

Dataset Deep Dive Webinar Series

WorldPop Gridded Population Datasets



Wednesday 9 December 2020

9 am New York | 3 pm Geneva | 5 pm Nairobi







The mission of the Centre is to increase the **use** and **impact** of data in humanitarian response.

Our focus areas



Data Services



Data Responsibility



Data Literacy

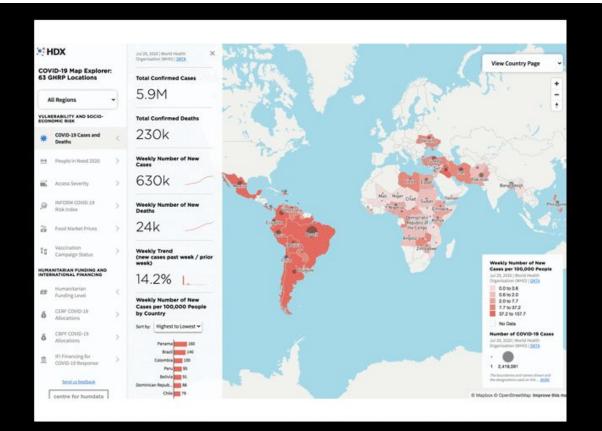


Predictive Analytics

COVID-19 Data Explorer

63 countries are considered most vulnerable to COVID-19, often because of ongoing humanitarian emergencies.

Our visualization brings together dozens of datasets from over 20 sources, supporting better decisions about aid allocation and distribution.



Today's agenda

15:05 – 15:15: Introduction to WorldPop

Andrew Tatem

15:15 – 15:20: How to Find and Use WorldPop Datasets on HDX

Godfrey Takavarasha

15:20 – 15:40: Overview of the WorldPop Gridded Population Datasets

Andrew Tatem, Alessandro Sorichetta, Maksym Bondarenko

15:40 – 16:00: Questions and Answers

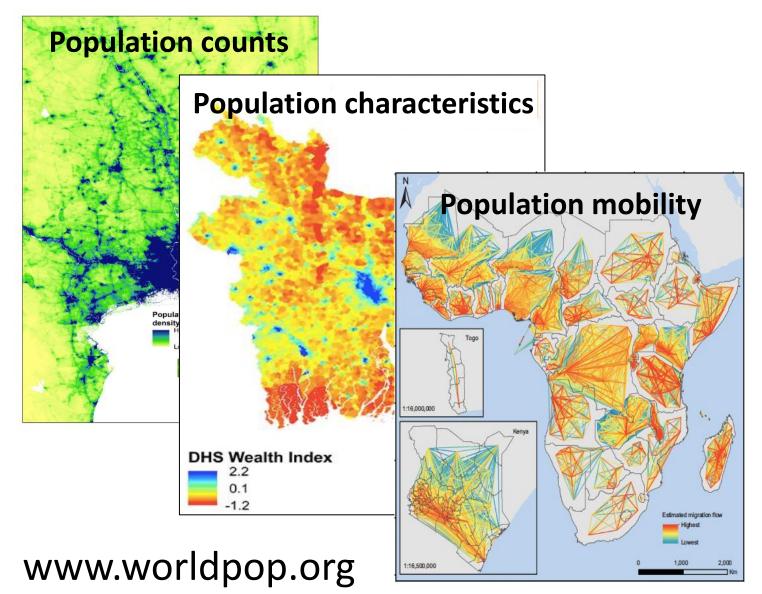
Introduction to WorldPop







Southampton



Applied research and implementation group

Mapping small area population distributions, demographics and dynamics

Open data, open peer-reviewed statistical methods, user engagement

Application in epidemiology, maternal/newborn health, childhood vaccination

Dynamic population mapping using mobile phone data

Pierre Deville^{a,b,c,1}, Catherine Linard^{c,d,1,2}, Samuel Martin^e, Marius Gilbert^{c,d}, Forrest R. Stevens^f, Andrea E. Gaughan^f, Vincent D. Blondela, and Andrew J. Tatemg,h,i

*Department of Applied Mathematics, Université catholique de Louvain, 1348 Louvain-la-Neuve, Belgium; bCenter for Complex Network Research and Physics Department, Northeastern University, Boston, MA 02115; 'Fonds National de la Recherche Scientifique, B-1000 Brussels, Belgium; 'Biological Control and Spatial Ecology, Université Libre de Bruxelles, 8-1050 Brussels, Belgium; "Université de Lorraine CNRS, Centre de Recherche en Automatique de Nancy, UMR 7039, 54518 Vandoeuvre-lès-Nancy, France Department of Geography and Geosciences, University of Louisville, Louisville, KY 40292; Department of Geography and Environment, University of Southampton, Southampton SO17 1BJ, United Kingdom; Fogarty International Center, National Institutes of Health, Bethesda, MD 20892; and Flowminder Foundation, 17177 Stockholm, Sweden

Edited by Michael F. Goodchild, University of California, Santa Barbara, CA, and approved September 15, 2014 (received for review May 8, 2014)

During the past few decades, technologies such as remote sensing. derive health and development indicators (3, 5, 26, 27). However.

ARTICLE

https://doi.org/10.1057/s41599-019-0242-9

Exploring the use of mobile phone data for nation migration statistics

Shengjie Lai 6 1,2,3, Elisabeth zu Erbach-Schoenberg 1,2, Carla Pezzulo 1, Nick W. Ruktanonchai 1,2, Alessandro Sorichetta^{1,2}, Jessica Steele¹, Tracey Li², Claire A. Dooley^{1,2} & Andrew J. Tatem^{1,2}

National population mapping from sparse survey data: A hierarchical Bayesian modeling framework to account for uncertainty

Douglas R. Leasure^{a,1}, Warren C. Jochem^a, Eric M. Weber^b, Vincent Seaman^c, and Andrew J. Tatem^a

*WorldPop, Geography and Environmental Science, University of Southampton, Southampton SO17 1BJ, United Kingdom; b National Security Emerging Technologies Division, Oak Ridge National Laboratory, Oak Ridge, TN 37830; and Global Development Division, The Bill and Melinda Gates Foundation,

INTERFACE

rsif.royalsocietypublishing.org



Fine resolution mapping of population age-structures for health and development applications

V. A. Alegana¹, P. M. Atkinson¹, C. Pezzulo¹, A. Sorichetta¹, D. Weiss², T. Bird¹, E. Erbach-Schoenberg¹ and A. J. Tatem^{1,3,4}

Open, peer-reviewed methods and datasets

Spatially disaggregated population estimates in the absence of national population and housing census data

N. A. Wardropa,b,1, W. C. Jochema,b,1, T. J. Birda,b, H. R. Chamberlaina,b, D. Clarkea,b, D. Kerra,b, L. Bengtsson^{a,b}, S. Juran^c, V. Seaman^d, and A. J. Tatem^{a,b,2}

Computers, Environment and Ordan Systems

Annually modelling built-settlements between remotely-sensed observations using relative changes in subnational populations and lights at night

Jeremiah J. Nieves^{a,b,*}, Alessandro Sorichetta^{a,b}, Catherine Linard^{a,c}, Maksym Bondarenko^{a,b}, Jessica E. Steele^{a,b}, Forrest R. Stevens^{a,d}, Andrea E. Gaughan^{a,d}, Alessandra Carioli^{a,b}, Donna J. Clarke^{a,b}, Thomas Esch^e, Andrew J. Tatem^{a,b}

Classifying settlement types from multi-scale spatial patterns of building footprints

© The Author(s) 2020 Article reuse guidelines DOI: 10.1177/2399808320921208

Warren C Jochem, Douglas R Leasure, Oliver Pannell (0), Heather R Chamberlain, Patricia Jones and Andrew J Tatem

WorldPop, School of Geography and Environmental Science, University of Southampton, UK





www.worldpop.org

WorldPop Demographics Portal

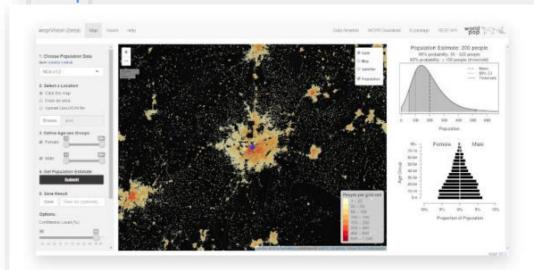
VV (

Popu

Global subnational age and sex structured estimates for 2020 obtained through integrating census, microdata and survey datasets in a spatiotemporal modelling framework.

