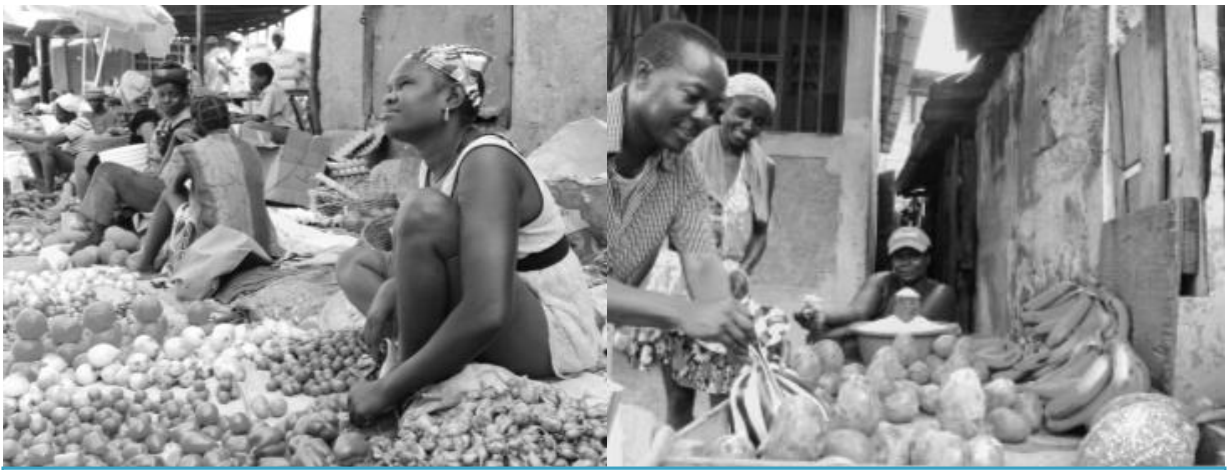


Adapting to an Urban World Phase II (2017)
Assessment Design in Urban Areas – Expert Consultations

1. Background

Managing urban areas is one of the major development challenges of the 21st century. While the increasing attention to the urban food insecure is evident, effective information gathering guiding humanitarian response and targeting remain a challenge to meet food security and other basic needs. In order to assess vulnerability to food insecurity in urban settings, the WFP Vulnerability Analysis and Mapping unit, the global Food Security Cluster and their partners are developing innovative approaches based on lessons learned and leveraging opportunities offered by new technologies as part of the second phase of the Adapting to an Urban World Initiative (2017).

As part of the planning phase, the World Food Programme (WFP) and the Global Food Security Cluster (gFSC) in partnership with Flowminder/WorldPop convened a technical meeting in Rome to brainstorm and discuss options to improve assessment designs in urban settings and identify innovative solutions for the way forward. Participating partners included the Food and Agriculture Organization of the United Nations (FAO), the GeoData Institute of the University of Southampton, REACH, the Humanitarian Open Street Map Team (HOT), Oxfam and the Massachusetts Institute of Technology (MIT). This report presents a non-exhaustive list of innovative tools and approaches to address identified challenges to assess food security in urban areas. This will be followed by the testing of tools and learning events towards the end of the year.

2. Key challenges for measuring overall and vulnerable populations in urban settings

Urban assessments present many specific and unique considerations and challenges. One of the most common challenges is to **measure the size, location, characteristics, and movement of urban populations**. The measurement of populations is closely related to representative survey sampling and it is used to provide target population estimates for programmes. The overall challenge of measuring populations can be broken down further into six non-exhaustive challenges:

Defining urban areas. While most countries have an official geographic classification of urban and rural areas, these are not always up to date or simply do not exist. Therefore, how do we validate existing defined urban/rural geographic areas? How do we correct them if necessary? How do we create them if they are non-existent?

Defining geographic strata/Identifying vulnerable areas. The official administrative boundaries are often not appropriate for defining survey strata and sampling frames. New technologies and approaches can help identify homogeneous areas and can suggest the information required to define and assess vulnerability.

