

### **FLOWMINDER.ORG**

Non-profit foundation working with data providers and international/government agencies to operationalize and scale applications in support of vulnerable populations and sustainable development.













### Key partners and donors:













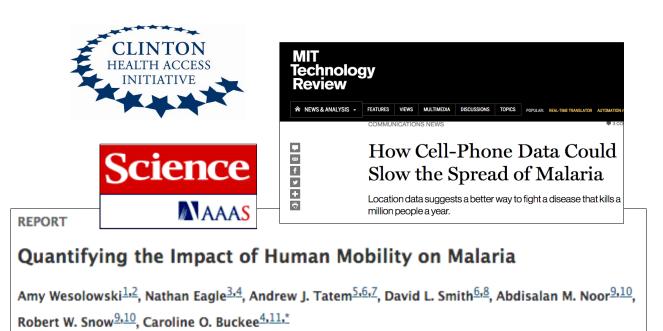


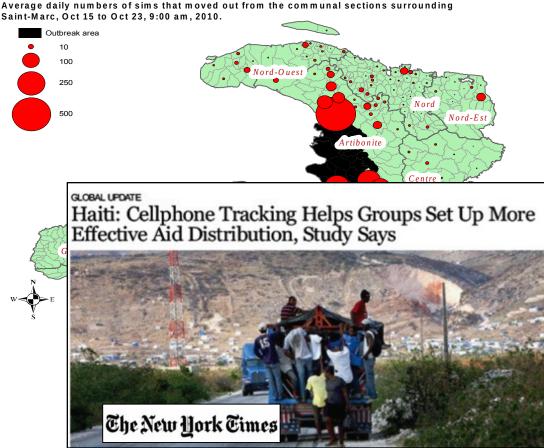






# Pioneered Anonymized Mobile Network Data for Infectious Disease: (2008 Zanzibar, Kenya, 2012 Haiti, 2013- Namibia, Indonesia) and Crisis Response (Haiti 2010 earthquake and cholera. Nepal 2015)







### WorldPop Project: Improving the spatial demographic evidence base for low and middle income countries

- We <u>develop scalable methods and models</u> for integrating ancillary datasources to complement and fill data gaps in census
- We <u>integrate new technologies</u>, including high resolution satellite imagery and cellphone data
- We <u>run training</u> for government agencies in these methods
- We publish fully documented, <u>peer-reviewed methods</u> and make outputs open access

### **Partners:**









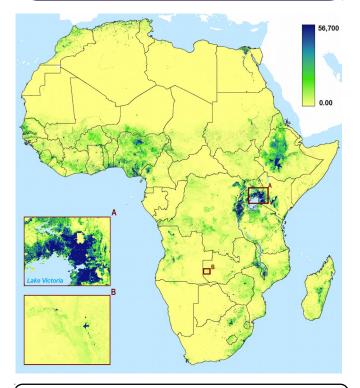






### Population characteristics

### Population dynamics









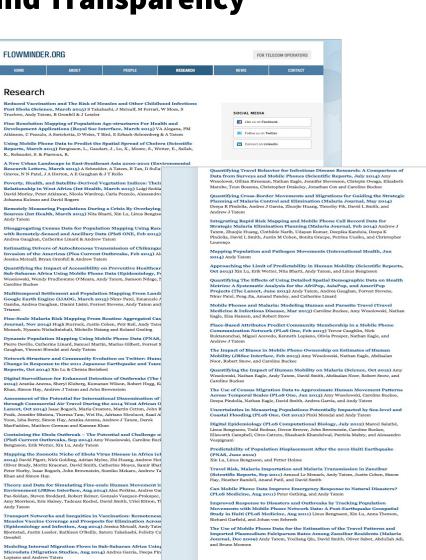


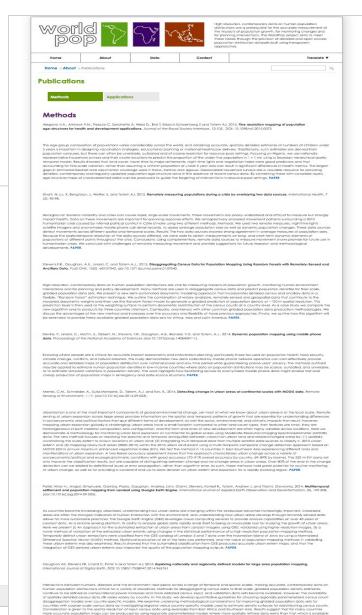


### All Methods are Open and Published in Peer Reviewed Journals for Validation

and Transparency

Quantifying Travel Behavior for Infectious Disease Research: A
Data from Surveys and Mobile Phones (Scientific Reports, July
Wesolowsi, Gillian Streaman, Nathan Eale, Jennifer Stevenson, Chrisbin Owaga, Elizabeth





and Andrew J. Tohem, 2014, Modeling Internal migration flows in sub-Saharan Africa using census

as the appliamicality of infectious disease. Quantifying and modeling further importance one content of map diversible evidence design diversifients for disease contrappings, consorting development, solicit from 10 countries in sub-disease. Affice with additional sport of additionals to develop models for models from 10 countries in sub-disease. Affice with additional sport of additional sport with the models on explaint value of production sport with the models on explaint value of production sport with the models of additional sport of

ments. International Health (doi:10.1093/inthealth/hu006) 6, 5-1

sociated from each other, and any reaceally tool there been extensive control between peoples. Both me and people of modern troot less unjerencements, with mannon modelly increasing in high receiving receive extraor and the foundation of the second section of territies disease, and from compressly new sector-vection and me diseases there canny one being section. Purchappers and their vectors can now move body however, we may have access to the most destilled and comprehensive additional on numeral and to consider these threats. This principle representation on control of these diseases, with not and to consider these threats. This principle representation on controlled or the controlled threats the controlled threats and the controlled or the controlled threats the controlled threats and the controlled threats the controlled or the controlled threats the controlled threats and the controlled threats the controlled threats and the controlled threats the controlled t

in A. Matthews Z. 2014. Mapping for maternal and newborn health: the distributions of women of nal of Health Geographics, 13:2

New John Sooties in Box Income Countries Ind Deern or lawy priority in public Institti strate the IMPAC determined and interest the Impact of Imp

Modelling spatial patterns of urban growth in Africa. Applied Geography, 44, 23-32.

If divers, driving exceptionally right unton apparets in their half-order agentican socio-exception or unton experiment in a been granted out of the full right or 1 this good level with a residency control or who is developed in a residency appropriate or or intermediate social in come to selectify factors in the extension of the selection o

Abdisalan M., Gaughan, Andrea E., Gilbert, Marius and Linard, Catherine, 2013, Mille women of childheadan one Bur? Providing Health Matter, 11, 11, 11

on exposition in approaches to deviling health method to measure progress through method recognition in the management of the method of the progress of the method progress of the method progress of the method progress of the method of the deviling the method of the

Igie, Nathan, Smith, David L., Garcia, Andres J. and Tatem, Andrew J., 2013. The use of census cross temporal scales. PLoS CNE, 8. (1), e52971.

somewith the deletery of terrificate, and the greet of of Infection diseases. Nowever, I removing peoply profitsuity for own formed countries. The most wide variabilities are registrated road to his more insoftway and the countries are received to a countries, but do not cepture the shorter-scale pollment, valuationed of well contributed abouts enterting may be removed the countries of the former scale, but his has have a good patter and for family any processed to enterior involvments between countries in 2000 on or many an extreme countries are considered on the countries of the countries in 2000 on or company an extreme countries in 2000 on or company an extreme countries of the countries of the countries of the countries of the countries of what countries are considered on the countries of the countries of the countries of the countries of a countries of the countri

plano, Kenneth K., Prosper, Olivia, Eagle, Nathan and Tatem, Andrew J., 2013, Place-based attitudes atten network, PLoS CHE, 8, (2), e55057.

corrected stooks, a properly knewn can make, buth, abcase disease. Internation, and some inferenmental production for some incidental for active policy a government or of the local works often incognorate his entitludes of individual people. Such as gender, entirely or cale social reviews, where soon near regioners or population of individual or locations, such cale social reviews, where soon near regioners are position of individual or locations, such as unknown. We describe the political of individually in a mobile private community as unknown, we describe the political of individually in a mobile private community and production, subcryotions, distances to the recent allows on any works of the production, subcryotions, distances to the recent allows on any work of the and production, subcryotions, distances to the recent allows.

vsk, A., Smith, D.L., Buckee, C.O., Noor, A.M., Snow, R.W. and Tatem, A.J., 2012 Human movement data for malaria of life Journal, 11: 205.

and committee and to the searches it imprecision in many matrix enteriors (curring, sounding to sold countries and enteriors (ingest, societies) for many population measured printly printing for countries in a relative to the present and interiors offered that the view accommend by the enteriors countries for interiors and an application of the product of the printing of the prin

e-scale spatial population databases in infectious disease research. International Journal of Health Geographics 1

Insuline and speed of infectious diseases are becoming interestingly another and spiriticistics, while globed in an opposite. We in design possibilities of the of disease estimates have experient possibilities and the possibilities globed on which are offered board on outdated and coarse residention date. Moreover, a visitery of different memoral trave are personal coarse. In this involver we estimate the mining placed and possibilities and the service of the properties of

Noor, A.M. and Totem, A.J., 2012. Population distribution, settlement patterns and accessibility across Africa in 2010.

is and antiferrors comes a country and their research production and accessibly from valors means are imported to and an odd production multiplier and to approxy from a production of the country resource of an odd production multiplier and to approxy from a production and approximate to the country resource of an odd production of the country of the country of the country and the country and the country of a set of the country of a set of the country of the coun

pulation mapping of poor countries, Nature, 474, 36.

justable for querifying populations of risk, such as those near huckes power plants (halow e.Z., 40,401; 201) lest must be conversibled. The carea load value in map controlled for for countries or exceed not desired for exempts, Angorsi and cerus, wis in 1970, believe down into just 18 daths. Strates of its current total resident principles for the properties of the current of the current total resident principles for define groups. If on lead to substantify developer estimate of population destructions only in an artificial properties for the control of the current strategy of the current of the

W., Snow, R.W. and Linard, C., 2011, The effects of spottal population dataset choice on population at risk of disease

If the time disease distrutions and grammics is increasing large understance for mach services partning and disease distrutions. When size the throughpost in page of diseased control properties are the throughpost in page of diseased control properties of the disease of the d

the dataseth vorted by more from 10 million people for some countries, even occounting for the fact that et class are used to cameric national habita in these destinates and can very by more than 25 for many low-knowne implies comprised more than 10% of the trotal national population. The destinate country-level assessments sugges are accurate than the others in estimating MAL. The size of such affectioned among modeled insuran populations.

involved process related to quality and an experience of the process related to quality and p

M., Snow, R.W. and Tatem. A.J., 2010, A high resolution spatial population database of Samalia for disease risk

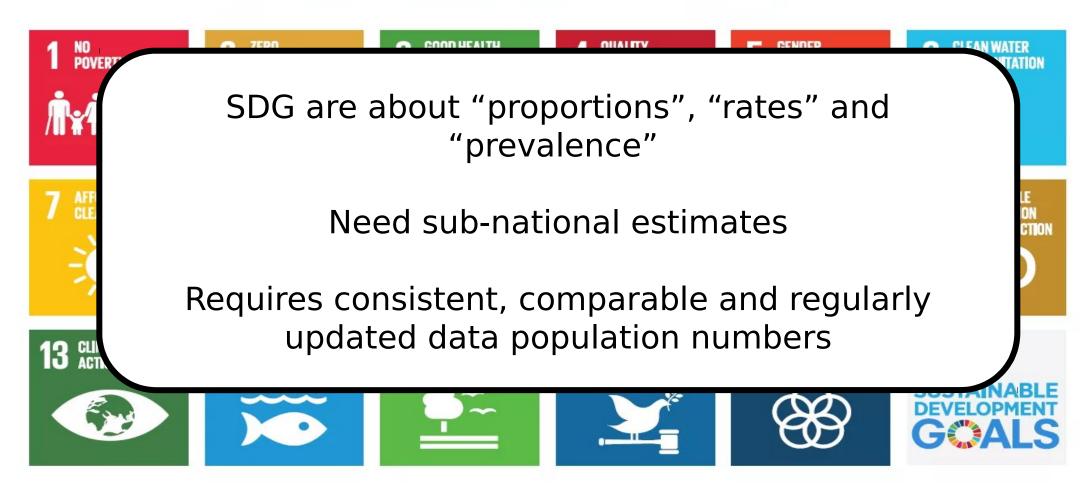
an deprine of boxic health enriches dun to the undebt particle direction of that country. Alternation are being for be stimulate the self-of-being direction of the self-of-being direction of the direction of the self-of-being population and the boxing, are of poor population distribution, in a law-income country such as Services, population data are boxing, and or poor another than the self-of-being direction of the self-of-being direction and self-of-being direction and the production of the self-of-being direction of the self-of-being direction of the self-of-being direction and production of the self-of-being direction of the self-of-b