DETAILS





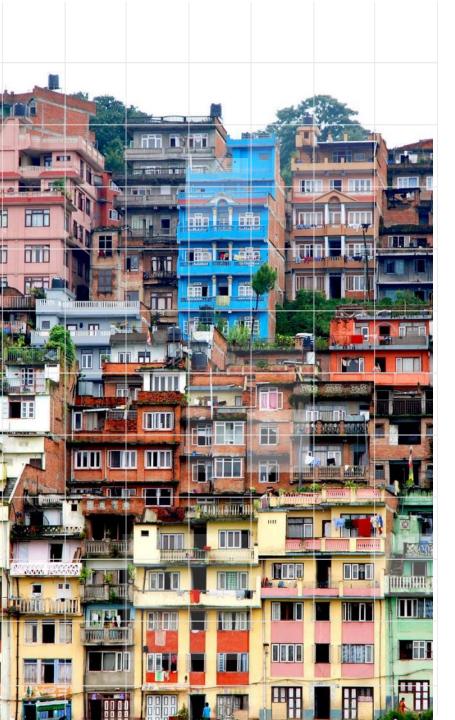
WorldPop Applications

Putting the power of population data into your hands.

DETAILS

Open datasets, code, applications



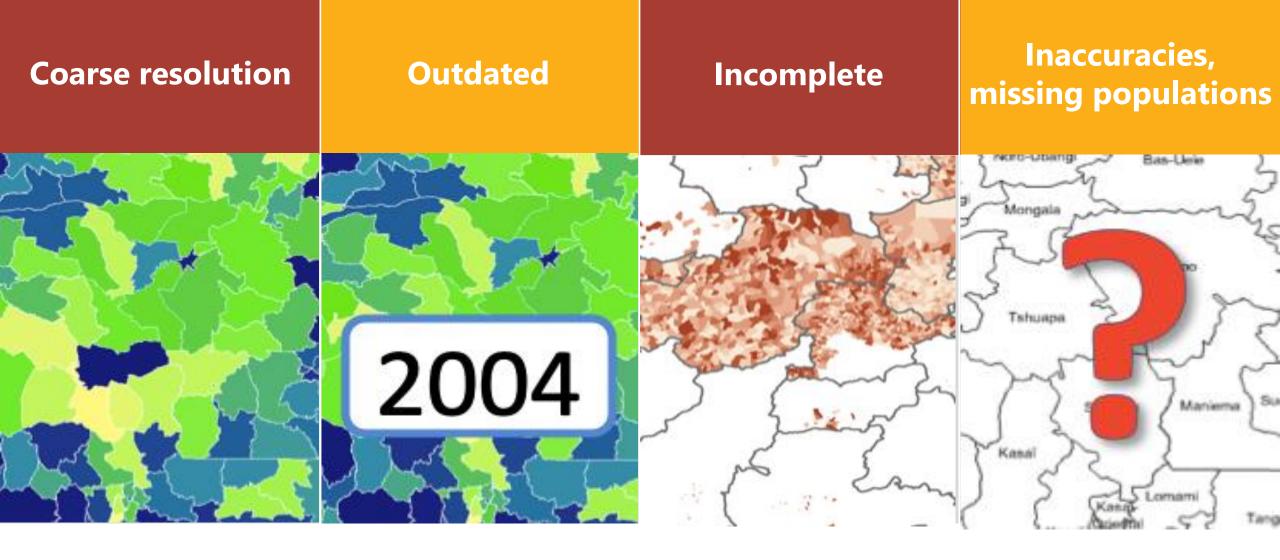


Uses of small area demographic data

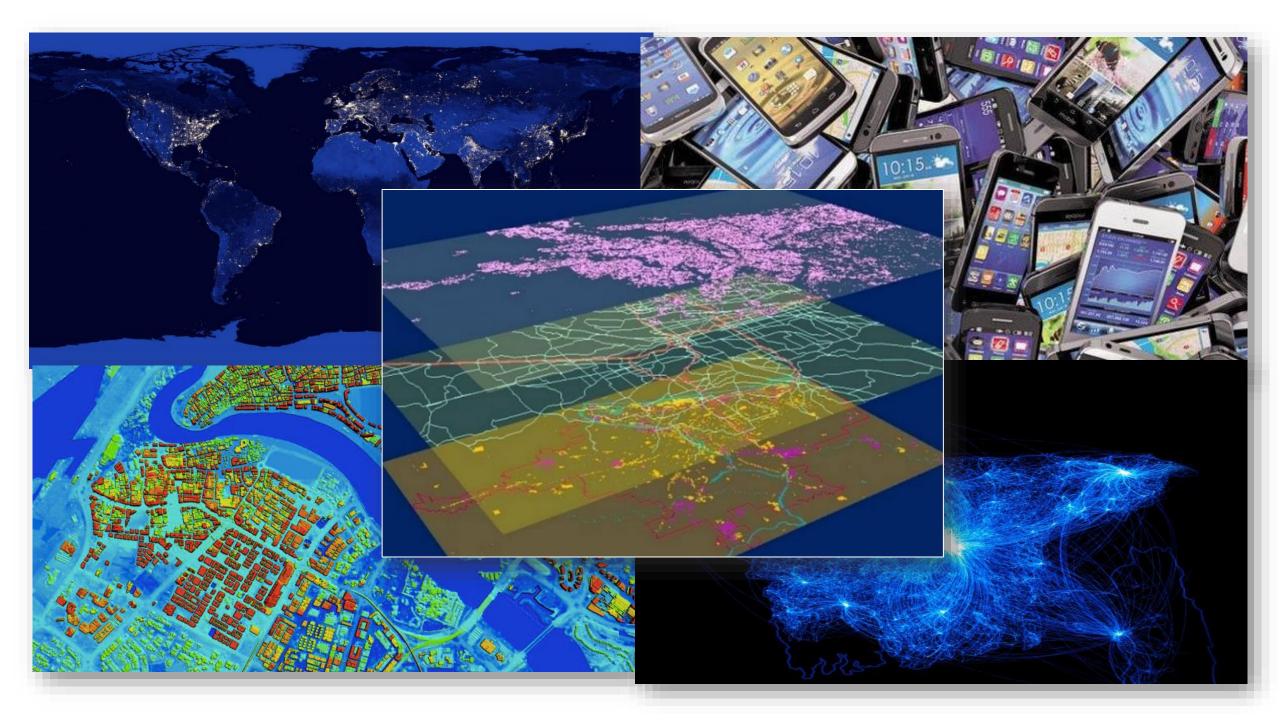
- Planning elections
- Calculating GDP
- Local governance
- Traffic planning
- Financial services
- Delivering utilities
- Agricultural subsidies
- Taxation
- Land use management
- Energy strategies
- Disaster response
- Humanitarian needs

