks such as seasonal flooding, lack of green space, constant noise, high crime, narrow paths lined with open sewers, and trash piles. A "slum household" in this setting may additionally be classified as such due to having a leaky tarp roof and limited living space. However, all households in a deprived neighborhood may benefit in certain ways by being well integrated in social and community networks, paying limited housing costs, and accessing charity programs that target vulnerable and informal areas.

Alternatively, a "slum household" — for example, a household which lacks adequate water and sanitation — and is located in a non-deprived neighborhood proximate to schools, health facilities, bus lines, and well-maintained roads may face different vulnerabilities including social isolation, cash flow problems due to higher costs of living, and thus financial barriers to the services located outside the front door. Different policies and interventions are needed for households located in deprived versus non-deprived areas, and thus it is imperative to define and map deprived areas in addition to "slum households."

 doi:10.20944/preprints201910.0242.v1

Figure 1. Four ways in which "slum household" and deprived area risks intersect

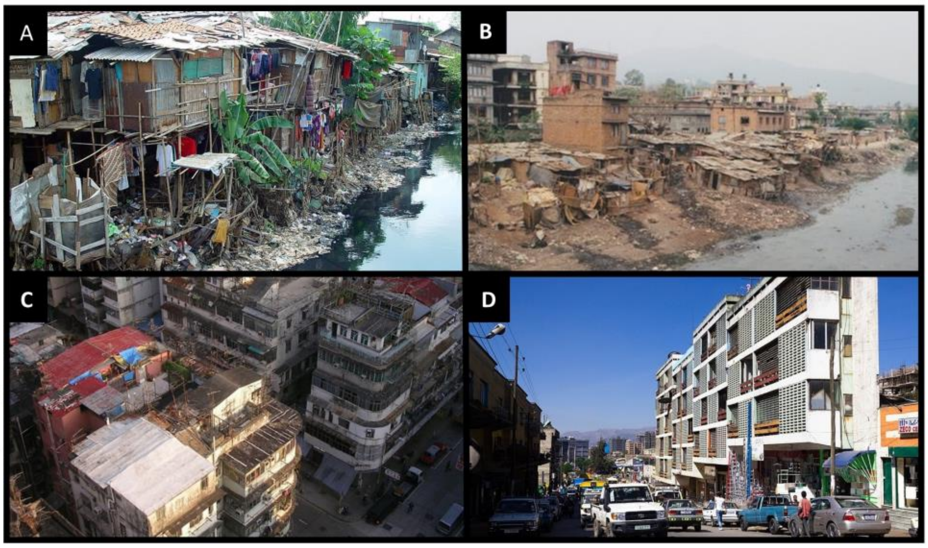


Image credits:
A - Jonathan McIntosh - Own work, CC BY 2.0, https://commons.wikimedia.org/w/index.php?curid=53838
B - Peter van der Sluijs - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=18905096
C - J. Carmichael - Own work, Public Domain, https://commons.wikimedia.org/w/index.php?curid=6441162
D - The Chroniclers - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=51369584

Requirements for area deprivation mapping

As mentioned before, no universal definition of a deprived urban area yet exists; however, the following seven requirements have been clearly articulated. Urban area deprivation maps need to be:

1) *Reflective of area physical characteristics.*

Deprived urban areas are often characterized by their morphology in the urban environment. Physical indicators of area deprivation reflect building characteristics such as their size, shape, and height; road and other access networks; building density; settlement shape; settlement location with respect to such environmental features as public green or blue spaces, steep slopes and flood zones; and neighborhood characteristics such as proximity to railways and high voltage power lines (Kohli et al., 2012).

2) *Reflective of area social characteristics*.

Deprived urban areas are characterized by a wide range of features in the social environment. Social indicators of neighborhood deprivation include presence of crime; presence and practices of law enforcement; coverage and quality of solid waste, water, sanitation, and power systems; proximity and accessibility to schools, health facilities, shops, employment, and public infrastructure; and social capital derived from community-based organizations and among neighbors with shared identities (Lilford et al., 2019).

3) *Context dependent*.

The physical and social characteristics that define a given deprived area differ across cities and countries and even within the same neighborhood (Kuffer et al., 2016). Furthermore, neighborhoods are not static in that the specific characteristics that define a deprived neighborhood at a particular moment in time change as the neighborhood evolves and policies and social forces unfold (Mahabir, et al., 2018).

4) *Comparable across cities and countries*.

To adequately support national planning and programs, and to be used in global initiatives such as the SDGs, a level of consistency in deprived urban area definitions are needed across cities and countries (Ezeh et al., 2017).

5) *Updated frequently with timely data*.

Deprived urban areas are highly dynamic and can be transformed over very short periods. As deprived neighborhoods transition through different development stages, from low- to high-density, and as they experience major shifts in population due to demolitions or "overnight invasions" of new residents, frequent updates to deprived area maps are needed based on very timely data (Sliuzas et al., 2010). Further, areas previously classified as deprived need to be able to be classified as non-deprived as infrastructure and services improve, sometimes because of gentrification.

6) *Protective of individual privacy, and vulnerable populations*.