KEY CHALLENGES

1. Defining urban areas
2. Defining geographic strata/Identifying vulnerable areas
3. Estimating population size
4. Measuring population movements
5. Sampling for representative household surveys
6. Feasibility of innovative solutions and challenges

Estimating population size. Official census data is often inaccurate or unavailable. Even in cases where census data is considered appropriate, it may not correspond to the geographic strata identified for the assessment. This is a particular concern for informal settings/slums, where populations are highly mobile and often not fully accounted for.

Measuring population movements. Population movements must be quantified to update population size in a given area and to recommend the most appropriate relief, recovery or structural interventions in support of the displaced.

Sampling for representative household surveys. Sampling in urban areas presents many challenges, such as agreeing on a common definition of an urban household and estimating their number and location in buildings that can contain many, a few or none of them. Sampling for representative surveys also relies on accurate geographic definitions of urban areas, geographic strata, and population distribution and size within these areas. Furthermore, representative sampling generally requires 'clusters' (e.g. enumeration areas) to sample from including their precise geographic boundaries and their population sizes. What can be done if the cluster data available is outdated or non-existent?

Feasibility of innovative solutions to challenges. Improved population measurements and sampling approaches must be feasible, i.e. technically within the capacity of the staff available to run the survey; financially within the available survey budget; and realistic in terms of time requirements. The latter is particularly important in the context of emergency assessments. As emphasized by Oxfam, agreeing on the targeting parameters is also a challenge. Selecting the right targeting approach combines financial, feasibility and vulnerability considerations, but also requires careful considerations of geographic and political dynamics. Another challenge highlighted was the importance of ensuring local ownership of any solution identified. Without that, results may not be accepted and longer-term sustainability undermined. As urban surveys usually depend on both qualitative and quantitative data as well as formal and informal data sources, challenges around data integration from various sources were also emphasized.

3. Solutions identified

Tools and solutions presented by partners were associated to the corresponding abovementioned challenges they address. Defining urban and strata boundaries were grouped under the wider key challenge of defining urban areas and sub-areas within urban settlements. Population estimation and population movement were grouped under the key challenge of estimating population sizes and movement. All tools essentially contribute to the overall challenge of sampling for representative surveys, which requires accurate geographic definitions of urban areas, geographic strata and population sizes.

Table 1: Tools proposed and corresponding challenges they address

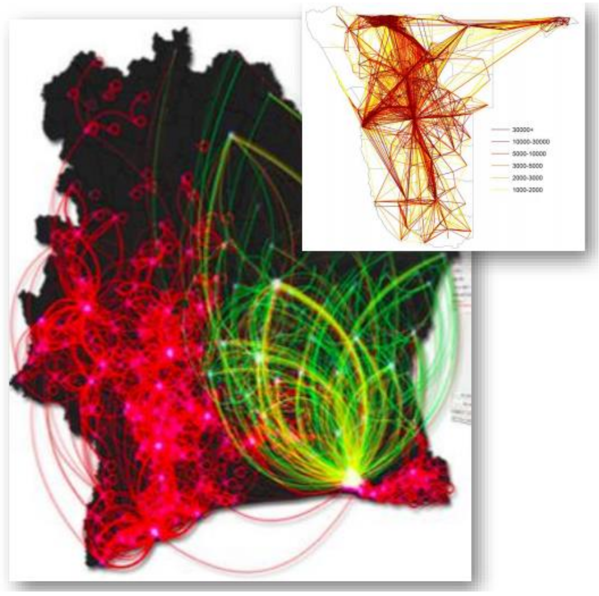
		Key challenge 1: Defining urban areas and sub-areas within urban settlements		**Key challenge 2:** Estimating population sizes and movements	
Organizations	**Potential solutions identified**	*Urban boundaries*	*Strata boundaries*	*Population estimation*	*Population movement*
Flowminder/ WorldPop	CDR-based poverty mapping		X		
	CDR-enhanced pop estimates, movements			X	X
	GridSample	X	X	X	
	Grids-to-"sensible" EAs		X		
	Top-down gridded population estimates				
	Satellite-enhanced gridded population			X	
FAO	City Region Food Systems Approach		X		
GeoData/University of Southampton	Slum detection algorithm (using satellite imagery)	X			
	Population estimation using secondary data			X	
REACH	Neighbourhood approach: map locally-recognized neighbourhood boundaries	X	X		
	Identification and trend monitoring of vulnerable urban territories				X
HOT	OpenStreetMap (OSM)	X	X	X	
	Insafe and Drainage Mapping		X		
JRC (presented by MIT)	Global Human settlements layer	X		X	X