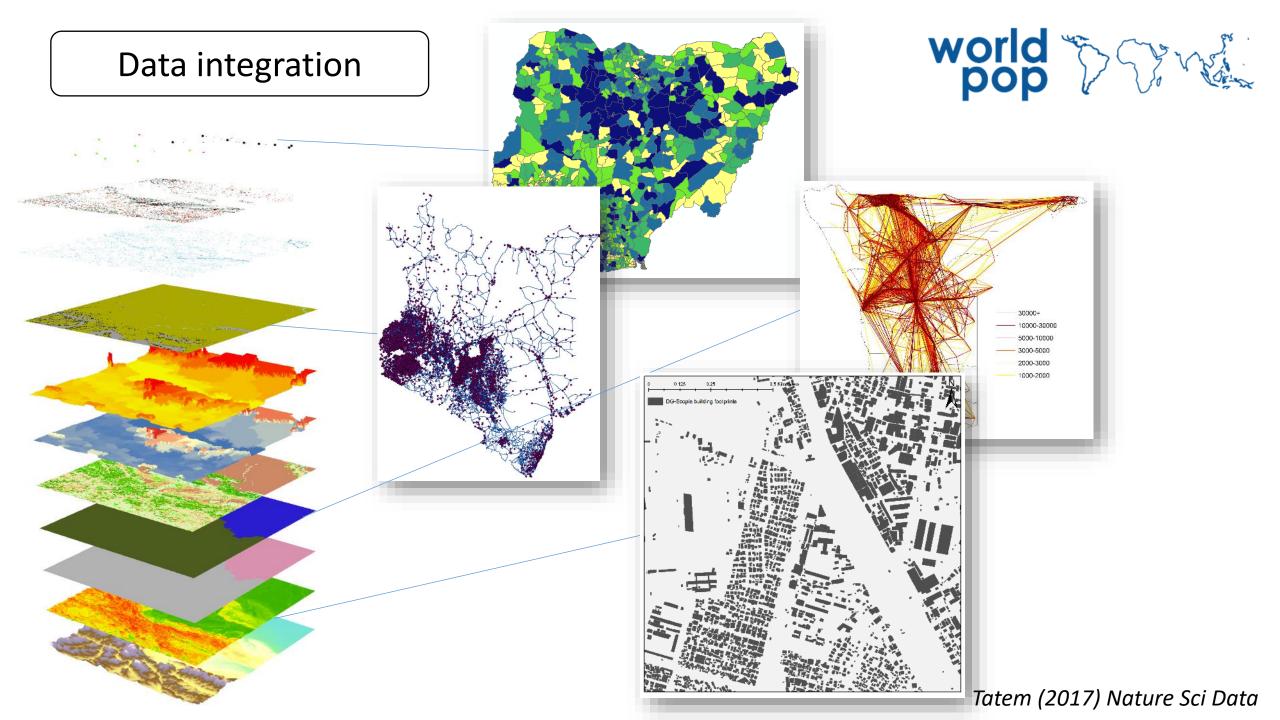
- Health system planning
- Supply chain management
- Health metrics
- Meeting SDGs
- Controlling infectious diseases
- Modelling disease spread and intervention

