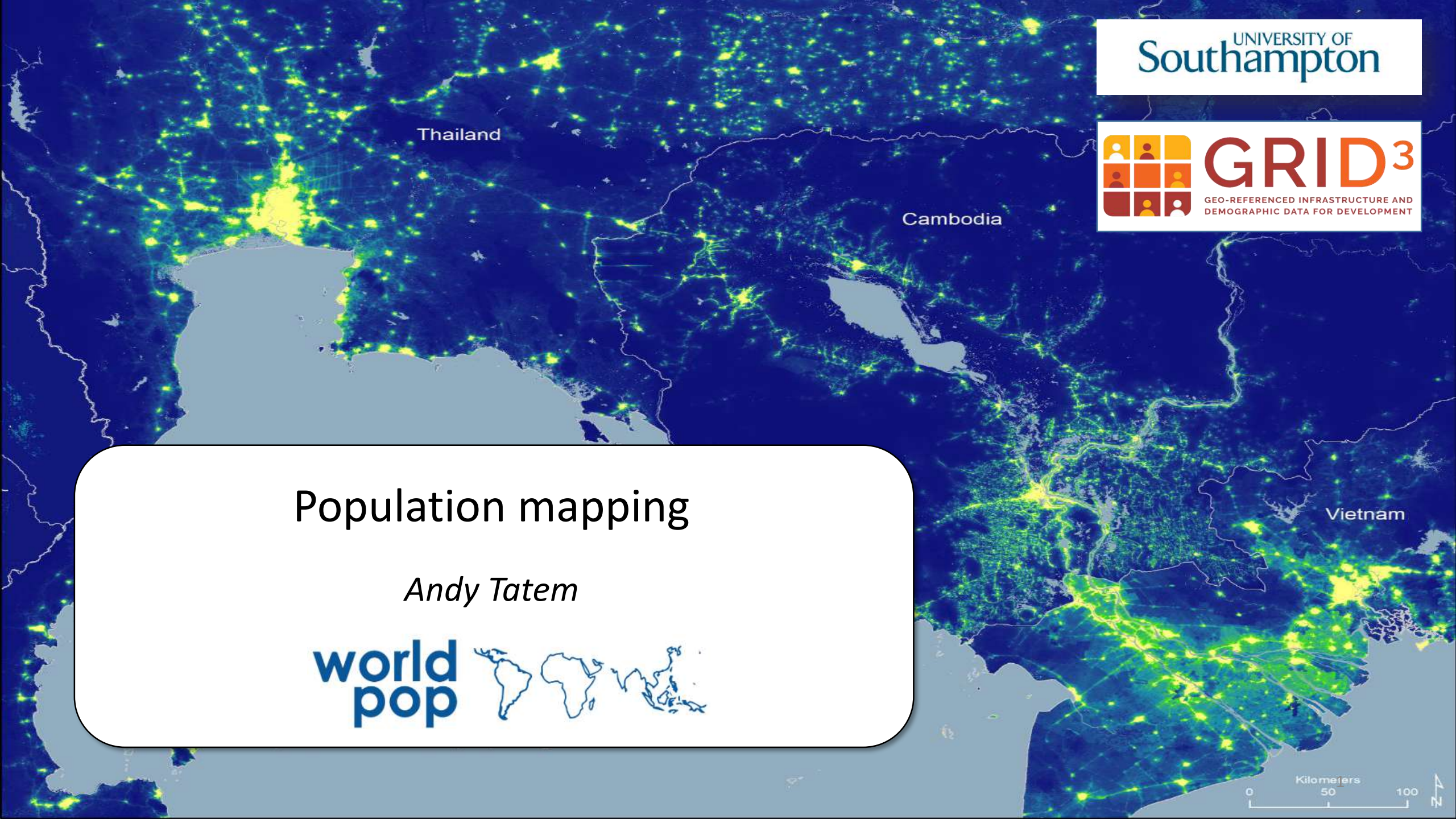


UNIVERSITY OF Southampton

GRID3
GEO-REFERENCED INFRASTRUCTURE AND DEMOGRAPHIC DATA FOR DEVELOPMENT

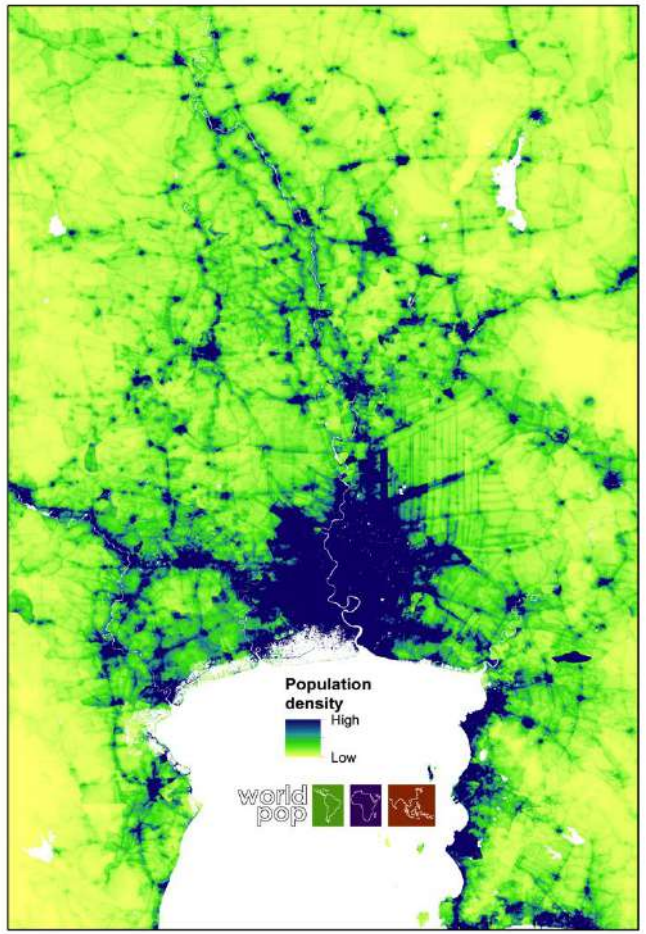
Population mapping

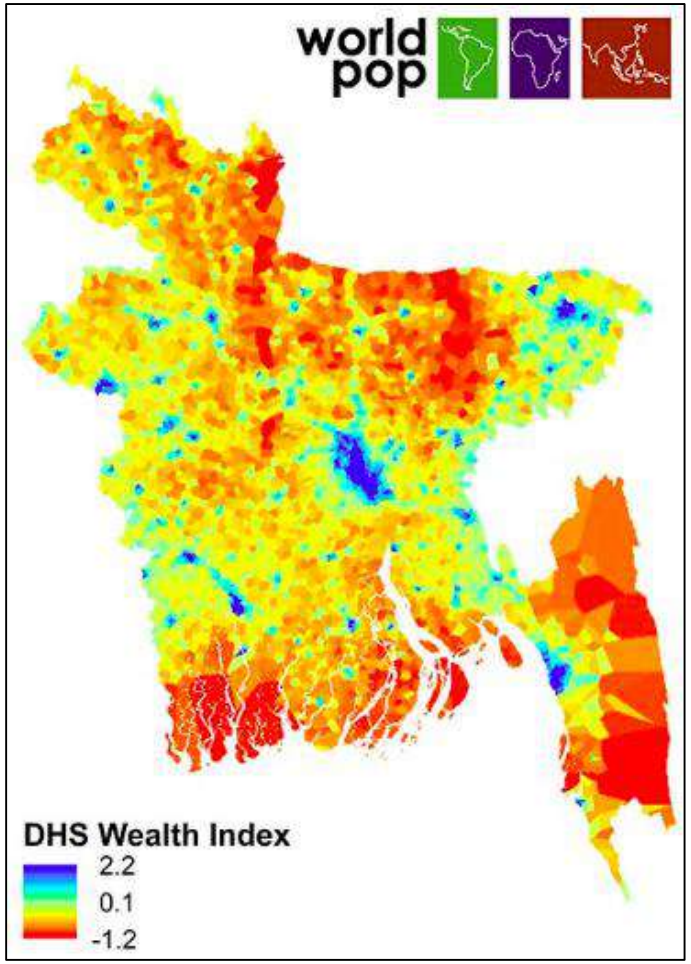
Andy Tatem

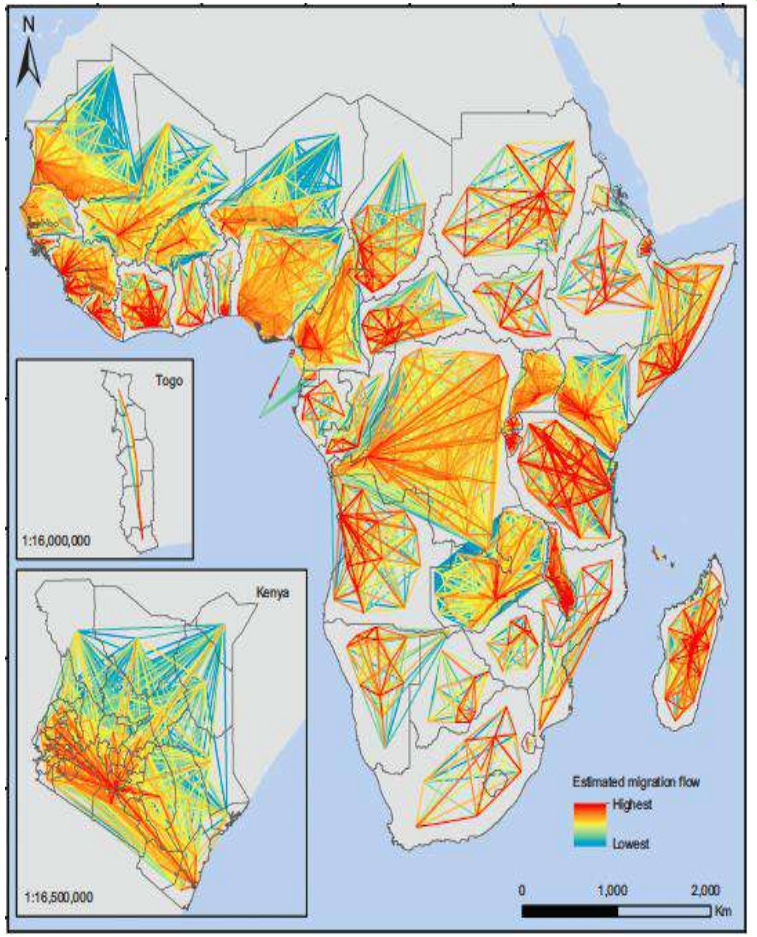
worldpop

Tatem (2017) Nature Sci Data

www.worldpop.org







Uses of small area population data

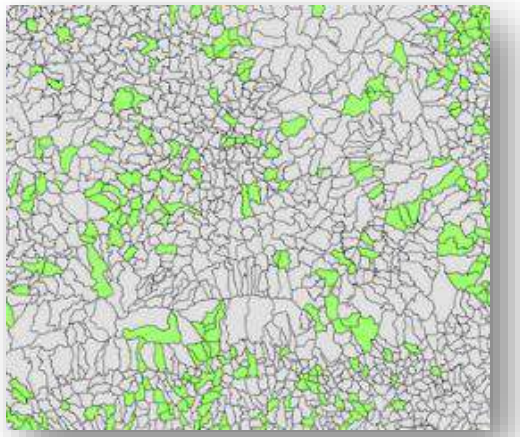
- Planning elections
- Calculating GDP
- Local governance
- Traffic planning
- Health systems
- Financial services
- Delivering utilities
- Controlling infectious diseases
- Agricultural subsidies
- Taxation
- Land use management
- Energy strategies
- Supply chain management
- Meeting SDGs

world pop

Tatem (2017) Nature Sci Data

Sources of small area population data

“Traditional” data sources

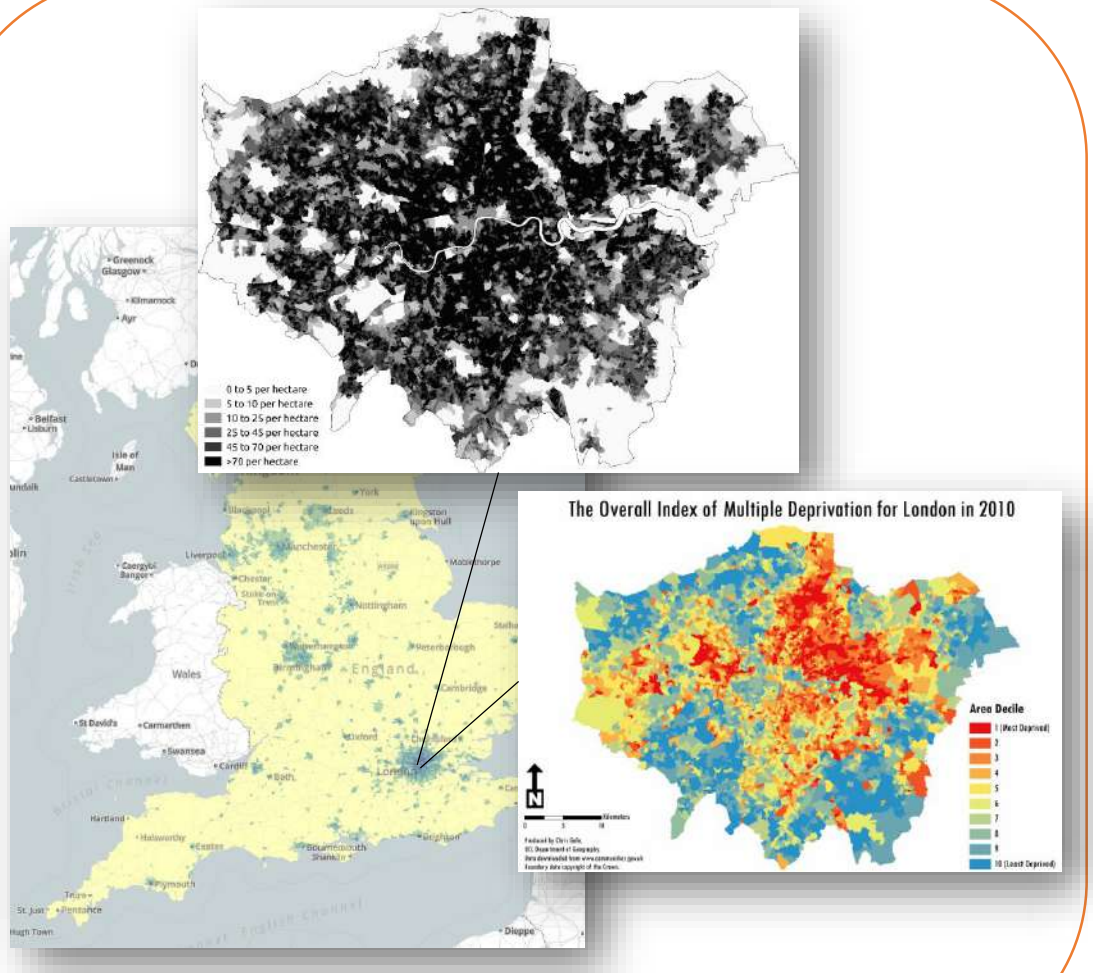


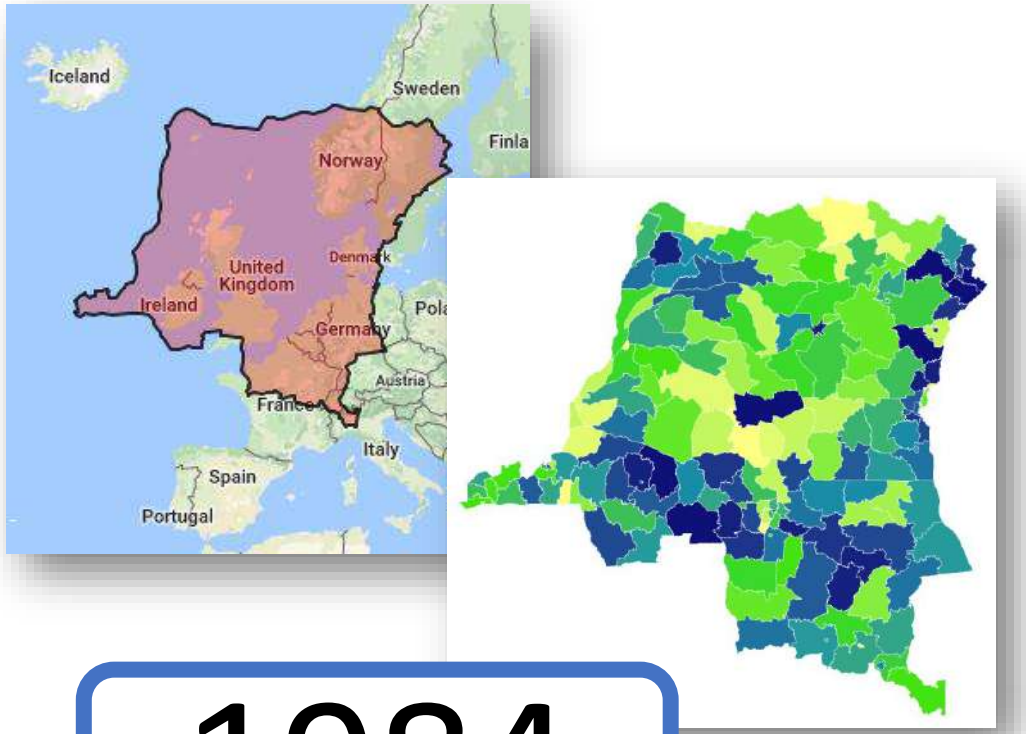
-Census
-Surveys
-Administrative/ registry data

world pop

Tatem (2017) Nature Sci Data

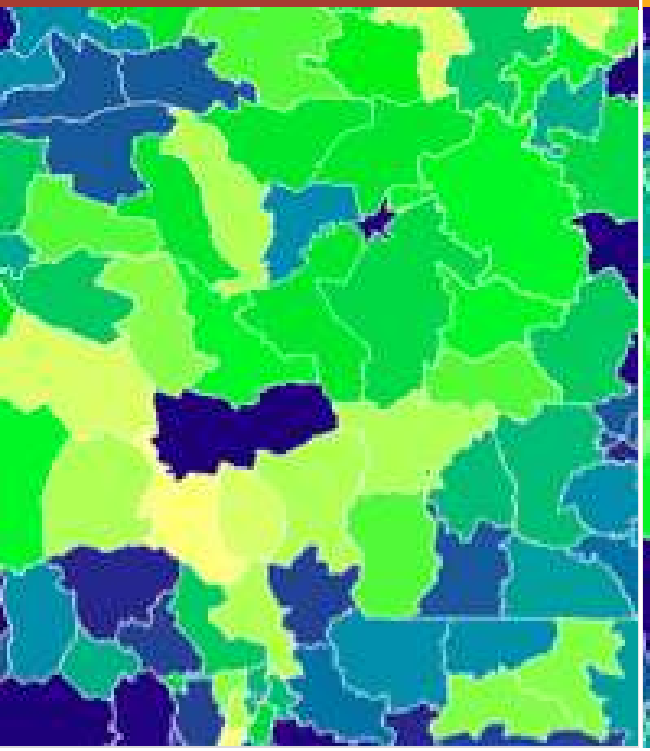
Country comparisons

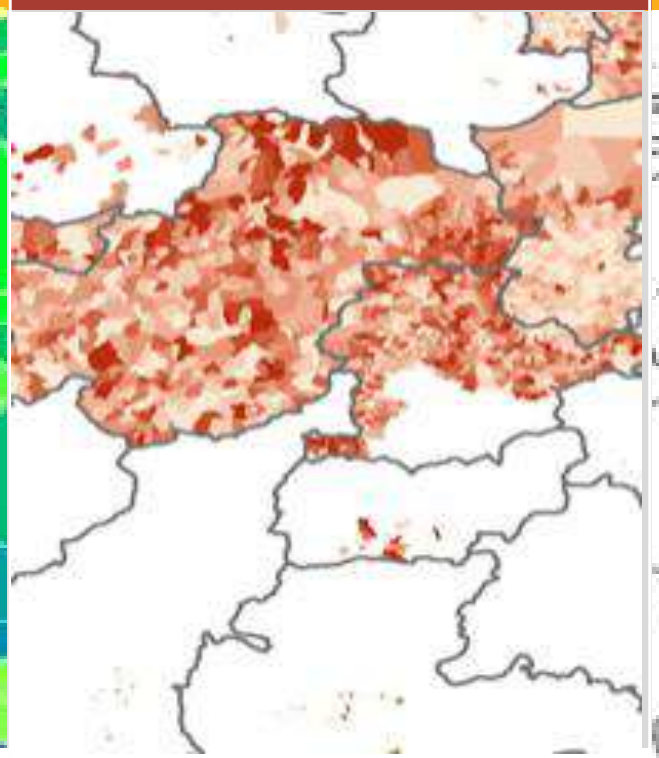




1984

Coarse resolution	Outdated	Incomplete	Inaccuracies, missing populations





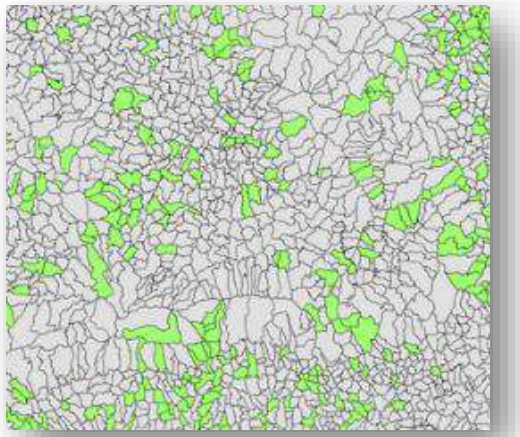
Demographic data challenges

world pop

Tatem (2017) Nature Sci Data

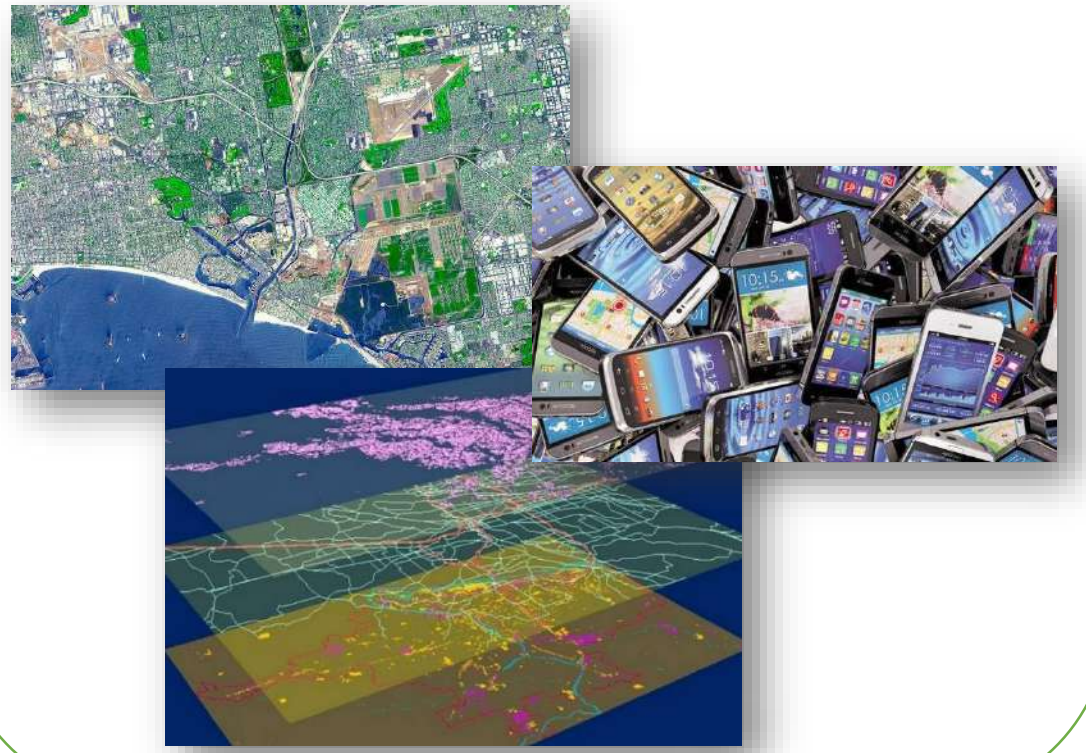
Complimentary data sources

“Traditional” data sources



-Census
-Surveys
-Administrative/ registry data

“Novel” data sources

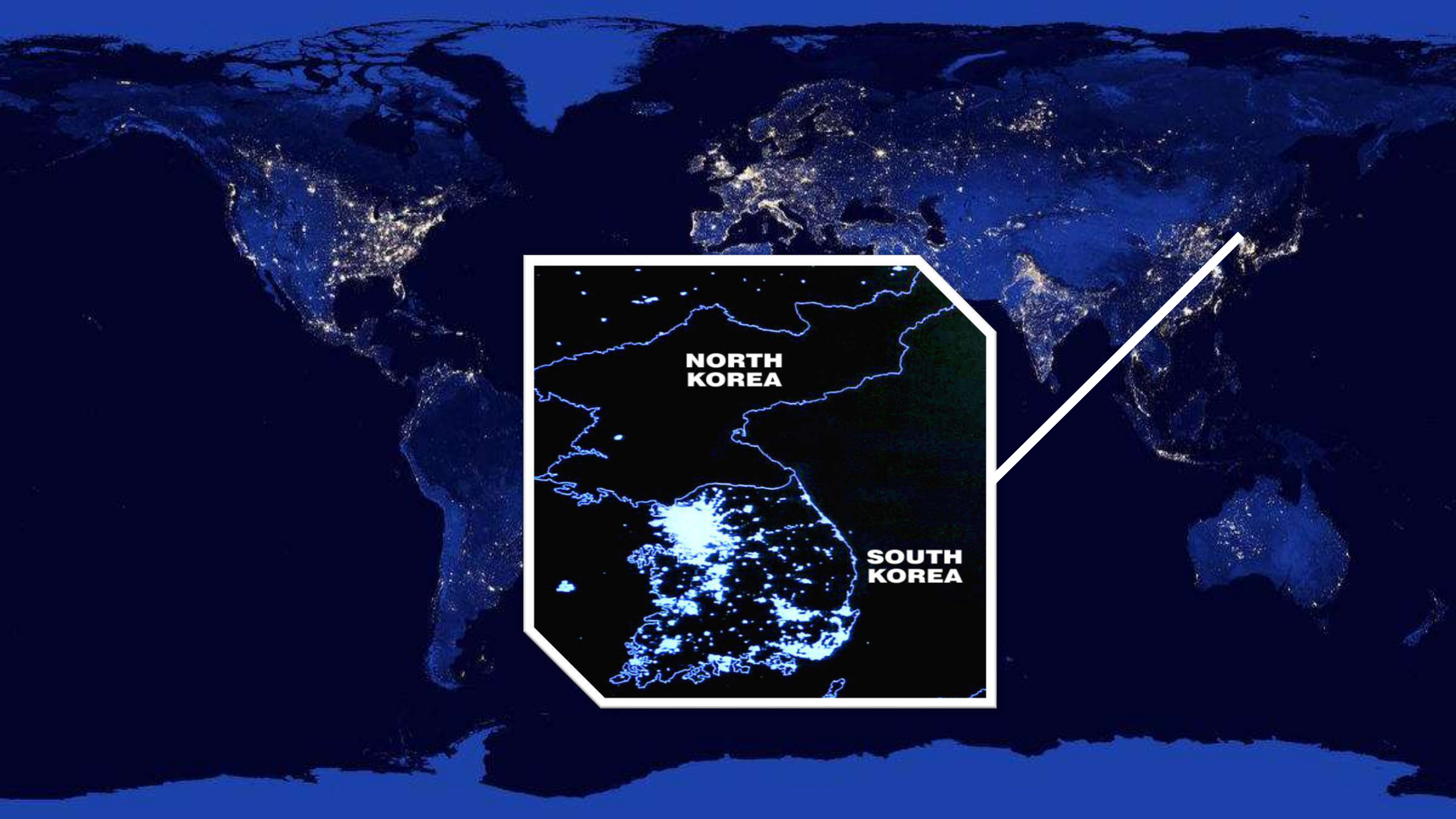


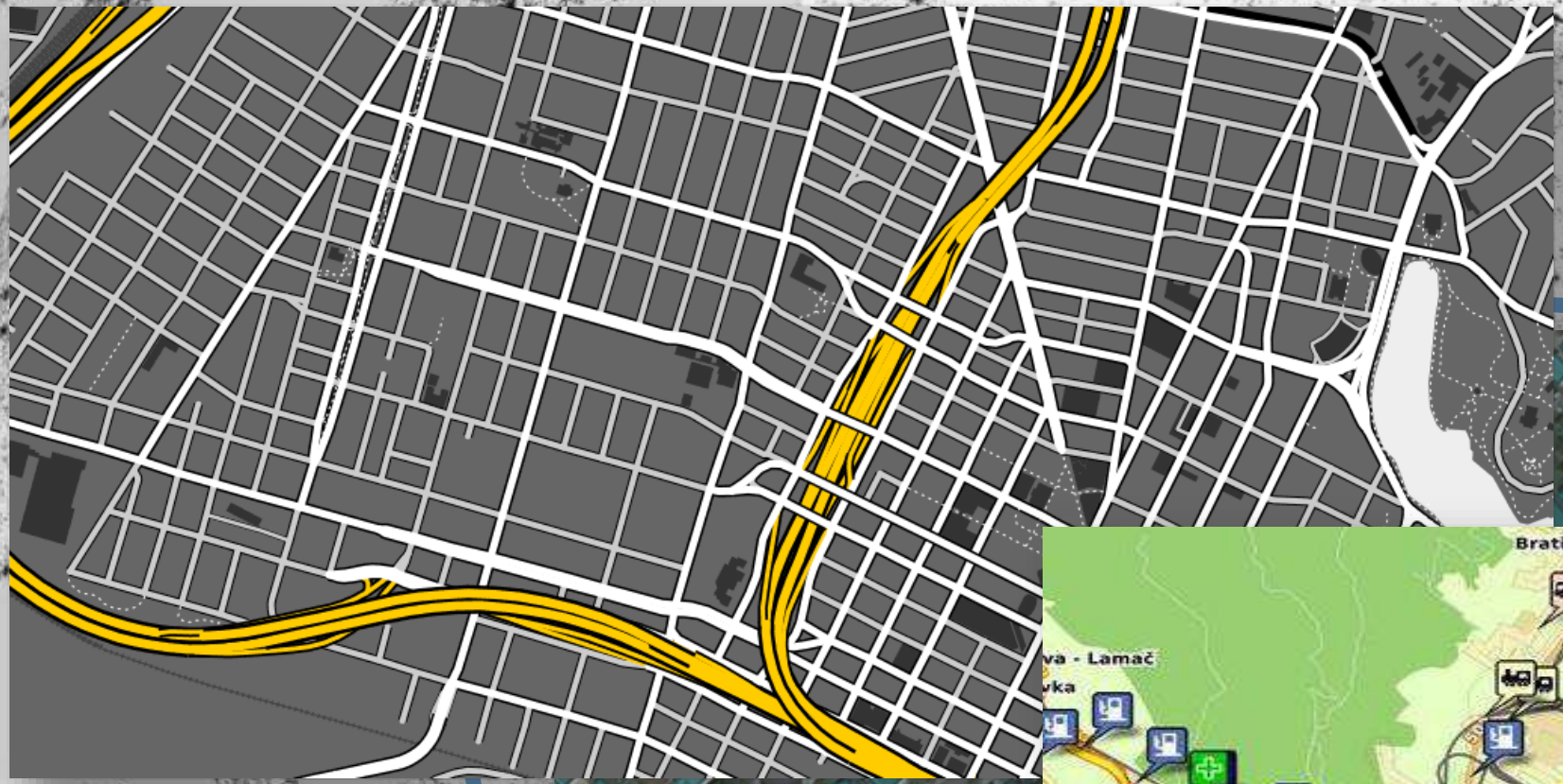
The New York Times | U.S.