2010. Assessing the use of global land cover data for guiding large area population distribution modelling

to an finding homeony use in a visit single of falls, including resource absorbing, disease burst entirelists on cover information control and und in control not will assist an information in the control to companie consists cover information control and cover to be improved across Africa, an afternative, control and companies control region. If gradee population and so the bit proposed across Africa, and anomalies, controlled will companie when the proposed for target under the controlled and cover detailed, controlled will companie and proposed for target and the controlled and cover detailed, controlled will controlled and the population distribution and in the effective controlled and the proposed for the controlled and the temporation distribution and in the effective controlled and the proposed for the controlled and the con

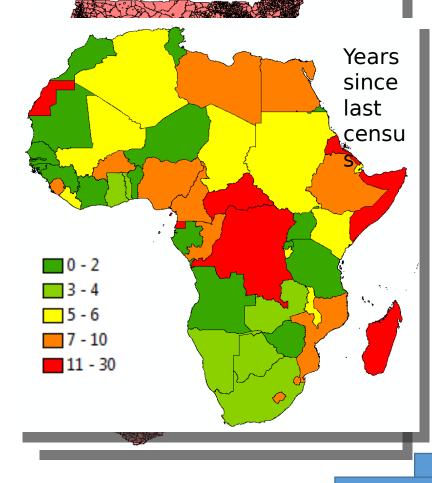


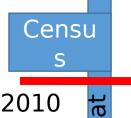




# The SDG Challenge: Measuring the Denominator

- National census data will continue to be our most important datasource
- Provides denominators and some numerators for the SDGs, and requisite subnational detail
- But, the 2015-2030 SDG period typically includes just one census datapoint
- And in some settings the situation is more challenging







2015



2020







2025



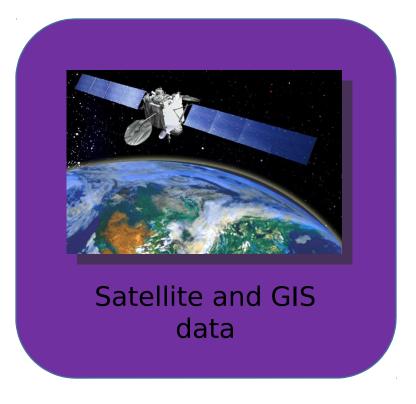




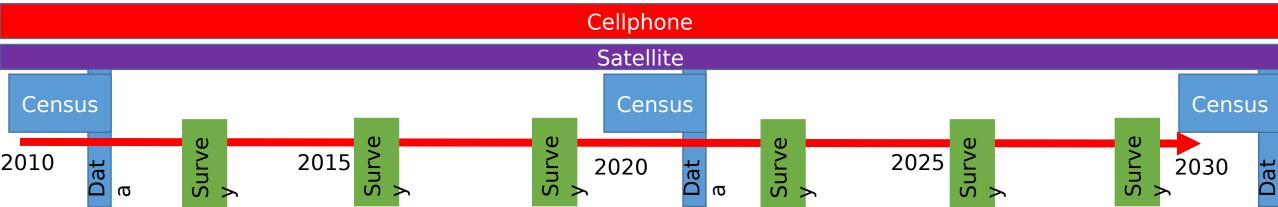
2030

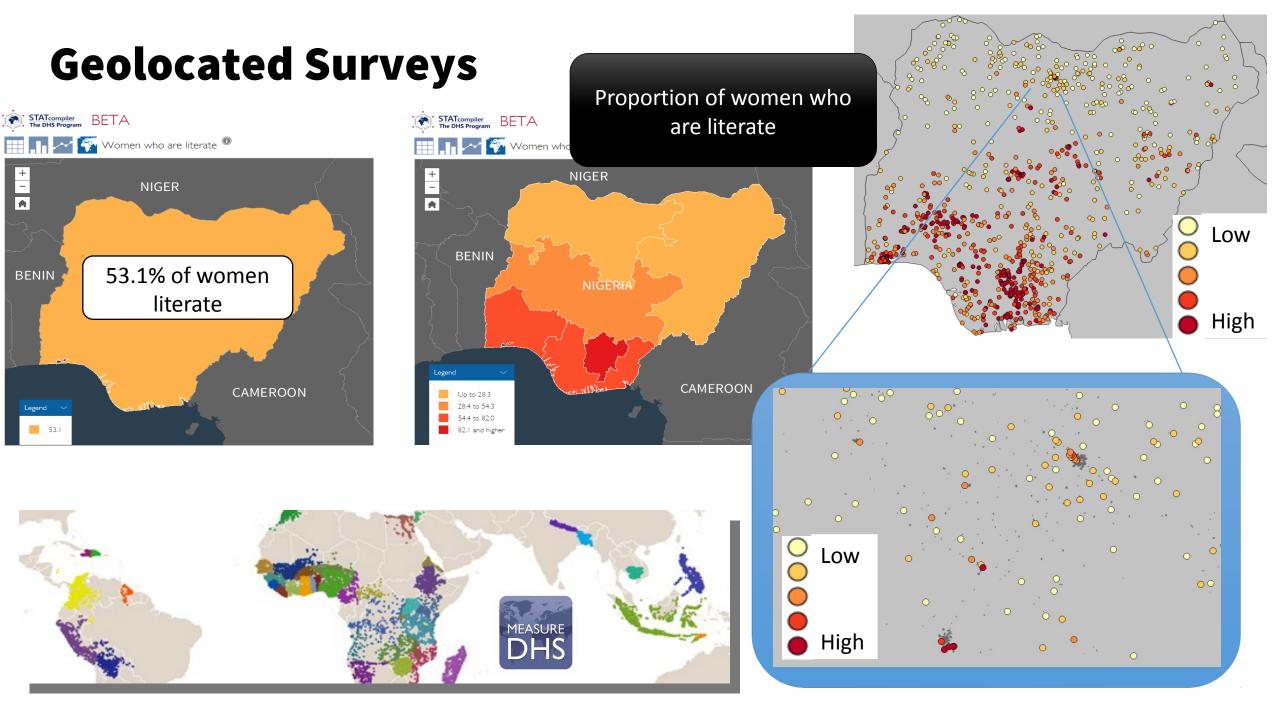
### What Do We Have to Help Us?



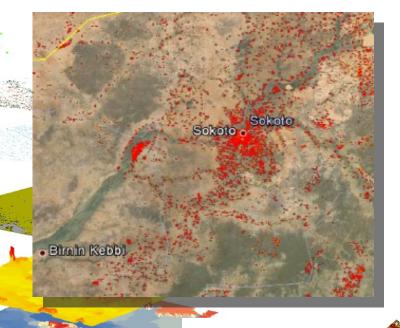


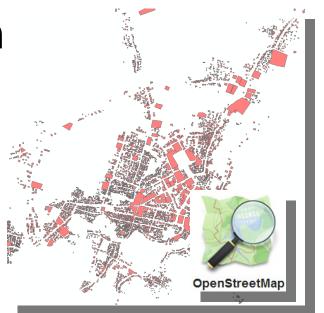


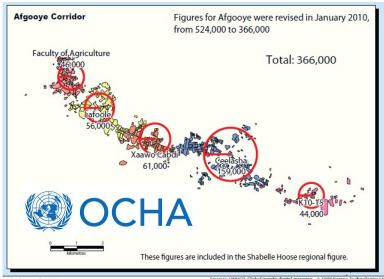




# **Satellite and GIS data**

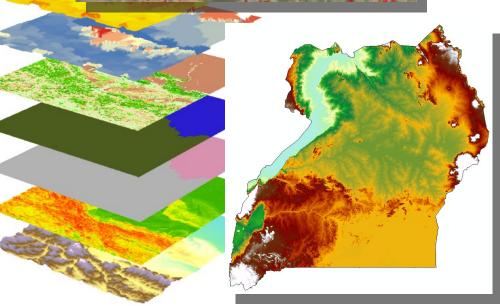


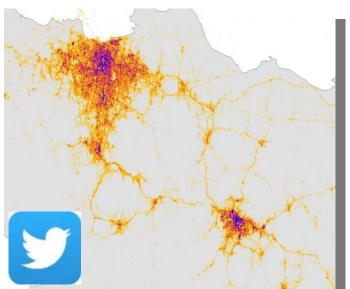


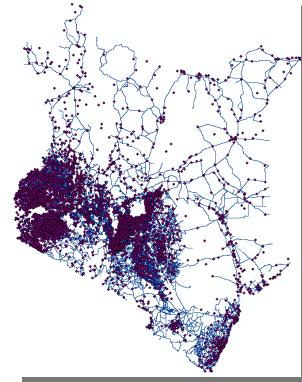


ce by the United Nations.

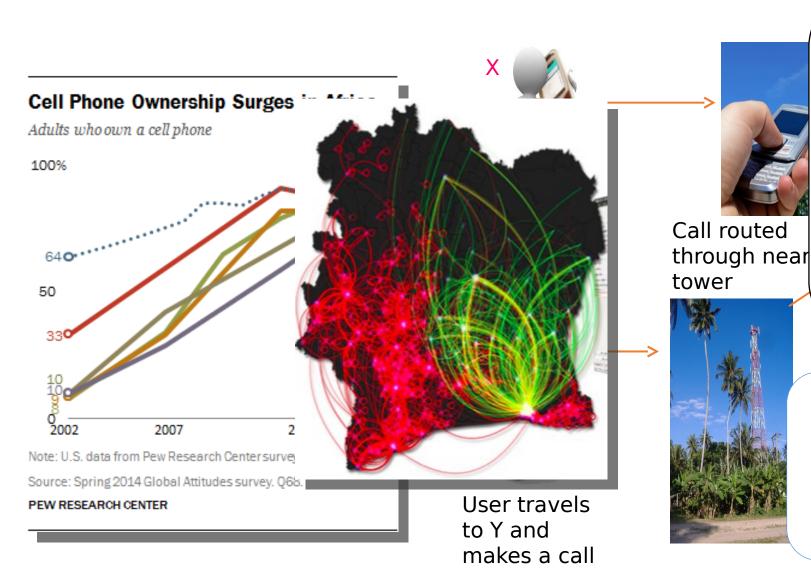
Sources: UNHCR, Global Insight digital mapping - © 1998 Europa Technologies I Statistics provided by UNHCR BO Somalia and OCHA Somalia.







# **Cellphone Call Detail Records**



### **Mobility:**

Changing densities, flows, seasonal/permanent migration

#### **Social networks:**

Number of contacts, calling patterns

#### **Consumption:**

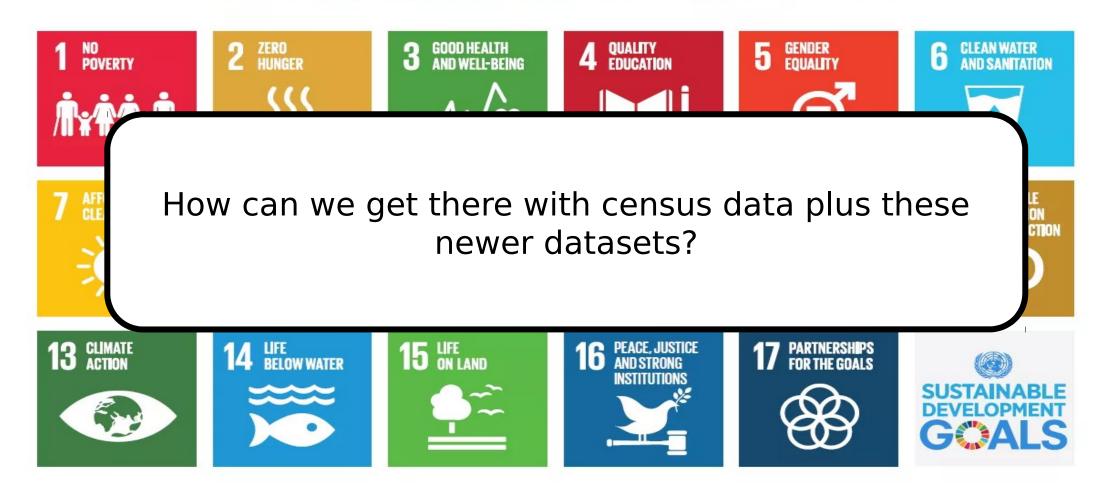
Credit purchase frequencies, top-up amounts

### **Protecting confidentiality**

-Aggregate summaries -Regulator approval -Raw data never leaves operator

distance (km)