Given the relatively high spatio-temporal resolution of neighborhood maps, approaches must ensure individual privacy in EO and other data, as well as transparency in the methods. There may additionally be a need to selectively filter vulnerable areas or obfuscate exact boundaries of deprived urban areas to protect already vulnerable populations (Thomson et al., 2019).

7) Developed via an inclusive multi-stakeholder process.

Existence of any deprived urban area reflects a story of social inequality, exclusion, and/or oppression. Urban "slums" do not emerge at random. For a deprived area to transition into a place that is "inclusive, safe, resilient and sustainable," the policies and social attitudes that permitted its formation need to be addressed. Neighborhood transformation requires involvement of communities, local authorities, and in many countries, national government representatives as well (Corburn and Cohen, 2012).

Existing approaches to area deprivation mapping

Existing efforts to map deprived urban areas follow one of four general approaches or a combination of these: (1) aggregation of "slum household" data; (2) field-based mapping by residents; (3) human visual interpretation of EO imagery (satellite, aerial, and drone-based imagery) ; and (4) semi-automatic classification of EO imagery with machine algorithms. These approaches have operated in parallel over the last two decades, largely in isolation, and each has strengths and limitations. Importantly, none of the existing approaches alone meet all requirements for area deprivation maps (Table 2).

Table 2. Strengths and limitations of existing approaches to area deprivation mapping

Requirements	Aggregated "slum households"	Field-based mapping	Human imagery interpretation	Machine imagery classification
1) Reflective of area physical characteristics	✖	✓	✓	✓
2) Reflective of area social characteristics	?	✓	?	?
3) Context dependent	✖	✓	?	?
4) Comparable across cities and countries	✓	✖	✖	✓
5) Updated frequently with timely data	✖	✖	✖	✓
6) Protective of individual privacy, and vulnerable populations	✓	✓	?	?
7) Developed via an inclusive multi-stakeholder process	✖	✖	✖	✖

1. Aggregated "Slum Households" Approach

The widely cited statistic - 1 billion slum dwellers globally - is calculated by classifying urban "slum households" in censuses or surveys, and then aggregating to country or sub-national region (UN-Habitat, 2003). Academics have similarly used the "slum household" definition to classify household survey data for statistical analysis, and interpret the results as representative of slum dwellers (e.g. Fink et al., 2014). Some experts from the social sciences recommend classifying census enumeration areas or survey clusters as "slum areas" when 50% or more of households meet the "slum household" definition (Lilford et al., 2017).

This approach has two major limitations. First, the indicators of a "slum household" do not reflect the social, environmental, and ecological factors that define deprived urban areas (Lilford et al., 2019). Second, this approach could exclude small pockets of deprived areas within a larger non-deprived area (Christ et al., 2016). A typical "slum area" is 1.6 hectares (less than two football fields) (Friesen et al., 2018). This can result in what geographers refer to as the modifiable areal unit problem (MAUP) in which a deprived area is arbitrarily divided across two or more census or survey units, resulting in a small portion of the deprived area in each unit and no units being classified as a "slum area" (Openshaw, 1983).

Key strengths of this approach are that it is easy to operationalize with the existing UN-Habitat "slum household" definition, it results in maps that are comparable across cities and countries, and exact boundaries of deprived urban areas are generally obfuscated by use of predefined census or survey units. However, classification of areas based on household-level data completely ignores area-level characteristics and important local contextual factors that define deprivation in a given city.

Furthermore, aggregation of "slum household" data is dependent on collection of censuses and surveys, ideally, every 10 years and 5 years, respectively. Publication of census and survey results usually take one to three years (Satterthwaite, 2002), preventing neighborhood deprivation maps from being updated frequently with timely information. What is more, routine censuses and surveys are not performed in all LMICs (Mahabir et al., 2018).

2. Field-based mapping

Field-based mapping is commonly performed by community NGOs such as Slum Dwellers International (Slum Dwellers International, 2016), and linked to advocacy for slum dwellers' recognition and rights (Karanja, 2010; Panek and Sobotova, 2015; Sen et al., 2003). In many cases, the approach is wholly participatory, where organized community members map and enumerate their settlement to gather invaluable planning data and catalyze community action and partnerships (e.g. Map Kibera Trust, 2009). When field-based mapping is performed by outsiders such as academics or government, the approach often begins with a review of EO imagery and identification of potential informal settlements before field validation with, or without, the involvement of community members (e.g. Cities Alliance, 2015). Many field-based approaches rely on handheld digital devices such as GPS units, and the collected data may be collated to reflect the, sometimes overlapping, land claims that exist in informal settlements (e.g. GLTN, 2017).

While field-based mapping strongly represents local context, area-level physical characteristics, and area-level social characteristics, the approach on its own is extremely difficult to upscale to whole cities and countries. In addition, field-based mapping results in area deprivation maps that are highly variable across cities and countries. Risks of fines, harassment, and eviction in vulnerable communities is often mitigated by advocacy efforts linked with the neighborhood mapping activities.

3. Human Imagery Interpretation Approach

Earth observation data are sometimes used to manually digitize informal settlements. Visual interpretation is generally based on very-high resolution imagery from satellites (up to 30 cm resolution), conventional aerial surveys (up to 10 cm resolution), or drones (up to 3 cm resolution), providing substantial detail of local physical conditions. This approach is typically based on *a priori* definitions of deprivation, for example, defining deprived areas only as informal settlements with high built-up density, irregular layout pattern, small or no internal access roads, small buildings and lack of green spaces. The use of imagery to identify and delineate informal settlements does not depend on predefined areal units and thus may approximate actual informal settlement boundaries (Lilford et al., 2019); however, the boundaries of more formalized deprived areas may be missed using this approach.