A MAP OF EVERY BUILDING IN AMERICA

By TIM WALLACE, DEREK WATKINS and JOHN SCHWARTZ

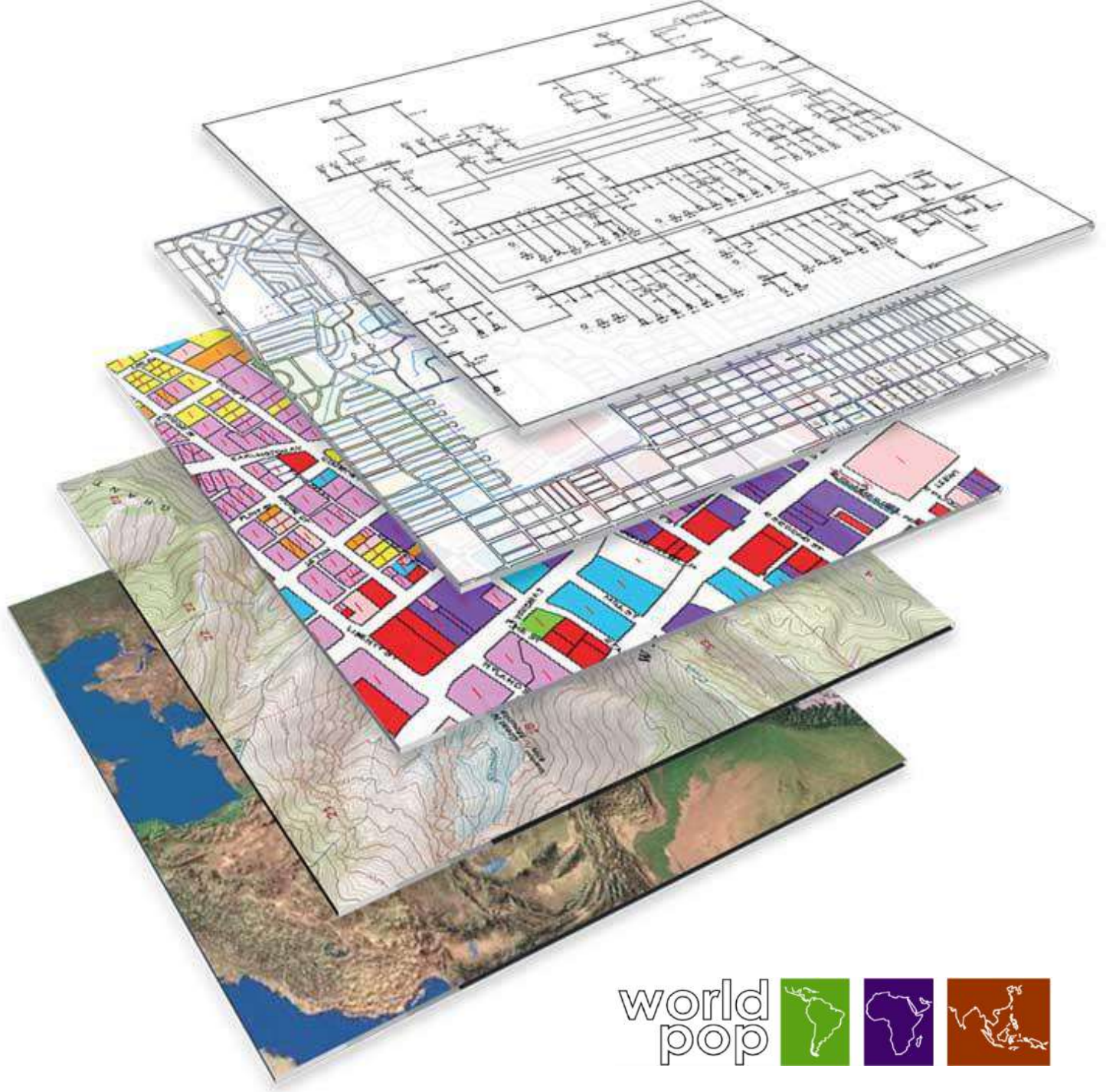
Oct. 12th, 2018


NORTH
KOREA
SOUTH
KOREA

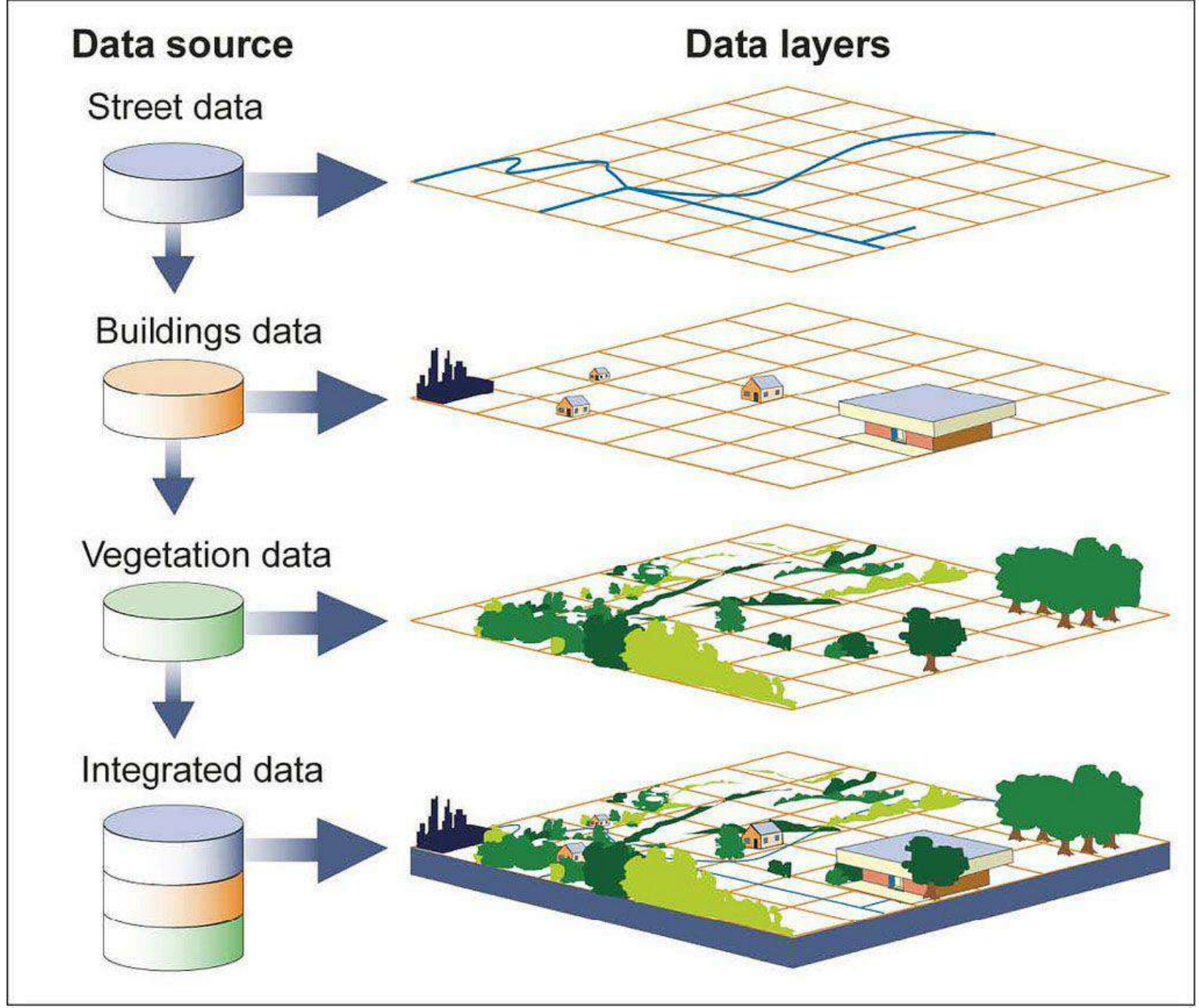



Bratislava - Vajnory
Karlova Ves
Bratislava - Staré Mesto
Bratislava - Vrakuňa
Most pri
Bratislava - Podunajské Biskupice

Geography as a framework for data integration



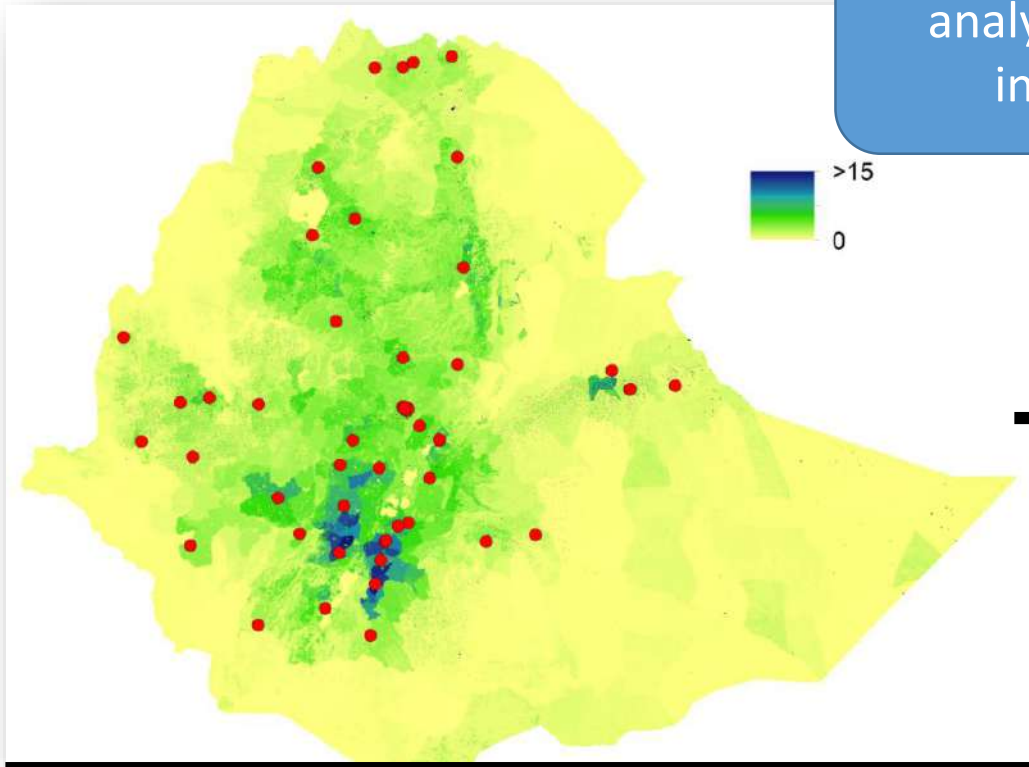
world pop


Data source
Data layers
Street data
Buildings data
Vegetation data
Integrated data

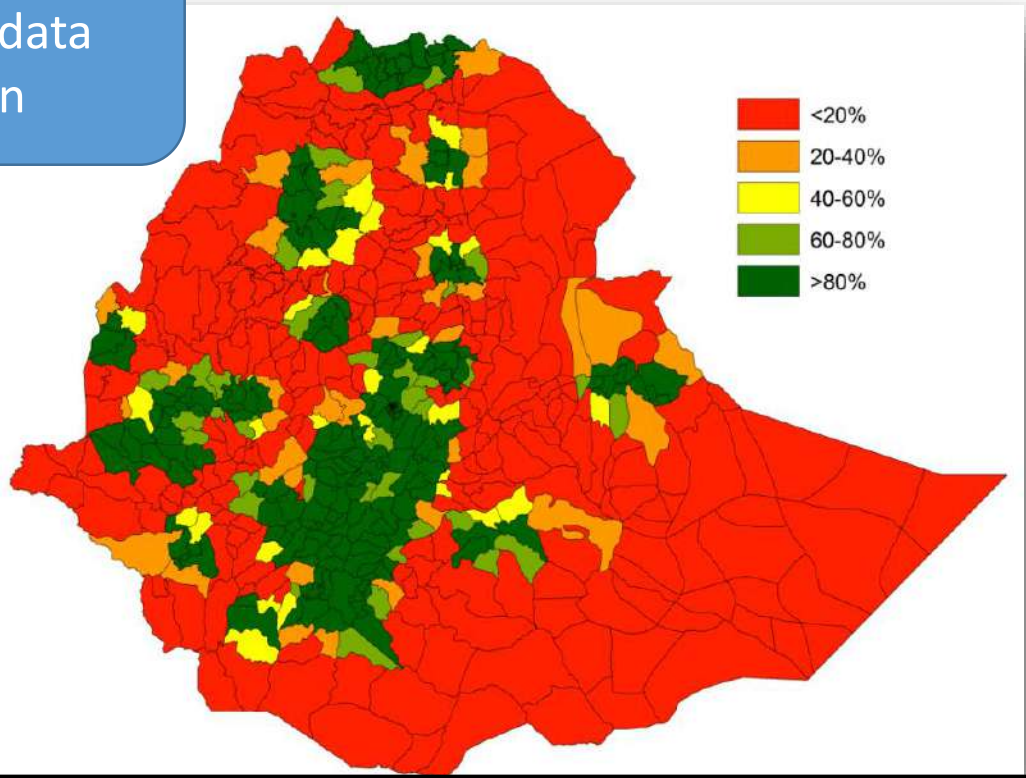
Source: GAO.

Gridded data

Grids: flexibility in analysis and data integration



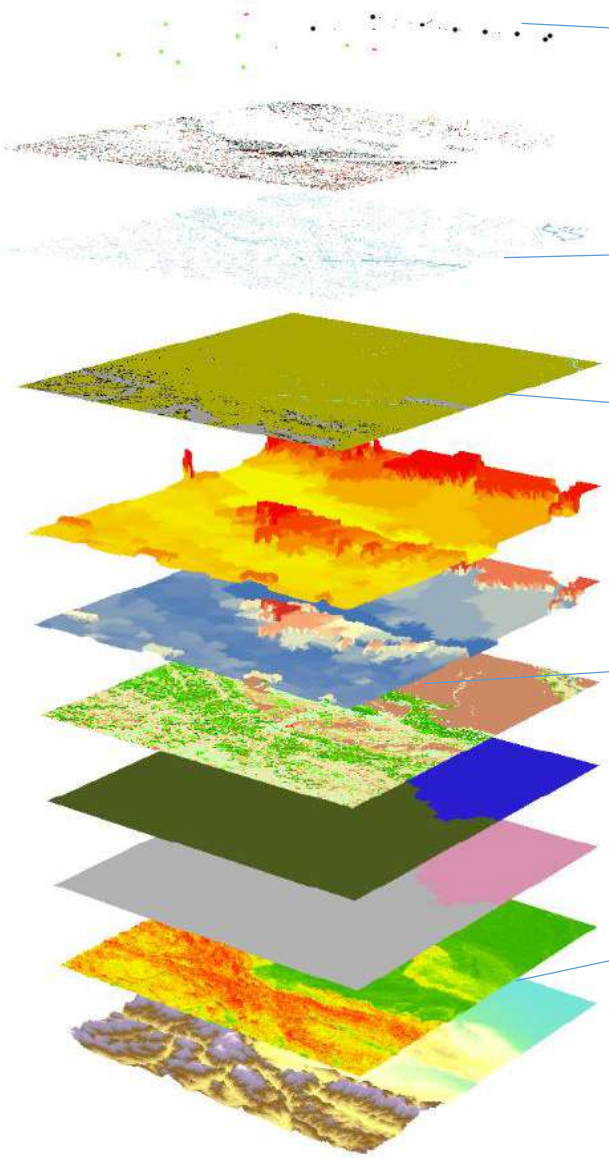
Comprehensive Emergency Obstetric and Neonatal Care (CEmONC) Facilities overlaid on grid of women of childbearing age

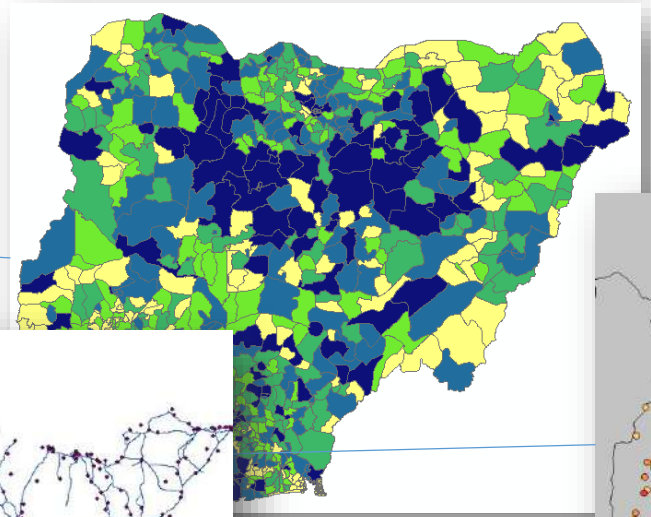


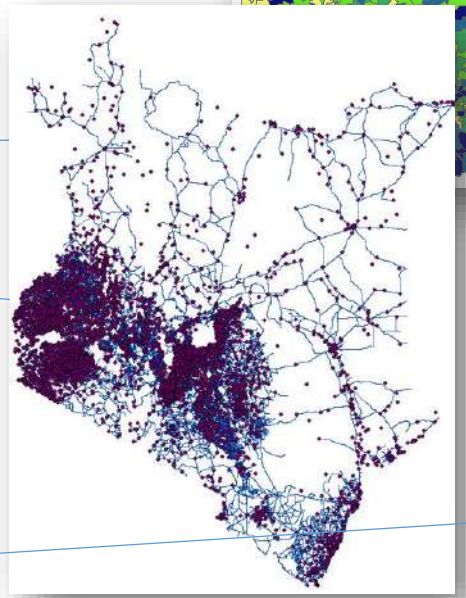
Percentage of women of childbearing age per woreda within 50km of a CEmONC