# SUSTAINABLE GENALS





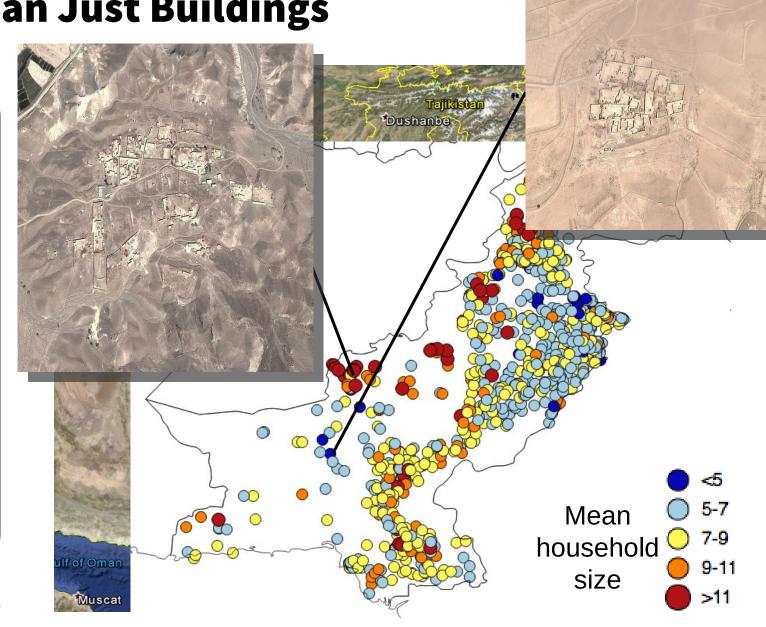
More Mation Than Just Buildings buildings/settlements

is a valuable initial step

But, buildings do not directly translate to people

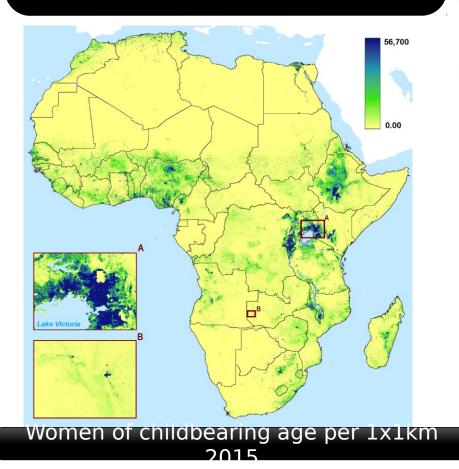
Multiple factors drive differences in distributions and demographics

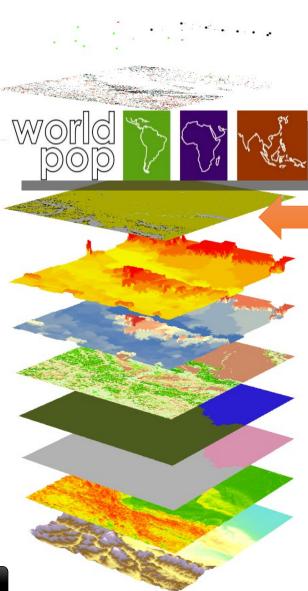
Importance of integrating multiple data sources

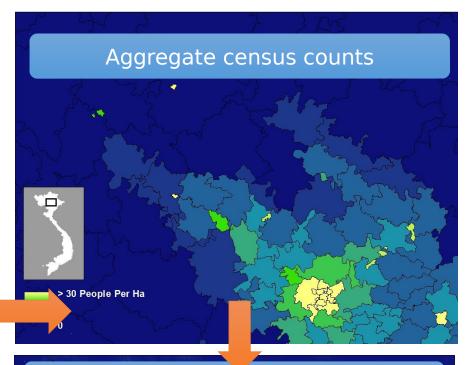


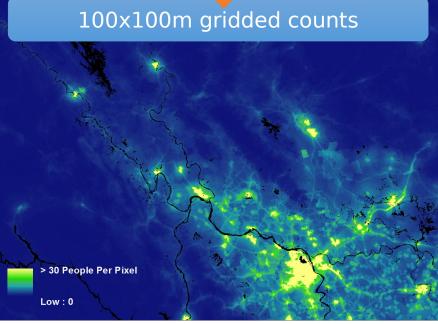
# **Census Data Disaggregation**

Integration with satellite/GIS data related to human population distribution patterns to disaggregate counts to regular grids

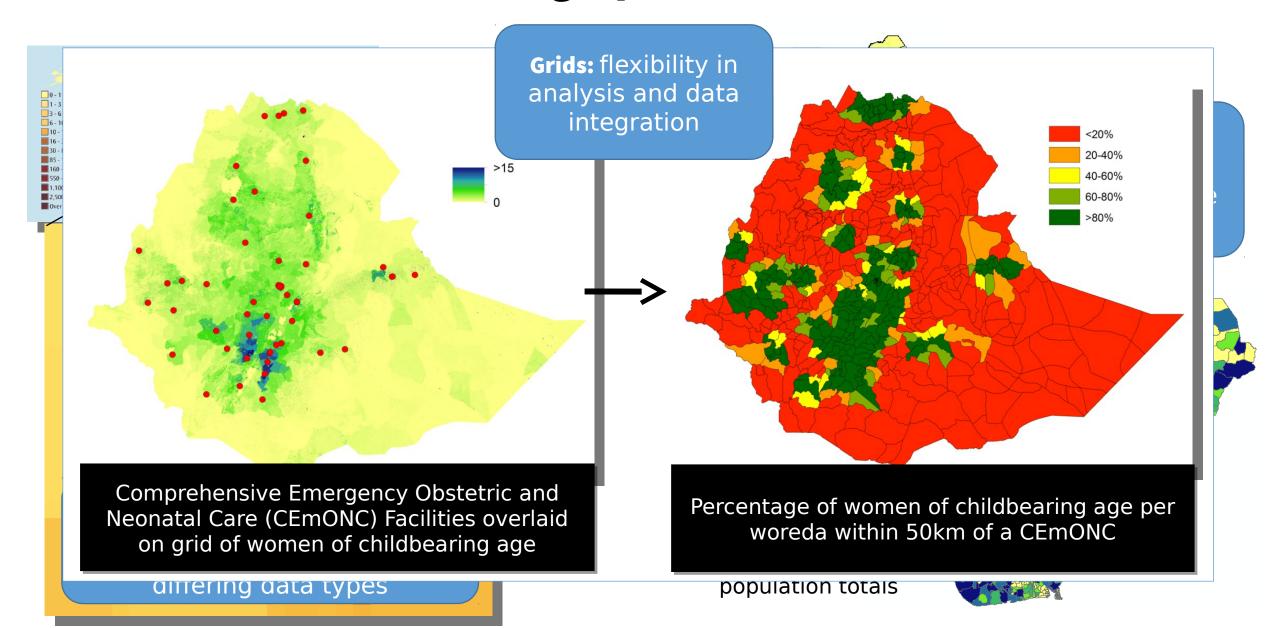






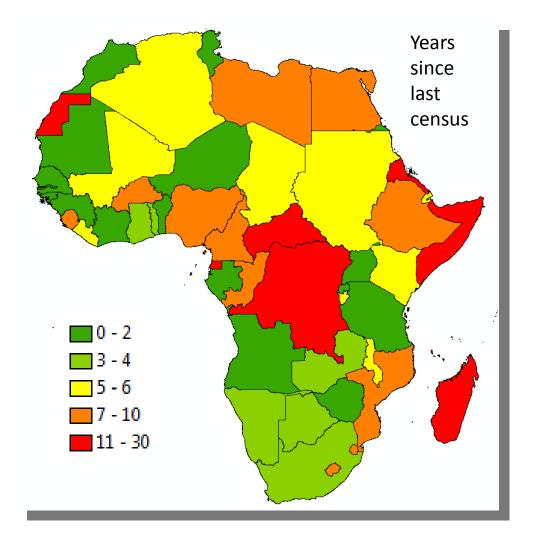


# **Benefits of 'Gridded' Demographic Data**

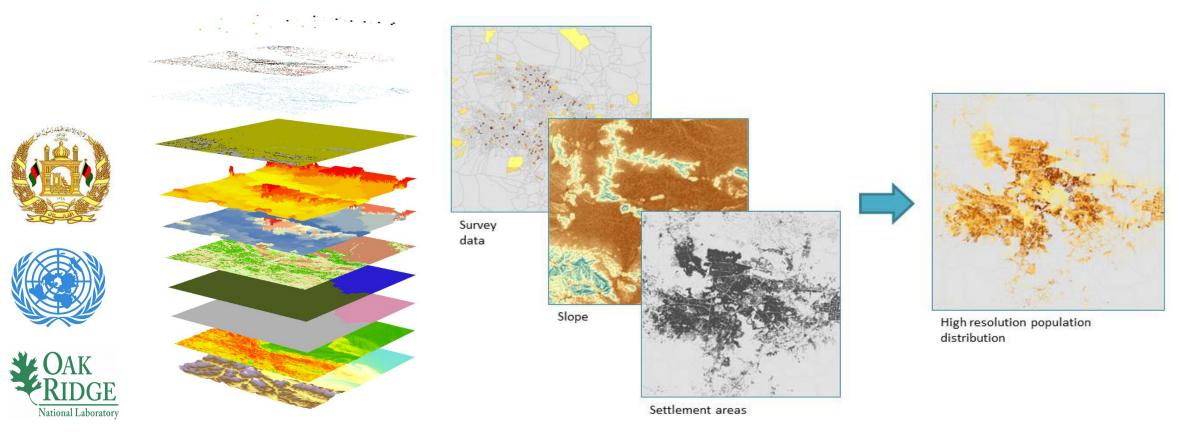


## **Census Data is not Always Trustworthy**

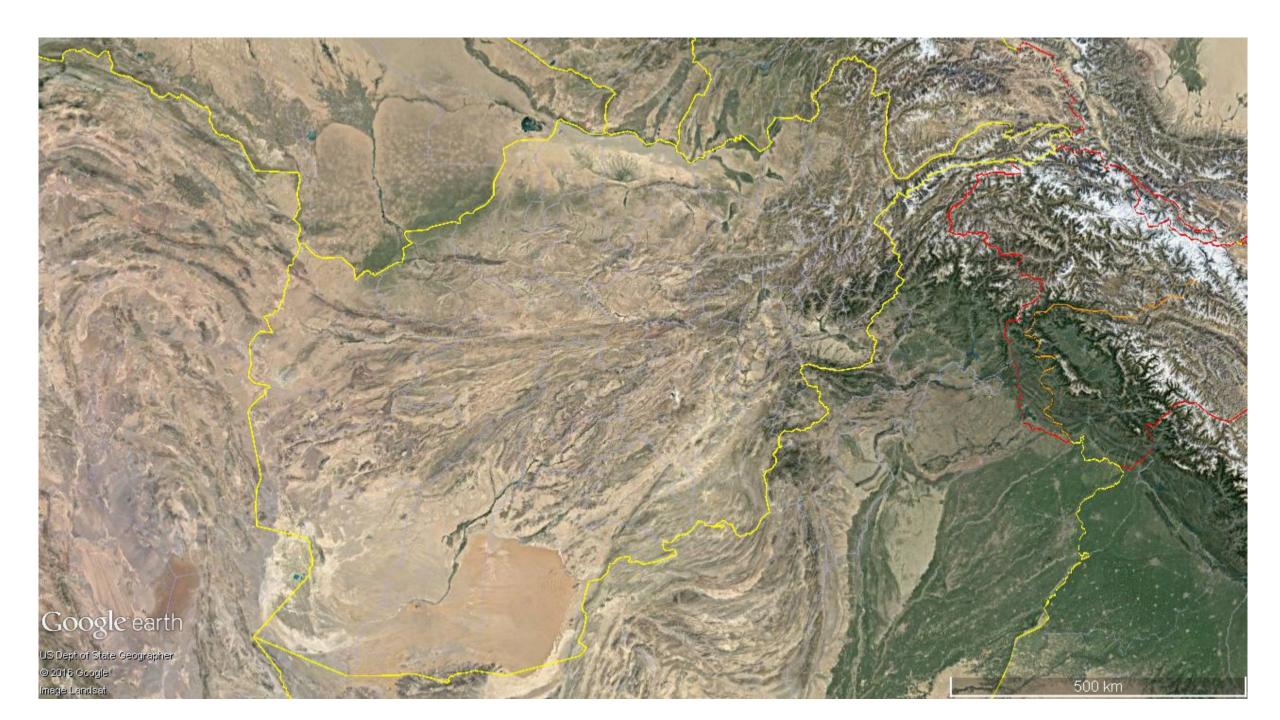
- Every ten years
- Released with delay
- Sometimes much older
- Sometimes manipulated