Such delineations may be performed by local (Angeles et al., 2009) or outside (Wurm and Taubenböck, 2019) experts, and are labor intensive but can provide high-quality, detailed maps required by planners. Manual delineation of informal settlements is sometimes performed to minimum requirements, and if done by several interpreters, might be inconsistent (Leonita et al., 2018). Furthermore, local experts might disagree in complex setting about the identification and delineation of informal versus formal areas (Kohli et al., 2016; Pratomo et al., 2017). Although local experts may be from the cities being mapped, delineation of informal settlements is generally performed without involvement of people living in those areas, ignoring local opinions, privacy, and geo-ethics. The degree to which human imagery interpretation reflects local context, including social information about deprived areas, depends entirely upon who is doing the interpretation and delineation.

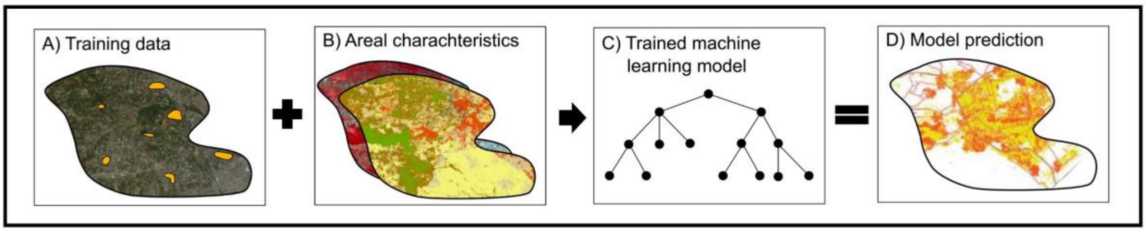
4. Machine Imagery Classification Approach

Semi-automatic "supervised" imagery classification is performed with satellite, aerial, and drone imagery (e.g. Gevaert et al., 2018; Verma et al., 2019; Wurm et al., 2019), as well as other spatial datasets such as road intersections (e.g. Jochem et al., 2018; Ibrahim et al., 2019) which allows the scaling-up of deprived area classifications. Developments in deep learning show that well-trained models can achieve classification accuracy of more than 90% (Kuffer et al., 2018). However, such methods require a large number of high-quality training data, expensive very high-resolution imagery, and are computationally demanding. Consequently, most machine-learning efforts are proof-of-concept studies that typically cover very small study areas within a single city.

Figure 2 shows a general framework for machine learning modelling. All supervised machine-learning models require training data (A). Most of the time, training data are manually digitized from imagery or field-referenced maps, which cover a limited part of the area of interest. Other areal characteristics (B) which cover the entire area of interest are used in a machine-learning model to (C) discriminate between informal/formal areas across the area of interest. These areal characteristics are also called model covariates, and overwhelmingly represent physical characteristics in practice, including building morphology, slope, and flood zone (Kuffer et al., 2016; Mahabir, et al., 2018). The machine-learning model is then used to predict the location of deprived neighborhoods, particularly informal settlements, across the area of interest (D).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0242.v1

Figure 2. Four general steps in machine imagery classification



A majority of image classification models result in binary or categorical outputs, that is, areas are classified as either formal or informal, or some other area type. In reality, however, informal settlements or deprived areas may not have sharp boundaries (Leonita et al., 2018). Further, a majority of image classification models do not account for disagreement among experts who delineate informal area training datasets (Verma et al., 2019). Both of these issues can be addressed with models that classify informal and other deprived neighborhoods on a continuous scale , that is, they classify tiny units such as grid cells by their degree of deprivation (Kohli et al., 2016).

The machine-learning imagery classification approach well reflects physical characteristics in deprived areas, and when performed by machine-learning models, it can produce results that are comparable across cities and countries. Given resources, computer-based models can also be updated frequently with recently collected EO imagery. One key limitation of this approach is that in practice it has overwhelmingly ignored area-level social characteristics, and thus better represents informal settlements only, rather than slums, informal settlements, and areas of inadequate housing. Datasets which reflect social services and social protections, and their lack thereof, such as trash pile locations, open sewer locations, crime rates, and land zone designation are likely important to distinguishing deprived and not deprived areas, however area-level social datasets have largely been untested in area deprivation machine learning models (Thomson et al., 2019).

Depending on the source of the area deprivation training data in the model, the machine imagery classification approach could be considered to represent local contexts, for example, if the training data were a representative sample of deprived neighborhoods in a city mapped by residents themselves. However, most existing machine learning models are trained on human imagery interpretations by geospatial staff located in a government or academic office, sometimes thousands of miles away. Typically, this means that local policies, histories, and definitions of deprivation are not well reflected in existing area deprivation maps.

Furthermore, many area deprivation maps based on machine imagery classification also result in sharply delineated informal settlement boundaries without the involvement of residents themselves, and may increase risk among already vulnerable communities if officials pursue fines, harassment, or eviction. In studies where degree of deprivation in tiny pixels are mapped (Schmitt et al., 2018; Wurm et al., 2019), deprived area boundaries can be obfuscated by enlarging the pixel size to limit unintended harm to communities.