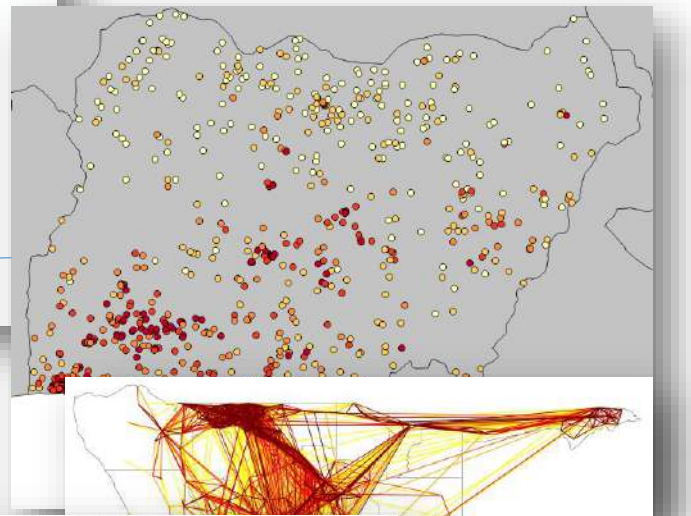
population totals

Data integration

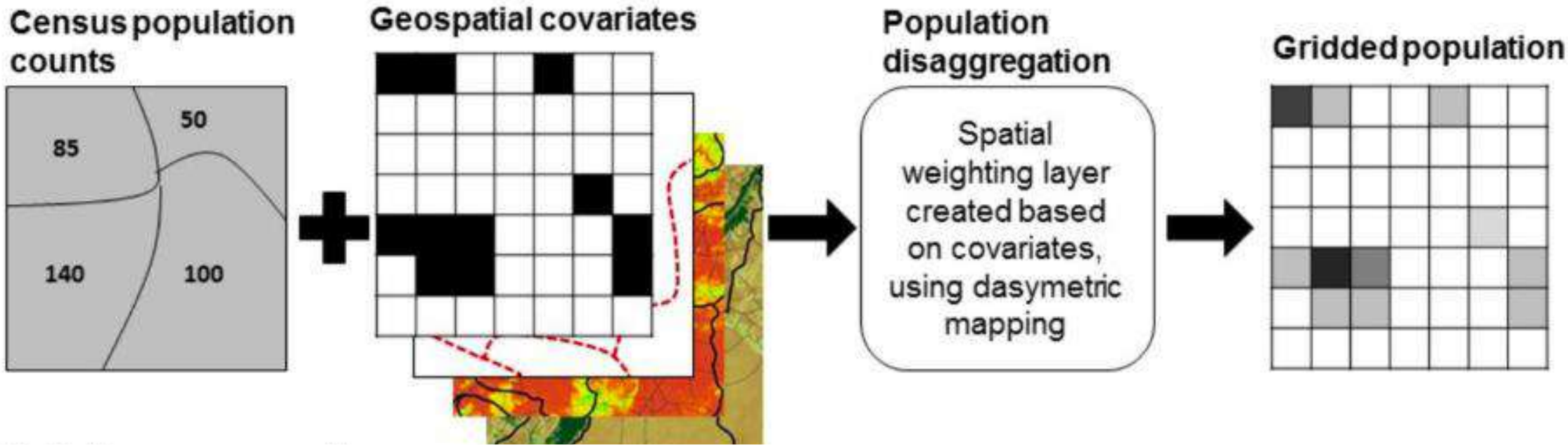




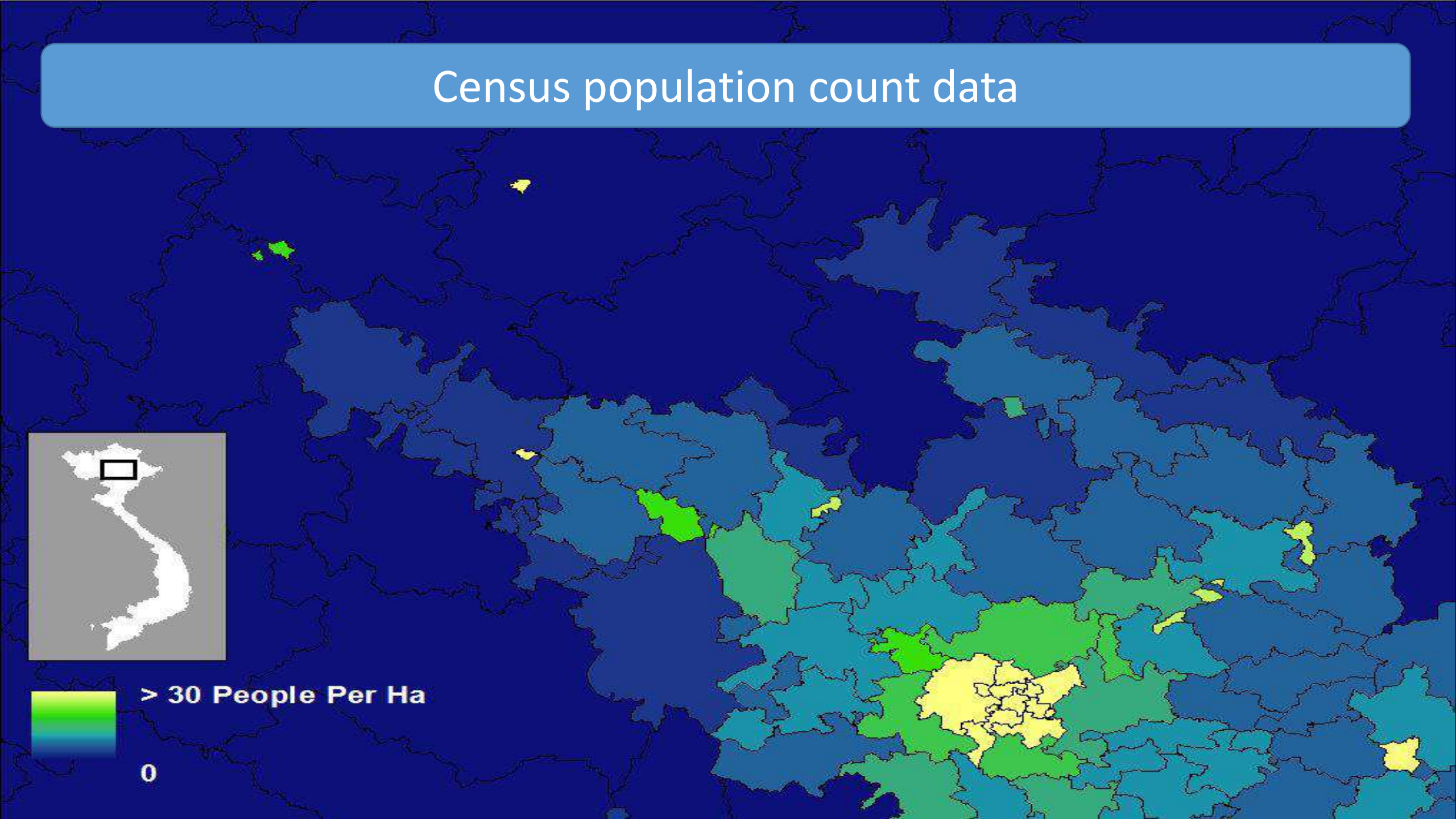




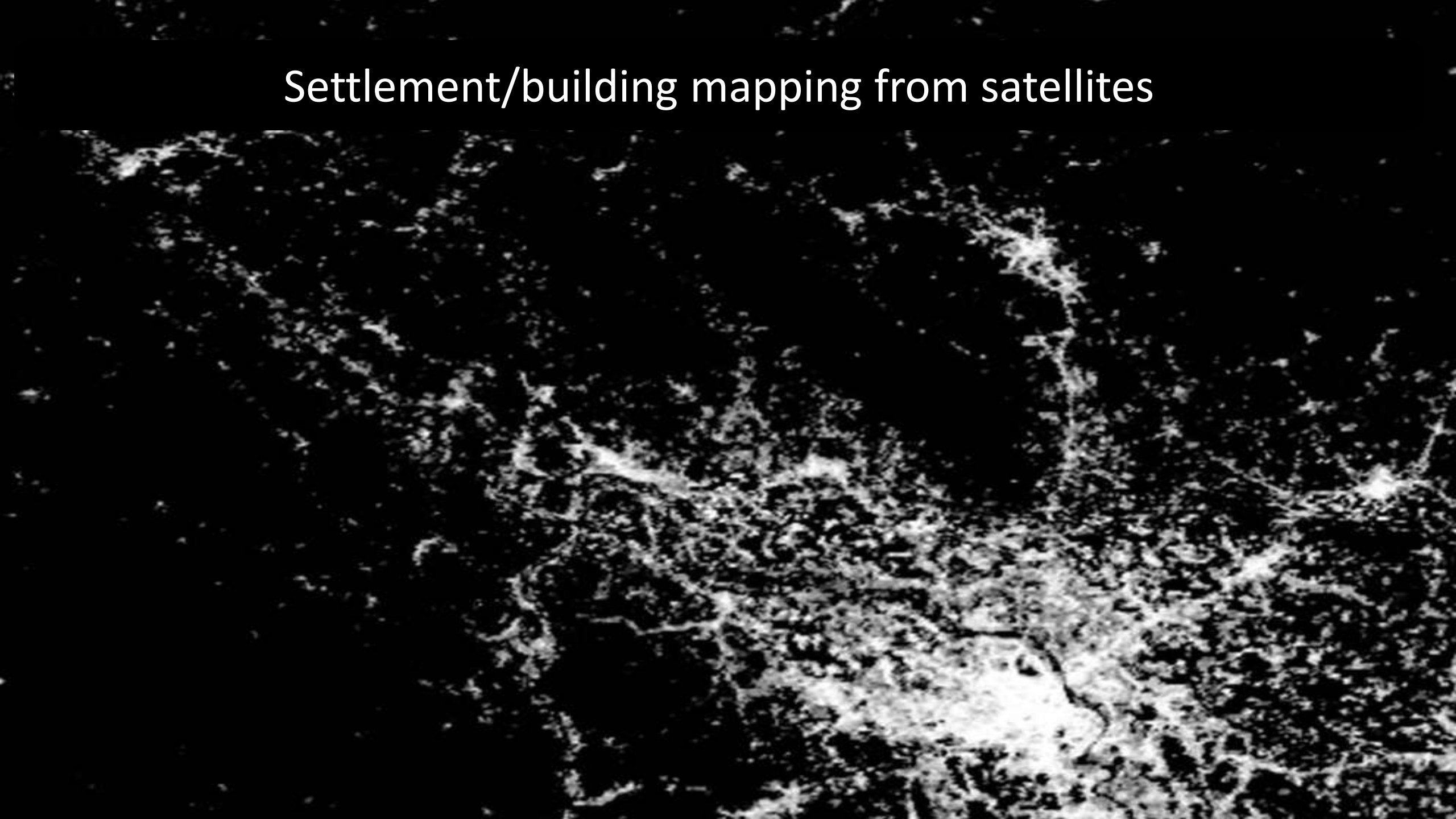
A Top down approach



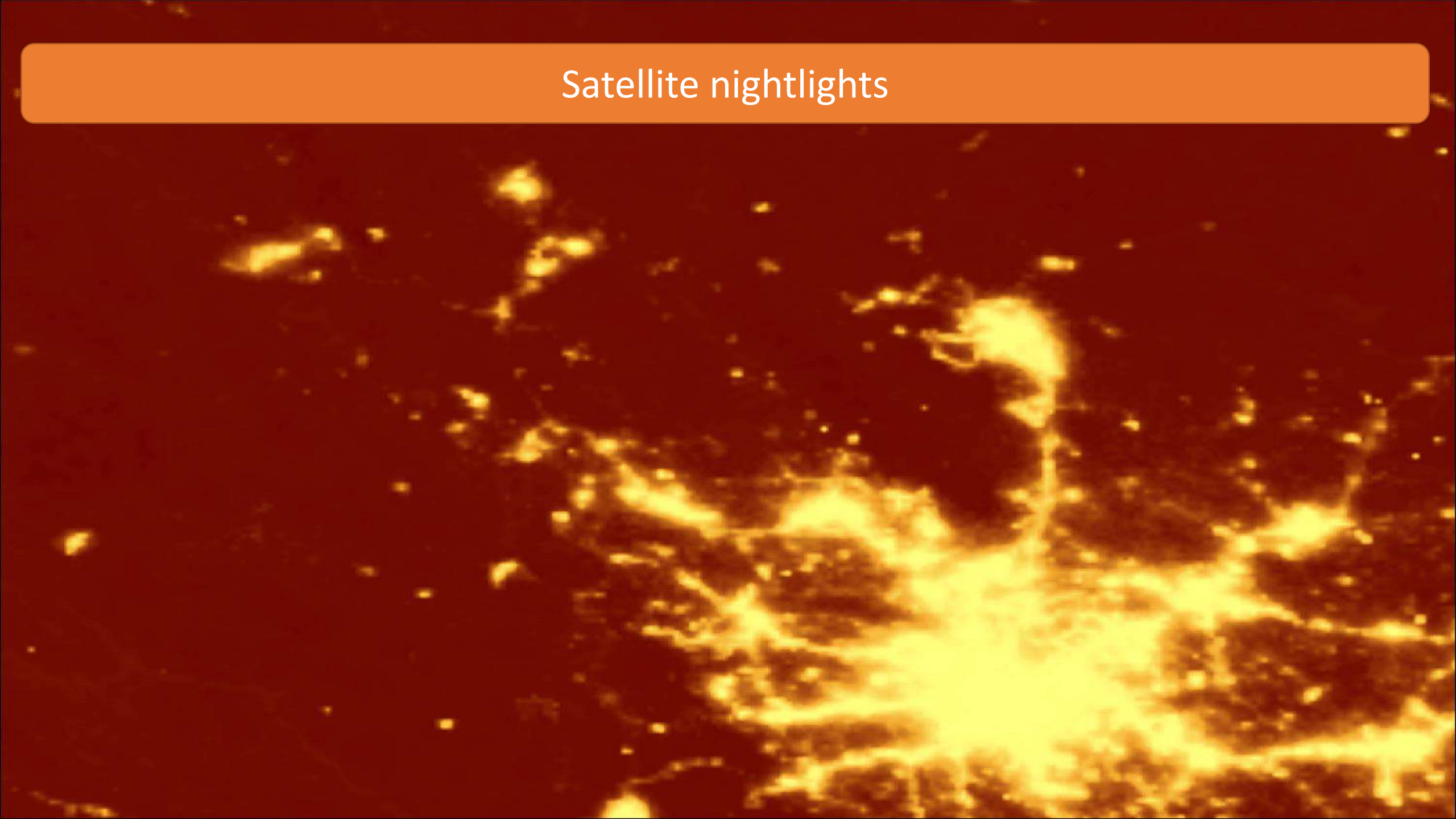
Census population count data



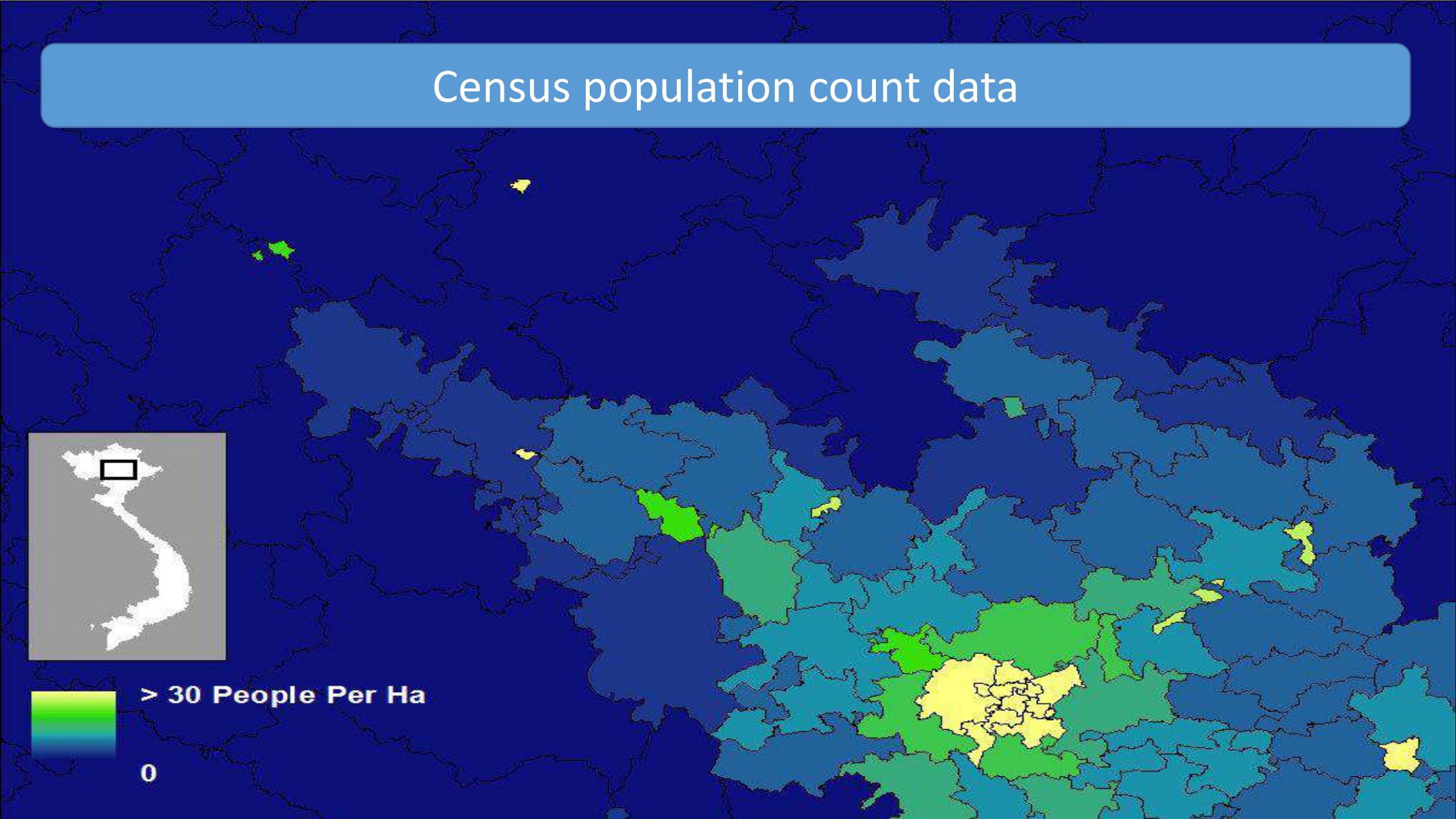
Settlement/building mapping from satellites



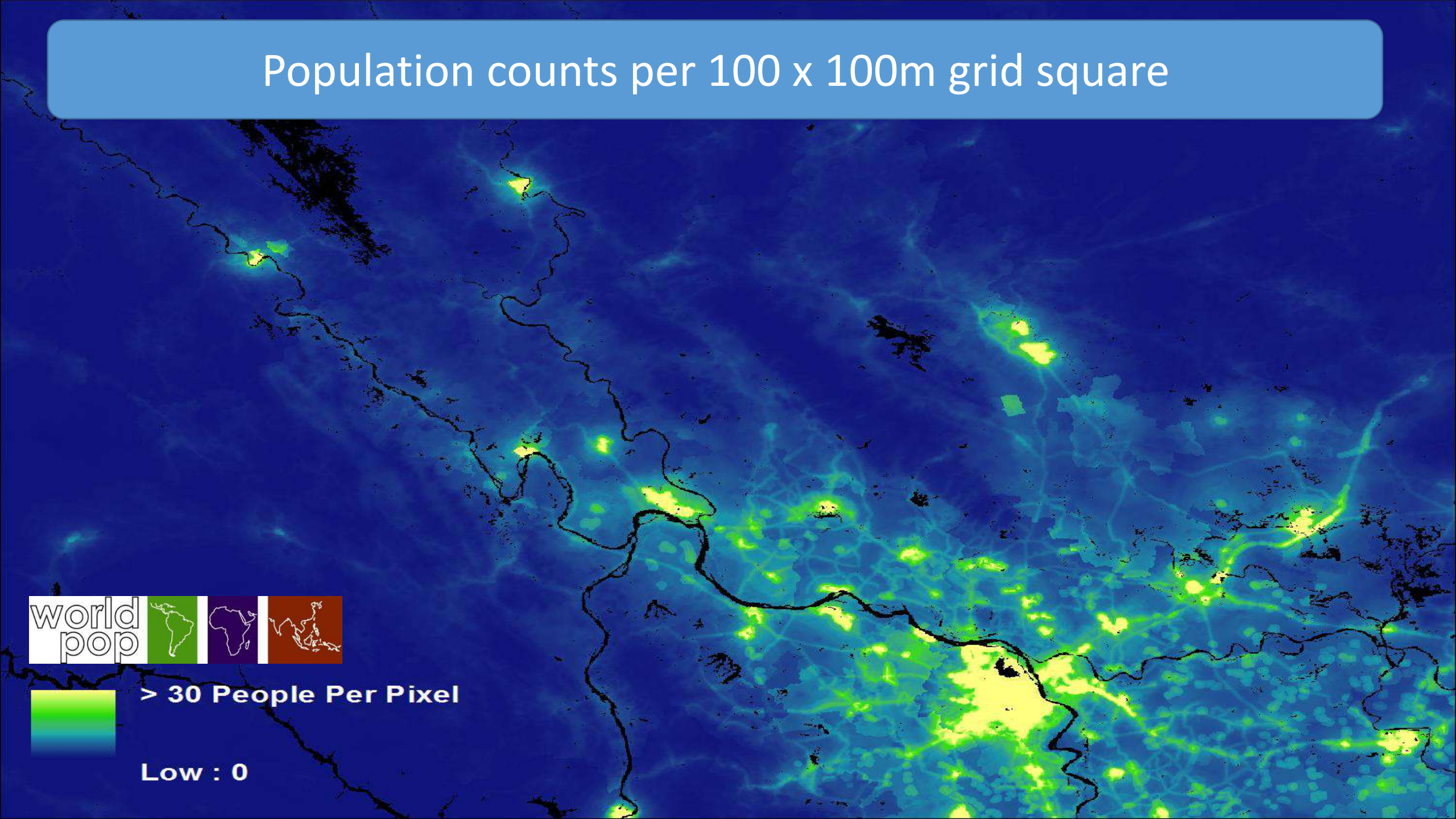
Satellite nightlights



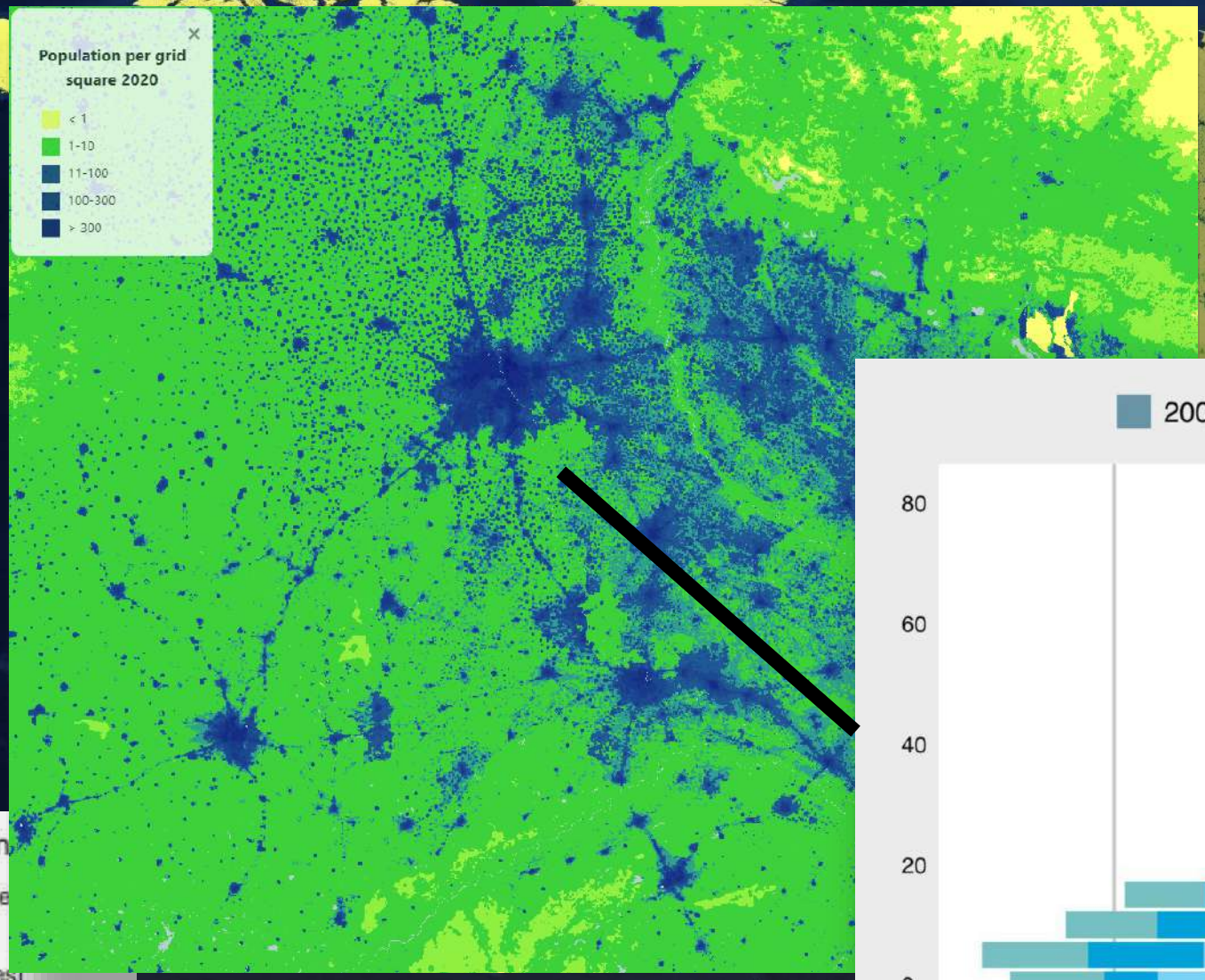
Census population count data

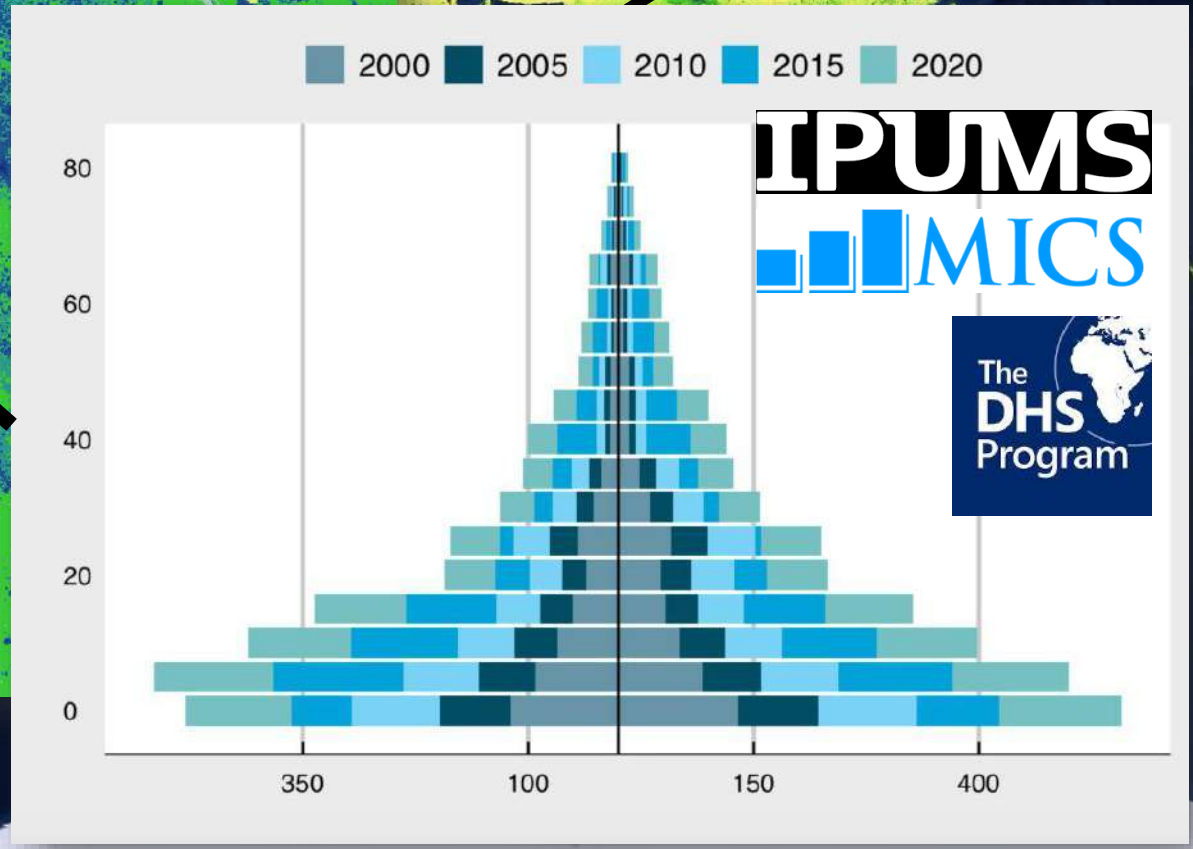


Population counts per 100 x 100m grid square



www.worldpop.org





Population
Highe
Lowest

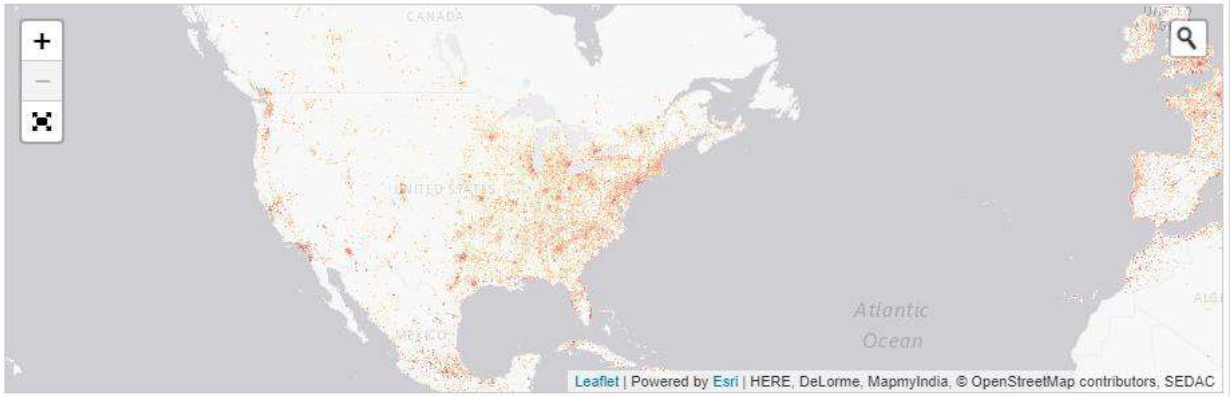
world pop

CIESIN

Sorichetta et al (2017) Nature Sci Data; Gaughan et al (2017) Nature Sci Data

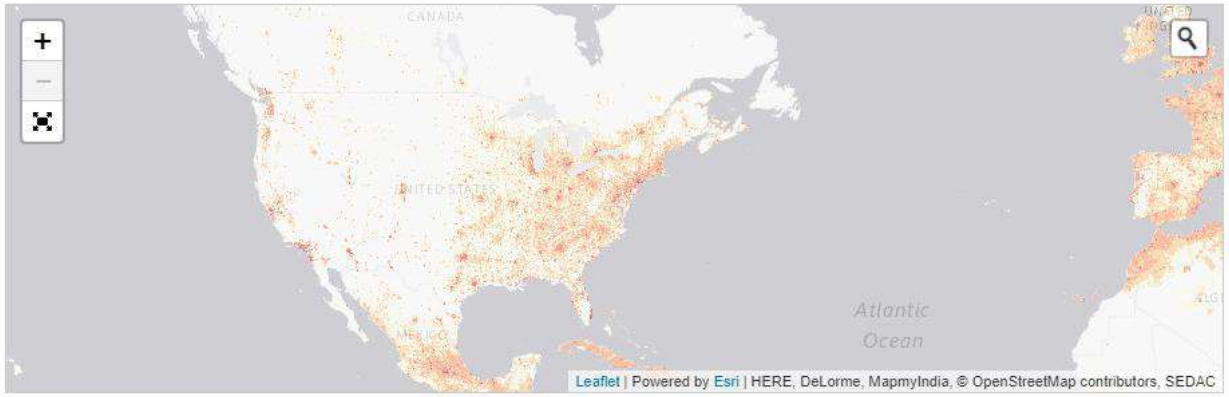
POPGRID Viewer

GHS-Pop 2015 Count



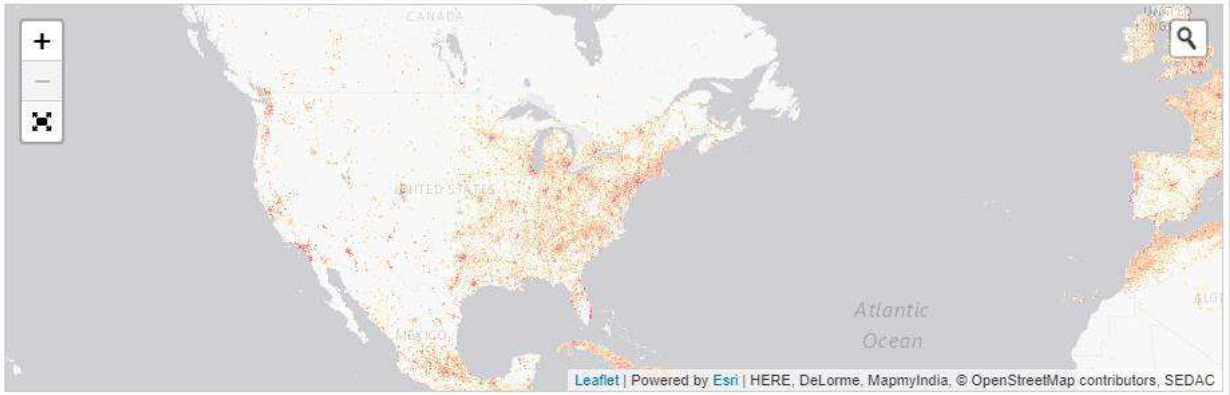
Sync Panels | Repeat Map | Zoom Out | Legend | Overview | Comparison View