Key challenge 1: Defining urban boundaries and sub-areas within urban settlements

CDR-based poverty mapping (Flowminder/WorldPop)

While providing essential information on population movements, mobile operated data also provides information on social networks and consumption. When a user makes a call using a mobile phone, the call is routed through the nearest tower and the network operator records the time and tower of call. Operators also store data such as credit purchase frequencies and top-up amounts, which may contribute to better understanding poverty situations in certain settings and help map out these areas. Moreover, data on number of contacts and calling patterns are also stored, which gives insight on people’s social networks. Analyzing this data could contribute to delimitating the boundaries of areas with higher poverty prevalence within urban settings.

Figure 1: Mapping mobility, social networks and consumption using CDR data



Mobility: Changing densities, flows, seasonal/permanent migration

Social networks: Number of contacts, calling patterns

Consumption: Credit purchase frequencies, top-up amounts

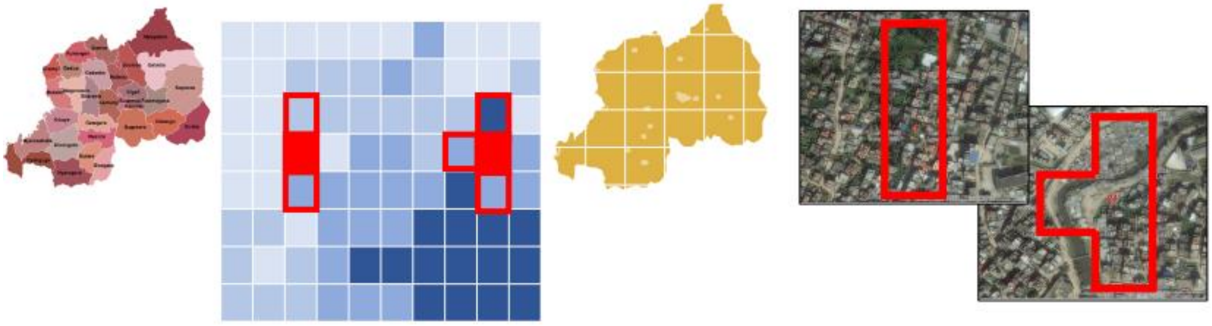
Gridded population sampling with GridSample (Flowminder/WorldPop)

In urban contexts, traditional sampling methods (e.g. two-stage cluster sampling) often do not work because sampling frames may be outdated, too aggregated or inaccurate. There may also be geographic exclusions due to government definitions of non-residential zones. Furthermore, traditional samples are not selected to be spatially representative. The GridSample tool[1] takes advantage of any gridded population dataset when traditional census-based sampling frames are not adequate, or when spatial-representation is needed in addition to population representation. A number of modelling techniques and data inputs are used to generate gridded population datasets. Generally, they integrate information about the population with satellite imagery and GIS to create a fine-scale raster representation of human population distribution. GridSample allows for sound random sampling designs that are analogous to the traditional sampling methods by sampling cells with probability proportionate to estimated population size, and allowing for oversampling of important sub-domains such as urban areas (in a national sample) or slum areas (in an urban only sample, if slum boundaries are mapped).

Optionally, cells can be grown into primary sampling units (PSUs) that have approximately the same number of people. These gridded PSU boundaries can be used directly, or more geographically sensible PSU boundaries can be manually defined within the gridded PSUs, reflecting the common practice of segmenting. GridSample summarizes population totals in PSUs and Strata to be used later to calculate sampling weights. The tool also supports rapid assessment non-probability sampling with spin-the-pen/random walk methods by applying a technique to identify a random dwelling within the PSU.