Proposing an integrated deprived area mapping system (IDeAMapS)

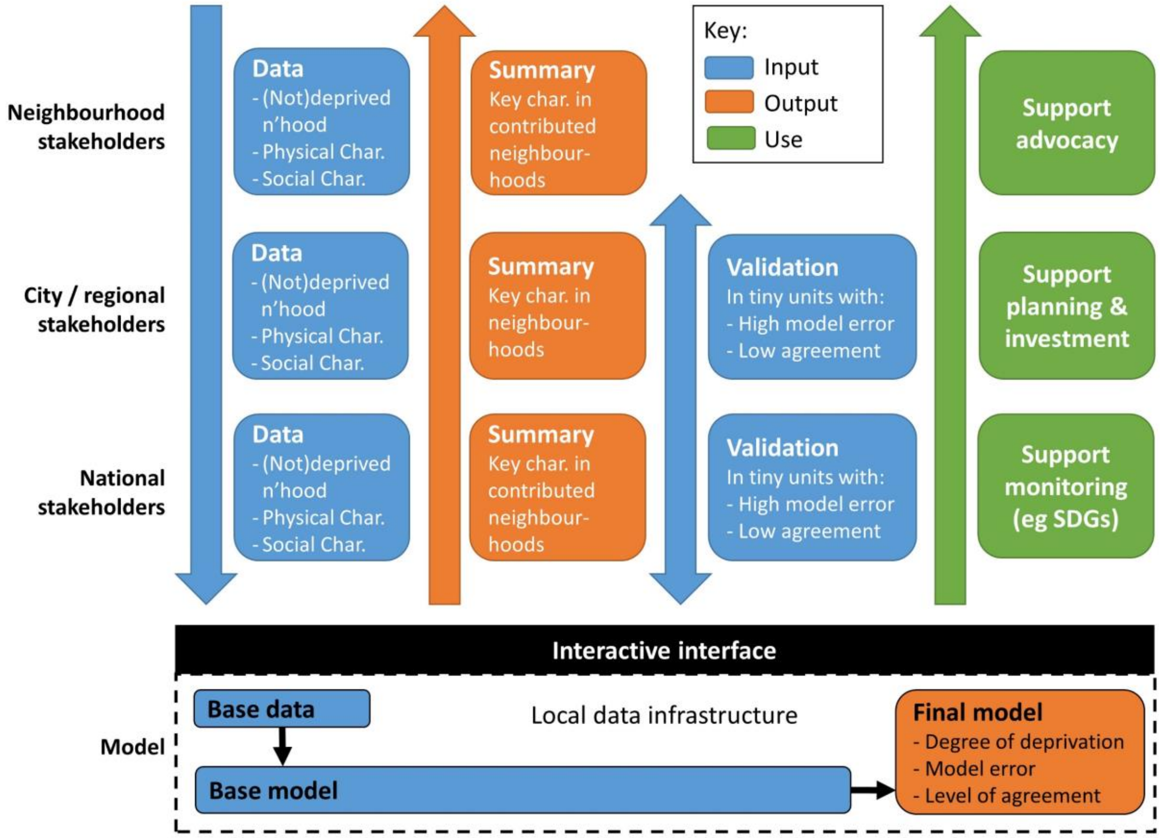
Alone, each of the current area deprivation mapping approaches has substantial limitations, however these approaches can be integrated to leverage their strengths and meet all of the area deprivation modelling requirements. Below and in Figure 3, we outline an integrated deprived area mapping system (IDeAMapS) that:

- leverages continual contributions of updated data from an ecosystem of national and local stakeholders,
- reflects the social and political realities on the ground that relate to deprived neighborhoods locally, and
- provides a simple interface with predefined geospatial models allowing users to decide which datasets are suitable to model neighborhood deprivation for their specific needs, generating an up-to-date custom map on demand.

The backbone of IDeAMapS might be a base area deprivation model and universal datasets embedded in a locally housed, open data infrastructure. A sizable amount of work would be needed up front to develop universal covariates that reflect both physical and social area-level characteristics. This would involve harmonizing national fine-scale spatial datasets of physical characteristics including land use and hazards (e.g., flood zones and unstable slopes), the shape and form of discrete settlement areas, and the components of settlements including building and road characteristics (Kohli et al., 2012). New social datasets would need to be created, for example, informal tenure by comparing real-estate website activity with population density (Mahabir, et al., 2018), or using feature extraction techniques to identify trash piles in very high resolution EO imagery (Thomson et al., 2019). Focusing on SDG area-level indicators such as literacy rates and quality of water bodies would simultaneously improve ability to measure SDGs while potentially improving the accuracy of deprived areas maps.