GPW4.10 UN Adjusted 2015 Count



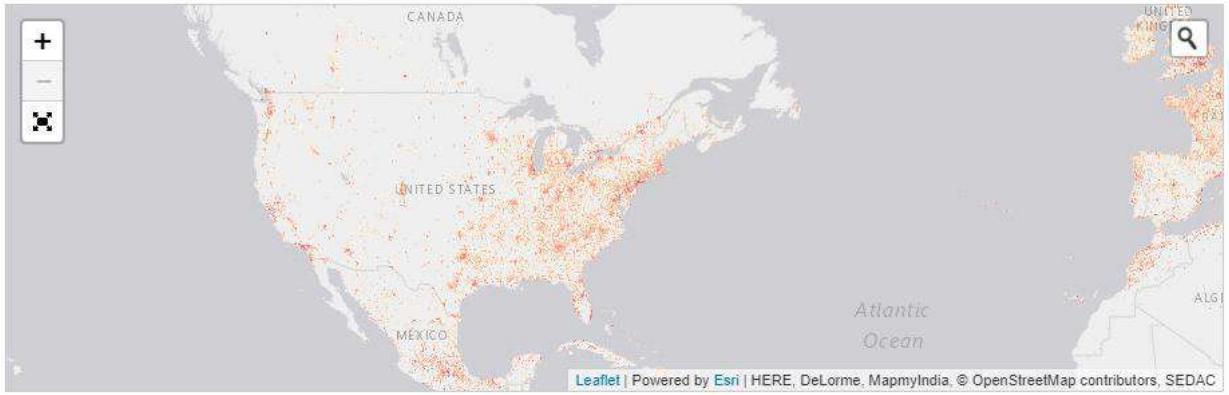
Zoom Out | Legend | Overview | Comparison View

Worldpop 2014 Count



Zoom Out | Legend | Overview | Comparison View

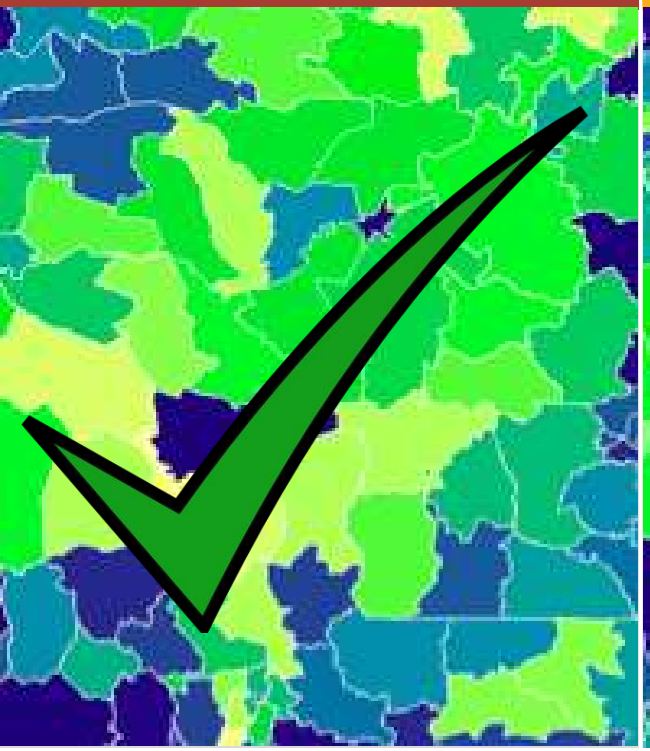
Esri's World Population Estimate 2016 Count

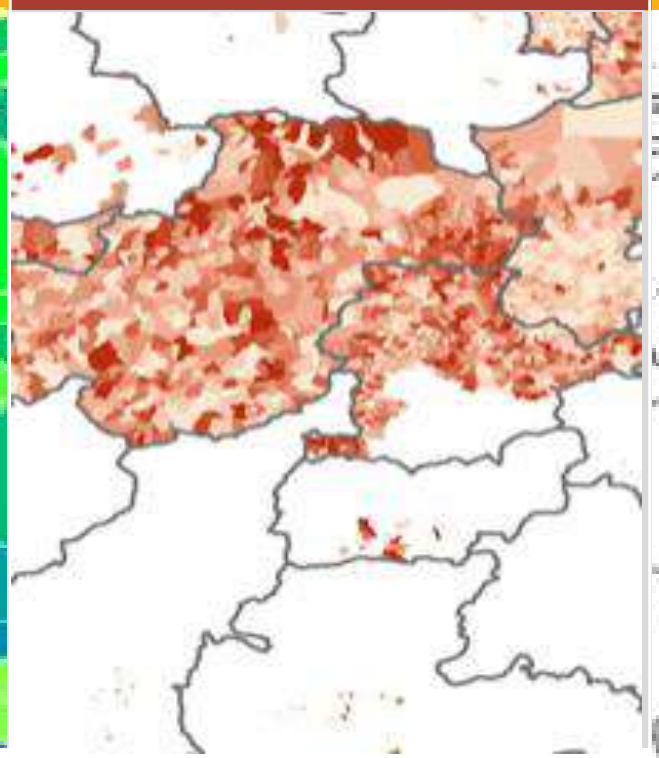


Zoom Out | Legend | Overview | Comparison View

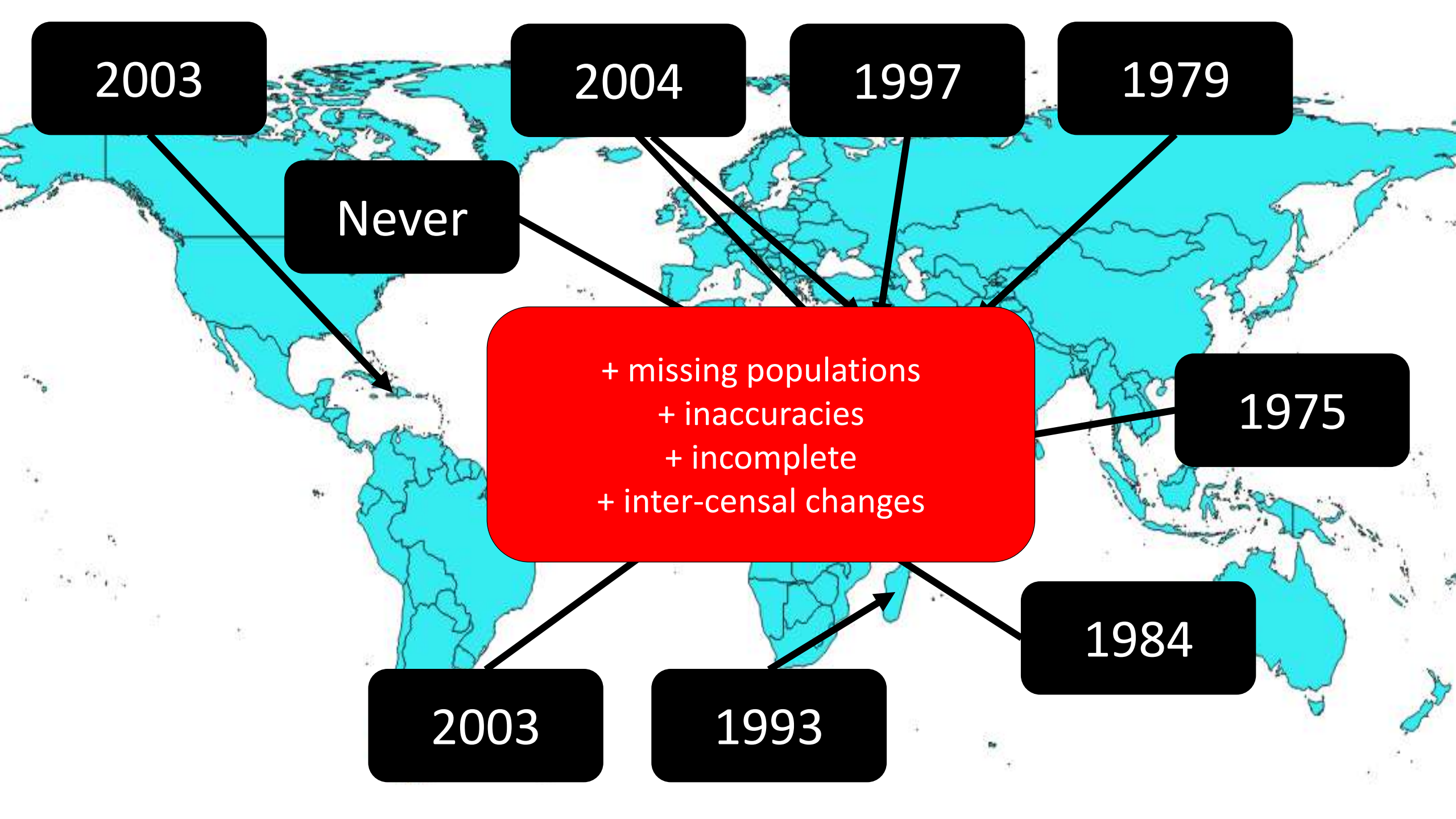
www.popgrid.org

Coarse resolution	Outdated	Incomplete	Inaccuracies, missing populations

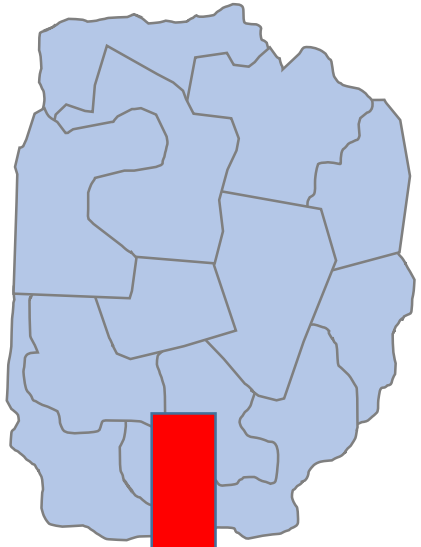


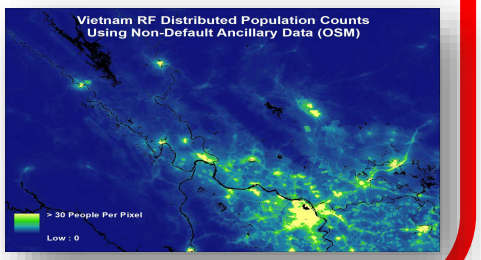


Demographic data challenges


2003
2004
1997
1979
Never
+ missing populations
+ inaccuracies
+ incomplete
+ inter-censal changes
1975
1984
2003
1993

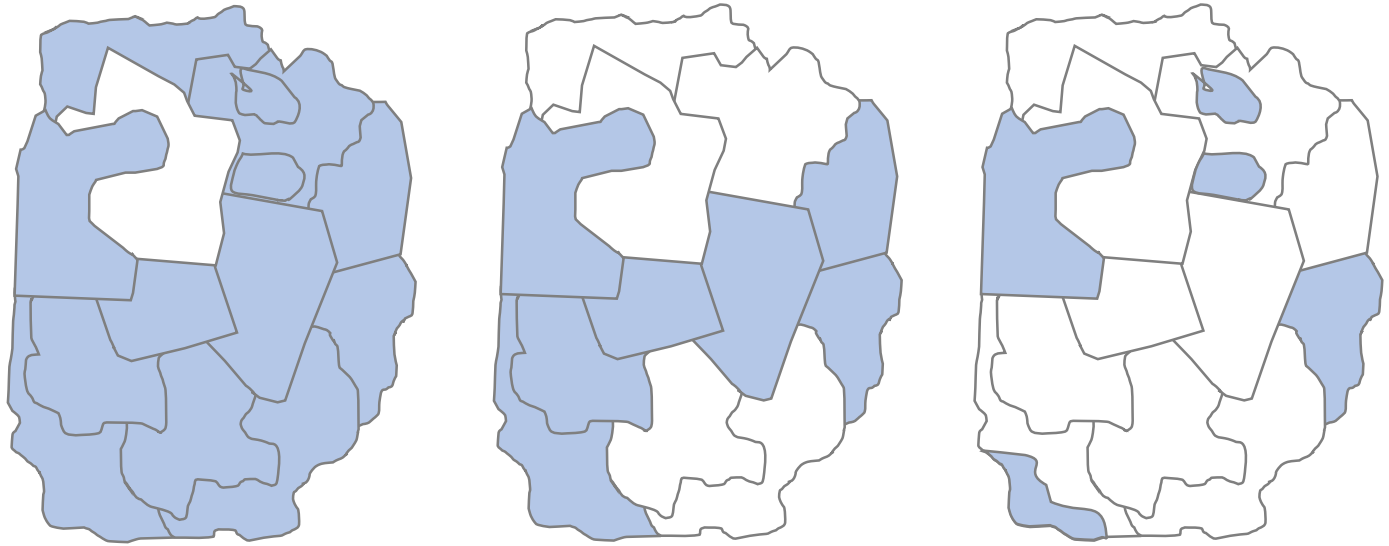
National Census/Projections /Estimates





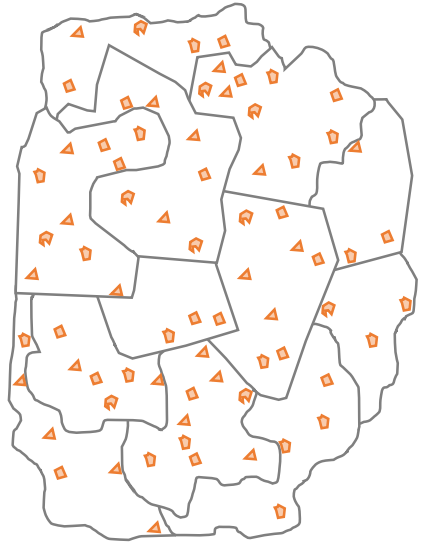
complete na[illegible]nal coverage

Hybrid Methods to Complete Census Count

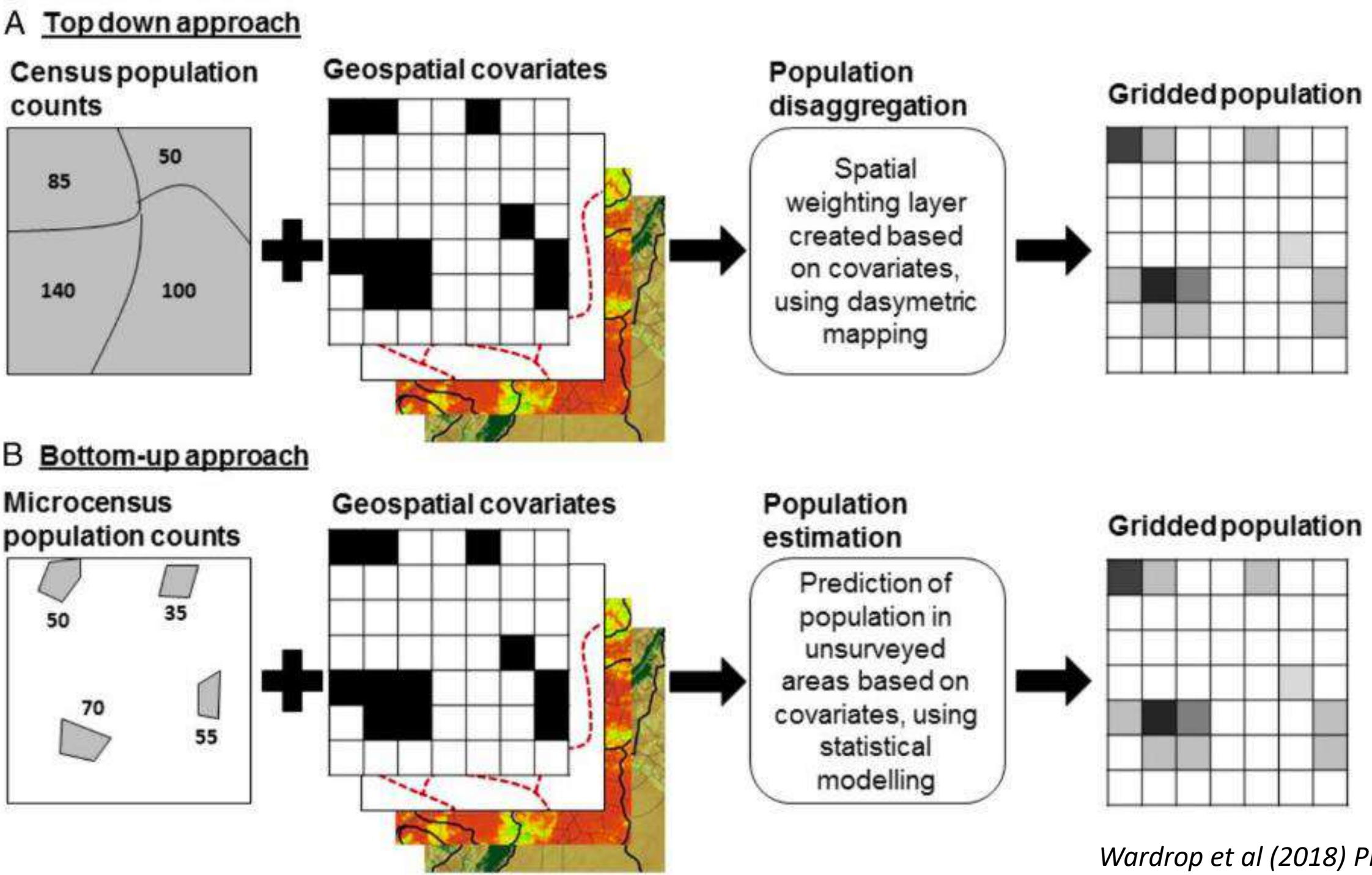


partial enumeration e.g. rolling census

Full Population Model



enumerated population in limited sampling units only

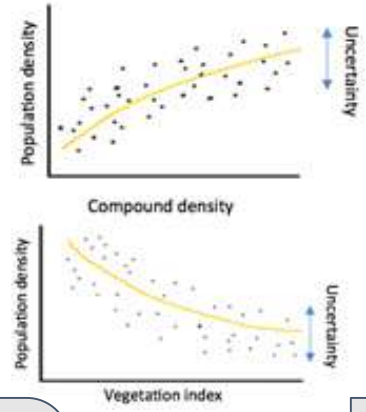


Wardrop et al (2018) PNAS

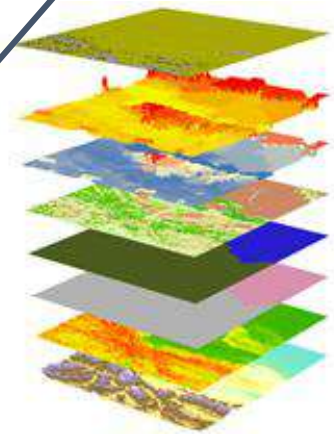
‘Bottom-up’ population modelling approach

Population Data

Settlement

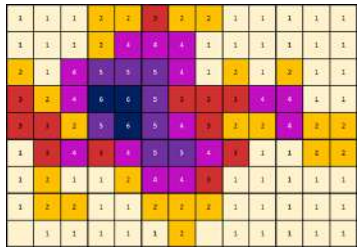


Geospatial Covariates



Modelling Processes

Population Estimates

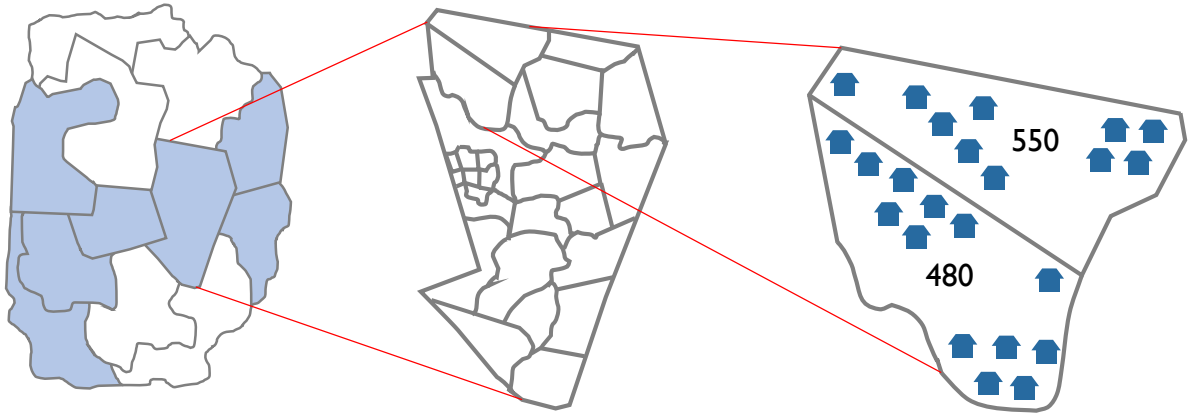


Wardrop et al (2018) PNAS

Bottom-up population modelling

Population Data

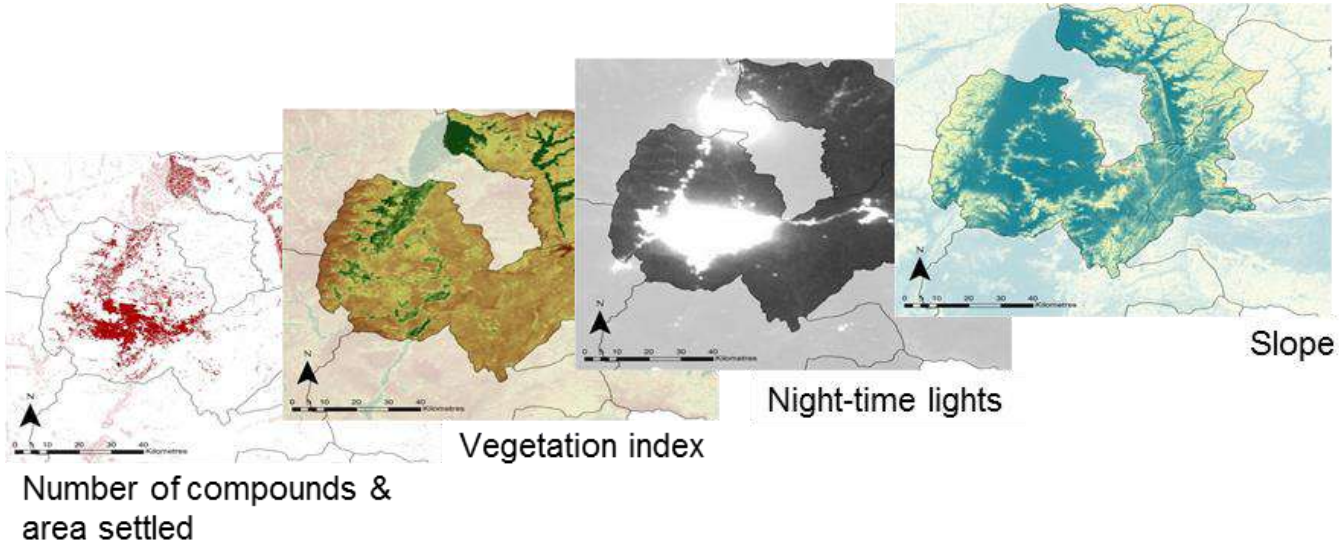
→ Full enumeration of population within a geographically well-defined area



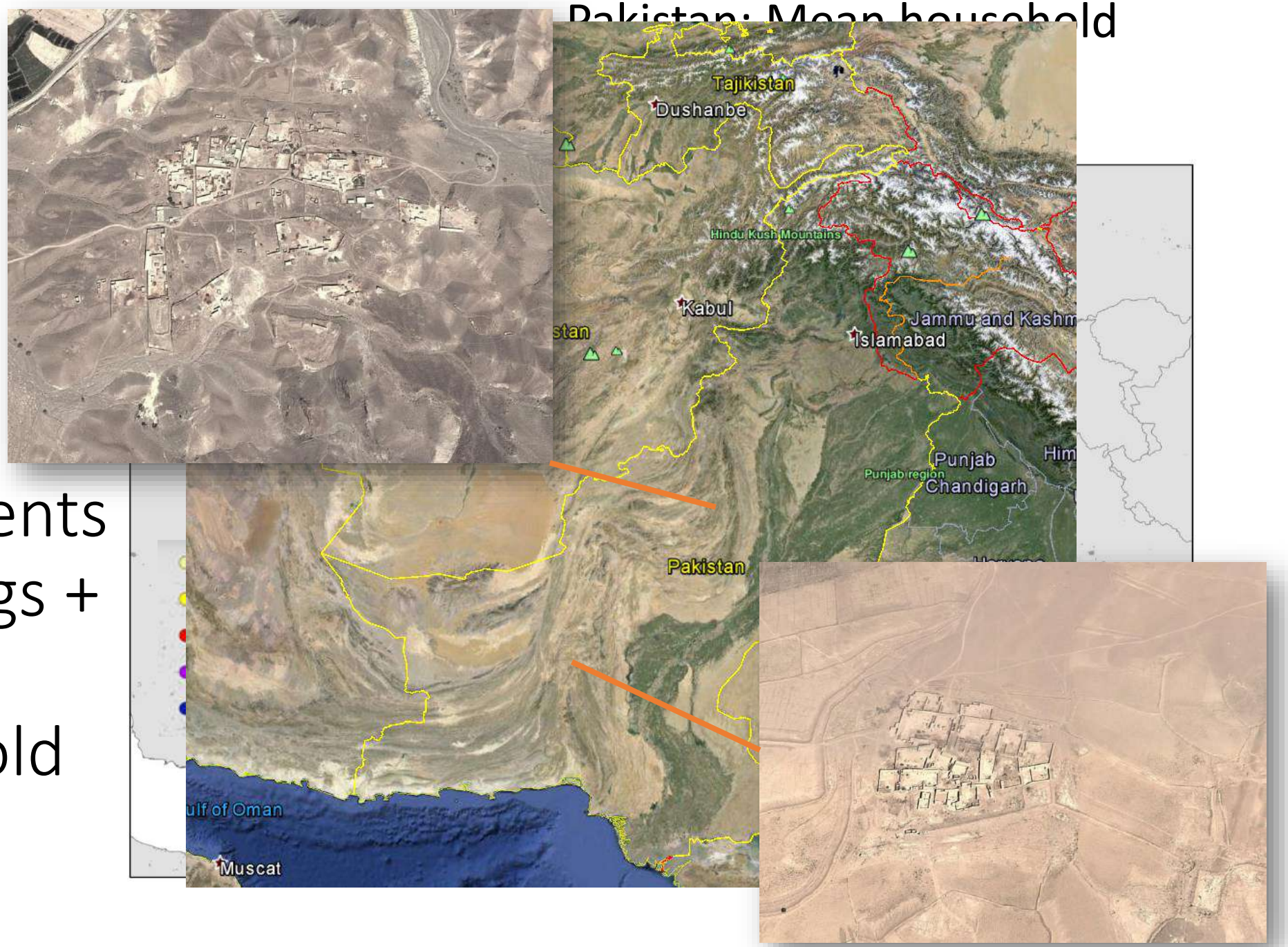
Settlement

Geospatial Covariates

→ Explanatory inputs: full coverage at the resolution of the final, gridded pop map

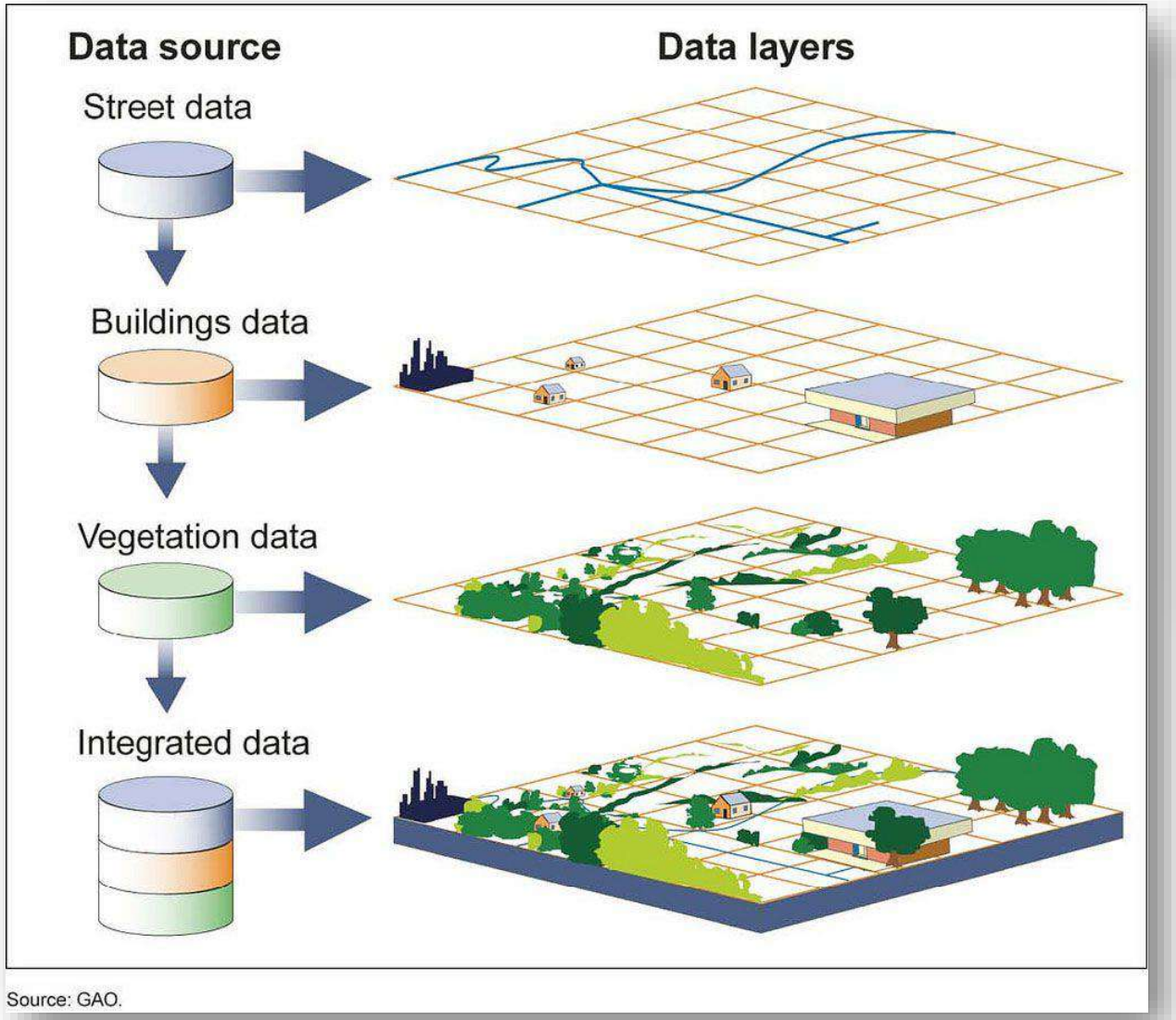


Pakistan: Mean household

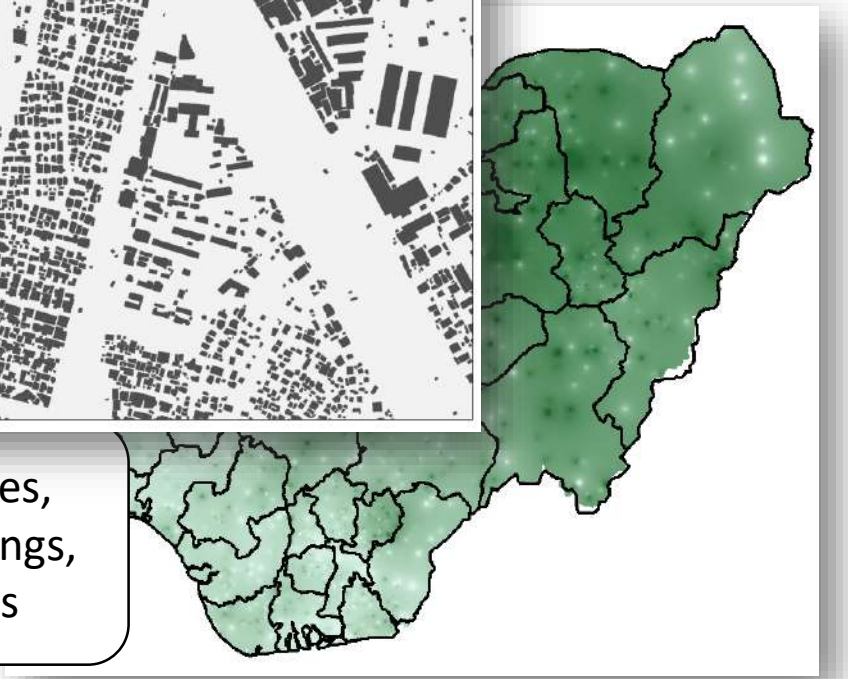


Settlements /Buildings + mean household size?