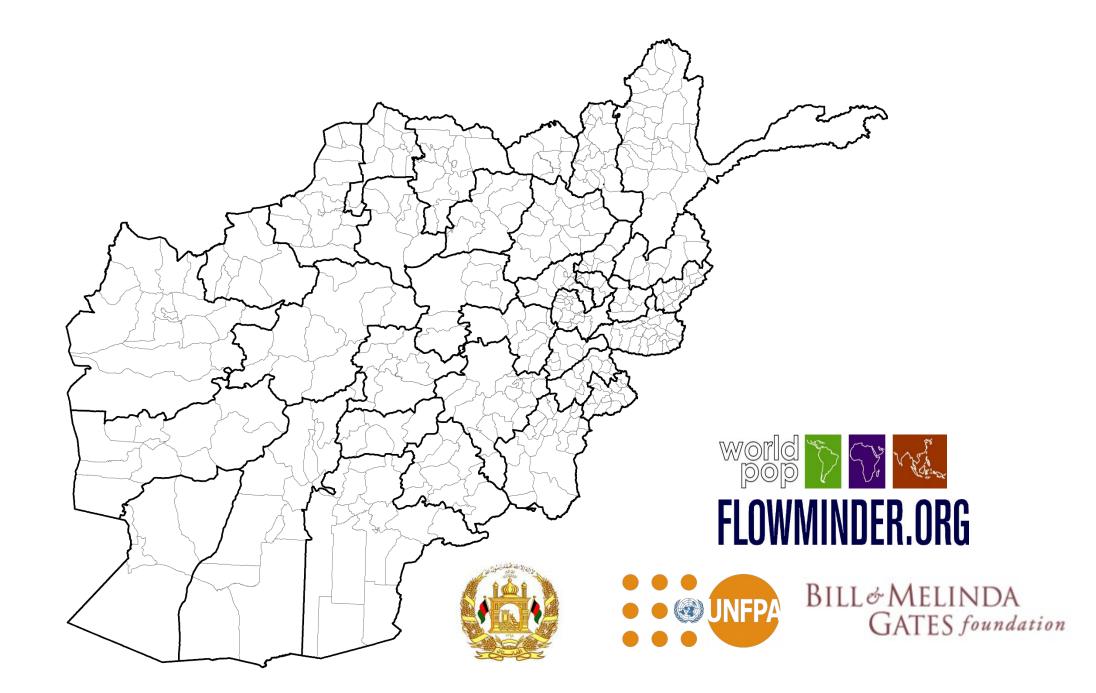


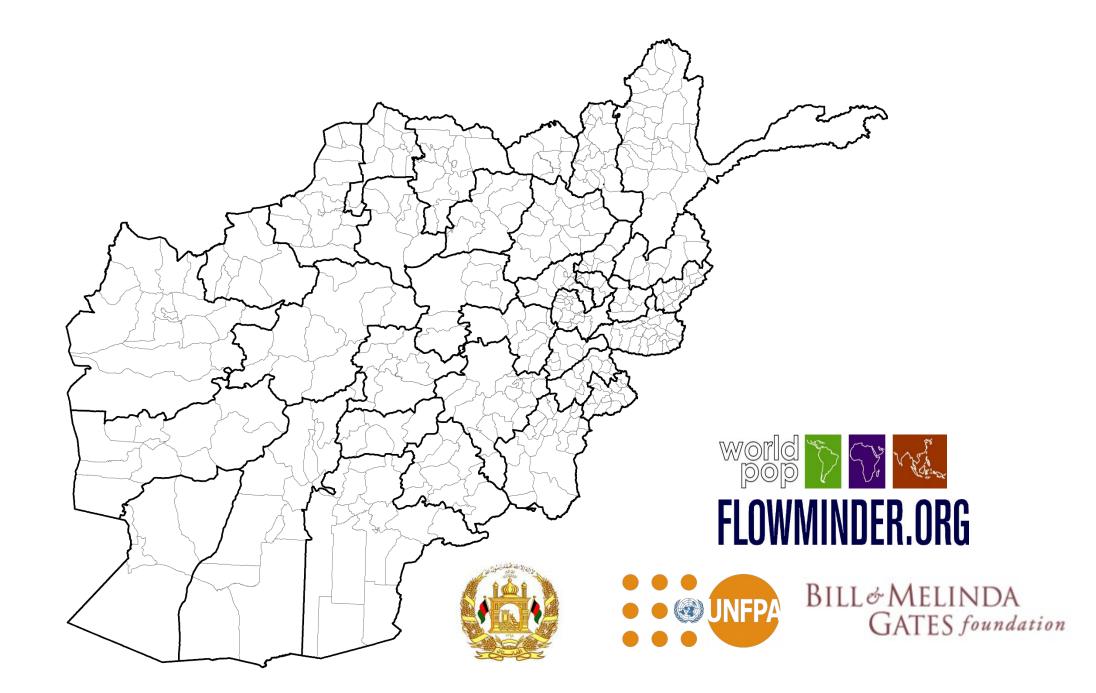
# **Bottom-up Mapping: Micro-Censuses in Selected Locations**