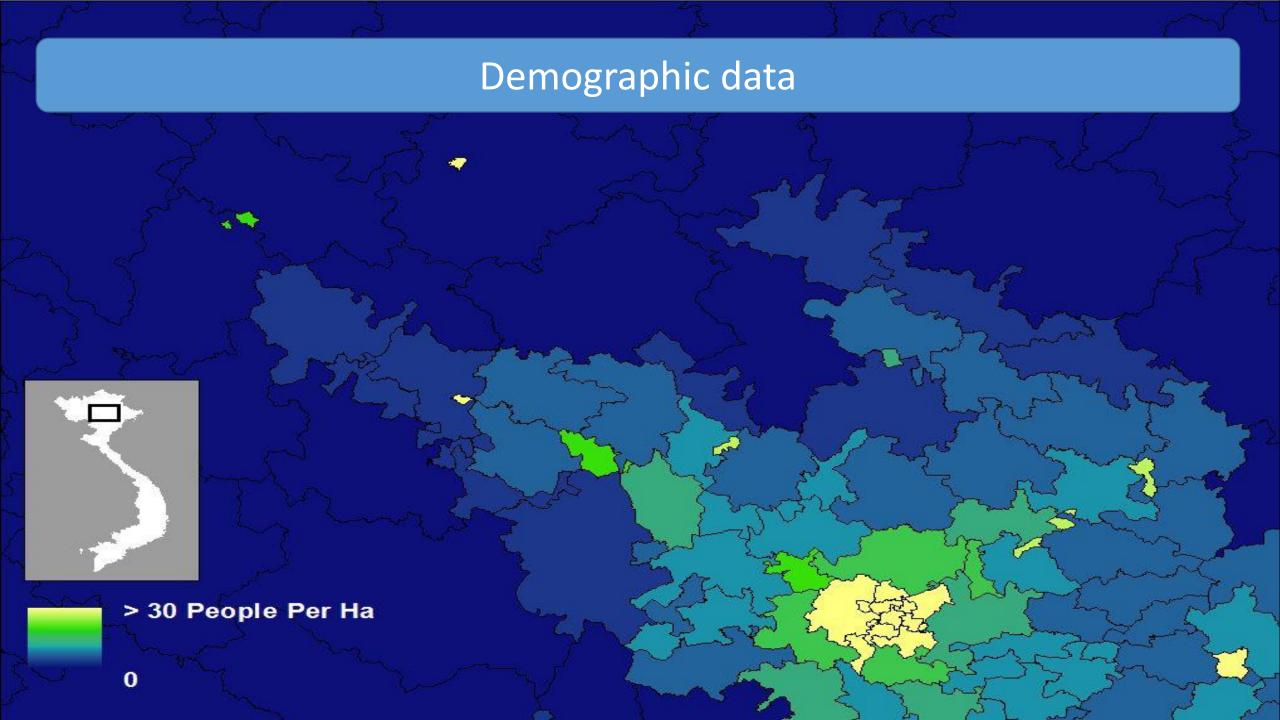
effects



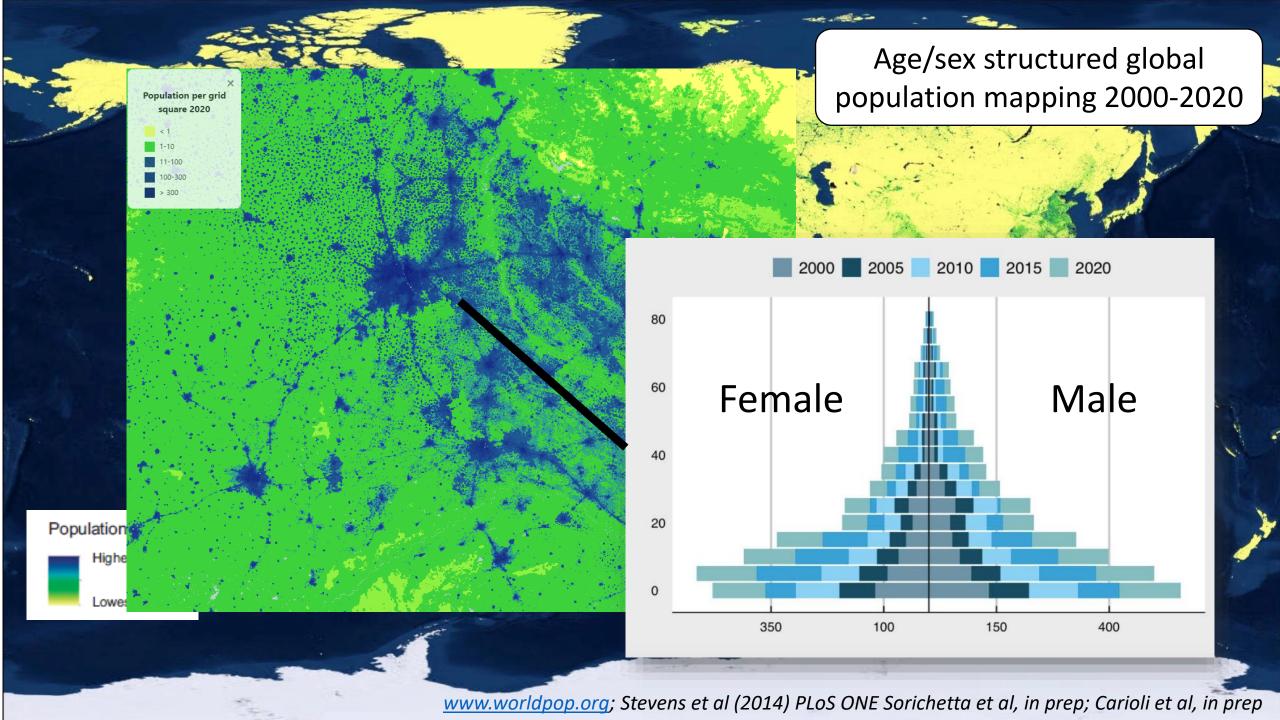
Demographic data challenges

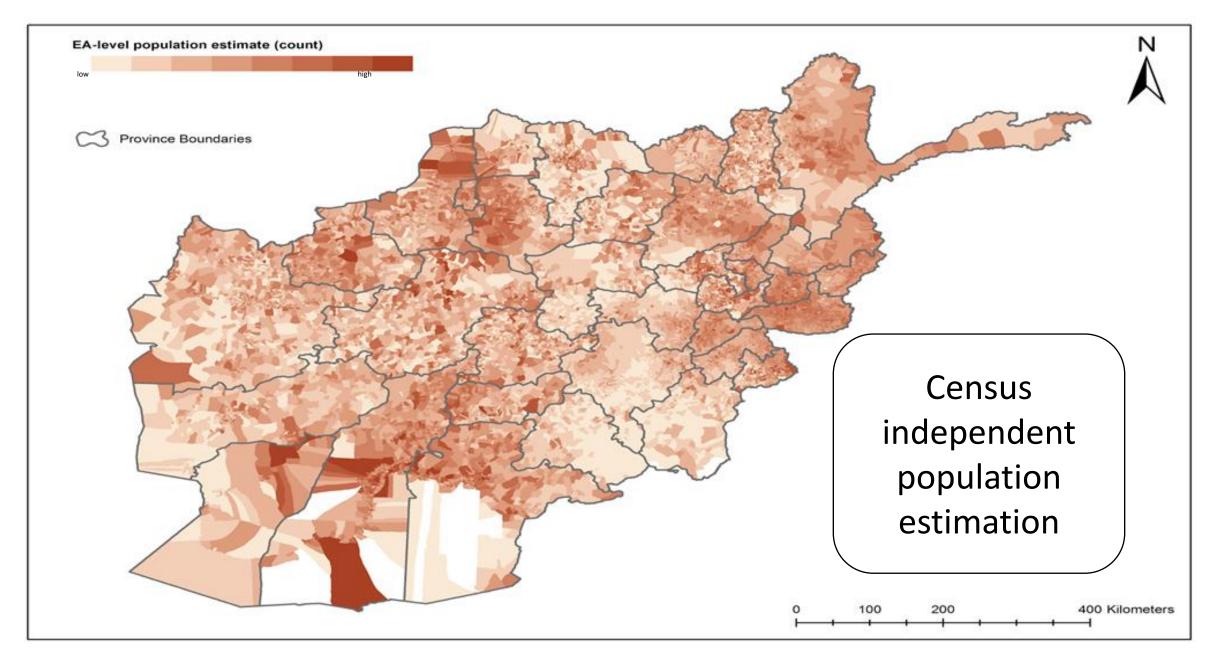






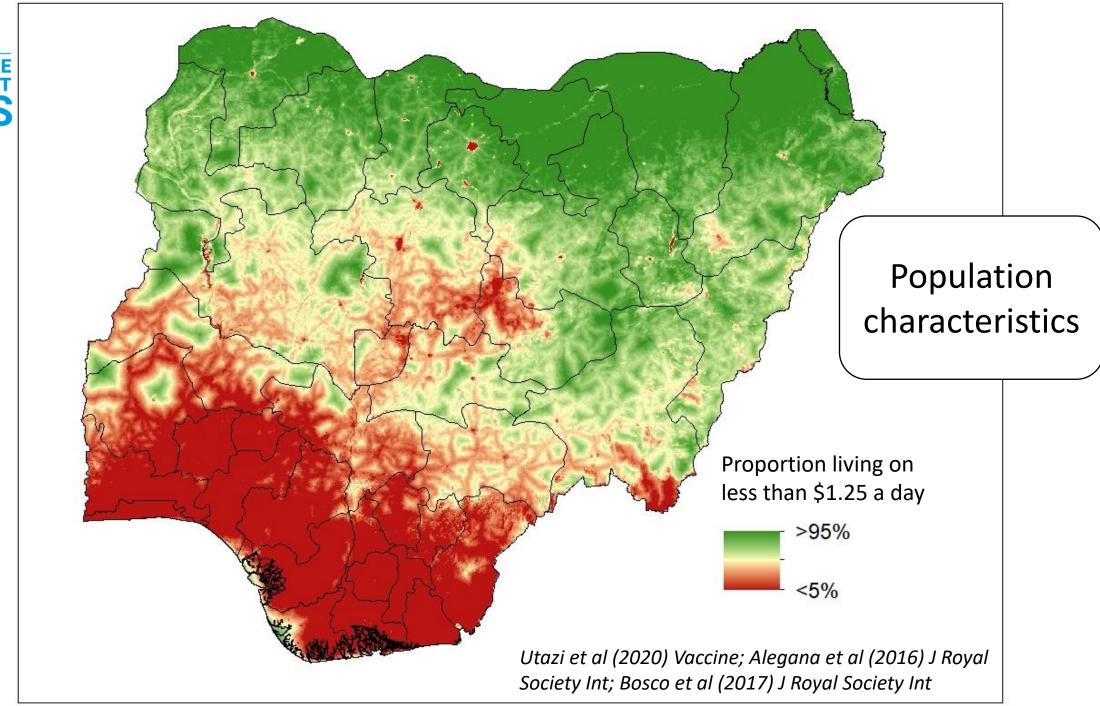
Population counts per 100 x 100m grid square > 30 People Per Pixel Low: 0

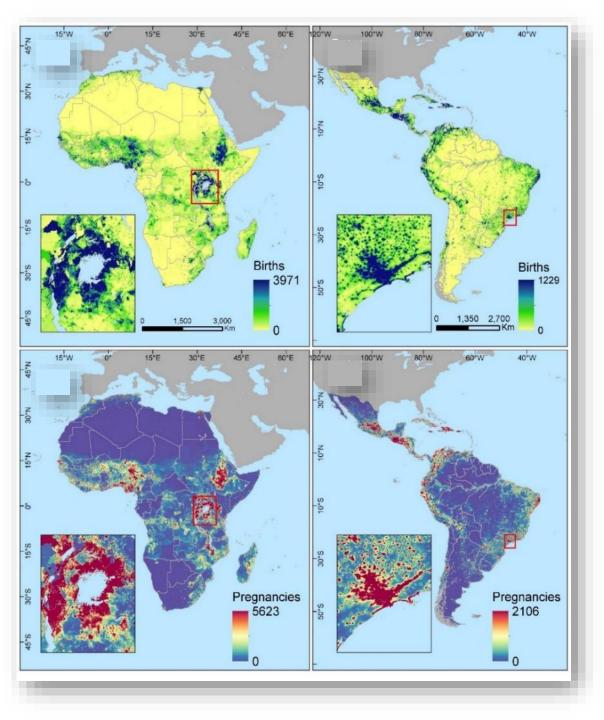




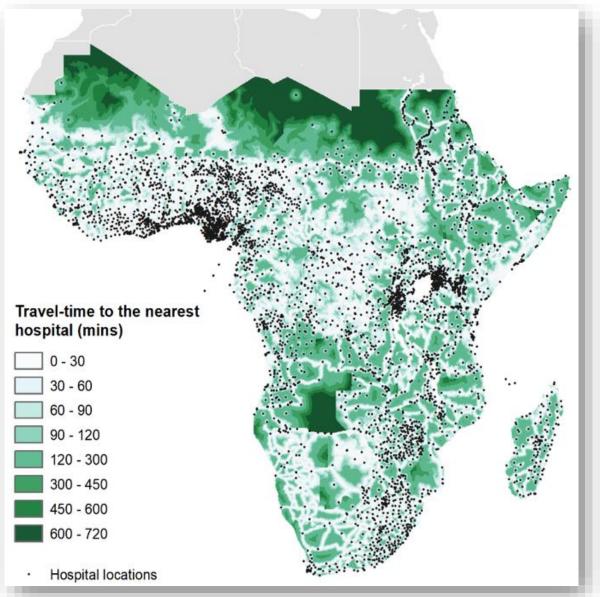
Wardrop et al (2018) PNAS; Chamberlain et al, in prep