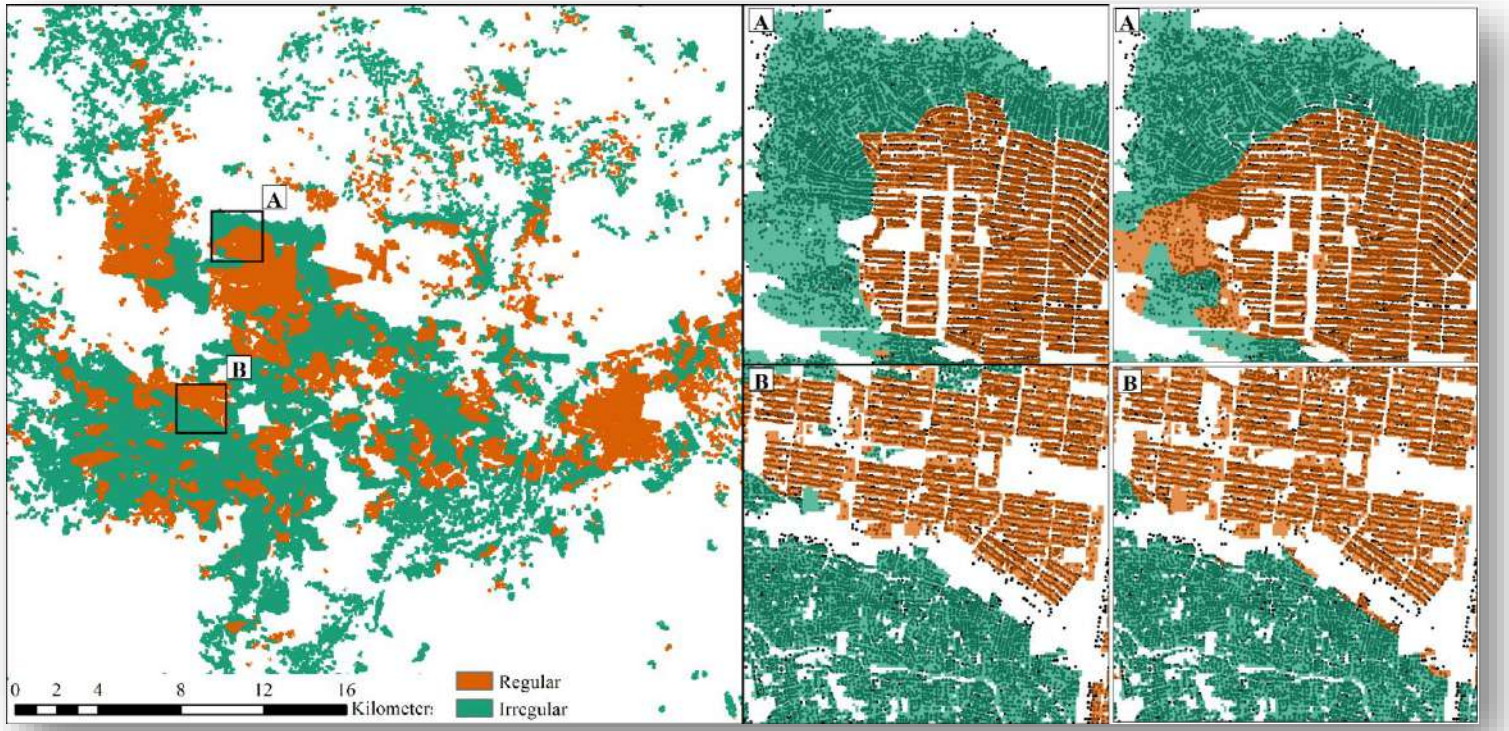
Capturing characteristics that determine variations in population density



Densities of schools, roads, market places, conflicts etc



Household sizes, regional groupings, poverty rates



Settlement/neighbourhood classifications

-Land use data; OpenStreetMap; Building heights; Mobile network data

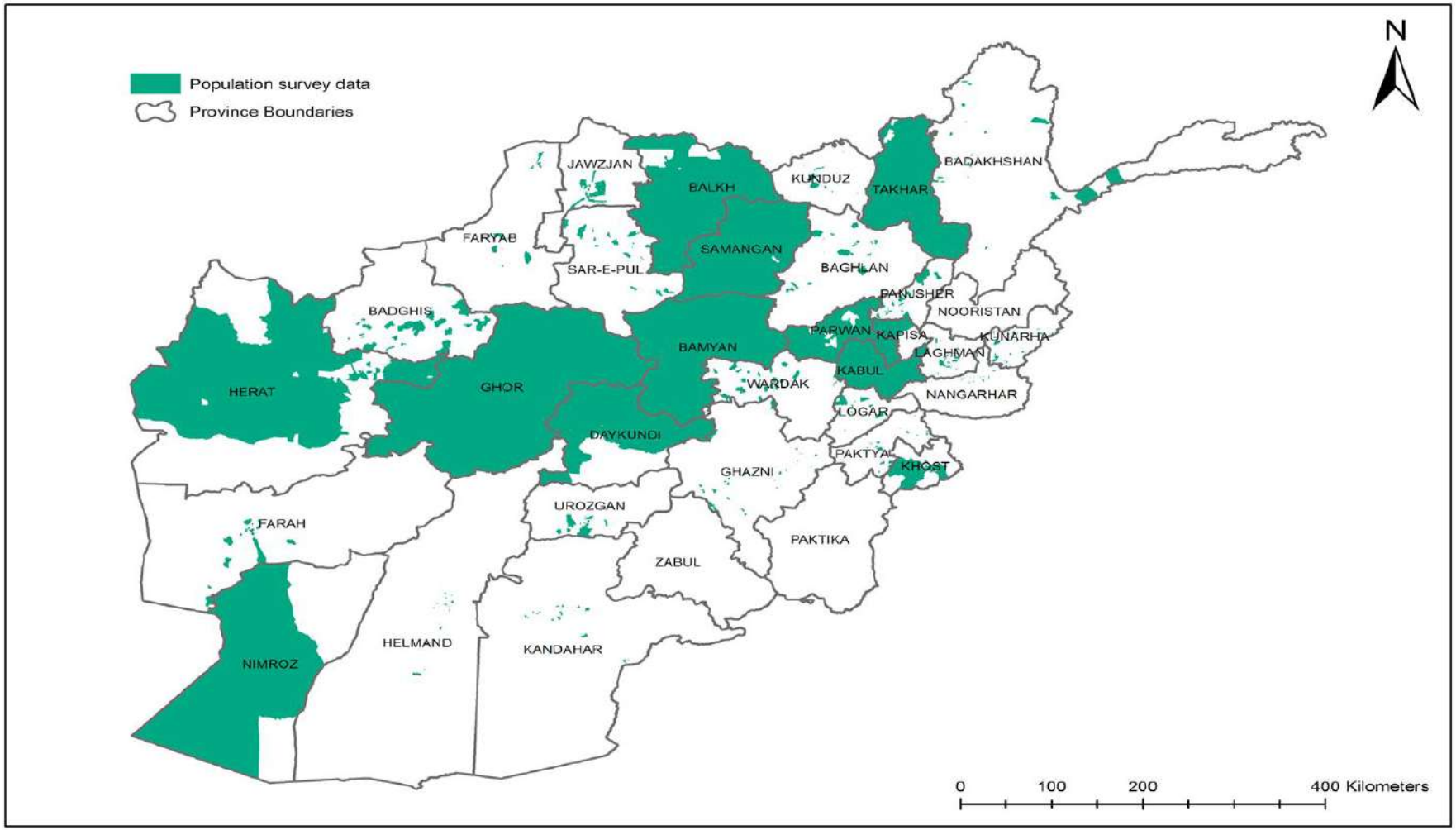
Jochem et al, CEUS (2018), in review (2019)