IDeAMapS would not only rely on universal datasets; it would also need **continual contributions** of custom, local covariates and classified neighborhood training datasets from a **range of stakeholders** at multiple levels. Contributions of deprived/not deprived area training datasets could be incentivized by providing **summary statistics for each contributed and classified area**. Summary statistics might include estimated total population from gridded population datasets, and percent of area covered by buildings, roads, water, or trash piles from feature extraction datasets. These summaries would provide value on their own. For example, a self-identified slum neighborhood groups might report trash pile coverage to officials to advocate for new municipal waste services. By allowing multiple stakeholders to contribute delineated and classified area boundaries, the system **eliminates the need for a single global deprived/"slum" area definition**, but rather accumulates a rich database of **context-dependent deprived area definitions** which can be used to train an area deprivation model in a local context.

Figure 3. Diagram of an integrated deprived area mapping system (IDeAMapS)



The output of IDeAMapS should be **formatted as a gridded dataset** in which the degree of deprivation is estimated for each grid cell. Gridded datasets are a highly flexible data format that allow the output to be aggregated to any number of spatial units such as census enumeration area or city wards. Furthermore, a sensibly sized grid cell (e.g. 50 x 50 meters) would allow for a high level of spatial detail across a city while obfuscating exact neighborhood or settlement boundaries - the appropriate size for grid cells to map deprived areas needs further study. Neighborhood names and specific geographic boundaries should never be publicly reported in this system to protect the privacy and security of residents in deprived areas. Many users will desire degree of deprivation to be translated into a classified deprived/not deprived (i.e. slum/non-slum) map, thus an option to classify cells based on a user-specified threshold of deprivation (i.e. “slumness”) should be included.

An important step in this integrated approach would be **iterating the model** by seeking additional training data from users depending on the results of the first model iteration. By running a first version of the model with the available universal and contributed dataset, grid cells in which the model performs poorly, and grid cells in which only one training dataset is available, could be sampled and presented to a locally-based user. The user could be asked to classify the cell as deprived/not deprived to feed back into the final model, both improving statistical certainty, and allowing for a measure of agreement among users about what is, and is not, a deprived area.

Users would need a simple **interactive interface** that is linked to a **locally-based data infrastructure.** Many government, NGO, and community groups may hesitate to contribute if their data will be extracted from the country. Additionally, contributors need control over their data, including the ability to validate, contest and revise contributed data. We envision this platform as a public good, freely accessible to national and local governments, community groups, NGOs, researchers, international agencies, and the public. Given the unique needs of national and local governments to produce an official "slum area" map for SDG and other official reporting, special support should be provided to national statistical agencies to generate area deprivation maps on a regular basis using approved covariates and training datasets. We recognize that this is an ambitious endeavor that requires clear terms of reference, sustained resources, commitment, and trust in the governance structure (see UTEP Consortium, 2019 for how this might work).

IDeAMapS has the ability to reflect both physical and social area-level characteristics, and multiple stakeholder definitions of deprivation. By integrating universal datasets and regularly contributed bespoke datasets - particularly field-verified boundaries of deprived/not deprived areas - in a geostatistical model, urban area deprivation maps could be updated frequently and more accurately. The role of community-contributed data would be highly valued in this system, and ensure a fair information exchange by providing a digestible summary of social and physical data for each contributed area boundary. The use of models also ensures the protection of vulnerable populations by formatting the output as a continuous deprivation value in small grid cells, thus not drawing hard boundaries around "slums" and other deprived areas.

Discussion

The authors of this commentary hail from the four existing approaches to area deprivation mapping - aggregated "slum households," field-based mapping, human imagery interpretation, and machine imagery classification. Through a series of workshops in 2018 and 2019, we came to understand the strengths and limitations of each other's approaches, and outlined this call for an integrated deprived area mapping system (IDeAMapS). We have summarized our thoughts here to stimulate discussion within and across our disciplines, and to connect with new and diverse stakeholders who share our goal to identify deprived urban areas in LMICs and improve the wellbeing of their residents. Our work together thus far has highlighted several important areas of understanding.

First, "slum households" and deprived areas, while related, are different phenomena. Deprived areas are defined by physical and social risks and outcomes such as absence of public services, while "slum households" are defined by risks and outcomes in households such as limited-income. To effectively target vulnerable populations with policies and programs, we need to locate both "slum households" and deprived urban areas, and understand the unique risks that face "slum households" in deprived, as well as not deprived, areas.

Second, a wealth of detailed maps of area-level physical characteristics exist in LMICs, however, limited maps are available of area-level social characteristics. Methods for area deprivation mapping that use satellite imagery or spatial data focus almost exclusively on small, disorganized buildings or streets; however, deprived areas are not synonymous with informal settlements (Nuissl and Heinrichs, 2013). Many of the risks and outcomes that define life in deprived areas are social in nature, and can co-exist with organized streets and permanent buildings. The creation of social area-level datasets such as areas of secure tenure or trash pile locations stand not only to improve the accuracy of area deprivation maps, but also serve as valuable decision-making tools on their own.

Third, area deprivation mapping can have both positive and negative effects on individuals who live in deprived areas. The mapping of deprived areas has been used to advocate for the rights of slum dwellers and help them access basic public services (Panek and Sobotova, 2015), as well as to fuel demolition campaigns in informal settlements and harassment of residents (Roy, 2009). Critically, it is involvement of residents in the process that determines the effect of such maps (Lilford et al., 2017; Panek and Sobotova, 2015). Community groups based in slums and other deprived areas must be central to any area deprivation mapping initiative, especially large-scale initiatives such as the one we propose.