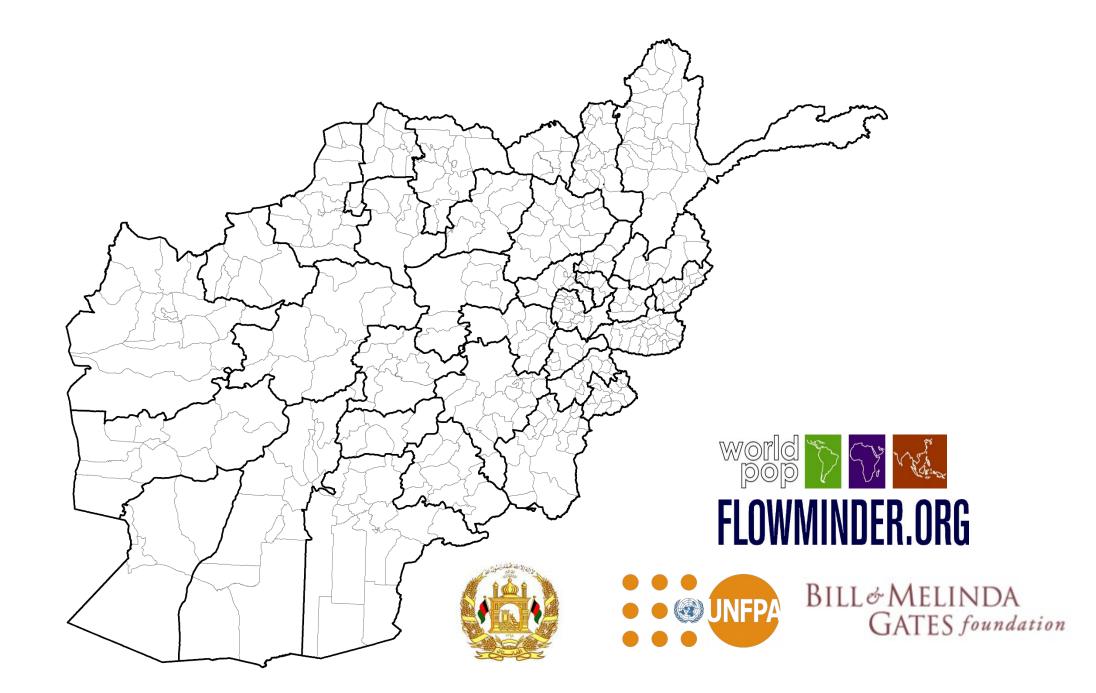


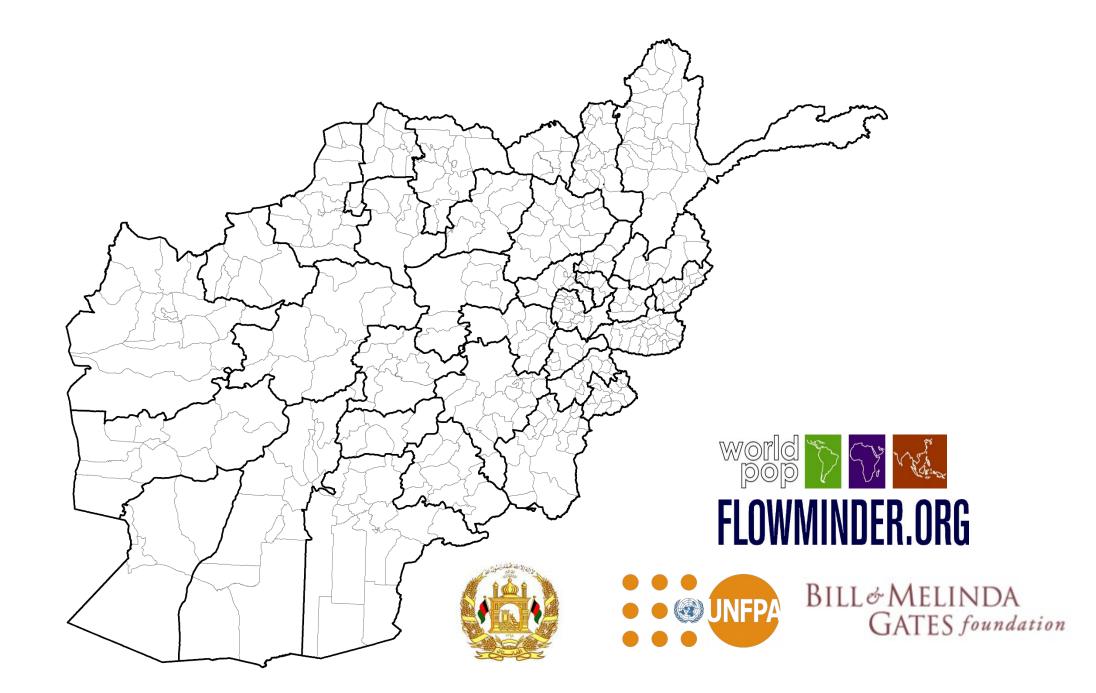


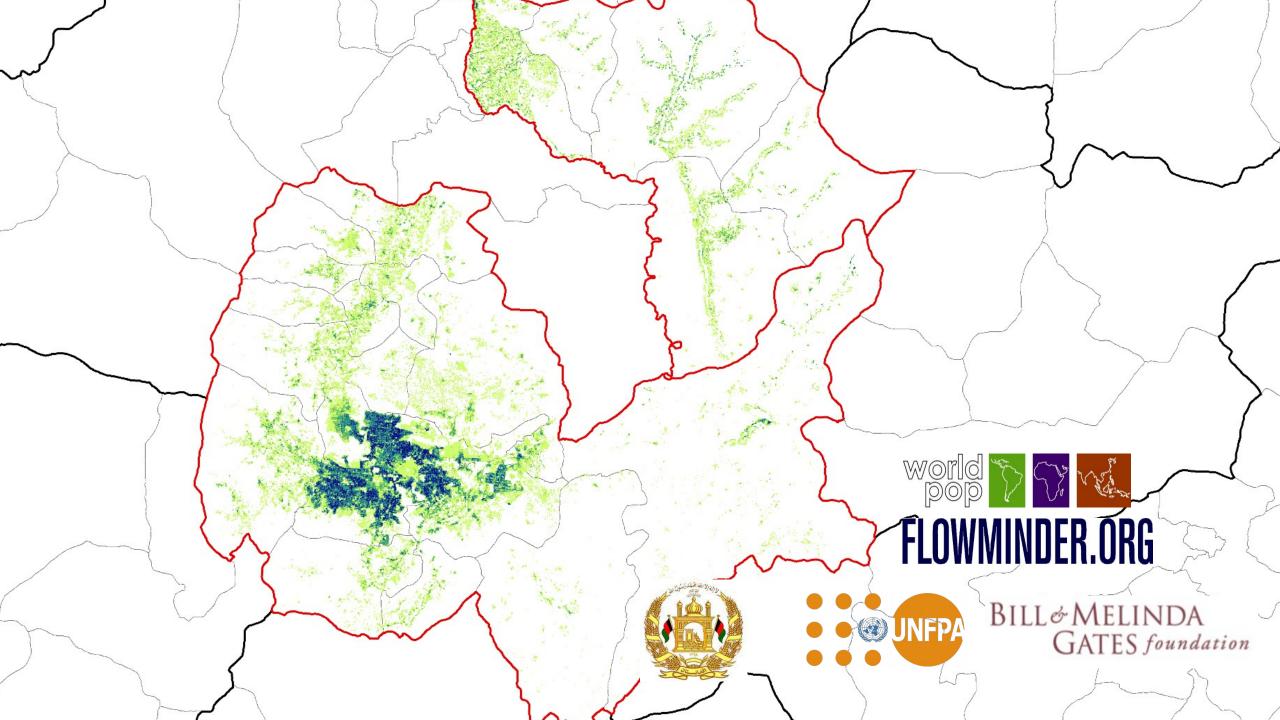


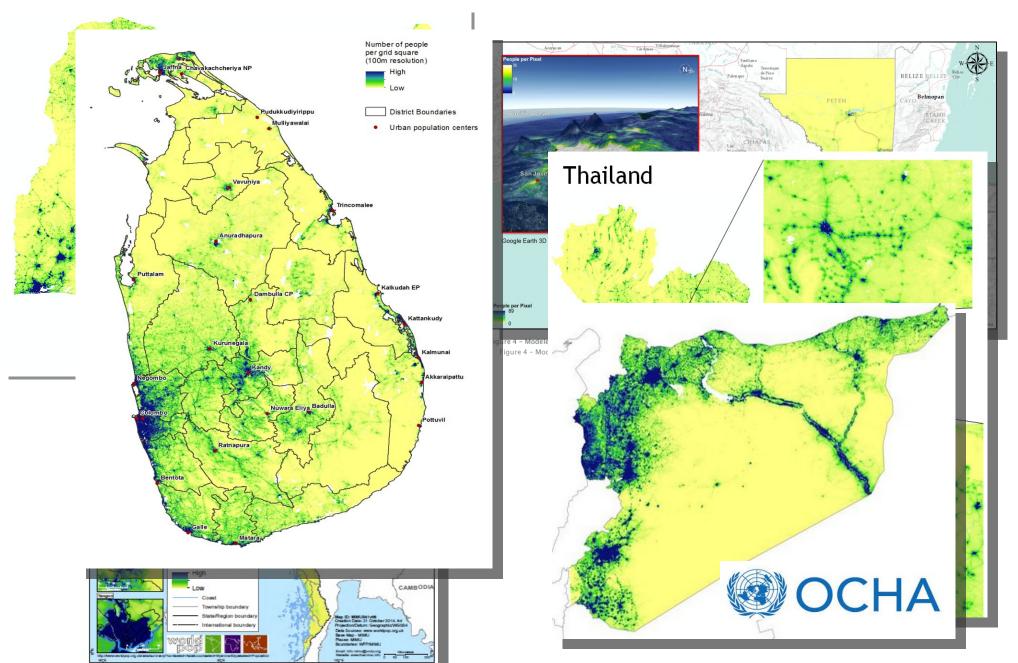






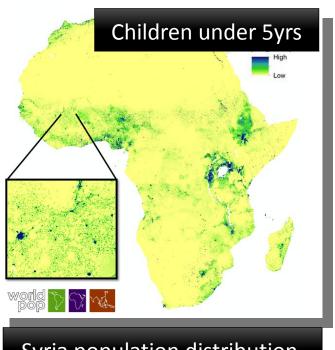


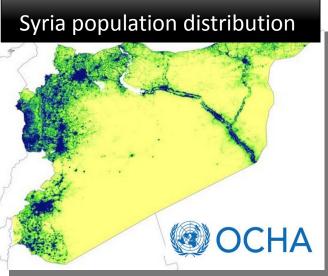




Refs: See www.worldpop.org

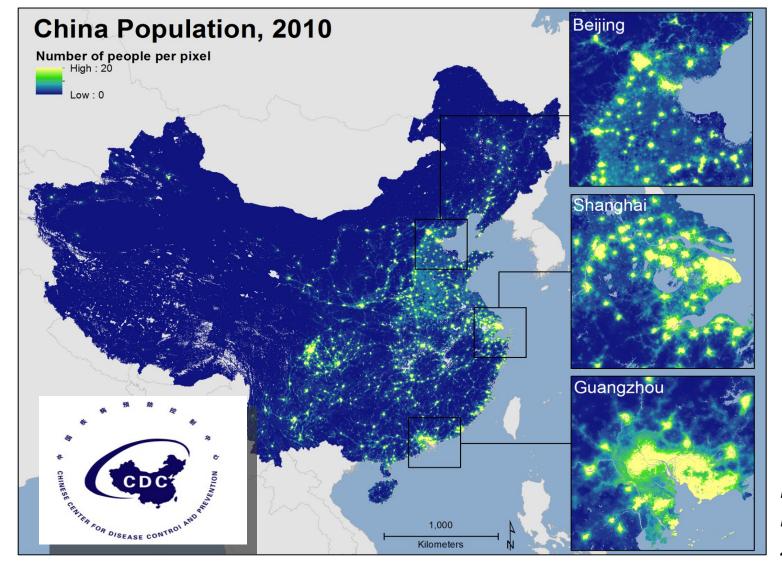
# Datasets, Methods in use Operationally





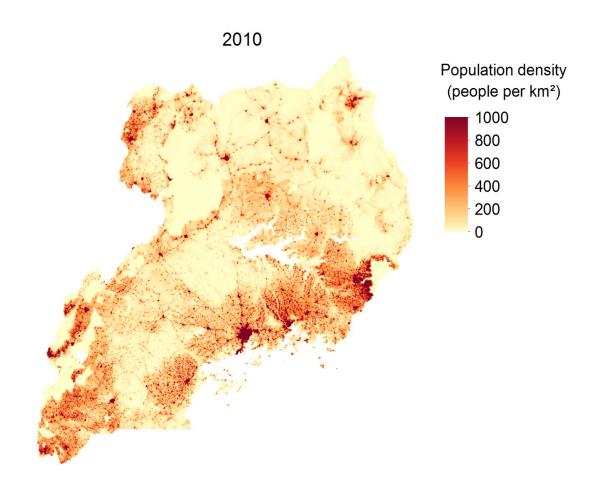


# **Temporal Change: Looking Back**



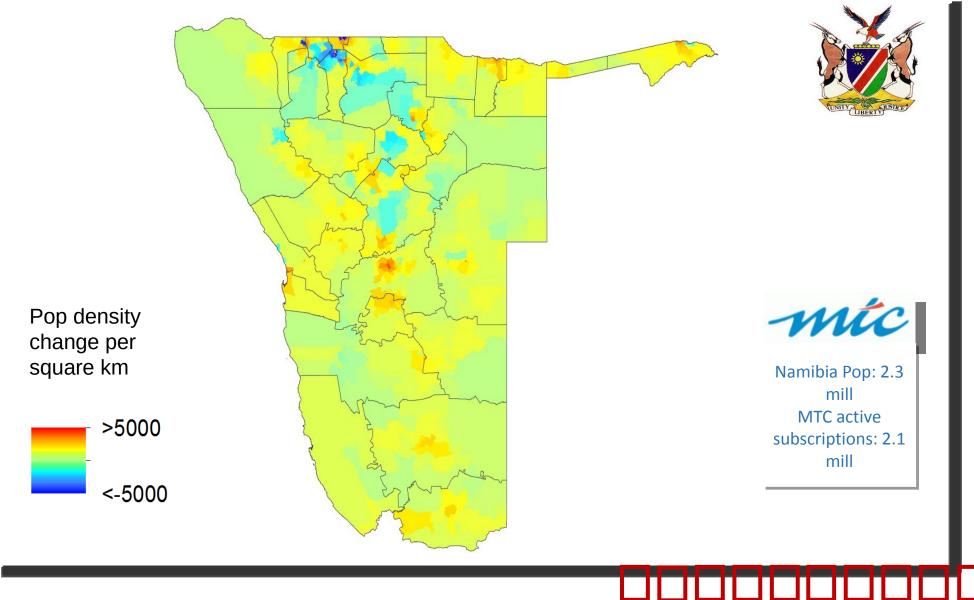
Ref: Gaughan et al. Nature Scientific Data. 2015

# **Temporal Change: Projecting Forward**



# **Populations Don't Stay Put...**







# **Measuring Targets: Population Characteristics and Coverages**

### **SDG** targets



1.1. Eradicate extreme poverty for all people everywhere



• 2.2. End all forms of malnutrition, including the internationally agreed targets on stunting and wasting in children under 5 years of age



• 3.7. Ensure universal access to sexual and reproductive health-care services

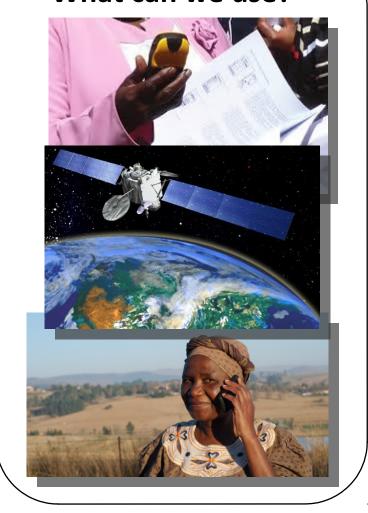


• 4.6. Ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy



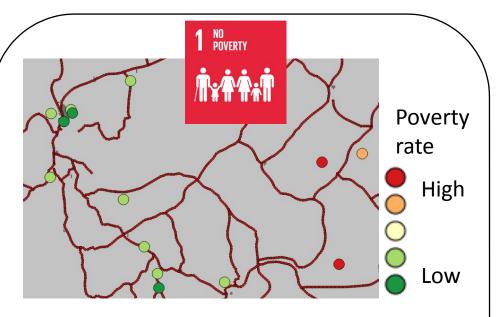
 6.2. Achieve access to adequate and equitable sanitation and hygiene for all

### What can we use?



# **Spatial Data Integration**

- Population characteristics measured in household surveys can be strongly related to features we can measure everywhere
- We can use these relationships to predict characteristics into unsampled locations using metrics from census, satellite and cellphone data to create maps of SDG-relevant indicators
- Importance of validation and the measurement and mapping of uncertainty

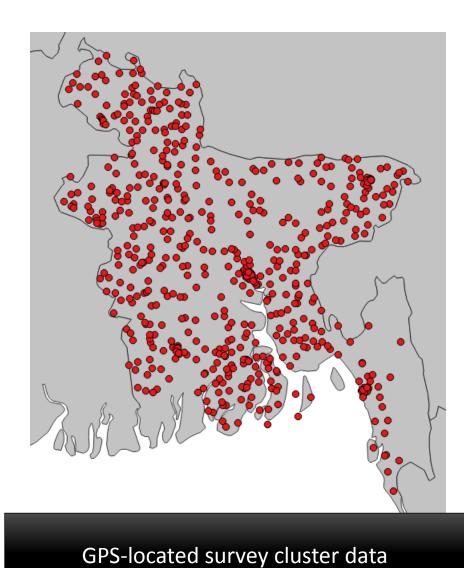


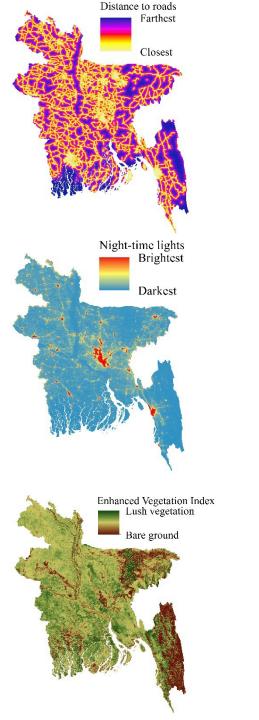
-Increasing distance from major roads =
 increasing poverty
 -Increasing urbanicity = decreasing
 poverty

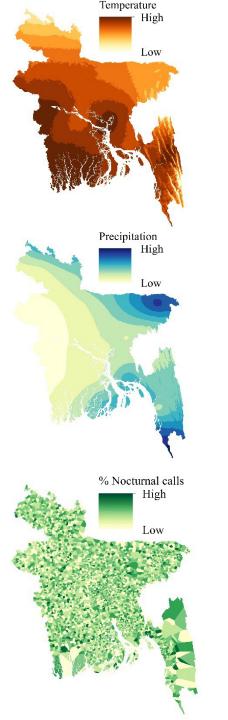


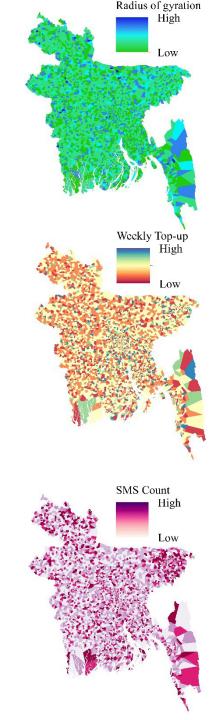
-Wider social network = lower poverty-Large, regular credit top up = lowerpoverty