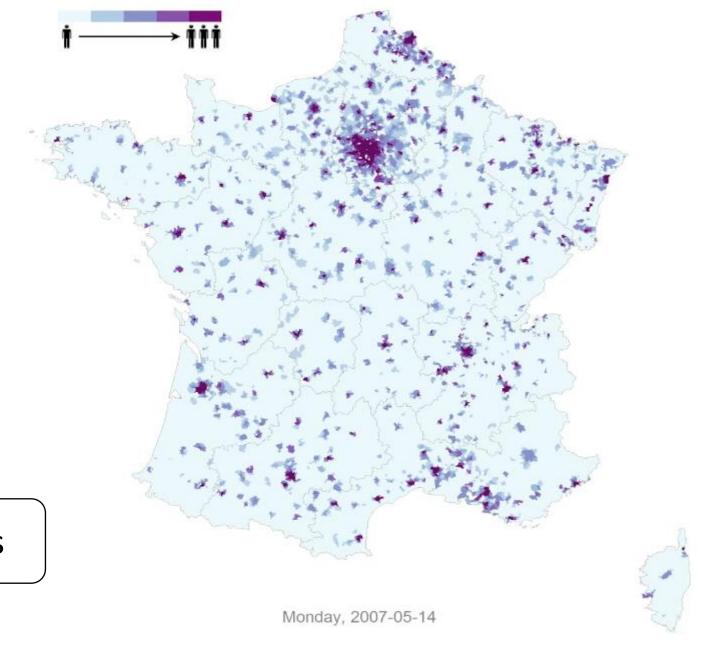




Scaling up data and insights



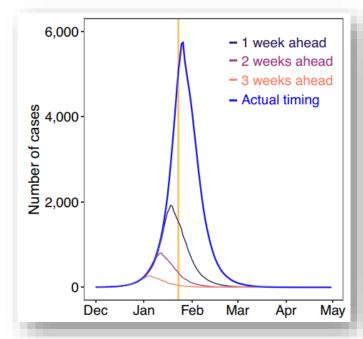
James et al (2018) Nature Sci Data; Wigley et al (2020) BMC Medicine



Mapping population dynamics

Accelerated Article Preview

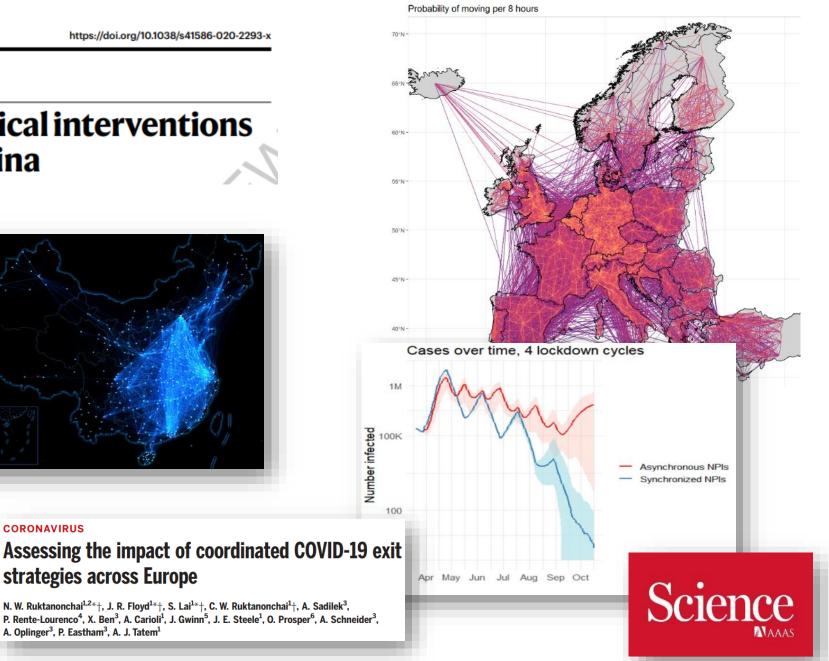
Effect of non-pharmaceutical interventions to contain COVID-19 in China





strategies across Europe

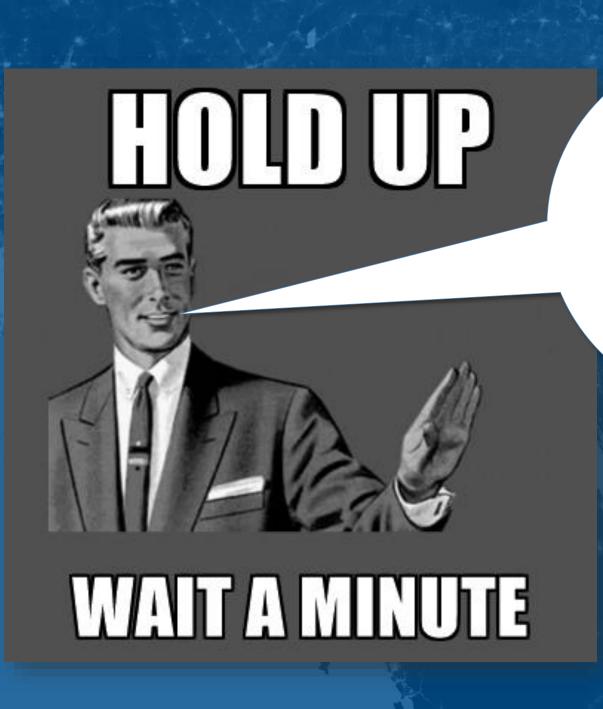
CORONAVIRUS



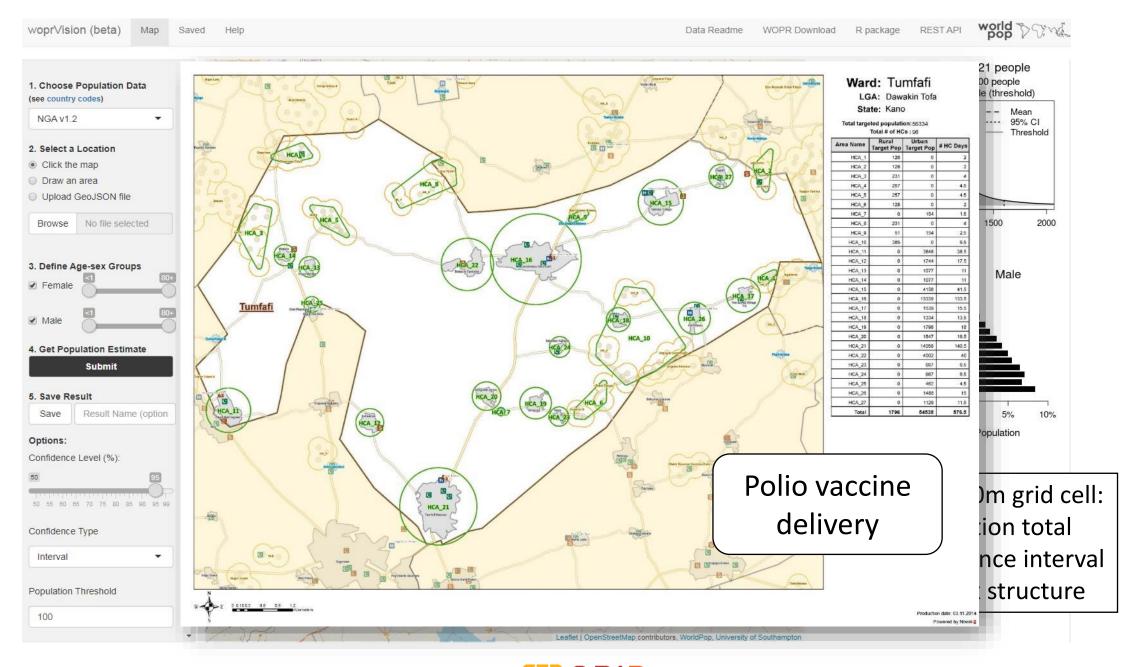
Spatial modelling of infectious diseases

P. Rente-Lourenco⁴, X. Ben³, A. Carioli¹, J. Gwinn⁵, J. E. Steele¹, O. Prosper⁶, A. Schneider³, A. Oplinger³, P. Eastham³, A. J. Tatem¹