Finally, existing evidence points toward seven basic requirements for area deprivation maps: (1) reflects physical risks, (2) reflects social risks, (3) is context dependent, (4) is comparable across cities and countries, (5) is updated frequently with timely data, (6) protects individual privacy, and vulnerable populations, and (7) is developed via an inclusive multi-stakeholder process. We believe all seven requirements can be achieved through IDeAMapS. The simple classification of deprived/not deprived areas enables reporting on slums, informal settlements and areas of inadequate housing for SDG 11, and provides the spatial information needed to disaggregate other population-based SDG indicators. An integrated mapping system further enables key dimensions of deprivation to be mapped to support critical budget and planning decisions for local and national governments. For example, IDeAMapS might separately identify areas of a city where pollution, or unplanned housing, or social risks are predominant problems. Self-identified slum communities who hold mapping campaigns can benefit from receiving data summaries of characteristics that have been mapped by others in their neighborhoods for use in planning and advocacy. Those deprived communities that do not have active mapping campaigns would benefit from being represented in national statistics and subsequent policies and programming.

References

Abbott J (2003) The use of GIS in informal settlement upgrading: Its role and impact on the community and on local government. Habitat International 27(4): 575–593. DOI: 10.1016/S0197-3975(03)00006-7.

Angeles G, Lance P, Fallon JB, et al. (2009) The 2005 census and mapping of slums in Bangladesh: design, select results and application. 8: 1–19. DOI: 10.1186/1476-072X-8-32.

Bramante JF and Raju DK (2013) Predicting the distribution of informal camps established by the displaced after a catastrophic disaster, Port-au-Prince, Haiti. Applied Geography 40: 30–39. DOI: 10.1016/j.apgeog.2013.02.001.

Chitekwe-Biti B, Mudimu P, Nyama GM, et al. (2012) Developing an informal settlement upgrading protocol in Zimbabwe - the Epworth story. Environment and Urbanization 24(1): 131–148. DOI: 10.1177/0956247812437138.

Christ K, Baier K and Azzam R (2016) Slums and informal housing in India: a critical look at official statistics with regard to water and sanitation. Water International 41(2). Routledge: 308–324. DOI: 10.1080/02508060.2016.1139656.

Cities Alliance (2015) DKI Jakarta: Informal Settlement Mapping Final Report. P121467. Washington DC USA. Available at: http://pubdocs.worldbank.org/en/408201444418421010/pdf/DKI-JAKARTA-INFORMAL-SETTLEMENT-MAPPING.pdf.

Corburn J and Cohen AK (2012) Why We Need Urban Health Equity Indicators: Integrating Science, Policy, and Community. PLoS Medicine 9(8): 1–6. DOI: 10.1371/journal.pmed.1001285.

Diaz MM, Engelman R, Klugman J, et al. (2017) State of World Population 2017. Erksen A (ed.). New York, NY: UNFPA. Available at: http://www.unfpa.org/swop.

Engstrom R, Ofiesh C, Rain D, et al. (2013) Defining neighborhood boundaries for urban health research in developing countries: A case study of Accra, Ghana. Journal of Maps 9(1): 36–42. DOI: 10.1080/17445647.2013.765366.

Ezeh A, Oyebode O, Satterthwaite D, et al. (2017) The history, geography, and sociology of slums and the health problems of people who live in slums. The Lancet 389: 547–558. DOI: 10.1016/S0140-6736(16)31650-6.

Fink G, Günther I and Hill K (2014) Slum residence and child health in developing countries. Demography 51(4): 1175–1197. DOI: 10.1007/s13524-014-0302-0.

Friesen J, Taubenböck H, Wurm M, et al. (2018) The similar size of slums. Habitat International 73(February). Elsevier: 79–88. DOI: 10.1016/j.habitatint.2018.02.002.

Gevaert CM, Persello C, Elberink SO, et al. (2018) Context-Based Filtering of Noisy Labels for Automatic Basemap Updating from UAV Data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11(8). IEEE: 2731–2741. DOI: 10.1109/JSTARS.2017.2762905.

Global Land Tool Network (GLTN) (2017) Social Tenure Domain Model (STDM). Available at: https://stdm.gltn.net/ (accessed 17 October 2019).

Gruebner O, Sachs J, Nockert A, et al. (2014) Mapping the Slums of Dhaka from 2006 to 2010. Dataset Papers in Science 2014(1): 1–7. DOI: 10.1155/2014/172182.

Harris-Fry H, Shrestha N, Costello A, et al. (2017) Determinants of intra-household food allocation between adults in South Asia - A systematic review. International Journal for Equity in Health 16(1). International Journal for Equity in Health: 1–21. DOI: 10.1186/s12939-017-0603-1.

Ibrahim MR, Titheridge H, Cheng T, et al. (2019) predictSLUMS: A new model for identifying and predicting informal settlements and slums in cities from street intersections using machine learning. Computers, Environment and Urban Systems 76. Elsevier: 31–56. DOI: 10.1016/j.compenvurbsys.2019.03.005.

Improving Health in Slums Collaborative (2019) A protocol for a multi-site, spatially-referenced household survey in slum settings: methods for access, sampling frame construction, sampling, and field data collection. BMC Medical Research Methodology 19(1): 109. DOI: 10.1186/s12874-019-0732-x.