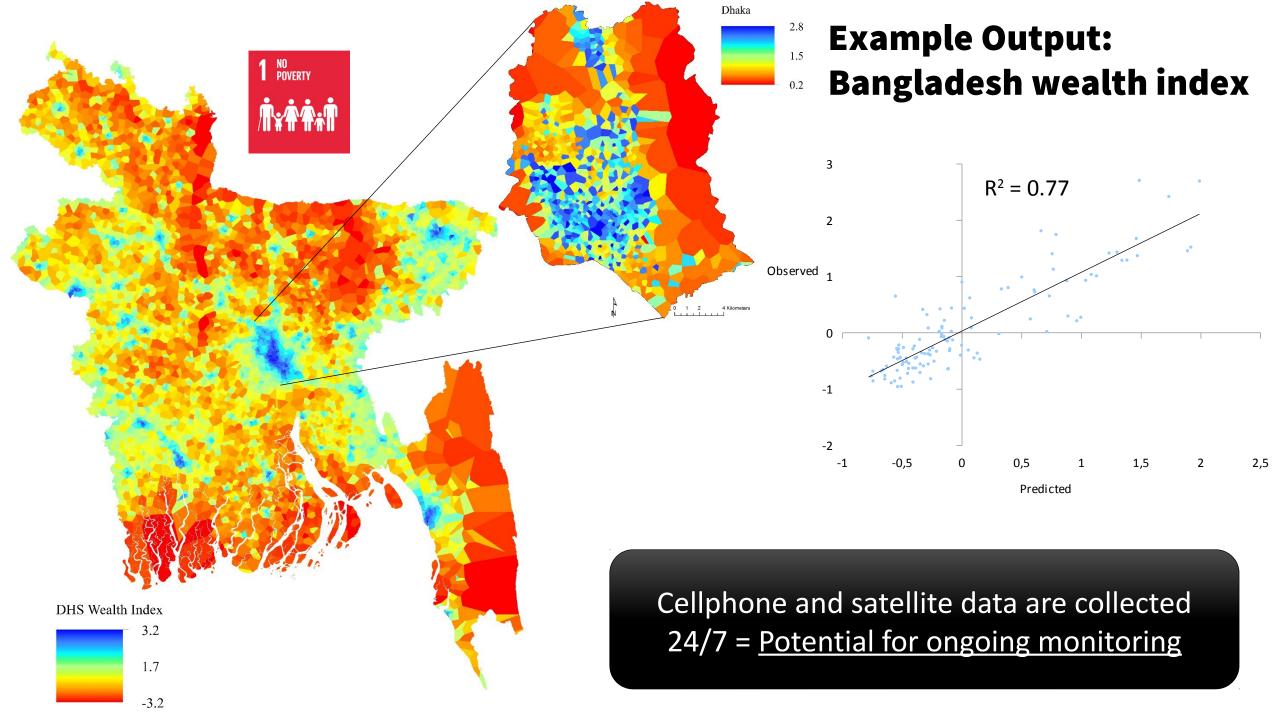
# **Improving Mapping of Socioeconomic Indicators**



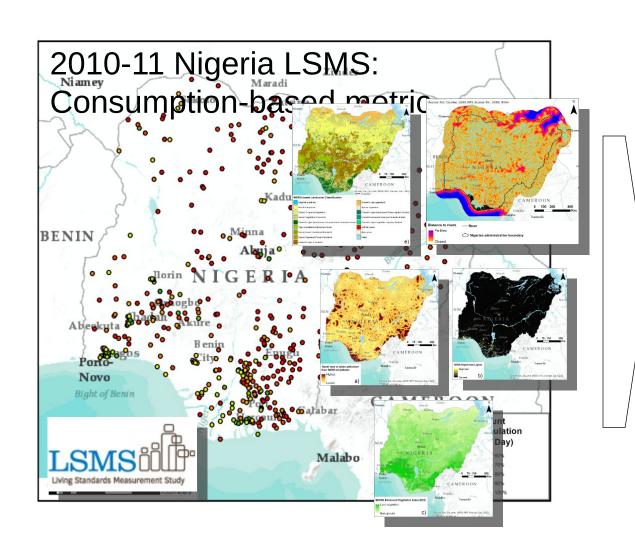


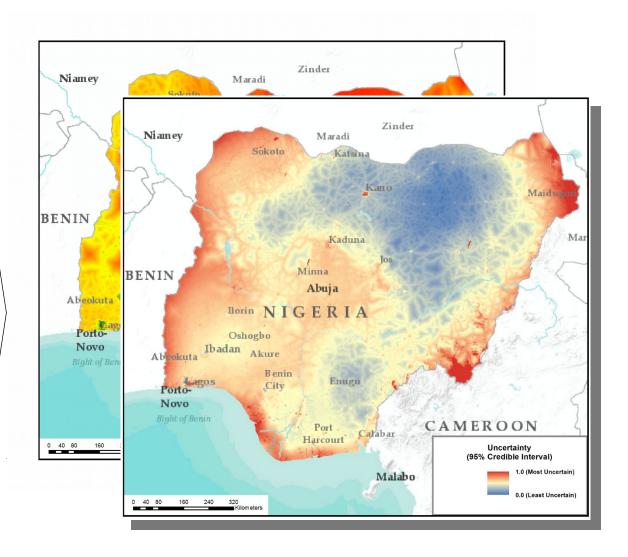




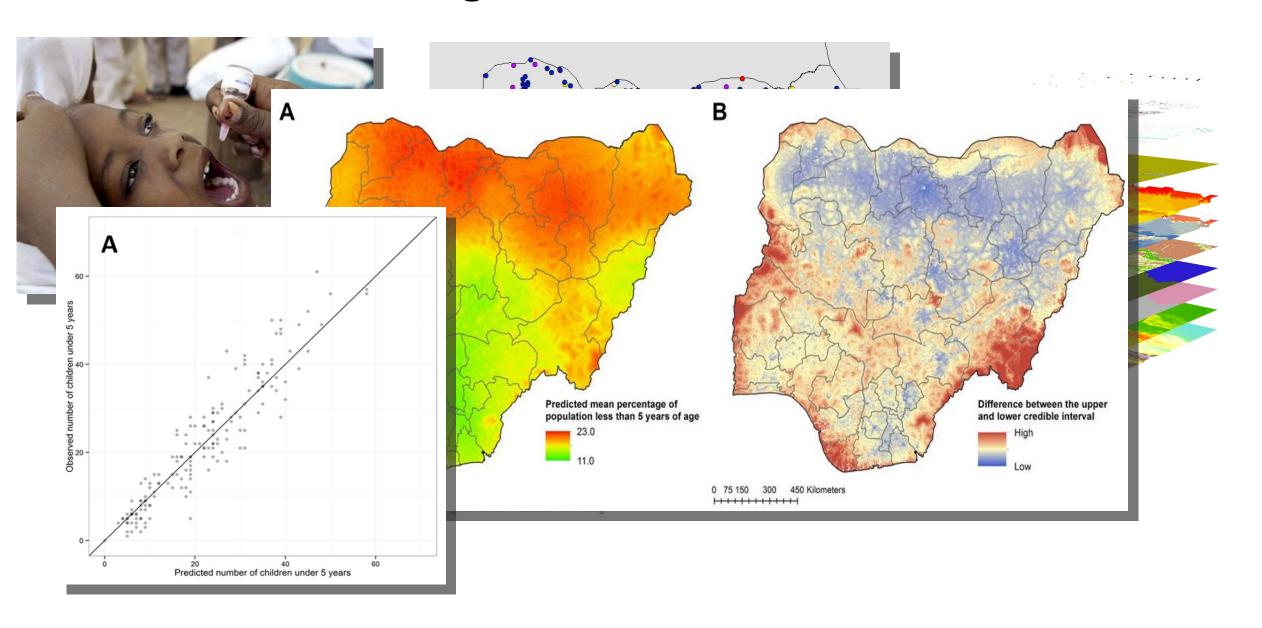


# **High-Resolution Poverty Maps**

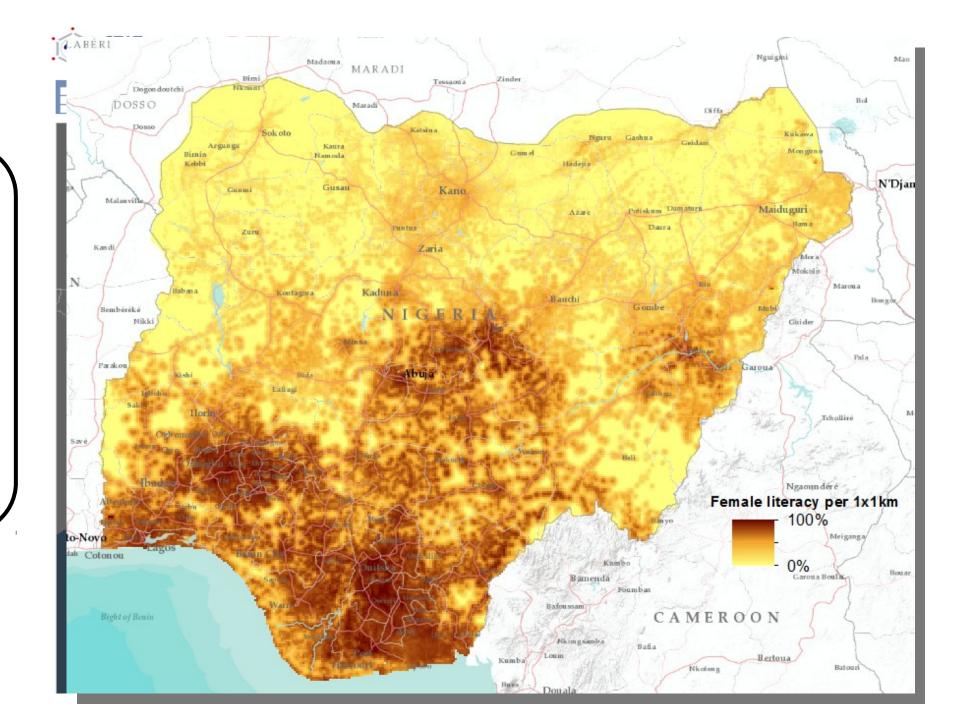




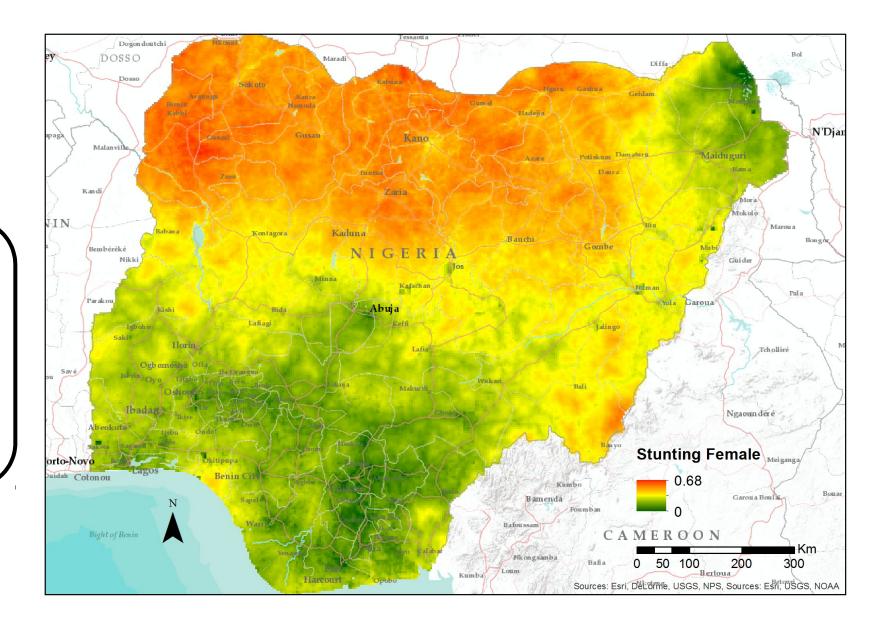
# Polio Vaccination in Nigeria: Where are the <5s?