N. W. Ruktanonchai^{1,2}*†, J. R. Floyd¹*†, S. Lai¹*†, C. W. Ruktanonchai¹†, A. Sadilek³,

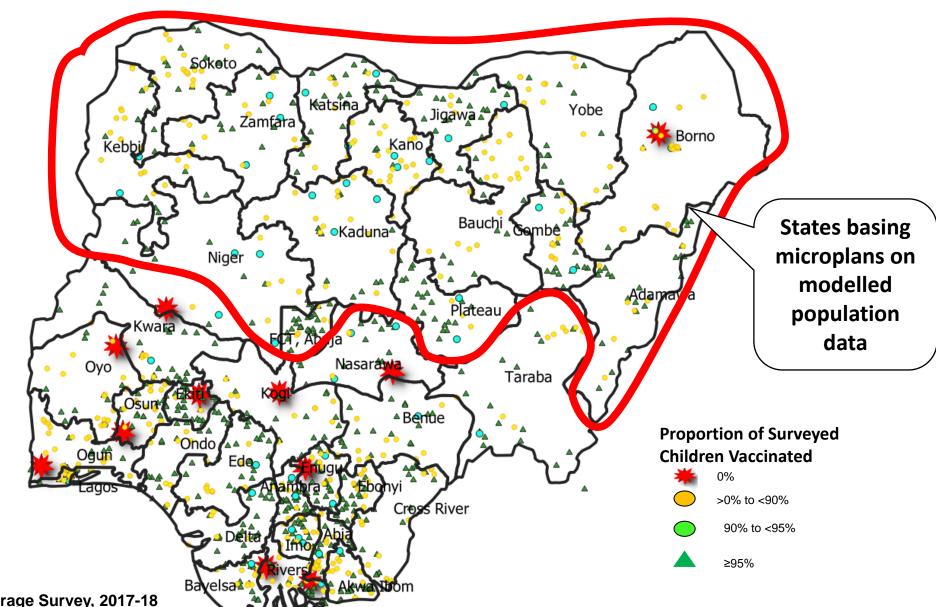


Is any of this finding its way outside of academia?