[1] Available free in R CRAN.

Figure 2: Gridded population sampling process



Transforming grids into "sensible" enumeration areas (Flowminder/WorldPop)

Flowminder/WorldPop is exploring a process to convert grid cells into 'sensible enumeration areas'. Essentially, it consists in analysing existing boundaries of "sensible" geographic/administrative units (e.g. old census enumeration areas) and updating population estimates by summing the contained gridded data. This technique ensures that these areas have more or less the same population count, which can be adjusted by aggregating or disaggregating neighbouring areas depending on their population size. Larger units can be divided using mapped features such as rivers, road and neighbourhood boundaries.

City Region Food Systems Approach (FAO)

The **City Region Food Systems approach** (CRFS) is a relevant additional element to consider when approaching urban sampling and fine-tuning tools, because it provides a breakdown of the complexity of variables and dynamics that affect availability and access to food, prices of food and purchasing power, role of actors, among others. These elements are key to define consumers' (as well as producers') vulnerability in city region settings (i.e. urban, peri-urban and rural hinterlands). **The main scope of CRFS is to understand the sustainability of food systems**. In order to achieve this, three steps are required:

- First, an **assessment phase** based on secondary data is conducted to get a clear understanding of the structure and nature of food system in all its aspects.
- The second phase requires a more **in-depth analysis of the key issues identified in the first phase** through primary data collection
- Finally, on the basis of the thorough assessment conducted, it is possible to **develop strategies** to address them with local institutions through consultative and participatory processes.

Figure 3: CFSR approach and indicator framework

Slum detection algorithm (University of Southampton/GeoData)

This solution is based on a study that was conducted in Hyderabad (India) and aimed to **identify slum areas using satellite-imagery and to provide an accurate estimate of population living in these areas**. The approach consists in applying an algorithm to analyse the mass of pixels on a satellite imagery and to identify slum settings based on the distinct morphology of roofing and density (called *Lacunarity*). Each pixel is associated with a binomial value (1 or 0). Based on the level of lacunarity typical of the slum, the algorithm identifies boundaries of slum areas. A visual verification from practitioners is then recommended to assess findings from the model. The model also goes one step further: it associates *lacunarity* and slum surface with ground-level data on population density from census or demographic surveys in order to compute the slum population.

Figure 4: Slum establishment process in Hyderabad (India) between 2003 and 2010

Identifying community-based neighbourhoods (REACH)

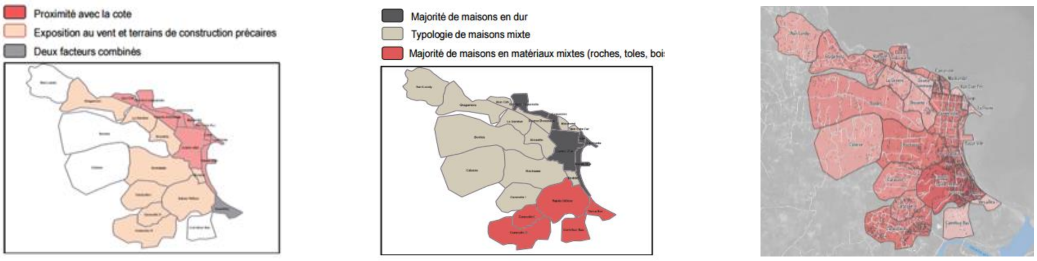
To address the issues of defining urban boundaries and strata boundaries, REACH uses a community-based neighbourhood approach. Through participatory mapping and/or transect walking and/or (focus) group discussions with local authorities and local communities, it is possible to better define the urban territory through recognized neighbourhoods as opposed to neighbourhoods defined through administrative boundaries, if different or not available. This community-based approach is relatively easy to conduct and communicate, and particularly relevant for rapid sudden onset as well as volatile humanitarian contexts. Moreover, a better understanding of neighbourhoods provides a better lens to understand the impacts of a crisis, the underlying factors of risks and vulnerabilities as well as post-crisis vulnerabilities, which can help direct interventions in cities.