Jochem WC, Bird TJ and Tatem AJ (2018) Identifying residential neighbourhood types from settlement points in a machine learning approach. Computers, Environment and Urban Systems 69. Elsevier: 104–113. DOI: 10.1016/j.compenvurbsys.2018.01.004.

Karanja I (2010) An enumeration and mapping of informal settlements in Kisumu, Kenya, implemented by their inhabitants. Environment and Urbanization 22(1): 217–239. DOI: 10.1177/0956247809362642.

Kohli D, Sliuzas R, Kerle N, et al. (2012) An ontology of slums for image-based classification. Computers, Environment and Urban Systems 36(2). Elsevier Ltd: 154–163. DOI: 10.1016/j.compenvurbsys.2011.11.001.

Kohli D, Stein A and Sliuzas R (2016) Uncertainty analysis for image interpretations of urban slums. Computers, Environment and Urban Systems 60. Elsevier Ltd: 37–49. DOI: 10.1016/j.compenvurbsys.2016.07.010.

Kuffer M, Barros J and Sliuzas R V. (2014) The development of a morphological unplanned settlement index using very-high-resolution (VHR) imagery. Computers, Environment and Urban Systems 48. Elsevier Ltd: 138–152. DOI: 10.1016/j.compenvurbsys.2014.07.012.

Kuffer M, Pfeffer K and Sliuzas R (2016) Slums from space-15 years of slum mapping using remote sensing. Remote Sensing 8: 1–29. DOI: 10.3390/rs8060455.

Kuffer M, Wang J, Nagenborg M, et al. (2018) The Scope of Earth-Observation to Improve the Consistency of the SDG Slum Indicator. ISPRS International Journal of Geo-Information 7(11): 428. DOI: 10.3390/ijgi7110428.

Lemma T, Sliuzas R and Kuffer M (2006) Participatory approach to monitoring slum conditions: An example from Ethiopia. In: Rambaldi G, Corbett J, Olson R, et al. (eds) Mapping for Change: Practice, Technologies and Communication: Proceedings of the International Conference on Participatory Spatial Information Management and Communication. London UK: IIED, pp. 58–66.

Leonita G, Kuffer M, Sliuzas R, et al. (2018) Machine Learning-Based Slum Mapping in Support of Slum Upgrading Programs: The Case of Bandung City, Indonesia. Remote Sensing 10(10): 1522. DOI: 10.3390/rs10101522.

Lilford R, Kyobutungi C, Ndugwa R, et al. (2019) Because space matters: conceptual framework to help distinguish slum from non-slum urban areas. BMJ Global Health 4(2): e001267. DOI: 10.1136/bmjgh-2018-001267.

Lilford RJ, Oyebode O, Satterthwaite D, et al. (2017) Improving the health and welfare of people who live in slums. The Lancet 389. Elsevier Ltd: 559–570. DOI: 10.1016/S0140-6736(16)31848-7.

Mahabir R, Crooks A, Croitoru A, et al. (2016) The study of slums as social and physical constructs: Challenges and emerging research opportunities. Regional Studies, Regional Science 3(1). Routledge: 399–419. DOI: 10.1080/21681376.2016.1229130.

Mahabir R, Croitoru A, Crooks A, et al. (2018) A Critical Review of High and Very High-Resolution Remote Sensing Approaches for Detecting and Mapping Slums: Trends, Challenges and Emerging Opportunities. Urban Science 2(1): 8. DOI: 10.3390/urbansci2010008.

 doi:10.20944/preprints201910.0242.v1

Mahabir R, Agouris P, Stefanidis A, et al. (2018) Detecting and mapping slums using open data : a case study in Kenya Detecting and mapping slums using open data : a case study in. International Journal of Digital Earth: 1–25. DOI: 10.1080/17538947.2018.1554010.

Map Kibera Trust (2009) Map Kibera. Available at: https://mapkibera.org (accessed 17 October 2019).

Nairobi City County (2018) Mukuru SPA Project. Available at: https://www.muungano.net/mukuru-spa (accessed 17 October 2019).

Nuissl H and Heinrichs D (2013) Slums: Perspectives on the definition, the appraisal and the management of an urban phenomenon. Journal of the Geographical Society of Berlin 144(2): 105–116. DOI: 10.12854/erde-144-8.

Openshaw S (1983) The Modifiable Areal Unit Problem. Norwick, England: Geo Books.

Panek J and Sobotova L (2015) Community Mapping in Urban Informal Settlements: Examples from Nairobi, Kenya. Electronic Journal of Information Systems in Developing Countries 68(1): 1–13. DOI: 10.1002/j.1681-4835.2015.tb00487.x.

Pratomo J, Kuffer M, Martinez J, et al. (2017) Coupling uncertainties with accuracy assessment in object-based slum detections, case study: Jakarta, Indonesia. Remote Sensing 9(11): 1–17. DOI: 10.3390/rs9111164.

Roy A (2009) Why India cannot plan its cities: Informality, insurgence and the idiom of urbanization. Planning Theory 8(1): 76–87. DOI: 10.1177/1473095208099299.

Satterthwaite D (2002) The Scale of Urban Change Worldwide 1950–2000 and its Underpinnings. In: Human Settlement Discussion-Urban Change. London UK: International Institute for Environment and Development, p. 42.