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





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

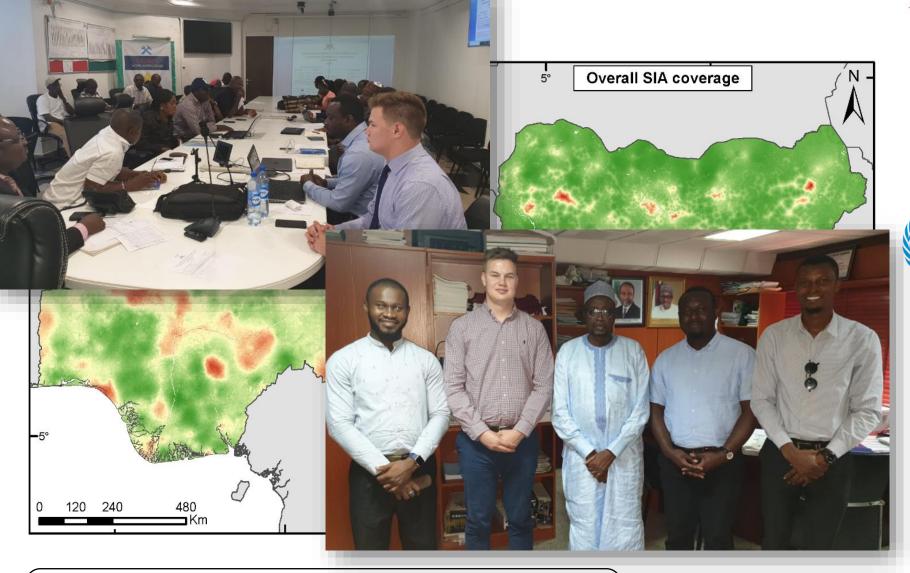




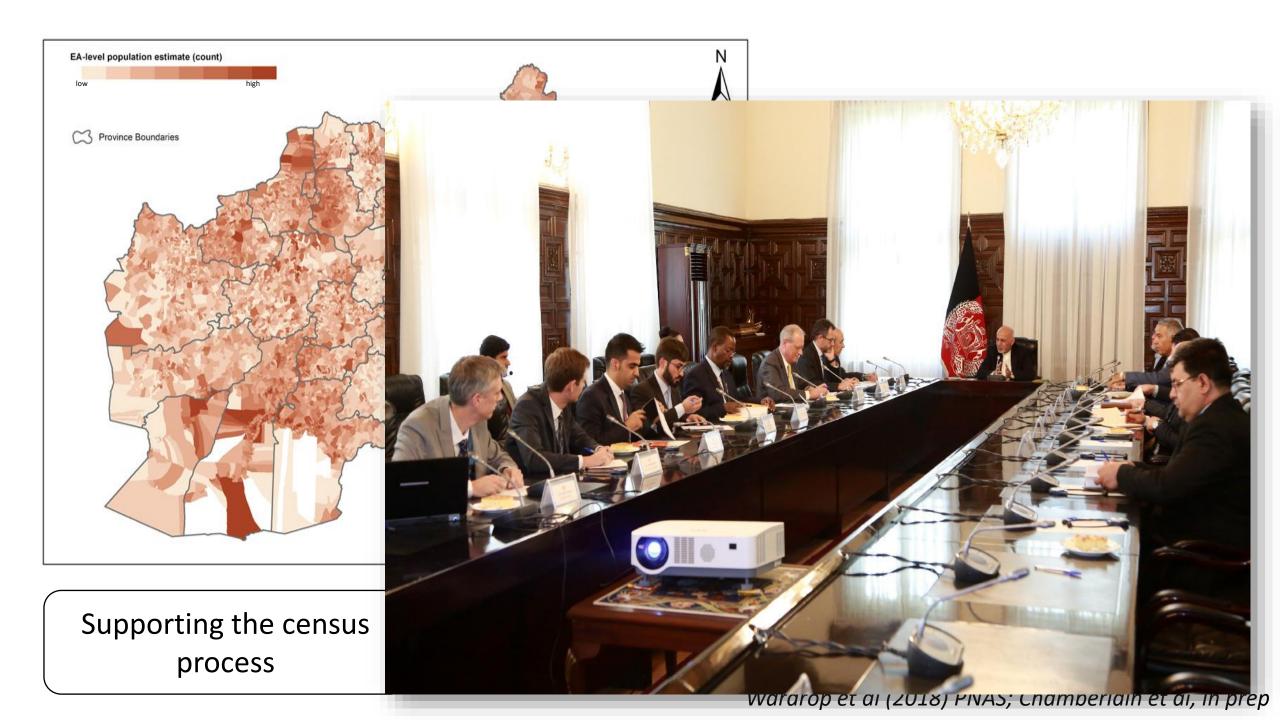


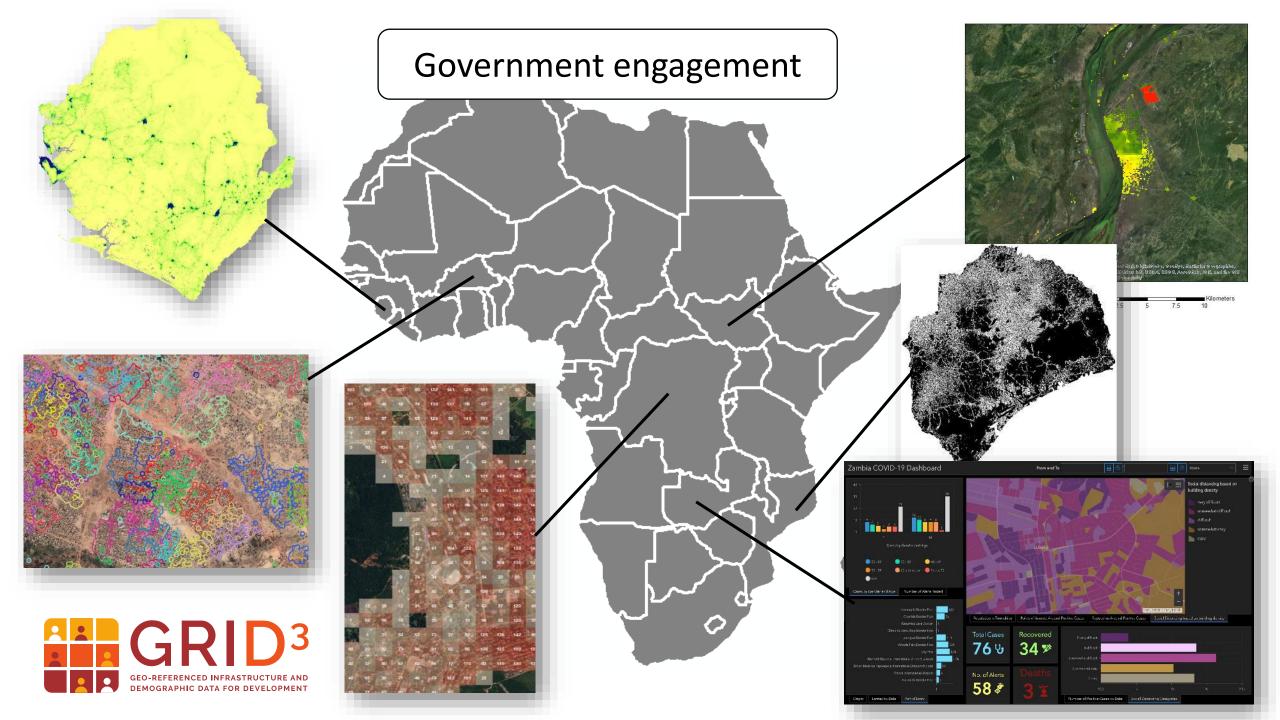


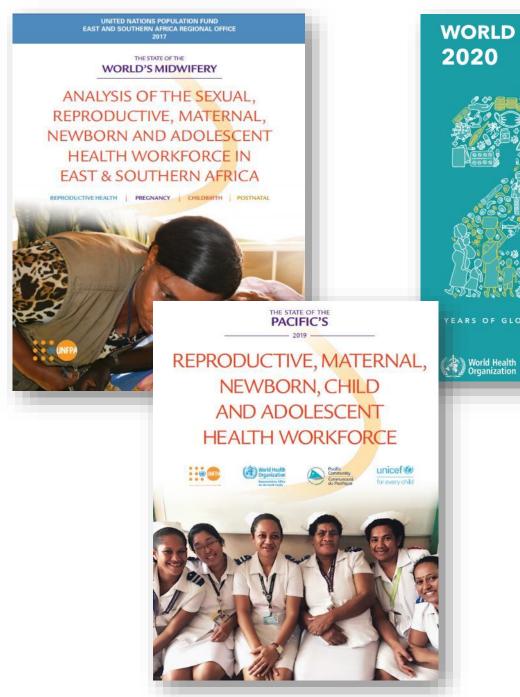


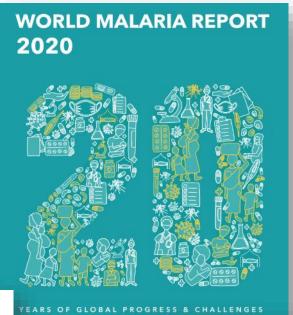


Childhood vaccination coverage mapping



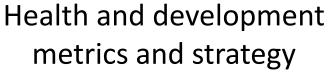






MALARIA ELIMINATION IN ZANZIBAR A FEASIBILITY ASSESSMENT | OCTOBER 2009

Gavi 🚱











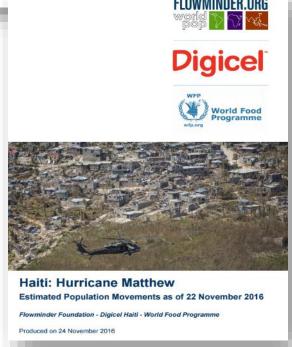




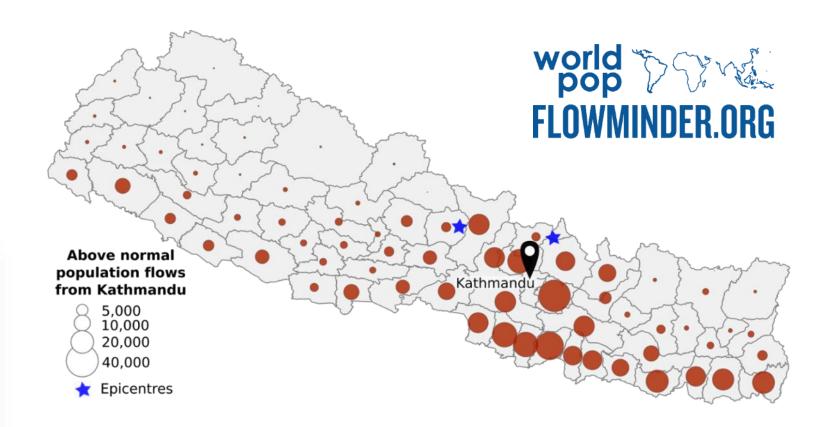




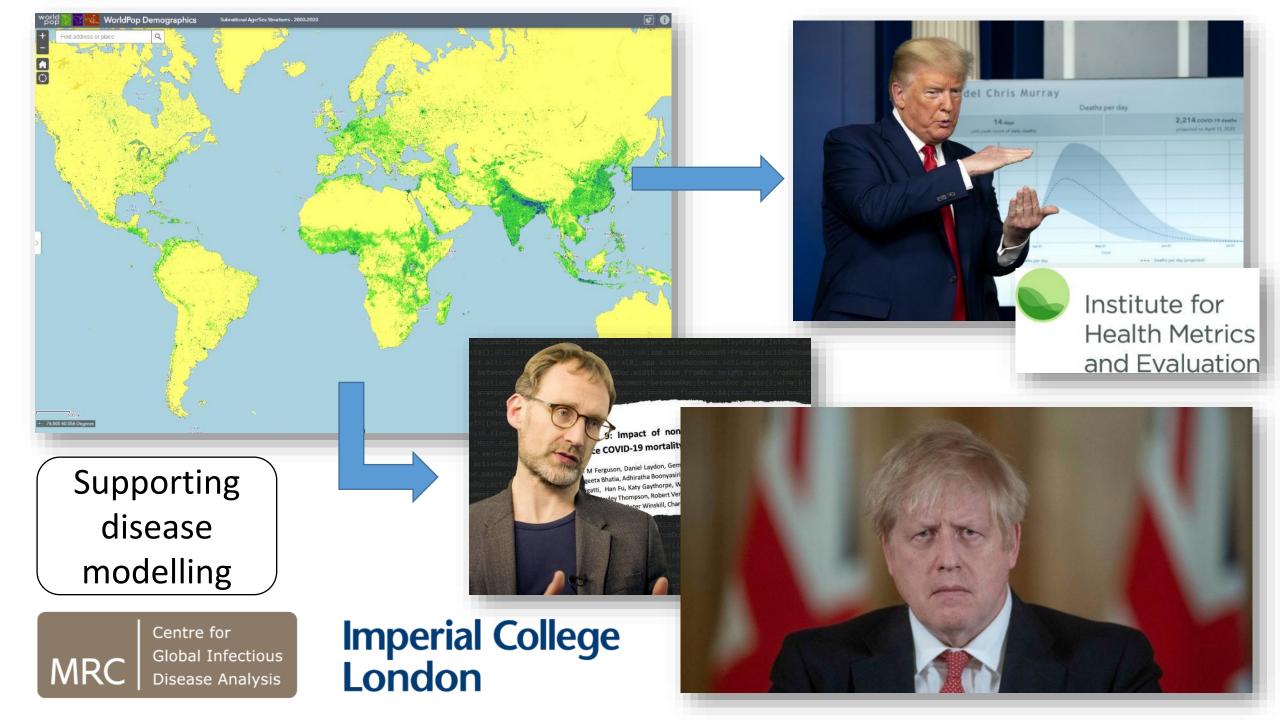




Disaster response







Further information



www.worldpop.org





www.grid3.org



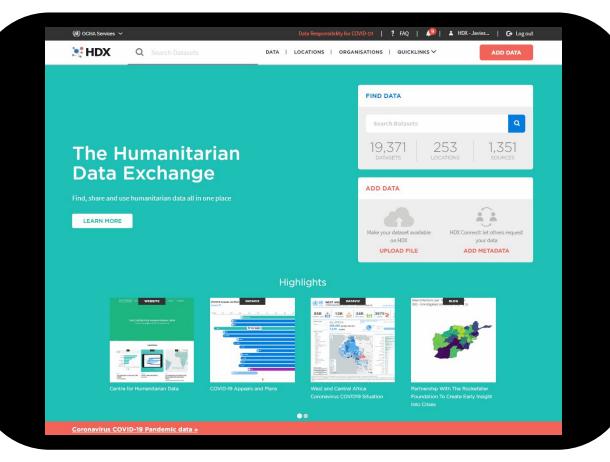
E-mail: A.J.Tatem@soton.ac.uk



How to Find and Use WorldPop Datasets on HDX

Humanitarian Data Exchange (HDX)