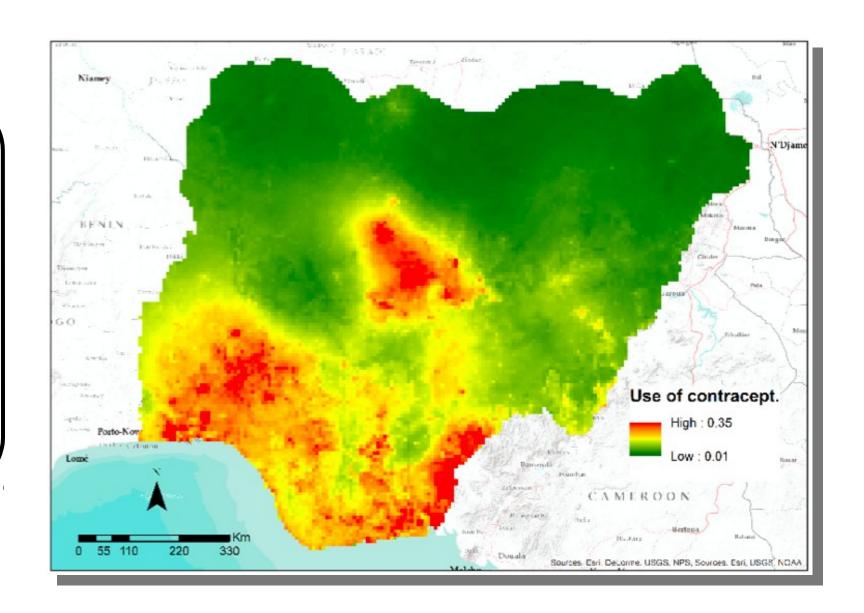
Nigeria:
Proportion
females
who are
literate



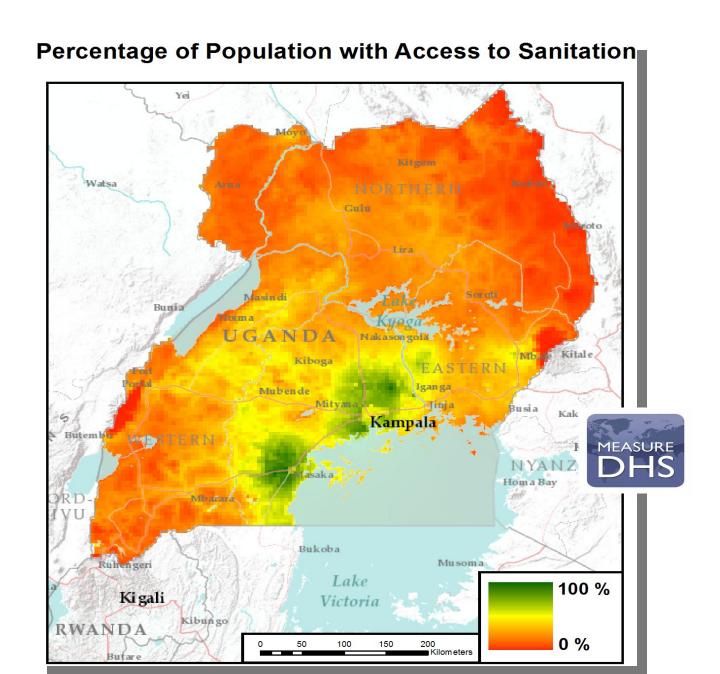
Stunting among girls Nigera



% women of reproductive age using modern contraceptive methods



Uganda: access to sanitation



# Mapping Displacement and Movements due to Climate Change and Disasters



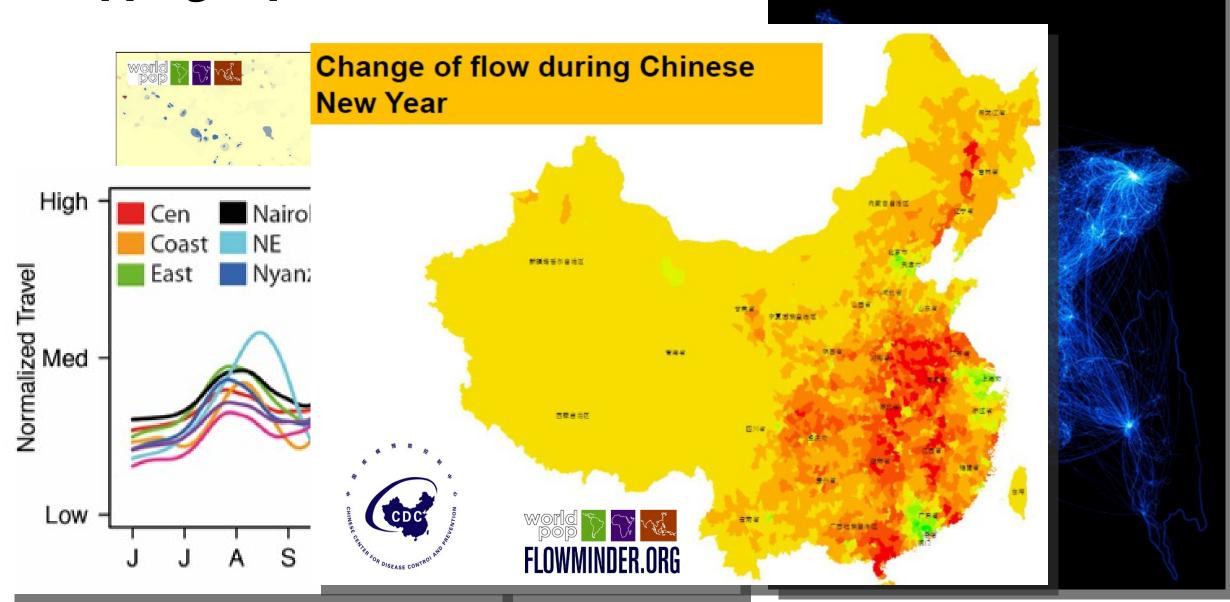




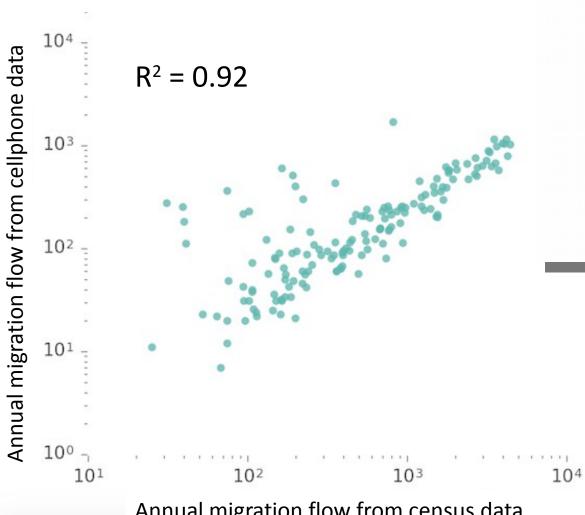




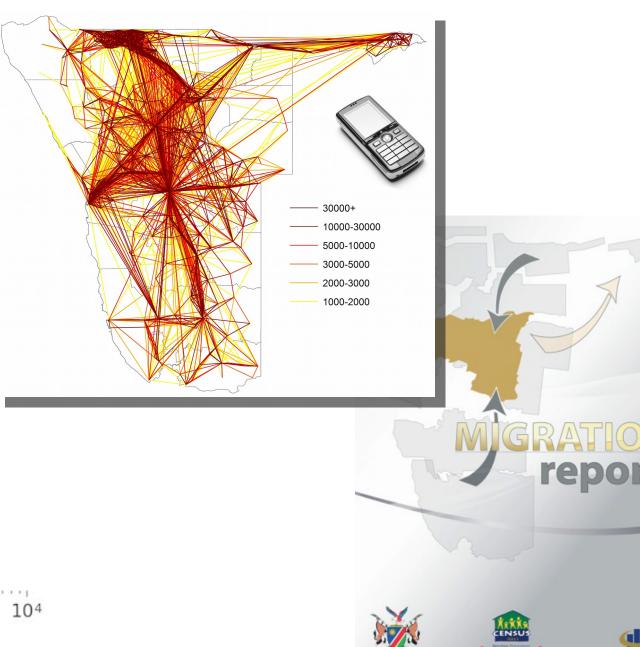
## **Mapping Population Movements**

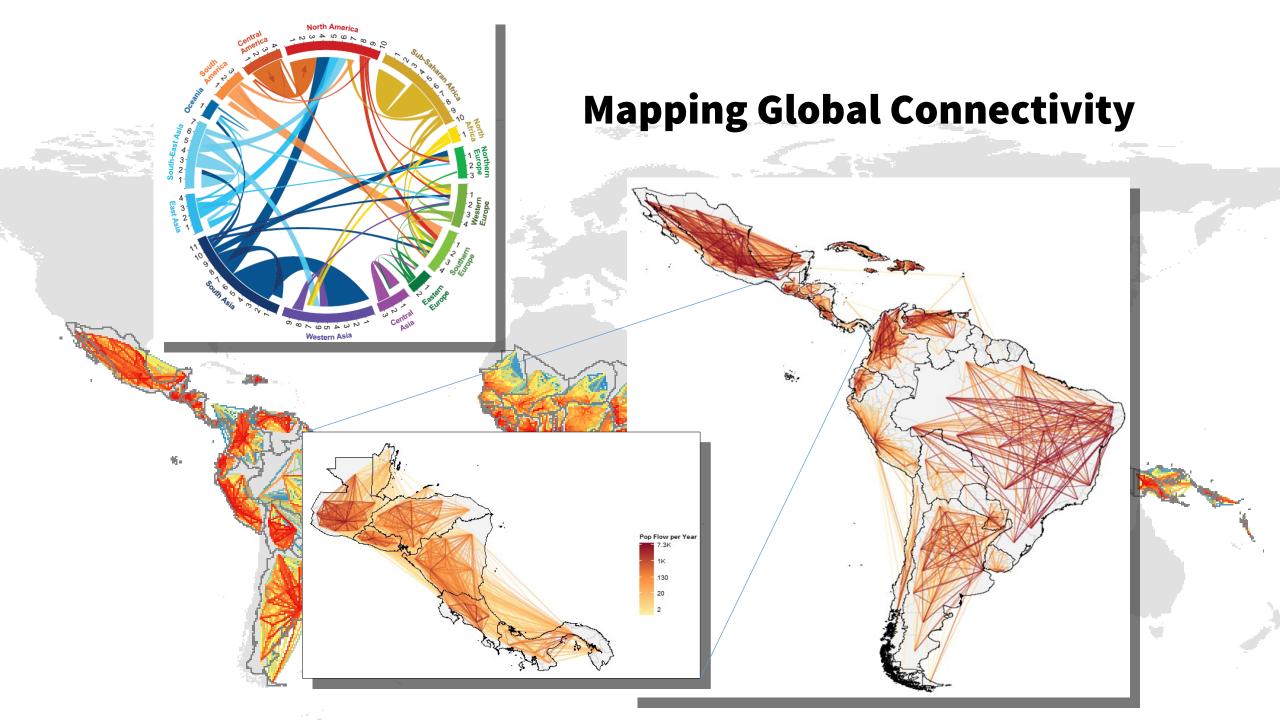


## **Measuring Migration**

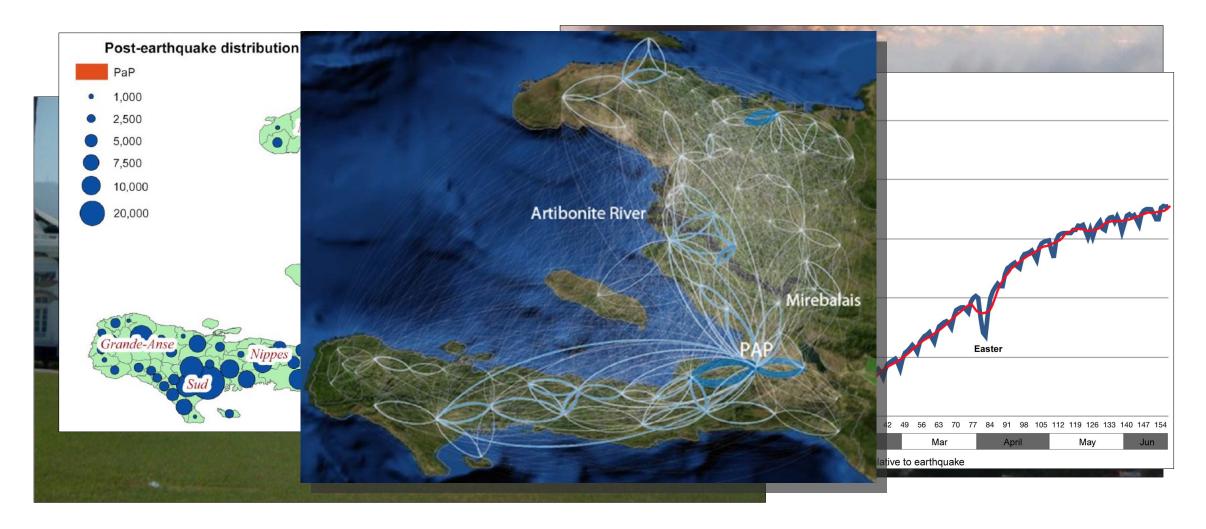








#### Preparedness and Response to Disasters: Haiti Earthquake 2010





#### **Hurricane Matthew: Estimated population movement, 24** October 2016

- The map shows the estimated distribution of people for whom their home Section Communale in the pre-hurricane period was in either Grande Anse, Sud or Nippes départment, and as of 24 October had moved to another Section Communale.
- Estimates are based on movements of de-identified SIM cards which made or received at least one call pre-hurricane and on 24 October 2016.
- The SIM card movements combined with available population data derived from estimates for the year 2015[1].
- · The table lists the locations with the largest number of arrivals

Location	Population pre-hurricane	Persons arrived	Ratio (%)
Port-au-Prince (Metropolitan area)	2870000	85700	3
Bourdet (Les Cayes)	71600	30100	42
Fond Rouge Daiyer (Jeremie)	27100	6950	26
Fond Rouge De Torbec (Jeremie)	26800	6790	25

[1] https://data.humdata.org/dataset/estimated-population-of-haiti-2015

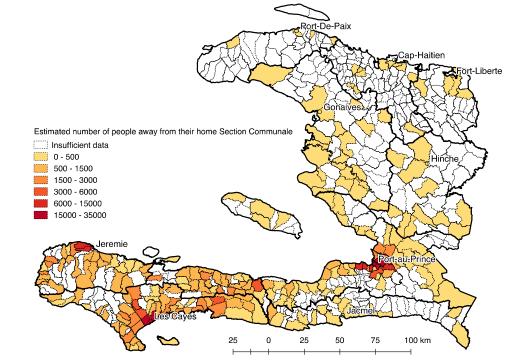
#### Contacts:

linus.bengtsson@flowminder.org +41 78 964 88 28 chris.brooks@flowminder.org +44 7815 944 012

#### Estimated population away from their home Section Communale<sup>[2]</sup>:

HOME DEPARTMENT:	GRANDE ANSE	SUD	NIPPES
POPULATION AWAY FROM HOME:	77500	132000	51000
% AWAY FROM HOME:	18%	17%	15%

24 October 2016, location of people away from their home Section Communale (out of those living pre-hurricane in Grande Anse, Sud and Nippes only)[3]



[2] Of the people normally resident within the given Départment, we estimate the Digicel<sup>®</sup> total number away from their home Section Communale on the given day.