Sen S, Hobson J and Joshi P (2003) The Pune Slum Census: Creating a socio-economic and spatial information base on a GIS for integrated and inclusive city development. Habitat International 27(4): 595–611. DOI: 10.1016/S0197-3975(03)00007-9.

Sliuzas R, Mboup G and Sherbinin A de (2008) Report of the Expert Group Meeting on Slum Identification and Mapping. Available at: https://www.alnap.org/help-library/report-of-the-expert-group-meeting-on-slum-identification-and-mapping (accessed 17 October 2019).

Sliuzas R, Kuffer M and Masser I (2010) Remote Sensing of Urban and Suburban Areas. In: Remote Sensing and Digital Image Processing. Dordrecht Germany: Springer, pp. 67–84. DOI: 10.1007/978-1-4020-4385-7.

Slum Dwellers International (2016) Know Your City BETA. Available at: http://knowyourcity.info/explore-our-data/ (accessed 17 October 2019).

Thomson DR, Linard C, Vanhuysse S, et al. (2019) Extending Data for Urban Health Decision-Making: a Menu of New and Potential Neighborhood-Level Health Determinants Datasets in LMICs. Journal of Urban Health: 1–23. DOI: 10.1007/s11524-019-00363-3.

UN Department of Economic and Social Affairs (UN-DESA) (2017) World Population Prospects 2017 Revision. Available at: https://esa.un.org/unpd/wpp/Publications/Files/WPP2017_KeyFindings.pdf.

UN Department of Economic and Social Affairs (UN-DESA) (2018a) World Urbanization Prospects: The 2018 Revision. Available at: https://population.un.org/wup/DataQuery/ (accessed 17 October 2019).

UN Department of Economic and Social Affairs (UN-DESA) (2018b) Sustainable Development Goals. Available at: https://sustainabledevelopment.un.org/sdgs (accessed 17 October 2019).

UN-Habitat, UN Statistics Division and Cities Alliance (2002) Expert Group Meeting on Urban Indicators: Secure Tenure, Slums and Global Sample of Cities. Nairobi Kenya.

UN Human Settlements Programme (UN-Habitat) (2003) The Challenge of Slums: Global Report on Human Settlements 2003. London UK: Earthscan Publications Ltd. DOI: 10.1006/abio.1996.0254.

UN Human Settlements Programme (UN-Habitat) (2007) Slums: Some Definitions. Available at: http://mirror.unhabitat.org/documents/media_centre/sowcr2006/SOWCR 5.pdf (accessed 22 March 2018).

UN Human Settlements Programme (UN-Habitat) (2017) Distinguishing slum from non-slum areas to identify occupants' issues. Available at: https://unhabitat.org/distinguishing-slum-from-non-slum-areas-to-identify-occupants-issues/ (accessed 22 March 2018).

UTEP Consortium (2019) Urban TEP. Available at: https://urban-tep.eu/#! (accessed 17 October 2019).

Verma D, Jana A and Ramamritham K (2019) Transfer learning approach to map urban slums using high and medium resolution satellite imagery. Habitat International 88. Elsevier: 101981. DOI: 10.1016/j.habitatint.2019.04.008.

Weeks JR, Getis A, Stow DA, et al. (2012) Connecting the Dots Between Health, Poverty and Place in Accra, Ghana. Ann Assoc Am Geogr 102(5): 932–941. DOI: 10.1080/00045608.2012.671132.Connecting.

Wurm M and Taubenböck H (2019) Detecting social groups from space - Assessment of remote sensing-based mapped morphological slums using income data. Remote Sensing Letters 9(1). Taylor & Francis: 443–456. DOI: 10.1080/2150704X.2017.1384586.

Wurm M, Stark T, Zhu XX, et al. (2019) Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. ISPRS Journal of Photogrammetry and Remote Sensing 150. Elsevier: 59–69. DOI: 10.1016/j.isprsjprs.2019.02.006.

Acknowledgements

Special thanks to the following people who contributed to key workshops and meetings that gave rise to this paper: Donatien Beguy, Daniel Githira, Julius Majale, Shipra Narang-Suri, Walter Oriedo, Kerstin Sommer, Daniela Di Filippo, and Faith Barorot from UN-Habitat; Kenneth Mubea from the Global Partnership for Sustainable Development Data (GPSDD), Sergio Freire and Alice Siragusa from the European Commission Joint Research Centre; Blessing Mberu from the African Population & Health Research Center; Andrew Tatem from the WorldPop Team at the University of Southampton; Steven Ramage from the Group on Earth Observations, Philipp Ulbrich and Godwin Yeboah from the Institute for Global Sustainable Development at the University of Warwick, and Robyn T. Emerson an urban development consultant and author.

Author contributions

All co-authors helped to conceptualize this paper, and read and approved the final version. DRT and MK performed the literature investigation, and DRT, MK, JPA, CK, GB, and TG wrote the initial draft. The following people contributed to visualizations: CA, JPA, GB, SC, TG, BH, CK, MK, NM, and DRT. The following people performed project administration: JPA, BH, CK, MK, DM, and DRT.

Competing interests

We declare no competing interests.

Funding

This work is unfunded.