Afghanistan example

Population Data

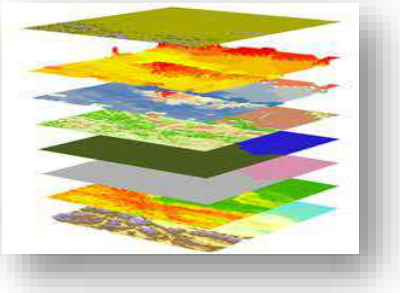
Settlement

Geospatial Covariates



Chamberlain et al (2019) in review

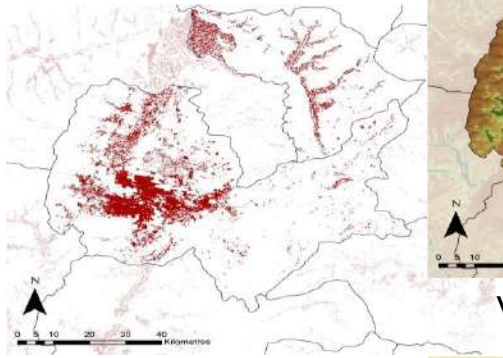
Afghanistan example



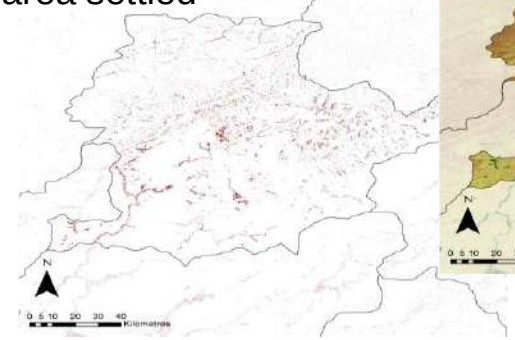
Population Data

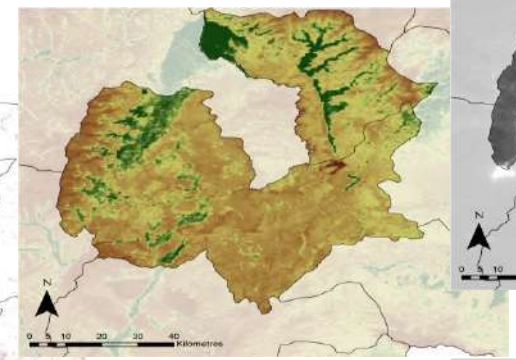
Settlement

Geospatial Covariates

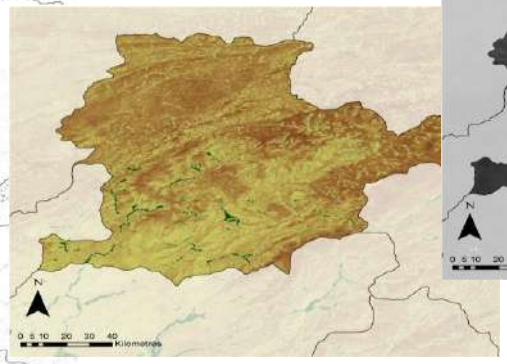


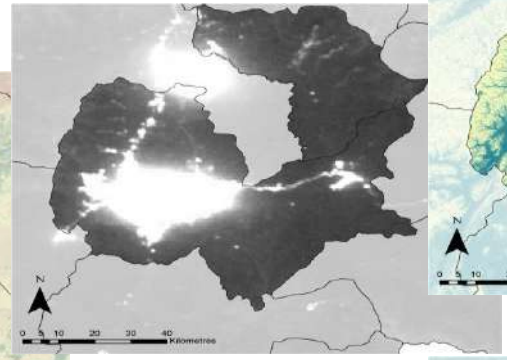
Number of compounds & area settled



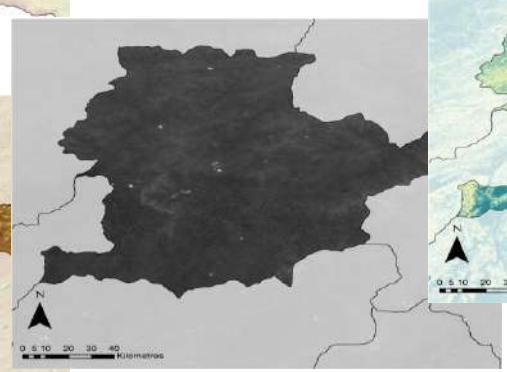


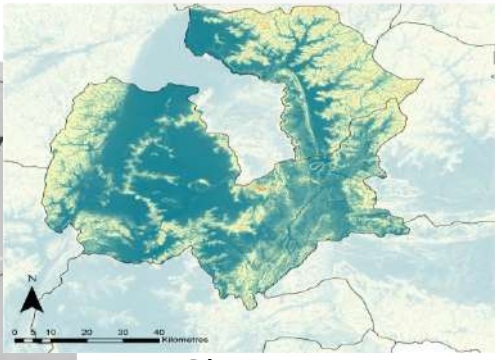
Vegetation index



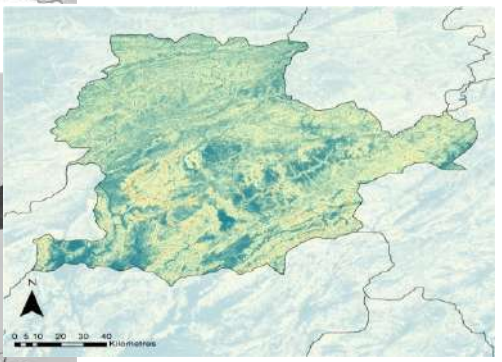


Night-time lights



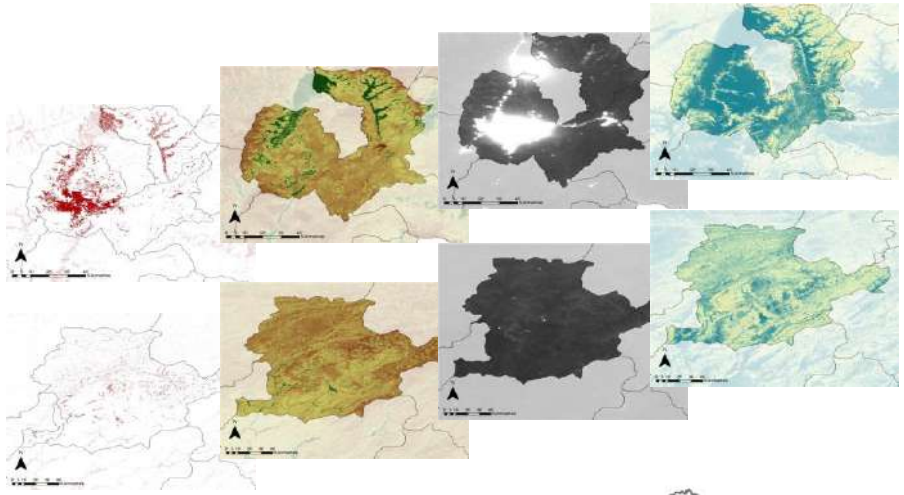


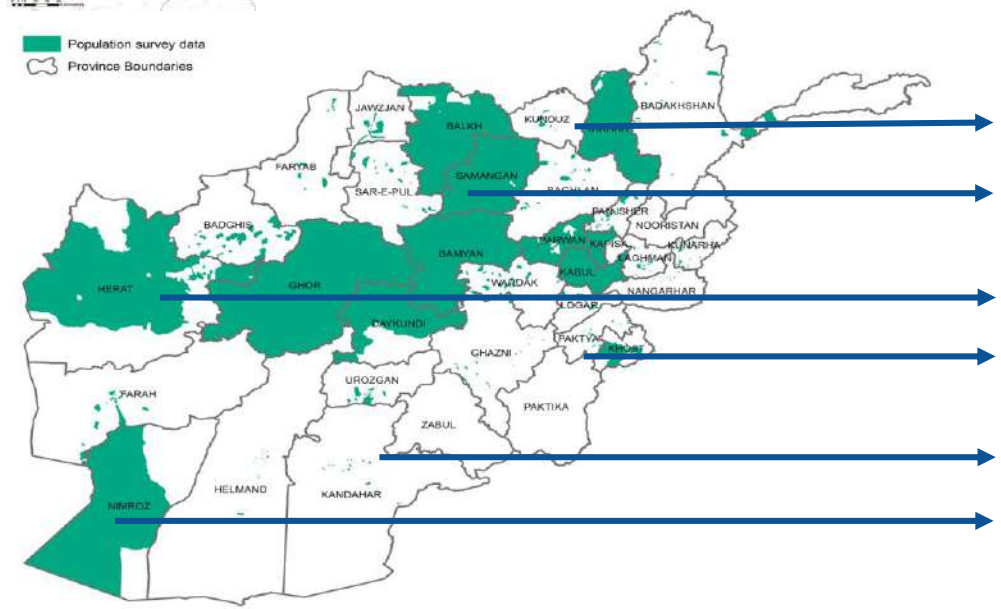
Slope

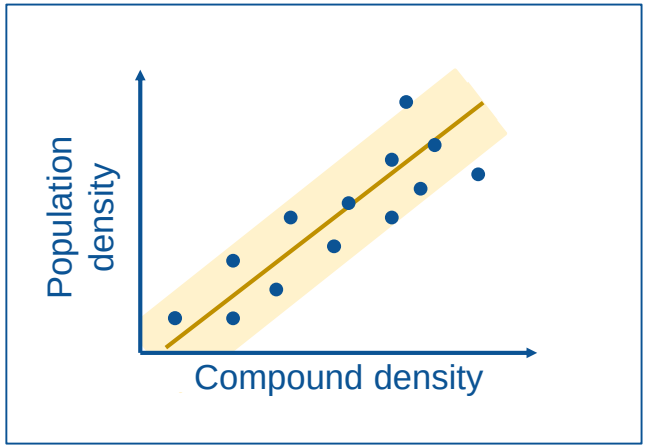


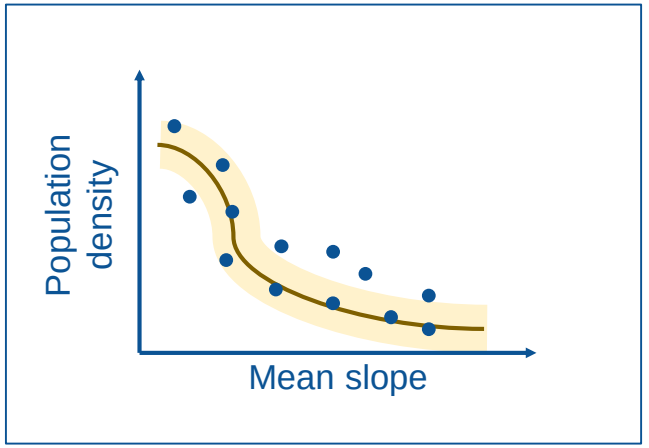
Chamberlain et al (2019) in review

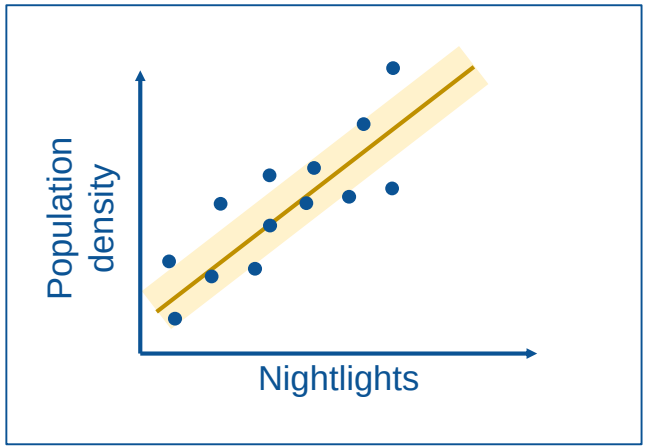
Afghanistan example

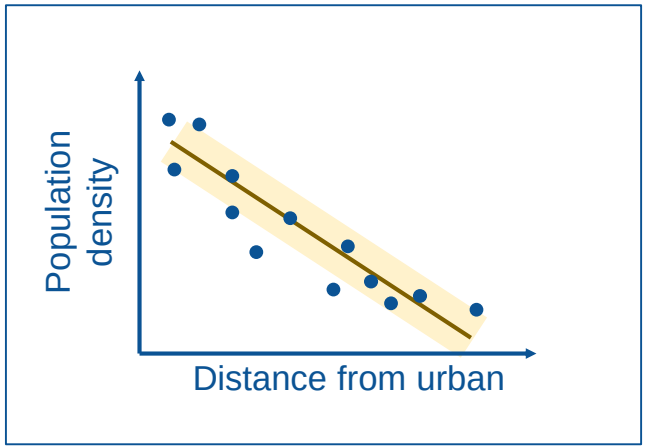












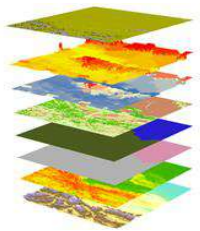
Chamberlain et al (2019) in review

Afghanistan example

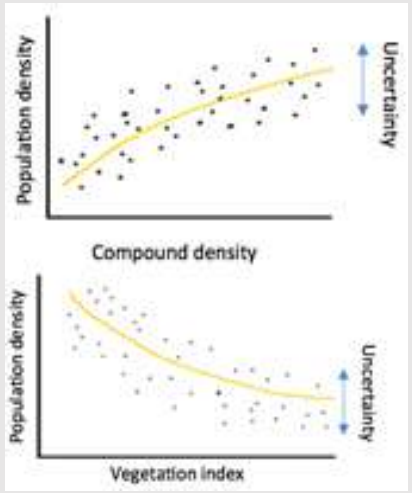
Population Data

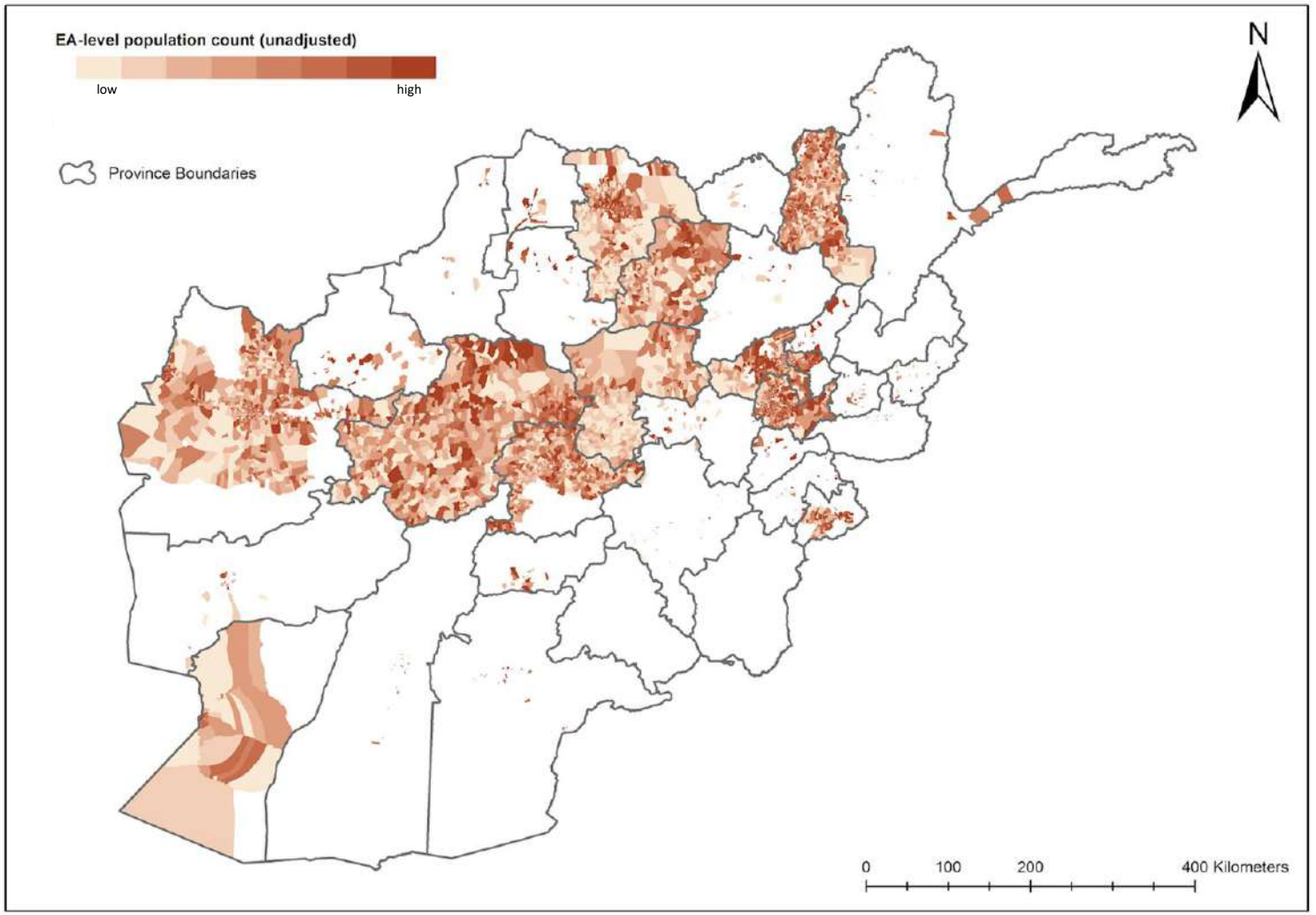
Settlement

Geospatial Covariates



Population model building and prediction





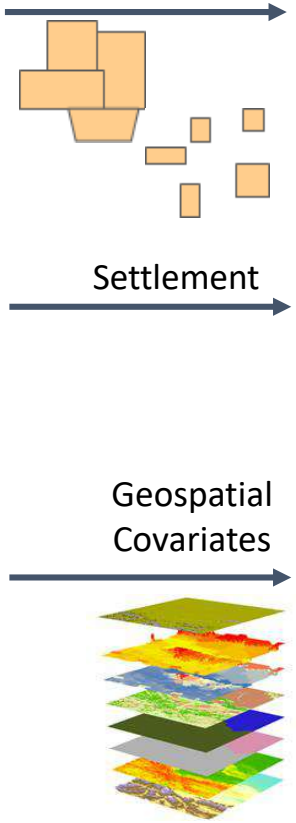
Chamberlain et al (2019) in review

Afghanistan example

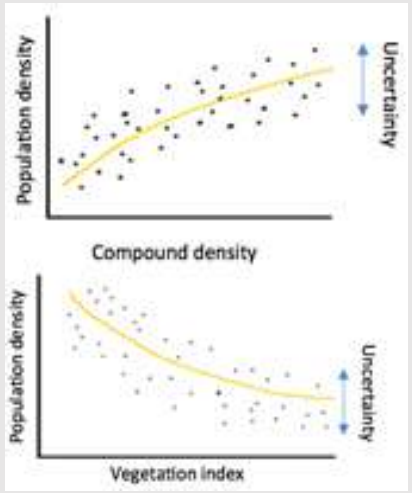
Population Data

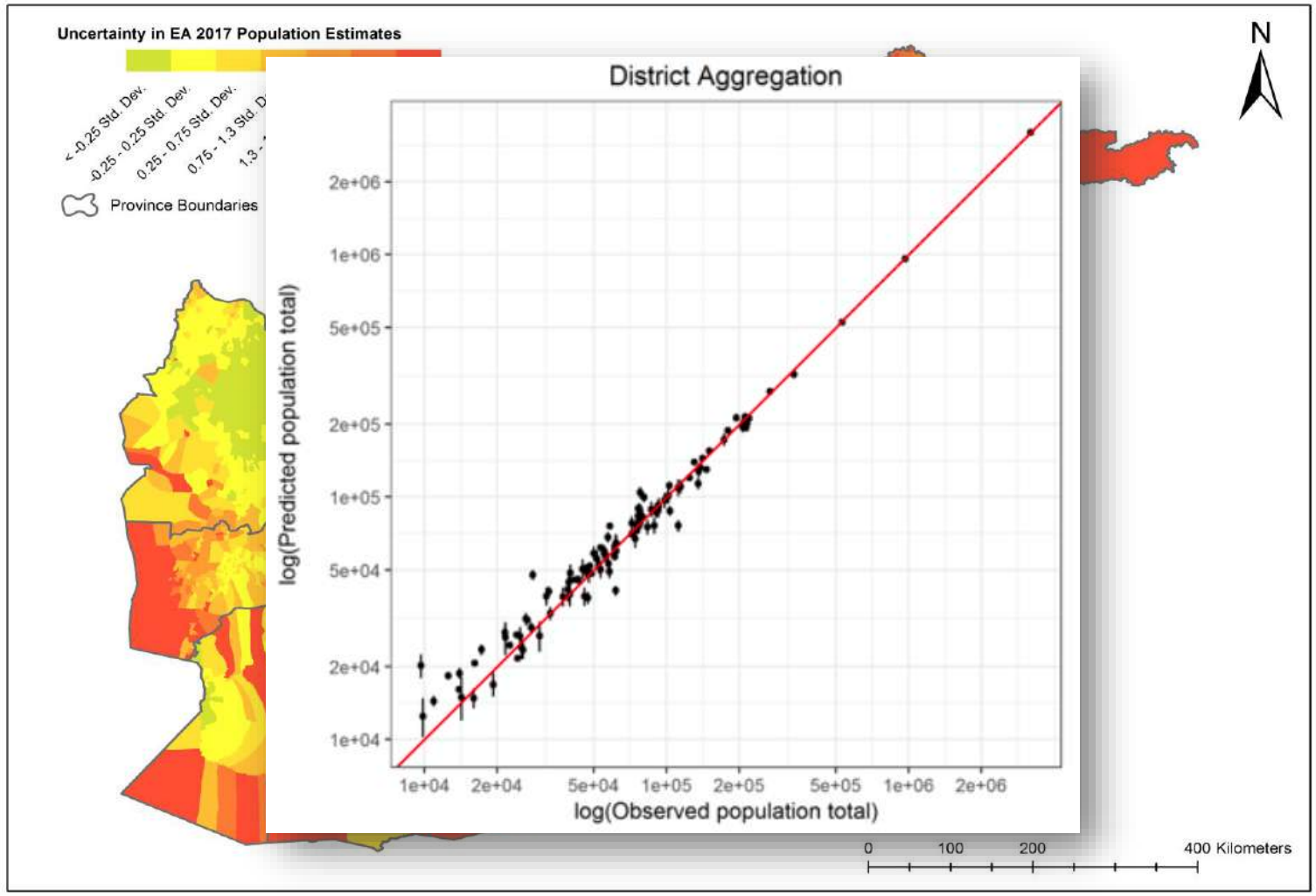
Settlement

Geospatial Covariates

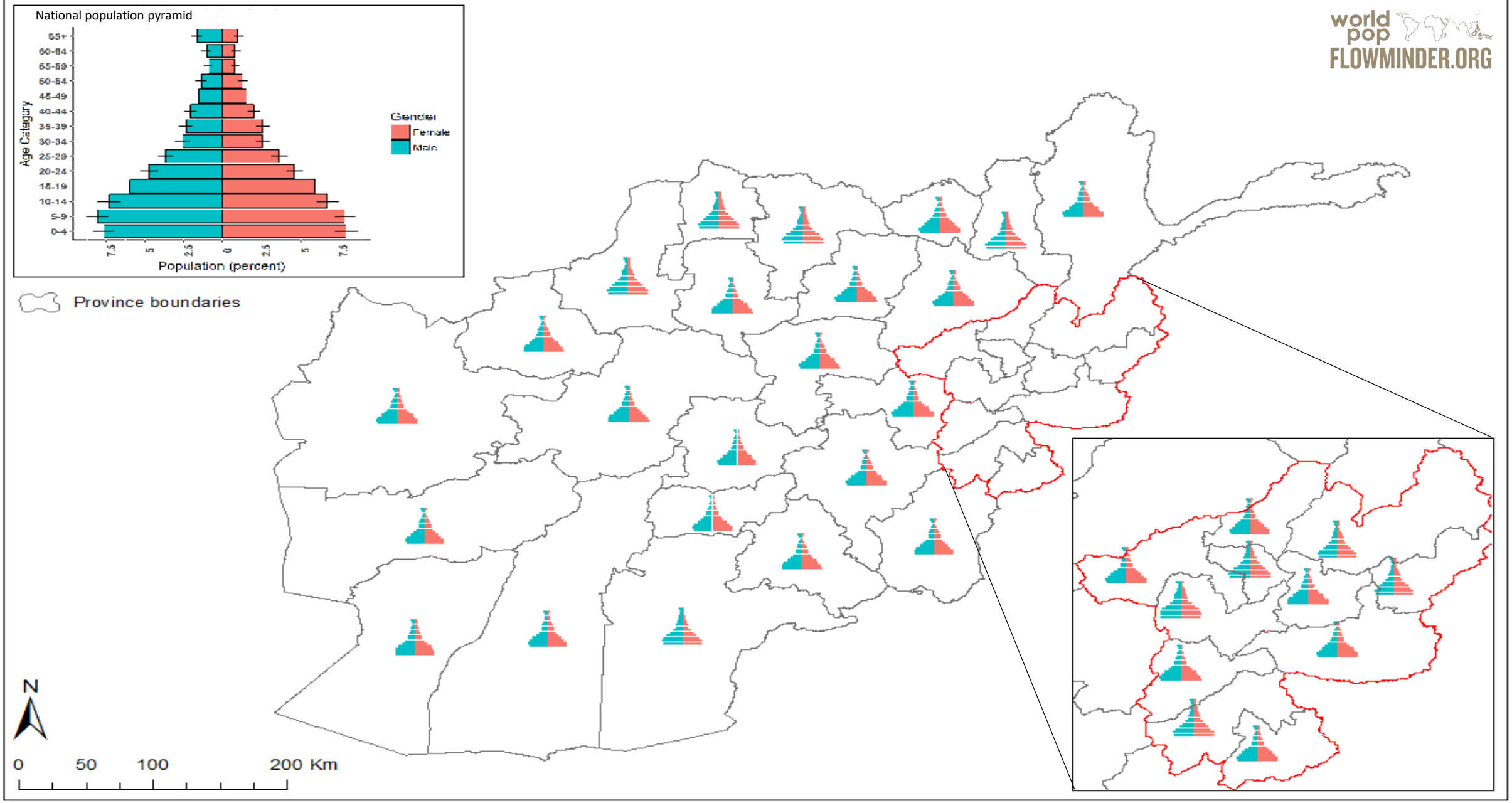


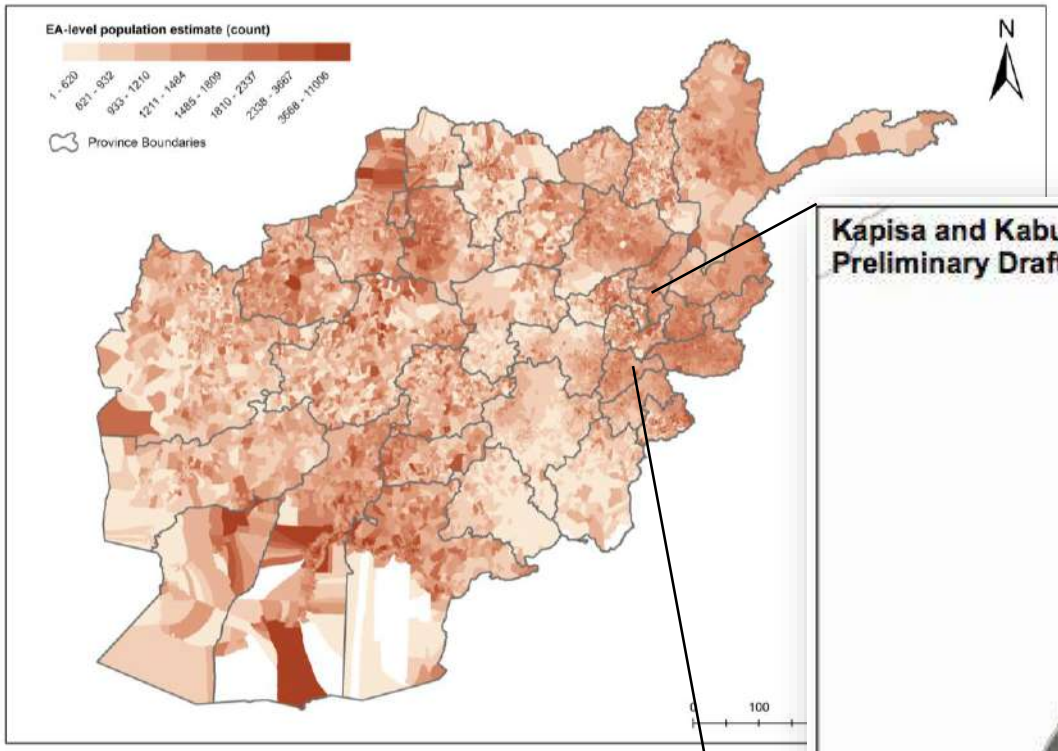
Population model building and prediction

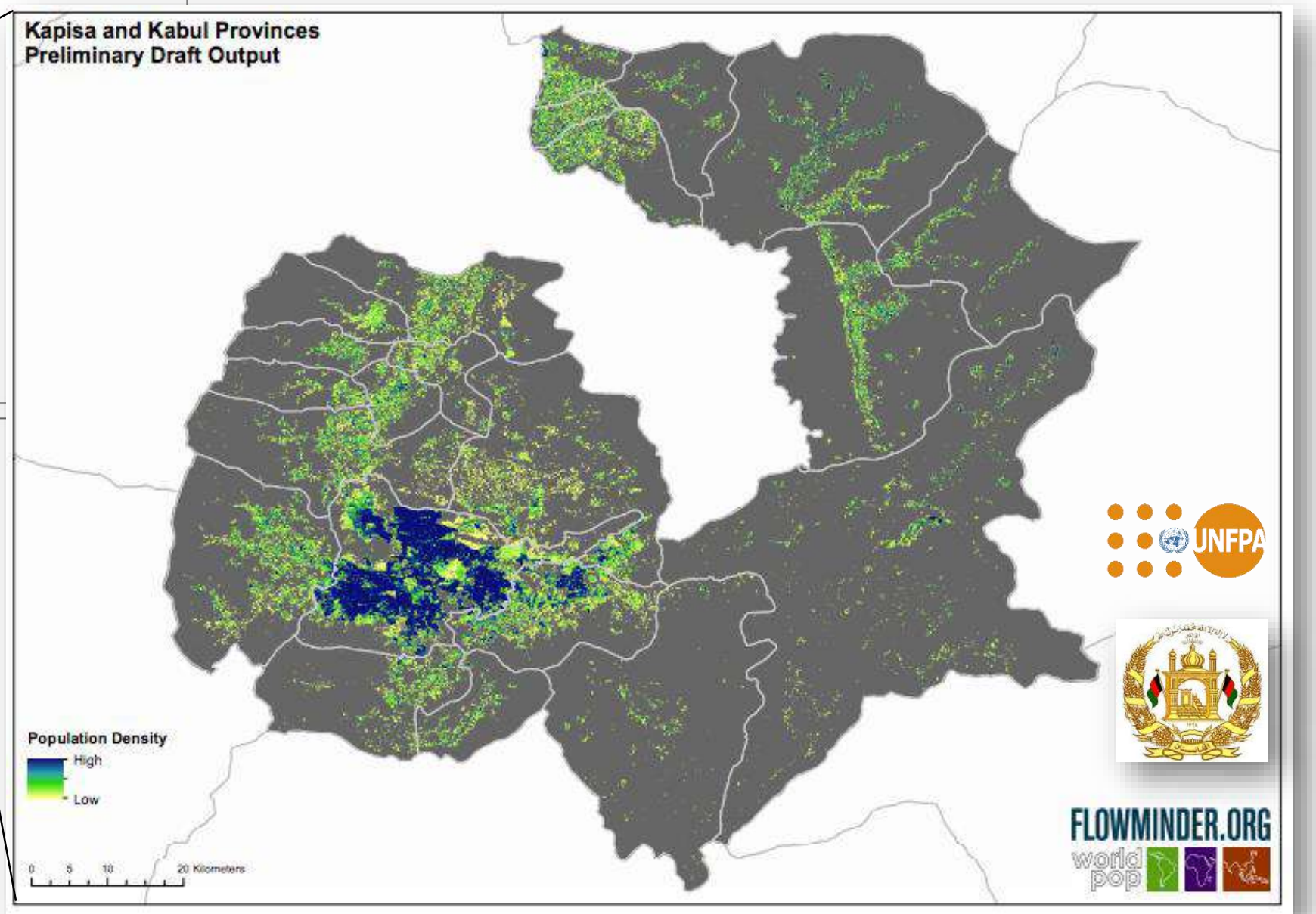




Chamberlain et al (2019) in review

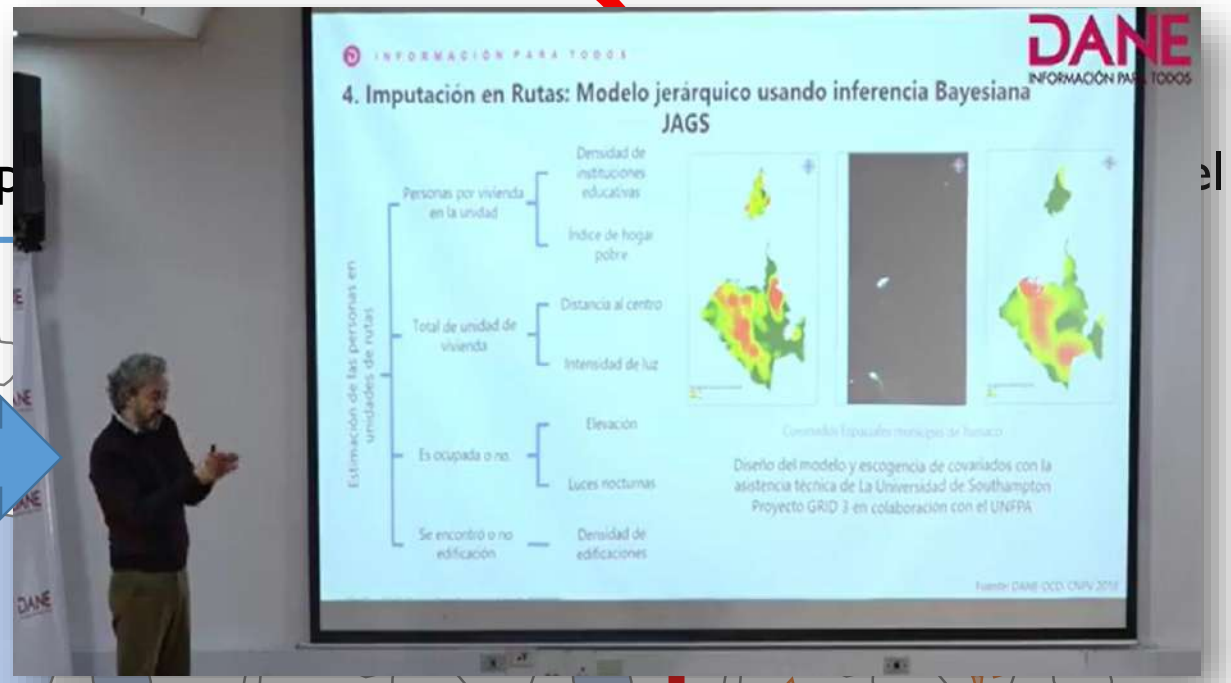

National population pyramid
65+
60-64
55-59
50-54
45-49
40-44
35-39
30-34
25-29
20-24
15-19
10-14
5-9
0-4
Age Category
7.5
5
2.5
0
2.5
5
7.5
Population (percent)
Gender
Female
Male
world pop
FLOWMINDER.ORG
Province boundaries
N
0
50
100
200 Km

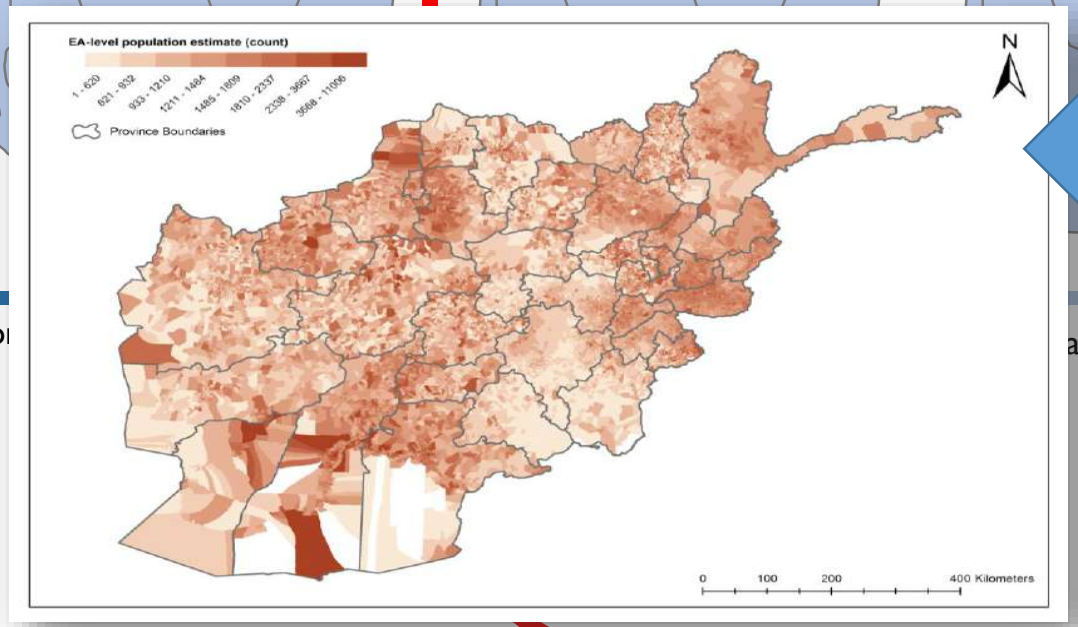

EA-level population estimate (count)
1 - 620
621 - 932
933 - 1210
1211 - 1484
1485 - 1809
1810 - 2337
2338 - 3667
3668 - 11006
Province Boundaries
N
100


Kapisa and Kabul Provinces
Preliminary Draft Output
Population Density
High
Low
0 5 10 20 Kilometers
UNFPA
FLOWMINDER.ORG
world pop