[3] Section Communales are left blank where insufficient data is available.

Flowminder.org is a non-profit organization registered in Stockholm, Sweden. Digicel is a mobile operator in Haiti

#### **Predictive Modeling**

devasta

Proceedings of the National Academy of Sciences of the United States of America

Scientific Reports 3, Article number: 2923 | doi:10.1038/srep02923





## First Project (MDEEP) on Mobile data and Climate Displacement: 2013 Bangladesh Cyclone Mahasen





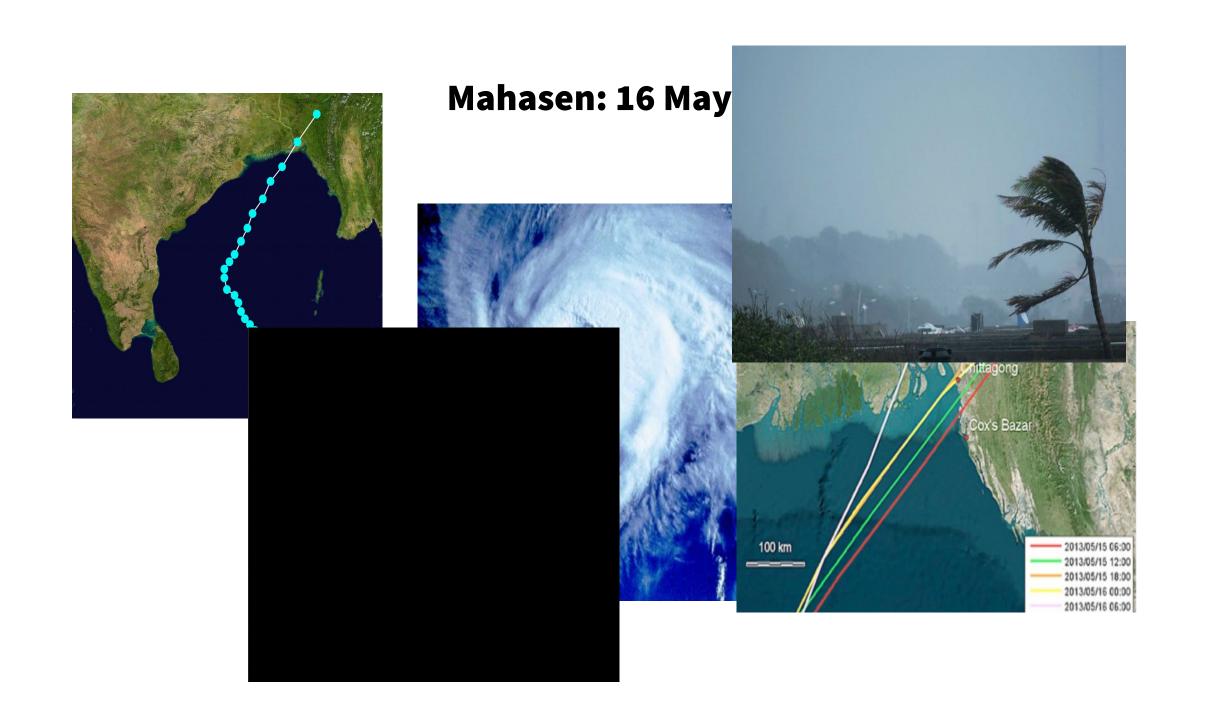
#### **FLOWMINDER.ORG**



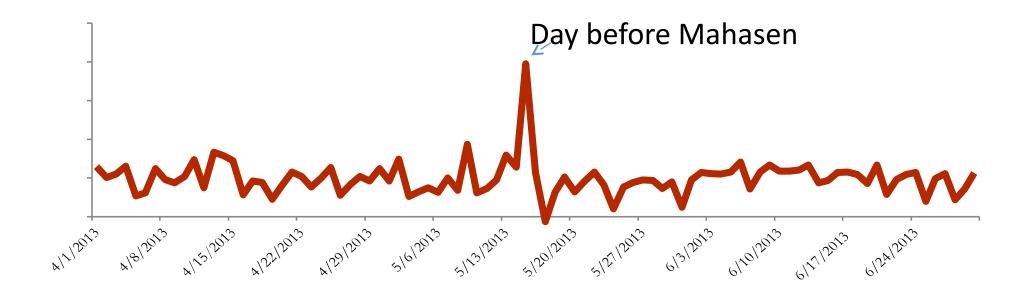


Institute for Environment and Human Security





## Large Spending on Airtime Before Mahasen: Ability to Communicate is Paramount in a Disaster







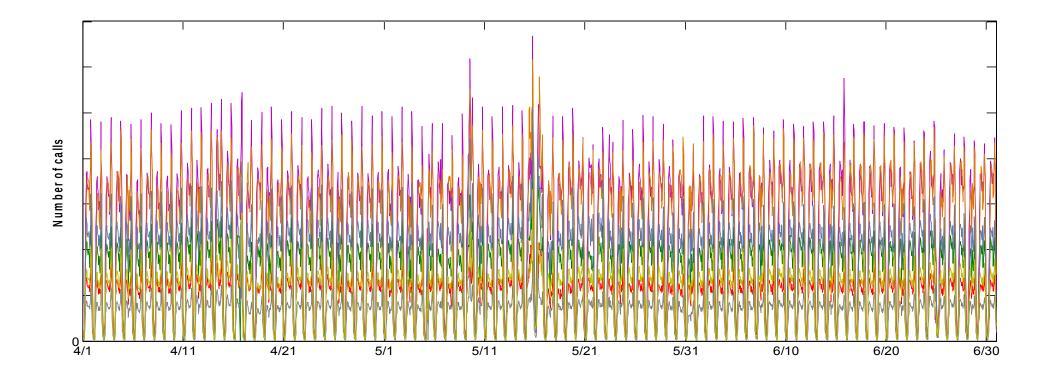






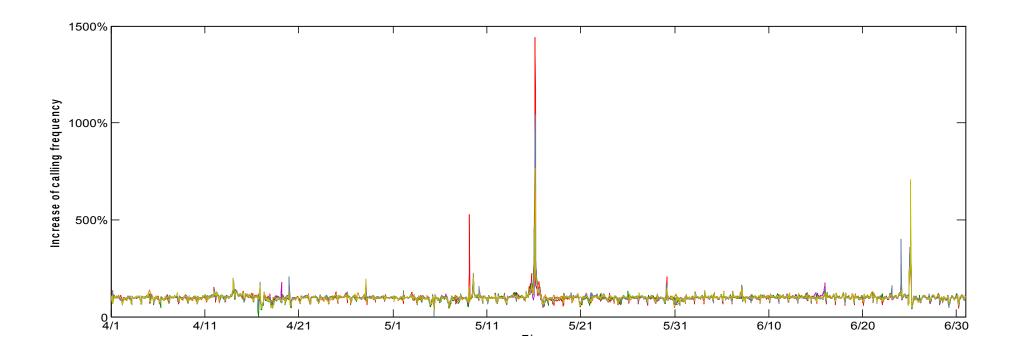
Institute for Environment and Human Security

#### **Call Frequencies Follow a Daily Rythm**



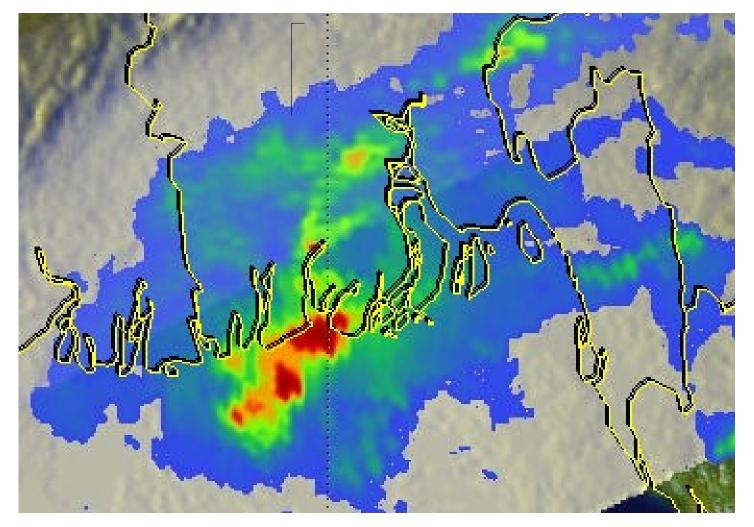
Ref: Lu et al. Climatic Change. 2016; Global Environmental Change. 2016

#### Call Frequency Increased 1500% During Mahasen



Ref: Lu et al. Climatic Change. 2016; Global Environmental Change. 2016

#### **Increase in Calling Frequency Likely Indicate Impact**



3.32 am on 16 May Ref: Lu et al. Climatic Change. 2016; Global Environmental Change. 2016

## Nepal Earthquake 2015: Setup & First Insights Within 14 Days

Nepal Population Estimates as of May 1, 2015

2.8m

**+390,000** (246,000~540,000)

- 247,000

 $(-155,000 \sim 339,000)$ 

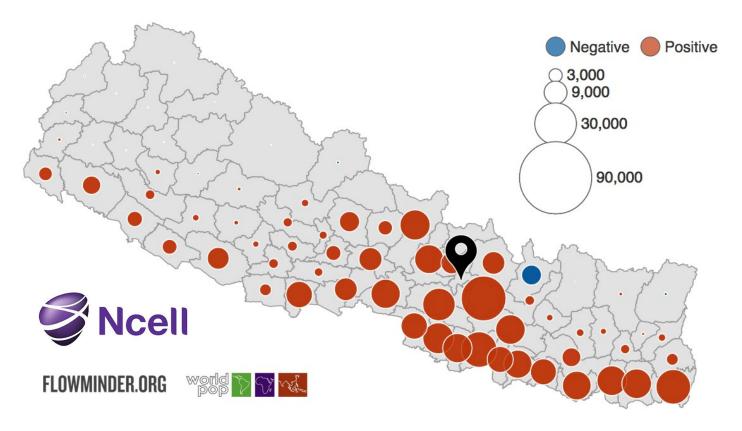
#### 2. Kathmandu Valley

Kathmandu Valley is here defined as the districts Kathmandu, Bhaktapur and Lalitpur. Kathmandu Valley is one of the most densely populated areas in Nepal and home to ca 2.8 m people [1].

#### Key findings:

- → An estimated 390,000 people more than normal had left the Kathmandu valley comparing May 1 with the day before the earthquake April 24 (ratio to the population: 14%).
- → An estimated 247,000 persons less than normal had come into the area during the same period(ratio to the population: 8.8%)
- People leaving Kathmandu Valley went to a large number of areas, notably the populous areas in the south and the Central and West Development Regions.

Above normal flows from Kathmandu Valley to other districts



Ref: Bengtsson et al. Sci Rep. 2015; Wesolowski et al. PLoS Currents 2014; Wilson et al. PLoS Current 201

#### **Global Mobile Industry Recognition (Feb 2016)**



"A brilliant example of how the application of big data analysis to mobile technologies can be used to accelerate emergency aid, and provide intelligence to help prepare for future disasters."



## The Star Trek Fallacy

- 1. Data is part of the tool, not the solution
  - Issue-driven vs. data-driven problem solving.
- 2. Remote sensing data and analytics can augment but not replace census and surveys
- 3. Very few studies of bias in mobile operator data
- 4. Mobile network data does not have stable properties:
  - Fundamental characteristics (subscribers) constantly changing.
  - Mobile data is heterogenous market/operators.
  - Representativeness what does a SIM card represent?
  - Realtime mobile data without validation = realtime mistakes

#### **Summary**

- In producing estimates for different geographical scales and time periods, the integration of multiple types of data to compliment traditional sources is often required
- Novel datasets (e.g. phones, satellite) are prone to biases, but each has advantages over census data in terms of the frequency of measurement and coverage
- Methods to account for biases, reporting uncertainties and providing clear metadata/documentation to inform users are all important
- Great potential in complimenting traditional sources to build strong demographic databases for measuring progress towards the SDGs





www.worldpop.org



# FLOWMINDER.ORG

www.flowminder.org