The following standard sampling approaches can be used to gather relevant information useful for the definition of neighbourhoods and estimates of vulnerable population:

- **Purposive or judgmental sampling**: in the case of a rapid assessment, this approach helps choose the most appropriate key informants and also better understand local community profiles.
- **Representative (random) sampling** through more accurate population estimates. This is especially the case when specific strata are required in order to draft an "as random as possible" sampling framework.

This approach can also **highly benefit from the use of secondary data**, both for identifying underlying vulnerabilities, as well as at the analysis stage. Main challenges to such approach is the usual reluctance of Governments and municipalities to review the urban boundaries when available, even more so if enumeration areas from censuses are perfectly over-layered to them. On the other hand, when such mapping has never been done before, local authorities and government agency are usually interested to inherit such mapping for future reference.

Figure 5: An example of community-based neighbourhoods



OpenStreetMap (HOT)

OpenStreetMap (OSM) is a tool, which is used to fill the gaps of data on maps. It is essentially an open geographical data set, accessible under open licensed The mapping process is divided in two phases:

- **Phase 1**: **Desk-based activities** can be done anywhere and by anybody. This includes mapping building footprints, road network and other basic features. It also includes using aerial imagery (satellite, lidar, and/or Unmanned Aerial Vehicle or UAV) to digitize. The HOT Tasking Manager tool was created to ensure an effective division of labour and to make sure that the work is not duplicated. Moreover, there is always a senior expert that reviews the data inserted to ensure its quality.
- **Phase 2**: **Field-based activities** to capture disaster vulnerability attributes, or other attributes depending on needs. Mobile data collection tools such as OpenMapKit, OsmTracker, Maps.ME, OsmAnd and Field Papers are used during field surveys – most are freely accessible and can be downloaded as Apps on common mobile smartphones.

Figure 4: Example of desk-based activities – "Mapathon" in Indonesia

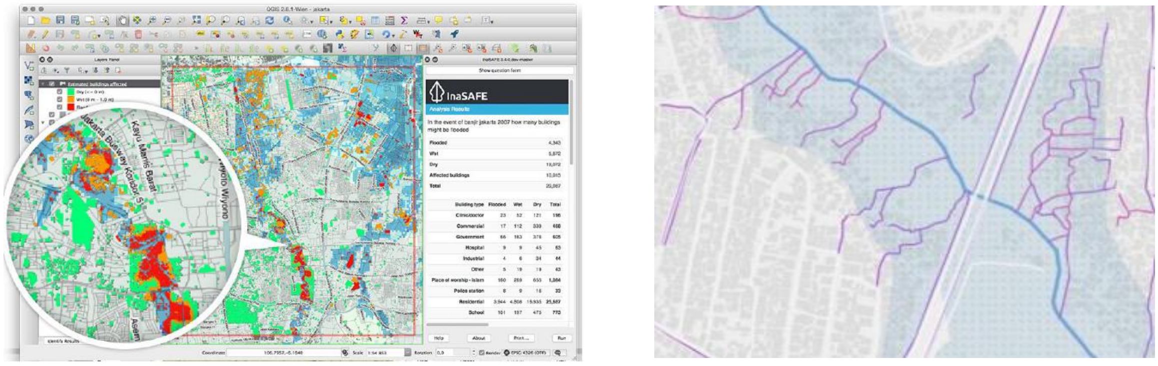
Discussion

Open access, a risk? Quality control risks are mitigated by the "peer" control of the local street map community itself that keeps an eye on their environment. There is also a senior expert controlling all data entered.

Inasafe and Drainage Mapping (HOT)

Inasafe is a tool used to create natural hazard impact scenarios. It works by uploading data from historical series for a given area and by selecting the types of recurrent hazards occurring in that area. With this information, the tool automatically analyses the areas and the people/households who are exposed to such hazards. Inasafe is currently being used in Indonesia in collaboration with the Government for emergency and preparedness purposes in the likely occurrence of natural disasters. This tool can be combined with drainage mapping, which is done through participatory mapping where people are asked to enumerate which areas are and/or subject to floods in their neighbourhoods.