National Census/Projections /Estimates

Hybrid Methods to Comp





co

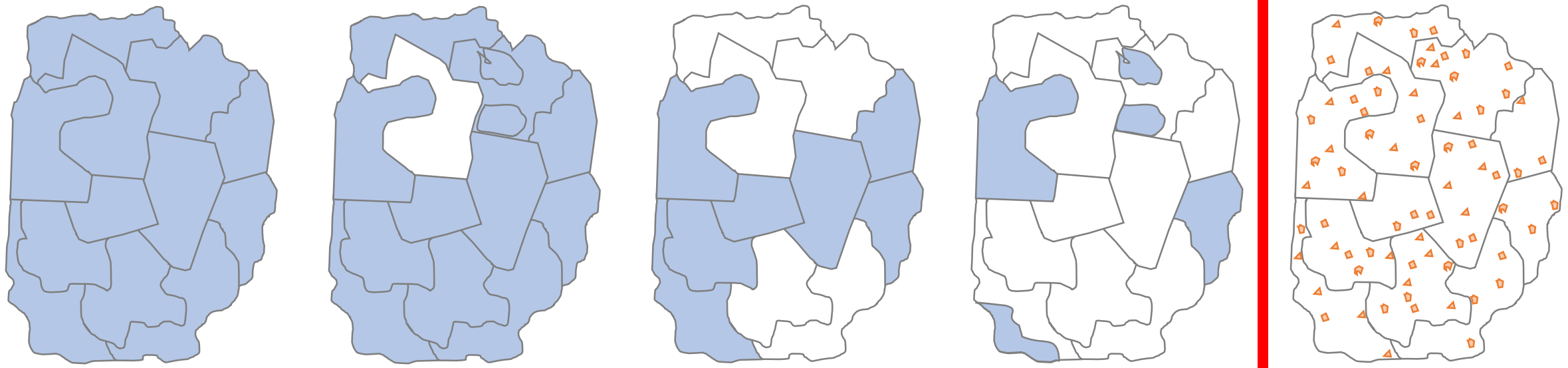
al enumeration e.g. rolling census

enumerated population in limited sampling units only

National Census/Projections /Estimates

Hybrid Methods to Complete Census Count

Full Population Model



complete national coverage

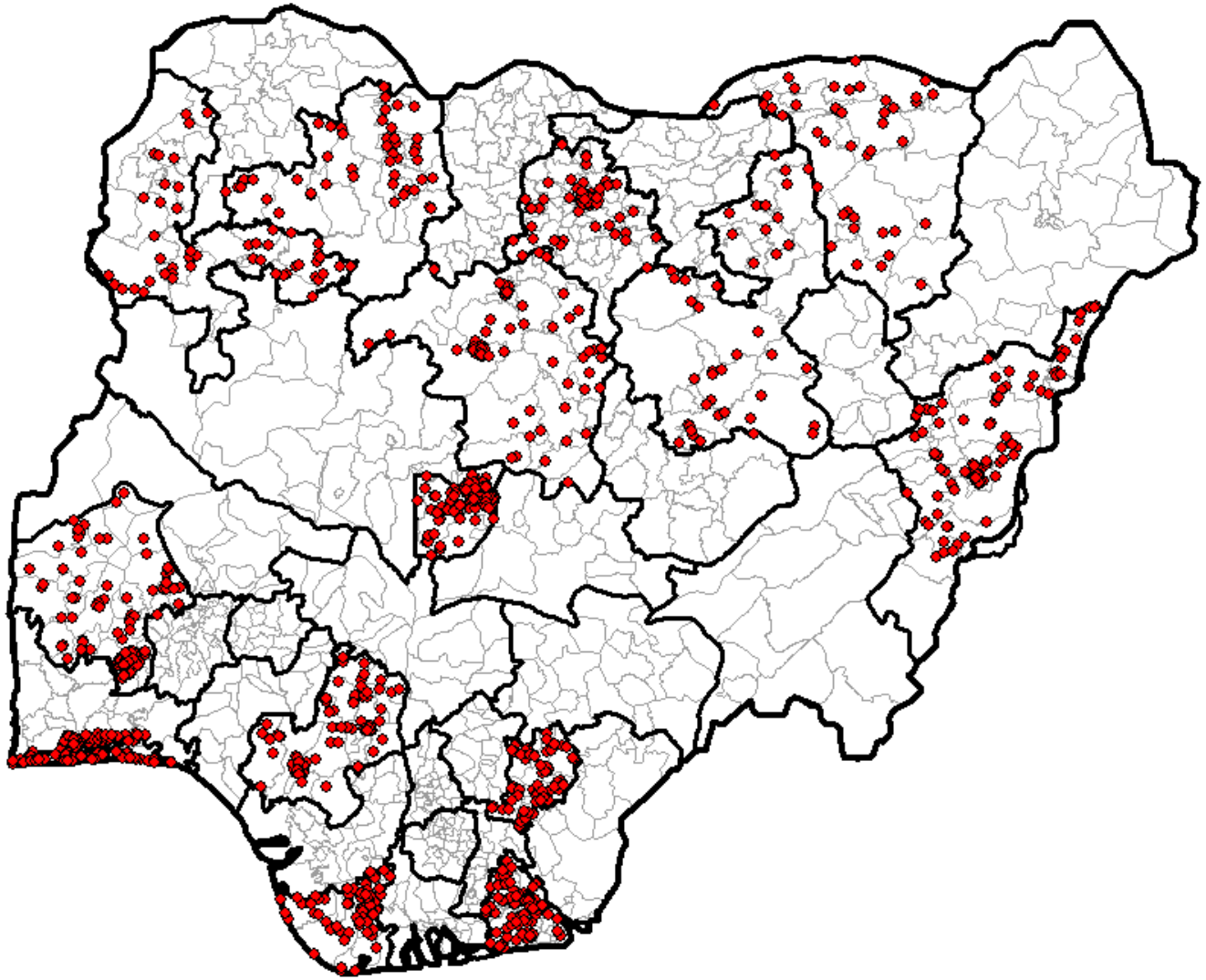
partial enumeration e.g. rolling census

enumerated population in limited sampling units only

Microcensus Surveys

1,143 microcensus locations

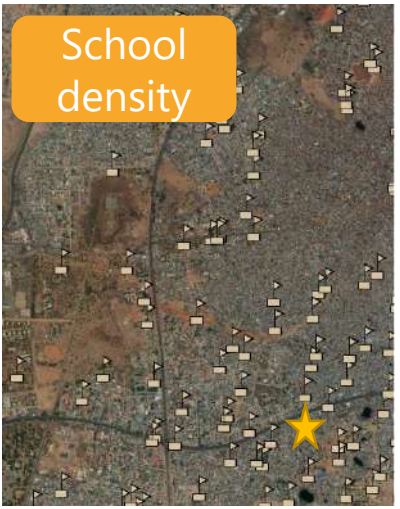
14 of 37 States surveyed



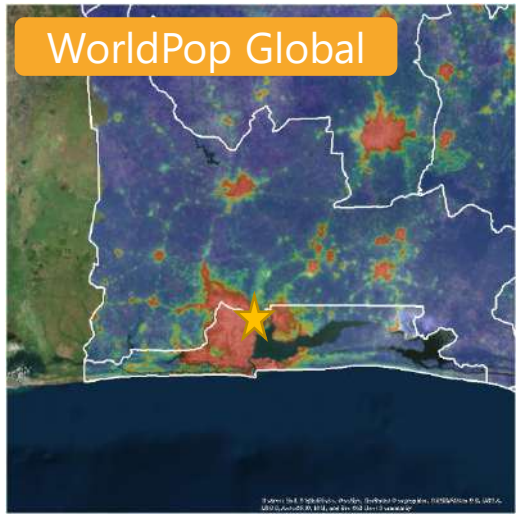
Leasure et al (2019) in review

Hierarchical Bayesian model

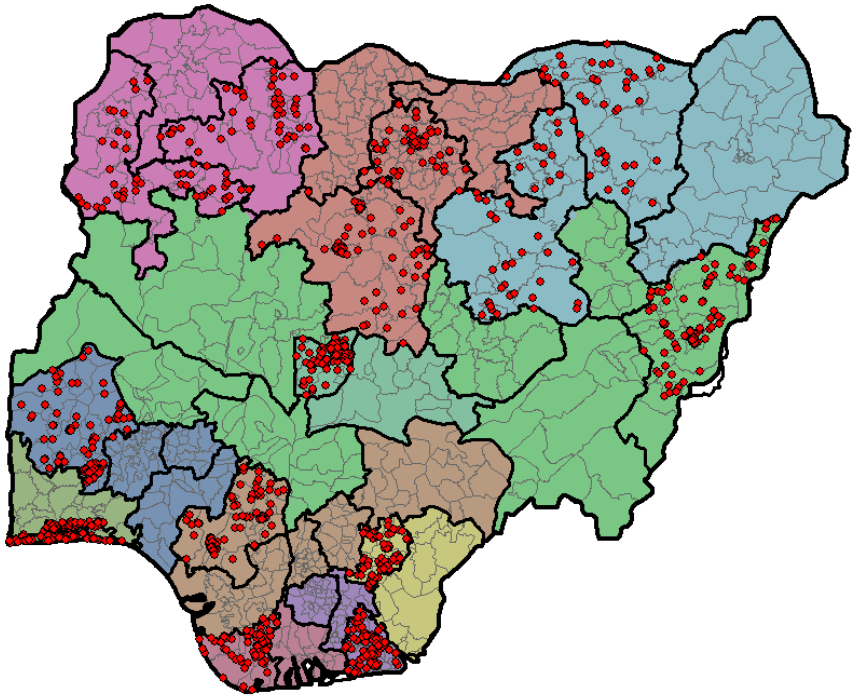
School density

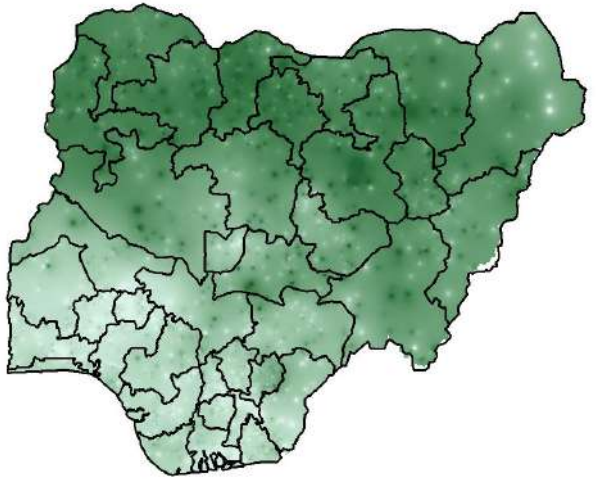


WorldPop Global

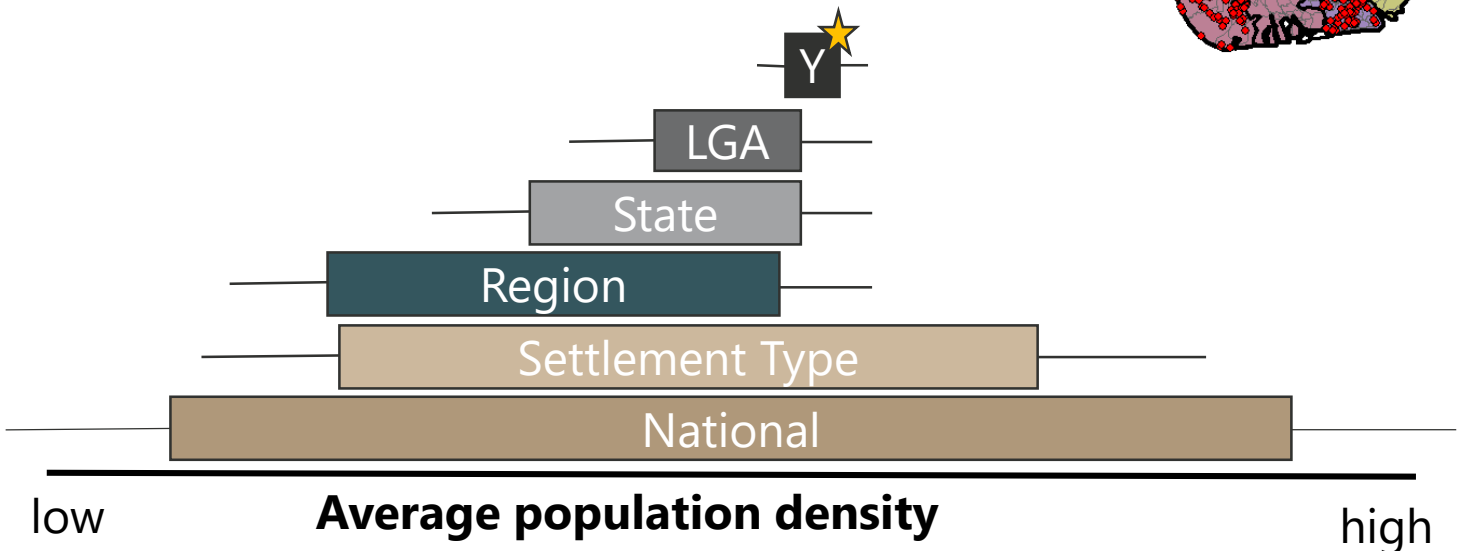


Admin boundaries: Region, State, LGA

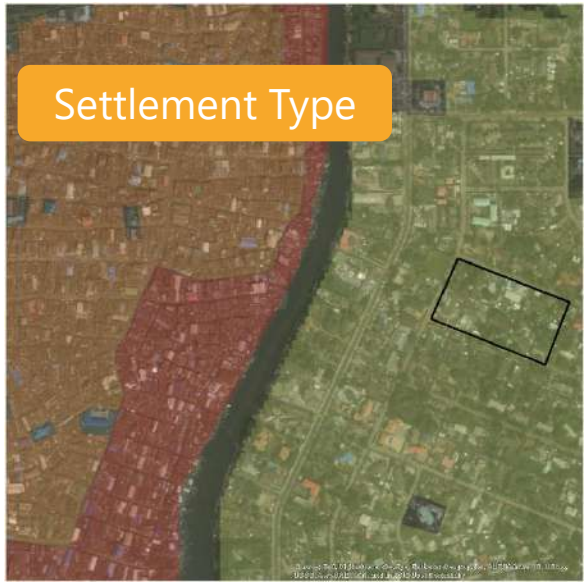




Household size

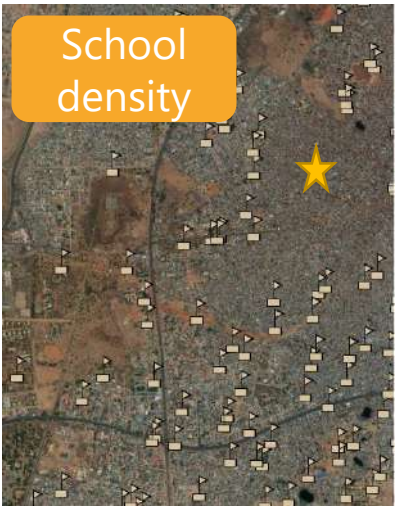


Settlement Type

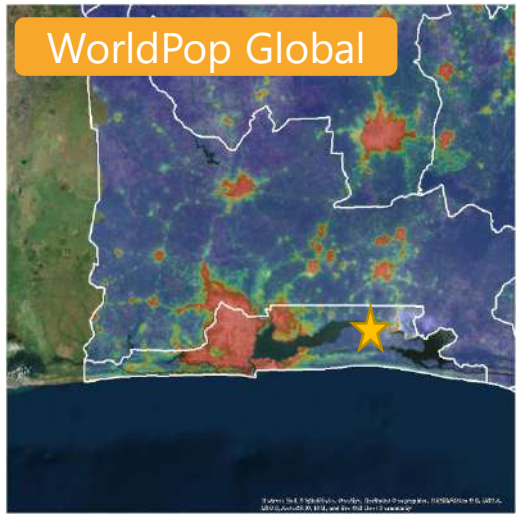


Hierarchical Bayesian model

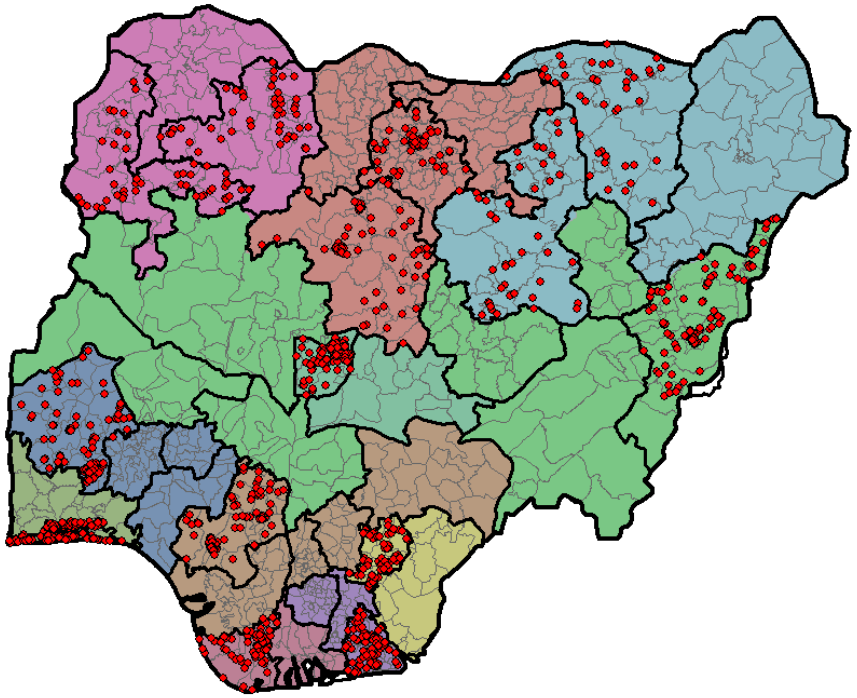
School density

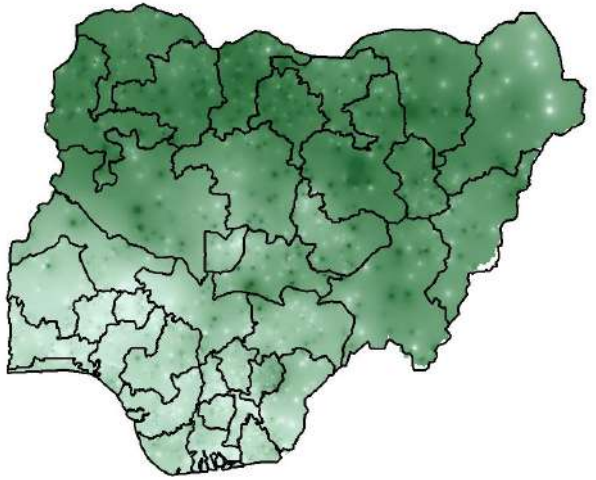


WorldPop Global

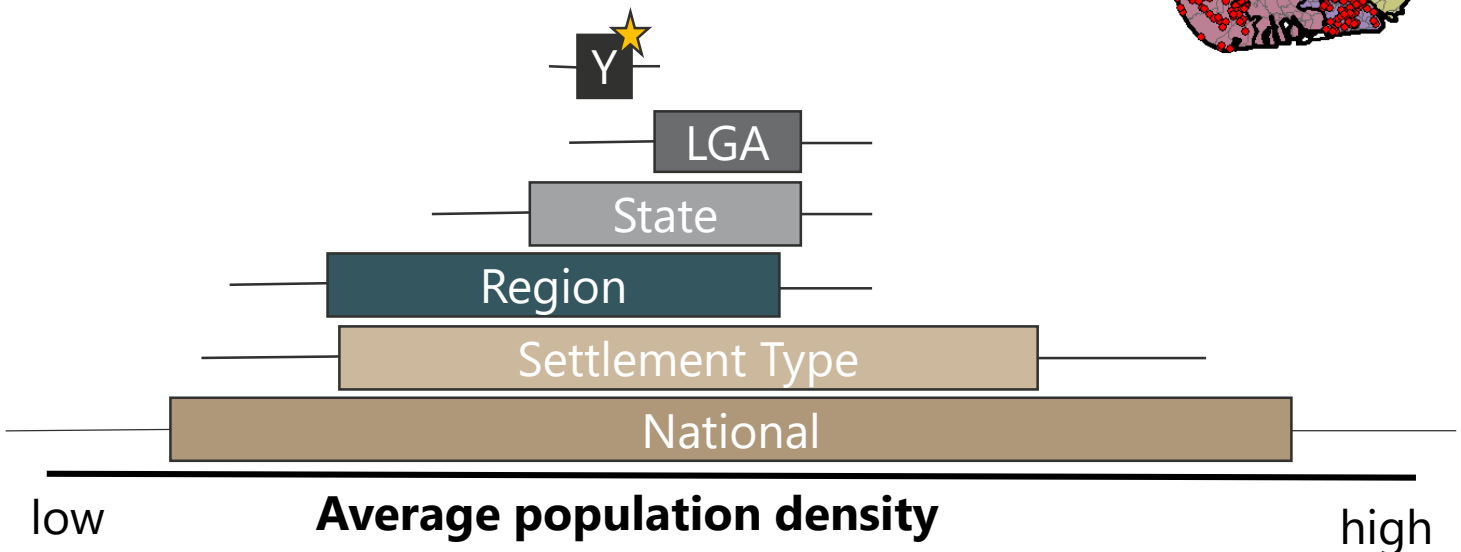


Admin boundaries: Region, State, LGA

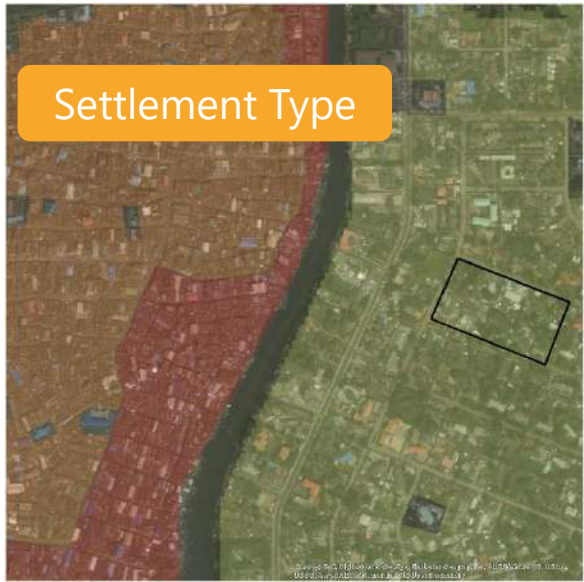


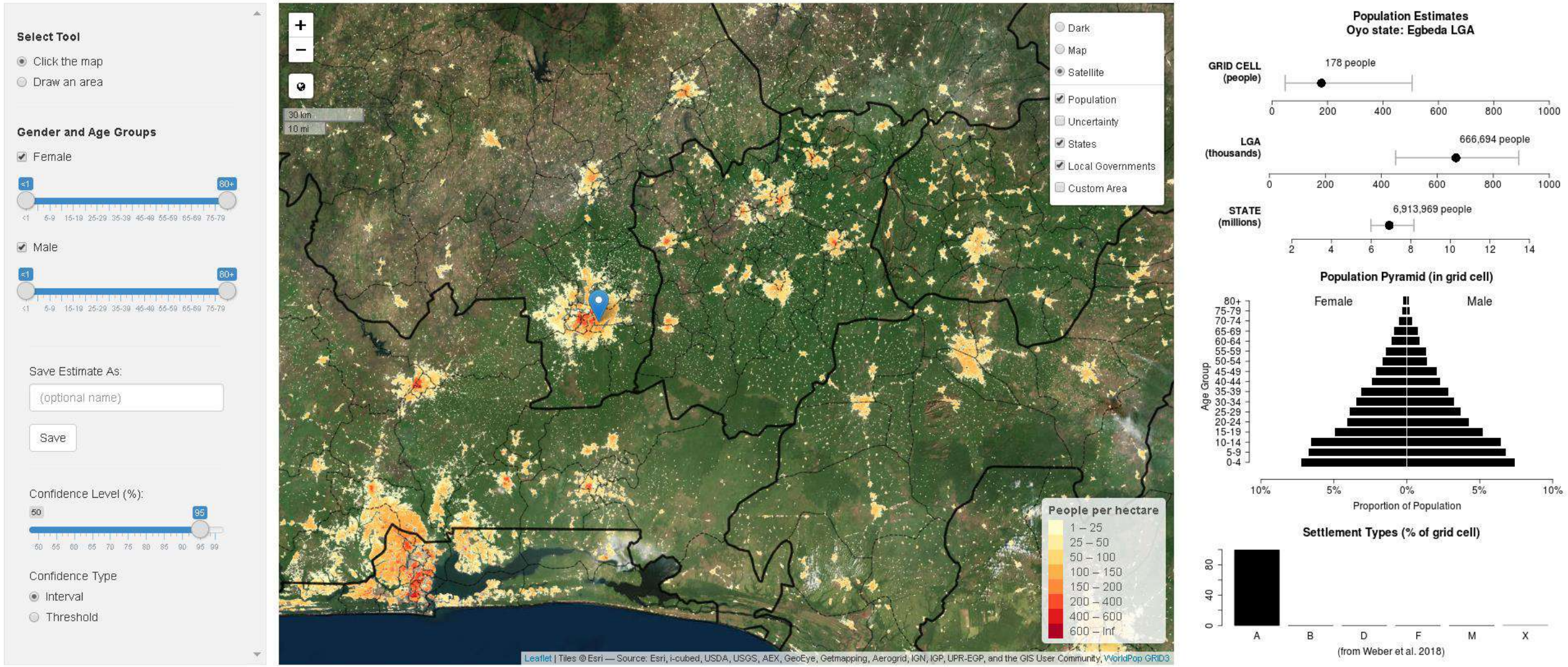


Household size

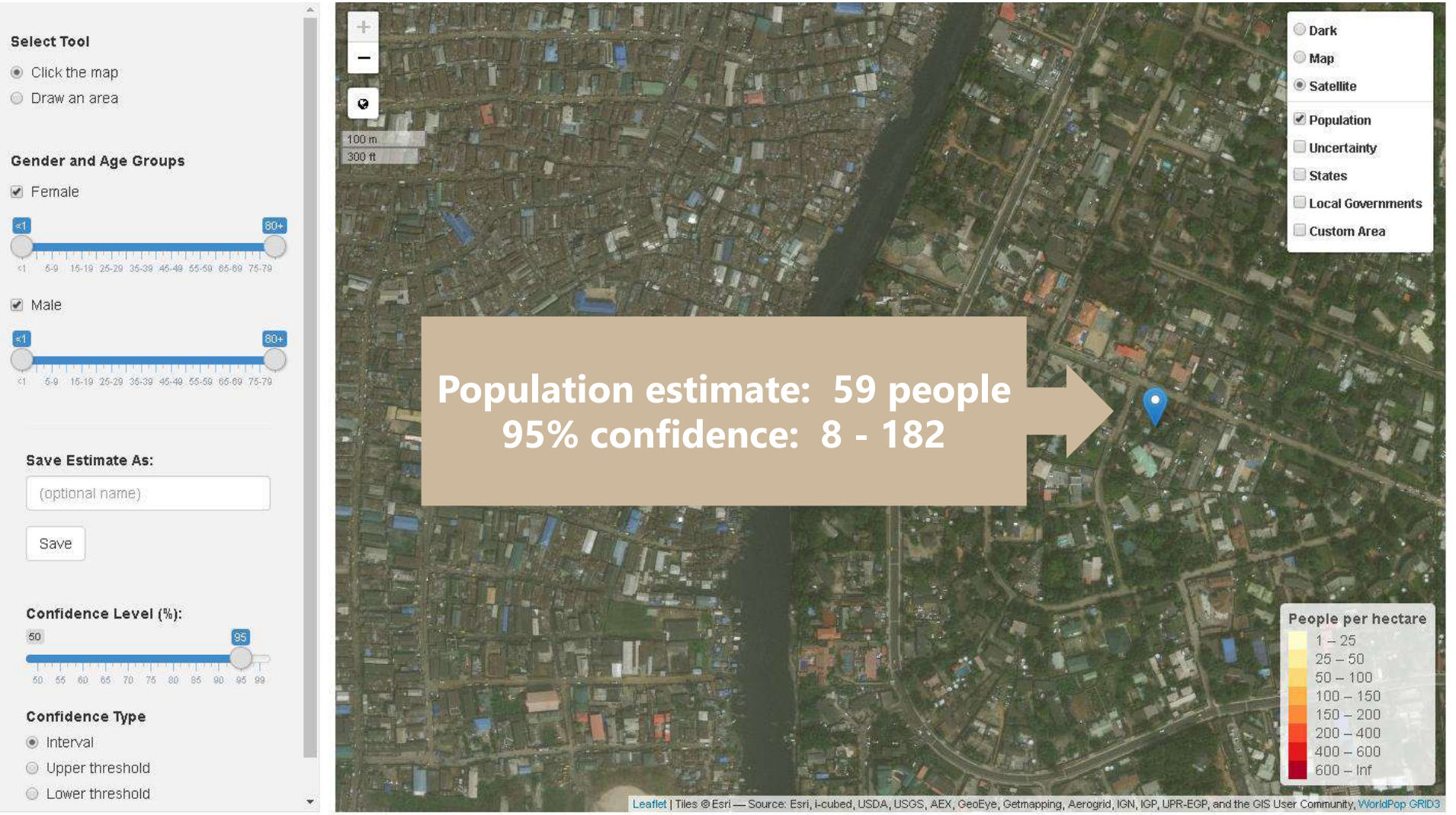


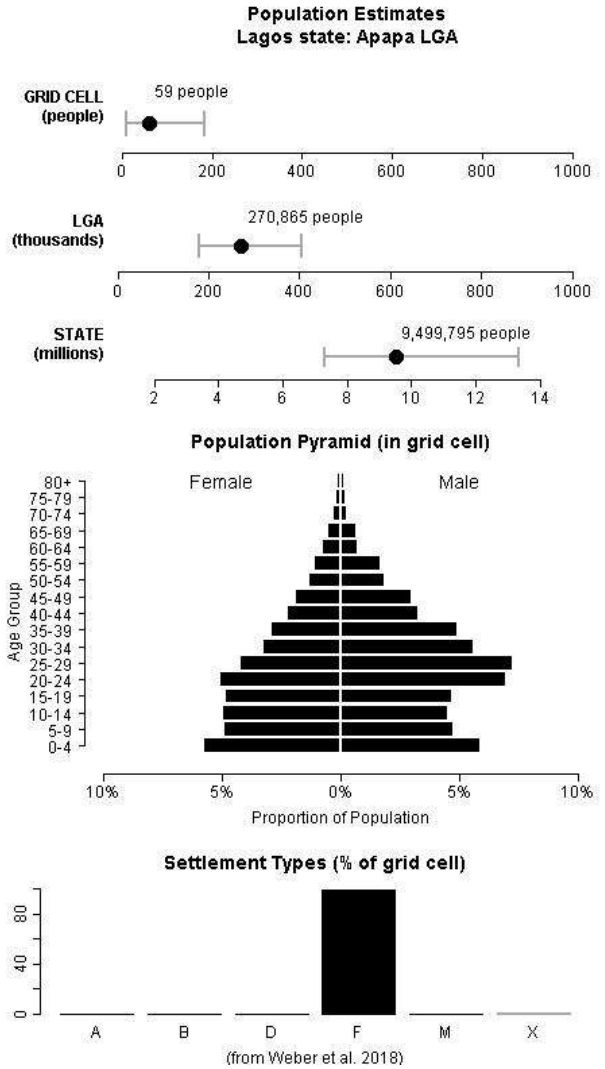
Settlement Type



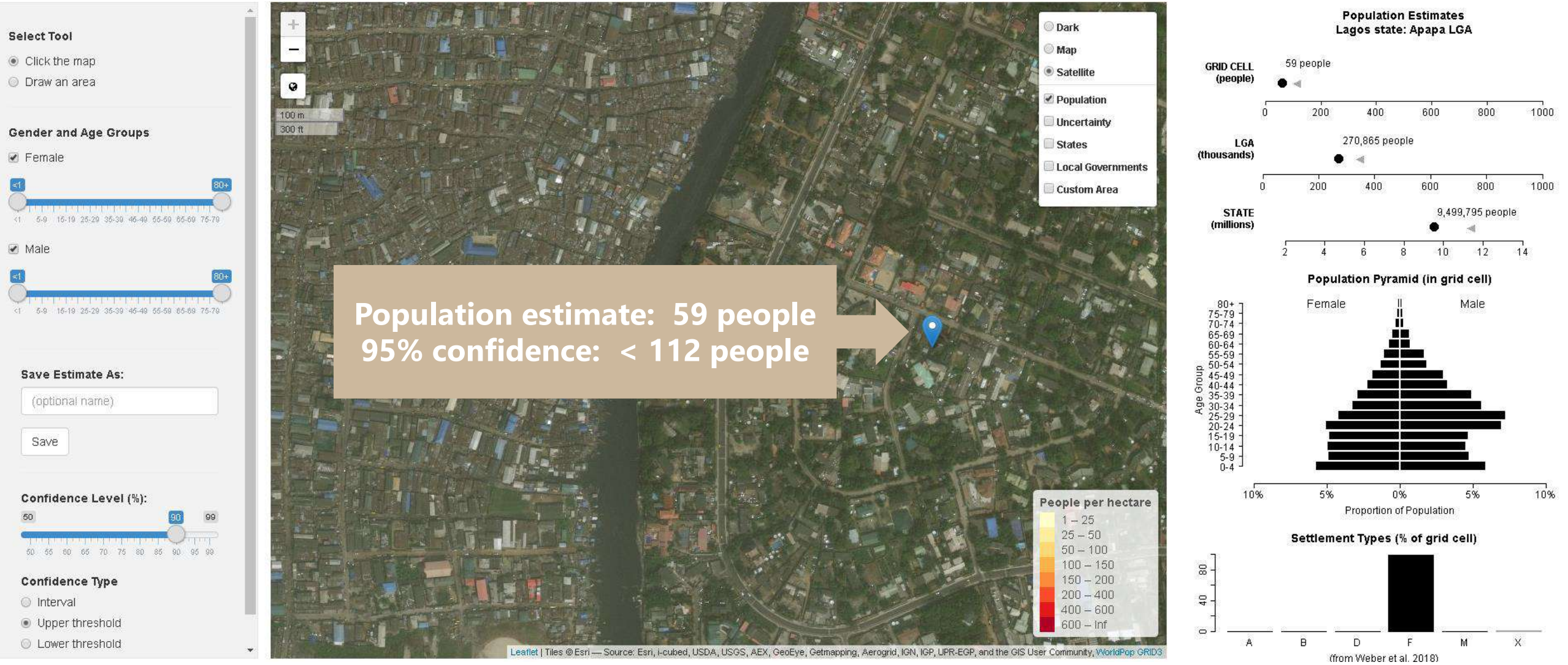


Leasure et al (2019) in review

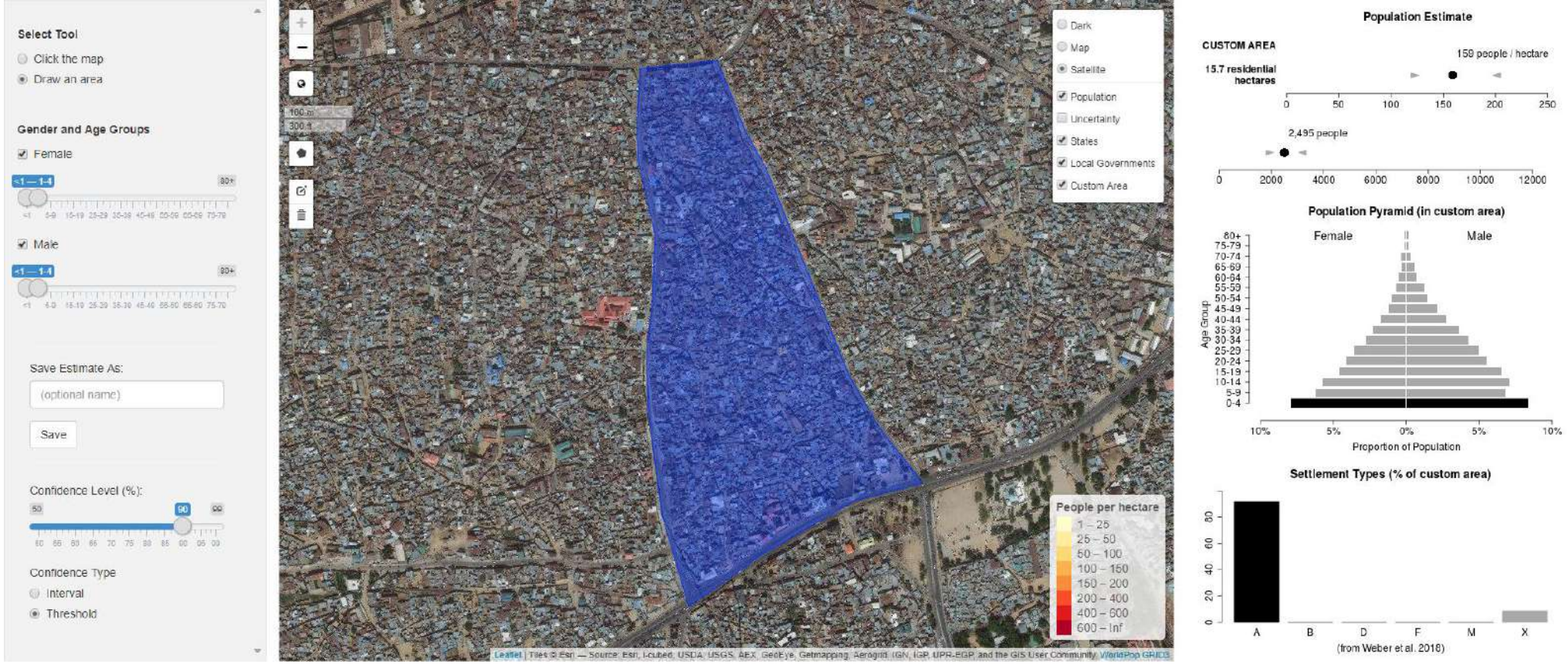




Leasure et al (2019) in review



Leasure et al (2019) in review



Using uncertainty

Requirement: vaccinate 90% of children under five years old

Our estimates indicate that this area in northern Nigeria (shown in blue) most likely contains 159 children under five years old. There is a 90% probability that no more than 212 children in this age group live here.

Leasure et al (2019) in review

VTS Map

Settlement/Ward/LGA/State

Go to XY

Long:

Lat:

Select population

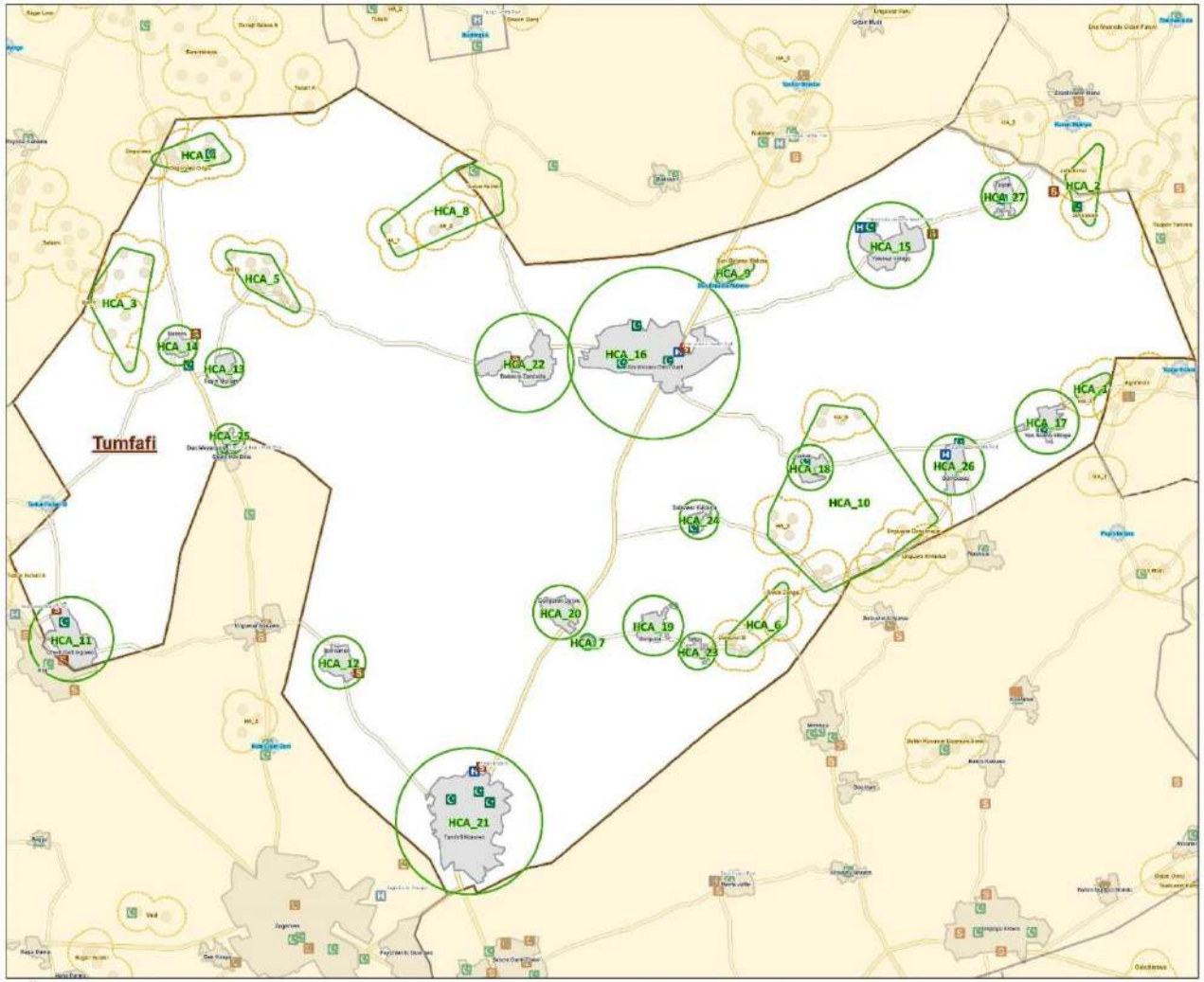
Age range [0 to 100]

Gender

buffer (m)

Male Female Total

Show Areas Table



Ward: Tumfafi
LGA: Dawakin Tofa
State: Kano

Total targeted population: 56334
Total # of HCs : 96

Area Name	Rural Target Pop	Urban Target Pop	# HC Days
HCA_1	128	0	2
HCA_2	128	0	2
HCA_3	231	0	4
HCA_4	257	0	4.5
HCA_5	257	0	4.5
HCA_6	128	0	2
HCA_7	0	154	1.5
HCA_8	231	0	4
HCA_9	51	154	2.5
HCA_10	385	0	6.5
HCA_11	0	3848	38.5
HCA_12	0	1744	17.5
HCA_13	0	1077	11
HCA_14	0	1077	11
HCA_15	0	4156	41.5
HCA_16	0	13338	133.5
HCA_17	0	1539	15.5
HCA_18	0	1334	13.5
HCA_19	0	1796	18
HCA_20	0	1847	18.5
HCA_21	0	14058	140.5
HCA_22	0	4002	40
HCA_23	0	667	6.5
HCA_24	0	667	6.5
HCA_25	0	462	4.5
HCA_26	0	1488	15
HCA_27	0	1129	11.5
Total	1796	54538	576.5

Production date: 03.11.2014
Powered by Novel-T

Pop (Under 5): 2429
Pop (Total): 12807
Zalau

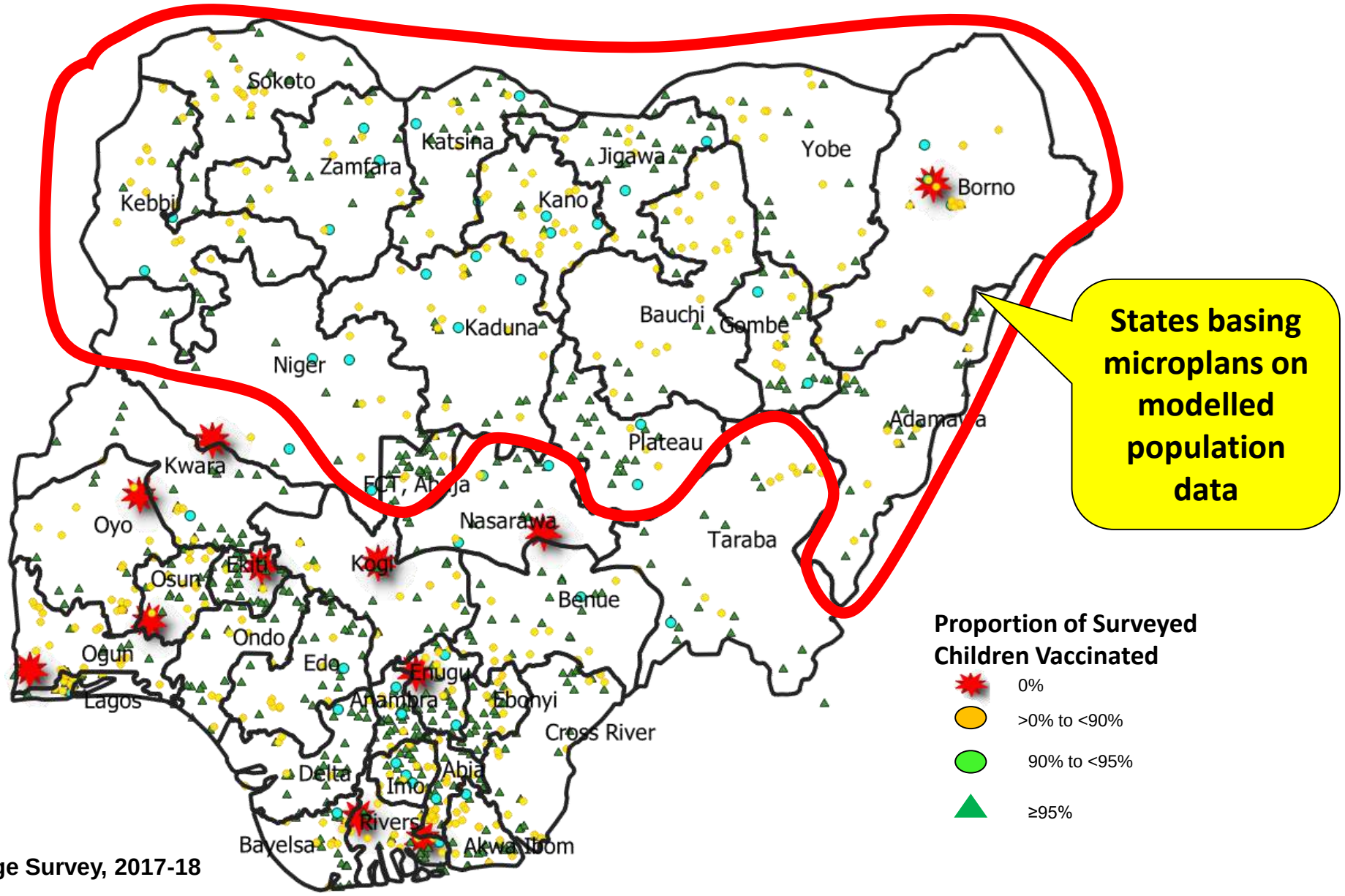
Pop (Under 5): 10965
Pop (Total): 55260
Zaranda

Google

0 2 4km

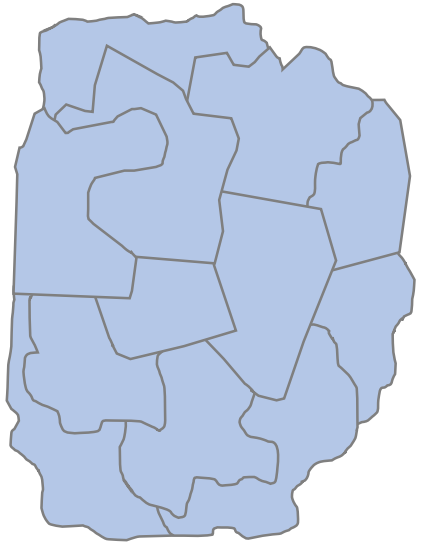
Powered by Novel-T

Measles Vaccination Post-Campaign Coverage Survey, by EA – Nigeria, 2017-18

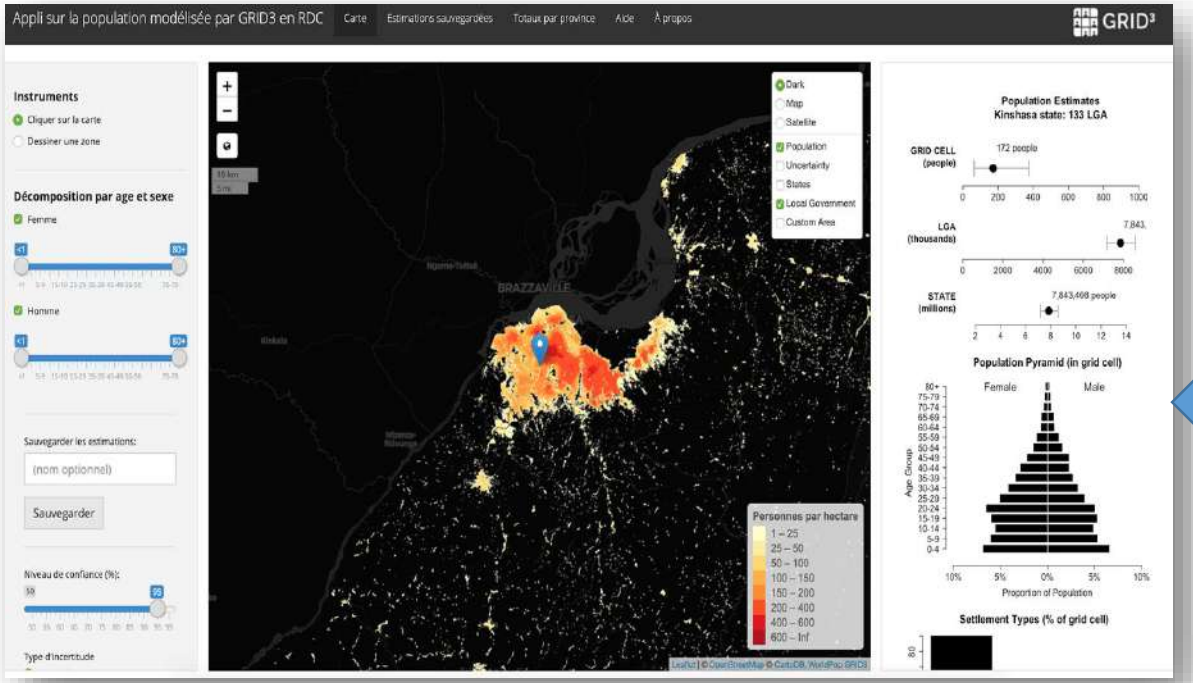


Source: Nigeria MVC Post Campaign Coverage Survey, 2017-18

National Census/Projections /Estimates

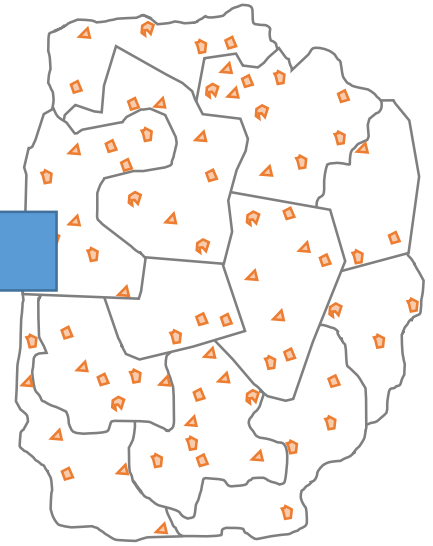


complete national coverage

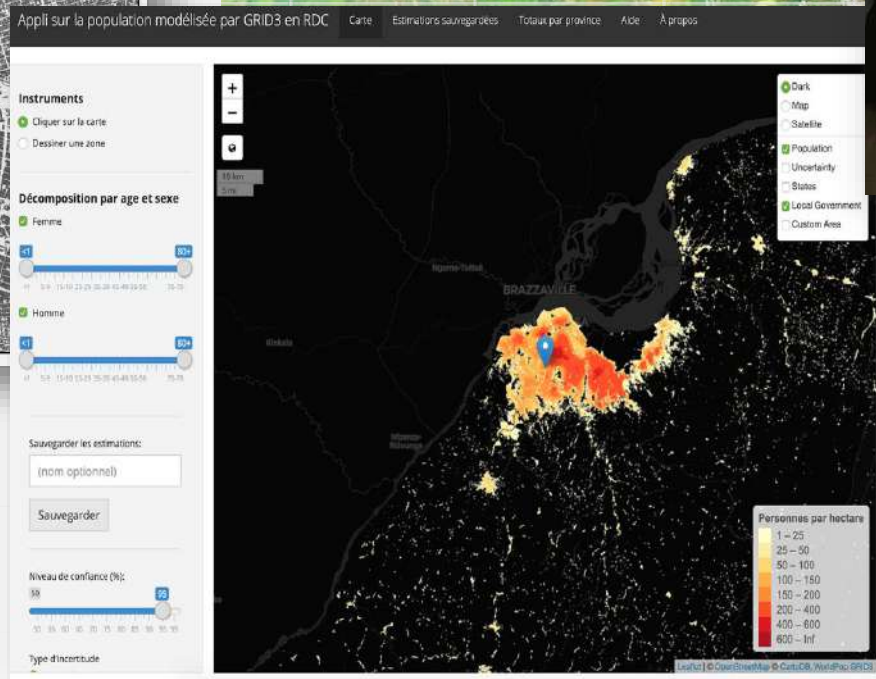


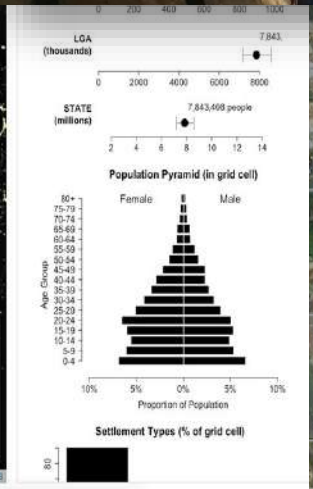
Nigeria, DRC, Zambia, S.Sudan

Full Population Model



enumerated population in limited sampling units only





GRID3

GEO-REFERENCED INFRASTRUCTURE AND DEMOGRAPHIC DATA FOR DEVELOPMENT

BILL & MELINDA GATES foundation

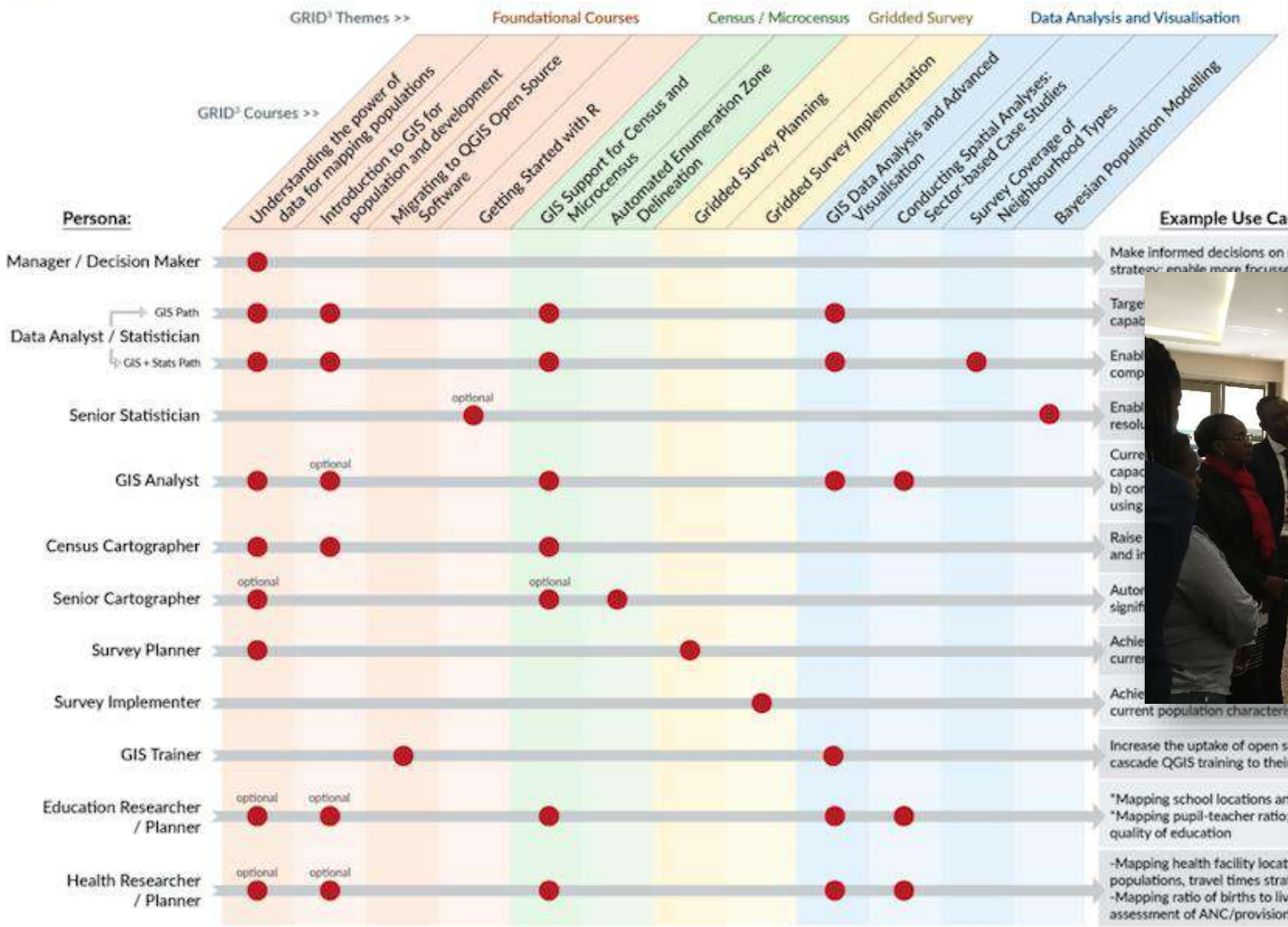
UKaid from the British people

world pop FLOWMINDER.ORG

UNFPA

Center for International Earth Science Information Network
EARTH INSTITUTE | COLUMBIA UNIVERSITY

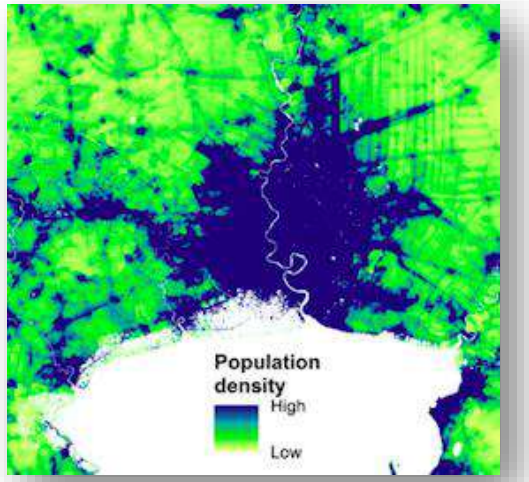
Capacity Strengthening - Training Pathways



Population estimation: Supporting the census process

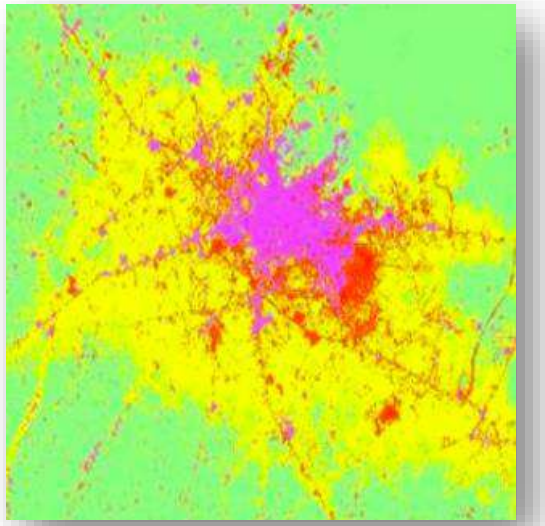
Census data dissemination and analysis

- Filling gaps
- Producing gridded outputs
- Integration with other data



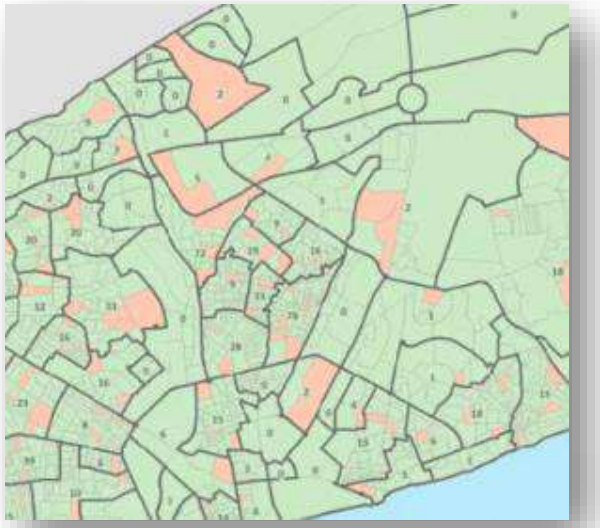
Inter-censal updates and estimates

- Capturing sub-national changes to inform inter-censal estimates



Survey/Census preparation

- Support for census planning
- Updating enumeration areas
- Defining sampling units



Considerations and limitations: population estimation modelling

- Not a replacement for a census!
- Still an area of active research – every setting is different
- Shows promise, but uncertainties can be large and accuracies low at fine spatial scales
- Are the areas being predicted into similar enough to those with the recent data used to build the models? Representative training and validation data is important – plus communicating assumptions
- Incomplete, inaccurate, outdated covariates
- Mobile populations, hidden populations, urban areas
- High level buy-in, close engagement with statistics office, need to be open on methods and limitations
- Communicating uncertainty

Some next steps

- Nigeria, DRC, Zambia, S.Sudan, Burkina Faso, Mali, Niger…..plus more
- Microcensus: Simulated population tests; Geolocated census tests; Household listings from surveys; Repeat surveys
- Models: Simulated population tests; Geolocated census tests; Alternative/hybrid models; Validation approaches
- Covariates: Residential/non-residential modelling; Building heights; Mobile network data; Neighbourhood mapping
- Supporting census/survey processes: Enumeration area delineation/update testing; Gridded population-based sampling
- Capacity strengthening: Population modelling workshops; User interface and tool development

Further information

www.worldpop.org

@WorldPopProject

GRID3

GEO-REFERENCED INFRASTRUCTURE AND DEMOGRAPHIC DATA FOR DEVELOPMENT

www.grid3.org

@GRID3Global

E-mail: A.J.Tatem@soton.ac.uk

@AndyTatem