Figure 5: Inasafe and Drainage Mapping



Discussion

How to reconcile informal boundaries with jurisdictional boundaries? When using innovative approaches to identify urban boundaries, there is always the challenge of reconciling these “informal” boundaries with the jurisdictional boundaries of an area. The neighbourhood approach notably raises this issue as it aims to define urban territories through recognized neighbourhoods in line with official boundaries (or enumeration areas) used by previous demographic surveys and censuses. Finding matching points can be quite challenging, but the key remains in identifying the right formula to making both data sets communicate and complement each other for a better understanding of the urban setting.

Is being on a map still too sensitive? HOT’s work in the field highlighted that in certain informal settlements, there is still fear of being identified on a map because of the consequences this may entail such as eviction by the local government. This fear may also be shared by local governments themselves who wish not to emphasize the presence of informal settlements on their territory. Although satellite imagery is not restricted - which makes informal settlements visible to all - the concern of being identified on an official map is a reality and may entail negative consequences. Very few ethic reviews are conducted on data collected.

Key challenge 2: Estimating population sizes and movements

Top-down gridded population estimates (Flowminder/WorldPop)

Various gridded population methods are available to disaggregate population data from administrative unit-based counts to much smaller grid cells based on the probability of human settlement derived from multiple satellite imagery and GIS data sources. WorldPop has produced gridded population estimates for 100m X 100m grid cells for a low- and middle-income countries using a random forests-based machine learning model[2].

Population estimation using satellite imagery (Flowminder/WorldPop)

Where census data are non-existent, outdated or unreliable, top-down gridded population estimates may be unreliable and thus WorldPop has been developing “bottom-up” mapping approaches to estimate population from satellite imagery and household surveys. The modelling happens in two phases. First, high resolution satellite imagery is used to map areas settled by people by extracting individual buildings or groups of buildings (using automated or crowdsource approaches). With this

[2] These gridded population estimates, plus metadata, are freely available for download and use at www.worldpop.org.

information, different population typologies are identified and classified. Second, a sample of the settled areas is visited and the population is counted. These population counts are used to train an algorithm that considers the settled typology and numerous additional satellite and GIS covariates to estimate population totals in all settled areas.

CDR-enhanced population estimates and movements (Flowminder/WorldPop)

Mobile operator data is very promising to understand changes in population distributions after a crisis. As it was briefly explained in the previous section, mobile operators record the time and tower of each call and SMS. When a person moves and makes a call or sends a message from a different location, this information is stored. Analyzing this data gives an insight the mobility of mobile users. Call data records (CDR) data, however, are not without limitation. There is not a one-to-one relationship between phone SIM cards and people: some individuals own multiple SIMs, some SIMs are used by multiple people, certain types of users might choose a certain provider, and while mobile coverage is close to 80% in many low- and middle-income countries, there are a substantial number of vulnerable people not represented in CDR data. Furthermore, people move without making calls, and therefore a lot of movements are missed on a small scale, this is a particular issues in areas with low density of tower. However, on an aggregate level, CDR data correlates with general population movements and densities. Research is underway to better understand the relationship of SIMs and people, and methods are being developed to, for example, separate movement of soldiers from citizens in conflict-affected areas. Access to CDR data requires lengthy negotiations and collaborative agreements with mobile operators, and substantial infrastructure to move large quantities of anonymized, aggregated data securely.

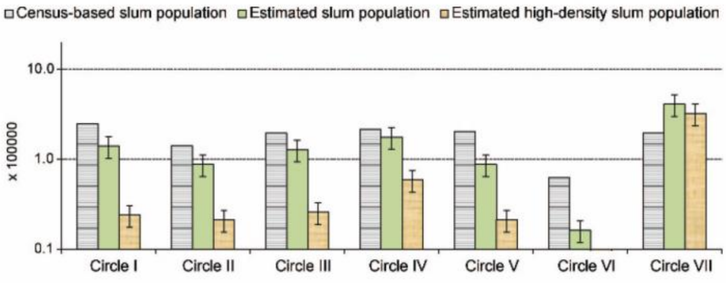
Figure 7: Estimating population movements using CDR

Discussion
Validation of WorldPop data in urban contexts? Many validation studies have shown that WorldPop is often more accurate than other data sets. However, there may be variabilities in an urban contexts. Planned upcoming validation studies will be paying close attention to this.

Population estimation using secondary data (University of Southampton/GeoData)

Once slum areas are identified through satellite-imagery and the algorithm of *lacunarity*, it is possible to estimate the population living in these areas using secondary data. The population number within a slum can be computed using local surveys and other secondary data. In the case of the study in Hyderabad, a partner was based directly in the field conducting surveys and population estimates, which were used as secondary data. To verify calculations, a comparison is done between the calculated estimate population number and the census estimate population number. One of the main findings of this study relevant to this issue is that 36% of the city's population is actually living in slums as opposed to 14% claimed by official statistics.

Figure 8: Improvement of slum population aggregates

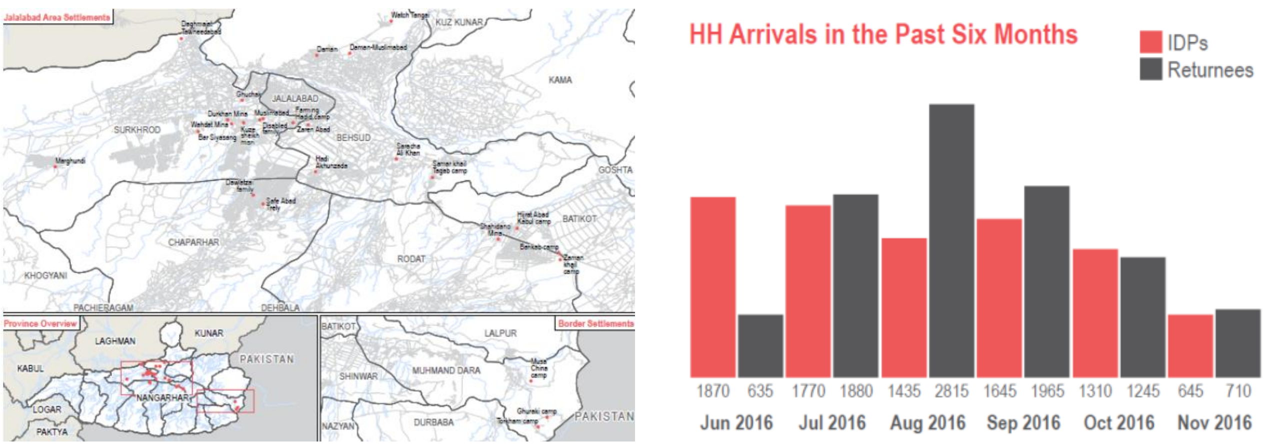


Discussion
How to minimize the margin of error? One option would be to combine this approach with gridded sampling to improve the correctness of population estimates in the slum.

Identification and trend monitoring of vulnerable urban territories (REACH)

The rationale behind this solution is that informal settlements in urban and peri-urban areas in low and middle-income countries are likely to host the most vulnerable urban populations. Along with the local/host communities, they are also often the recipients of IDPs, Refugees and Returnees (both documented and undocumented). Although informal settlements tend to provide free or cheap low quality accommodation, and informal livelihood opportunities, access to minimum quality basic services is often a challenge. **Monitoring vulnerable territories in urban areas can inform about the size and characteristics of the population moving in and out of those areas**. Identifying trends in population movements is done through a qualitative approach, which allows a "quick and dirty" estimation of population groups' composition and vulnerabilities. From the moment informal settlements are defined, it is possible to identify the stakeholders that live there and who have a good knowledge of the community. Regular interactions and meetings with these key informants every month can give an accurate overview of the demographic evolution in that area. This approach was used to track informal settlements in Kabul and Nangharar districts (Afghanistan) where it help inform food security assessment design in the most vulnerable urban areas and identify the most likely vulnerable population.

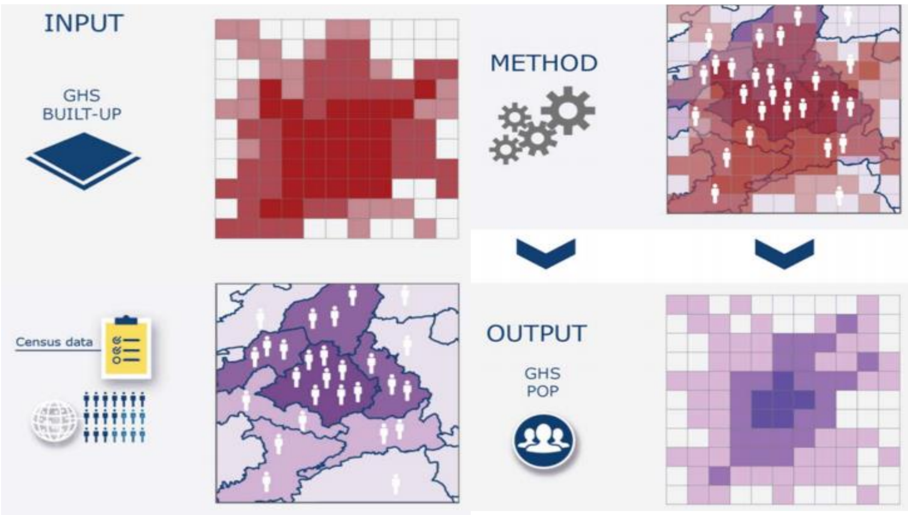
Figure 9: Map of informal settlements and population movement data in Kabul, Afghanistan



The Global Human Settlement Layer (JRC)

The Global Human Settlement Layer (GHSL) is a new open and free tool for assessing human presence on the planet. Satellite imagery helps define a built-up area (or cell) expressing the proportion of building footprint. Such information is then combined with population censuses-related data to build a population estimate into grid-cells of one Km of resolution. The combined information results into a new layer, which disregards administrative boundaries and shows presence, density and absolute number of inhabitants by cell. GHSL can be used for population estimates where and when census data is infrequent or misleading. Moreover, it can be used where irregular/non-admin shaped boundaries are available (e.g. livelihood zones). Finally, GHSL can help tracking population movements over space and over time.

Figure 10: From built-up area to population grid



4. Key learnings

While it is commonly recognized that there are vast challenges to assess and respond to urban insecurity, the next steps of the *Adapting to an Urban World* initiative include identifying potential solutions to these challenges. Some of the key learnings in this specific area are summarized below, which will guide the way forward:

KEY LEARNINGS & OPPORTUNITIES

1. A combination of ground-based and satellite tools can help understand rapidly evolving and complex urban settings. These will be extremely useful to estimate population numbers and to define the most appropriate sampling design, including the identification of vulnerable neighbourhoods in urban settings.

2. Gridded sampling can be a solution to have a sound sampling design when boundaries and population estimates are outdated.

3. The volunteers' approach promoted through OpenStreetMap is key to visualize a more granular picture of the reality on the ground in terms of population distribution and other dimensions. However, practitioners must be wary of any bias or interference due to the sensitivity of data collected.

4. A continuous interaction between technical solutions and more qualitative programmatic approaches must go hand in hand to overcome the challenges encountered in the urban space.

5. As all innovations are part of the solutions, bringing them together and combing them smartly (e.g. high resolution mapping, grid sampling, CDR, OpenStreetMap), will enhance the usefulness and precision of the tools applied.

6. Successful pilots and good communication can help to advocate solutions involving national institutions from the onset.

Contacts**:**

claudia.ahpoe@wfp.org
sergio.regi@wfp.org
marina.angeloni@wfp.org