HDX makes data easy to find and use for analysis.



HDX at a Glance

100,000+

USERS PER MONTH 17,800+

DATASETS SHARED

1,300+

DATA SOURCES

275+

ACTIVE ORGANISATIONS

25,000

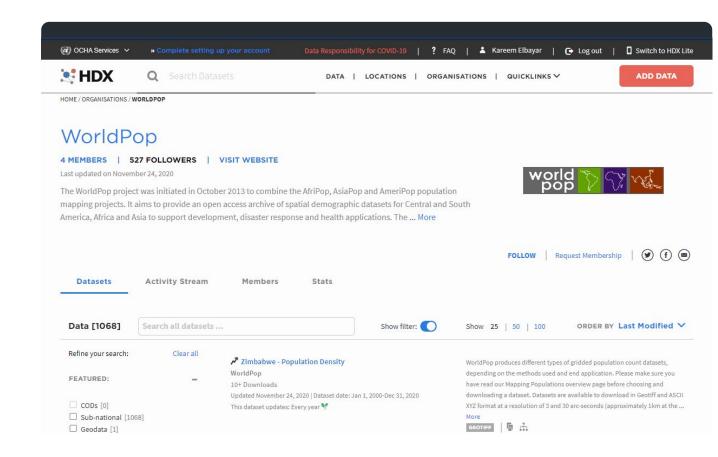
UNIQUE DOWNLOADS PER MONTH

Why join HDX?

- → Find and Share Data Fast Be data-ready for the next crisis.
- → Get **Insight** into Crises around the World Explore and understand data through interactive visualisations.
- → Join a Global **Community**Be part of the mission to connect data from organisations around the world.

WorldPop on HDX

Access by searching for 'WorldPop'



https://data.humdata.org/organization/worldpop

Data Grids

Our Crisis Data Grids examine availability and freshness of key datasets across six categories and 27 indicators.

HDX LOCATIONS ORGANISATIONS QUICKLINKS Y ADD DATA

HOME / DASHBOARDS / OVERVIEW OF DATA GRIDS

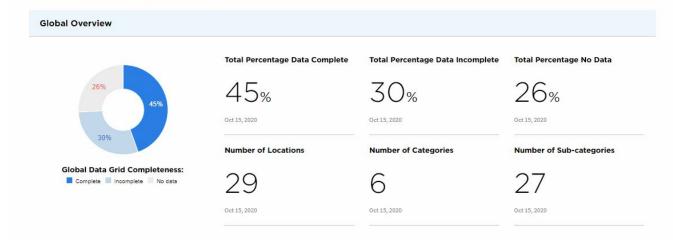
Overview of Data Grids

The Data Grid places the most important crisis data into six categories and several sub-categories. Relevant data is included in the Data Grid if it is sub-national, in a common format, and timely.



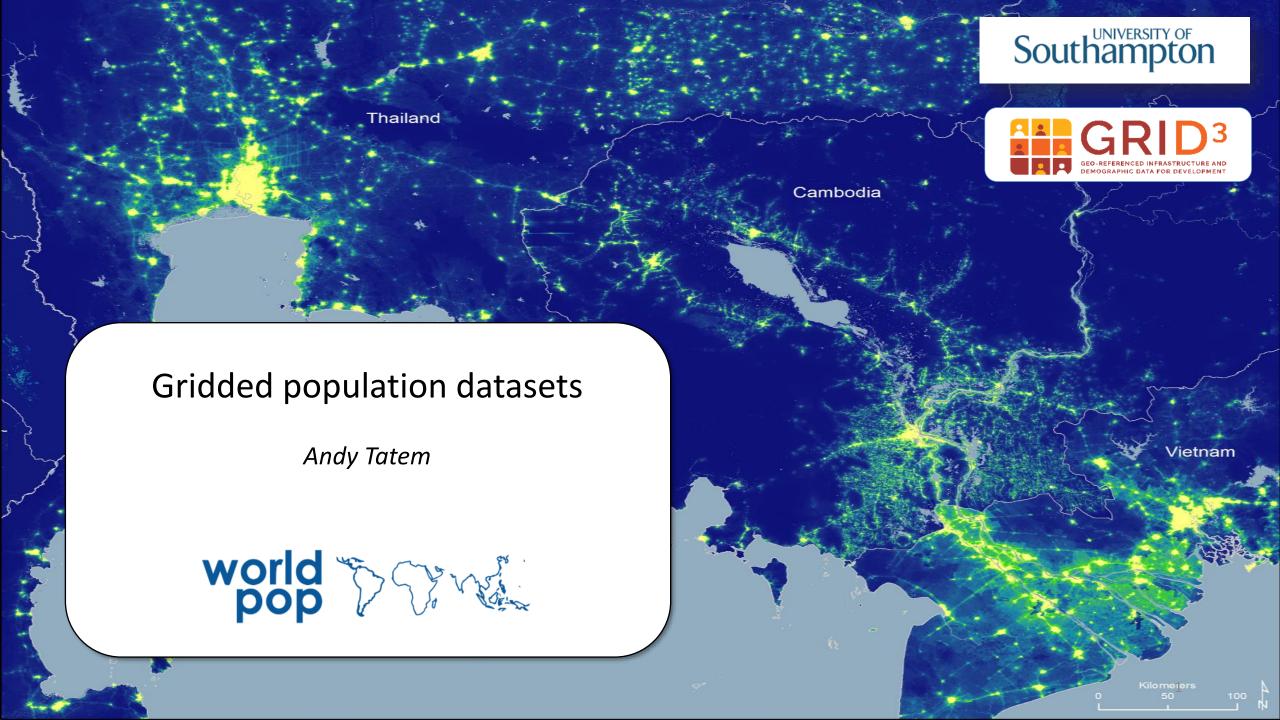


If at least one dataset meets all criteria, that subcategory is considered 'complete'. If at least one dataset meets some of these criteria, the sub-category is considered 'incomplete'. If a dataset does not meet the criteria or does not exist on HDX, the sub-category is considered empty or as having no data.



https://data.humdata.org/dashboards/ overview-of-data-grids

Overview of the WorldPop Gridded Population Datasets

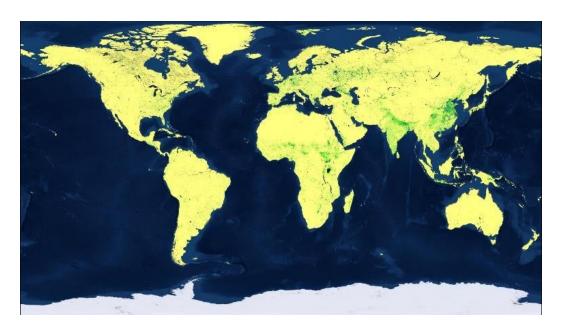




Gridded population datasets/methods

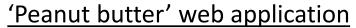
'Top down' global estimates

- 'Unconstrained' 2000-2020
- 'Constrained' 2020





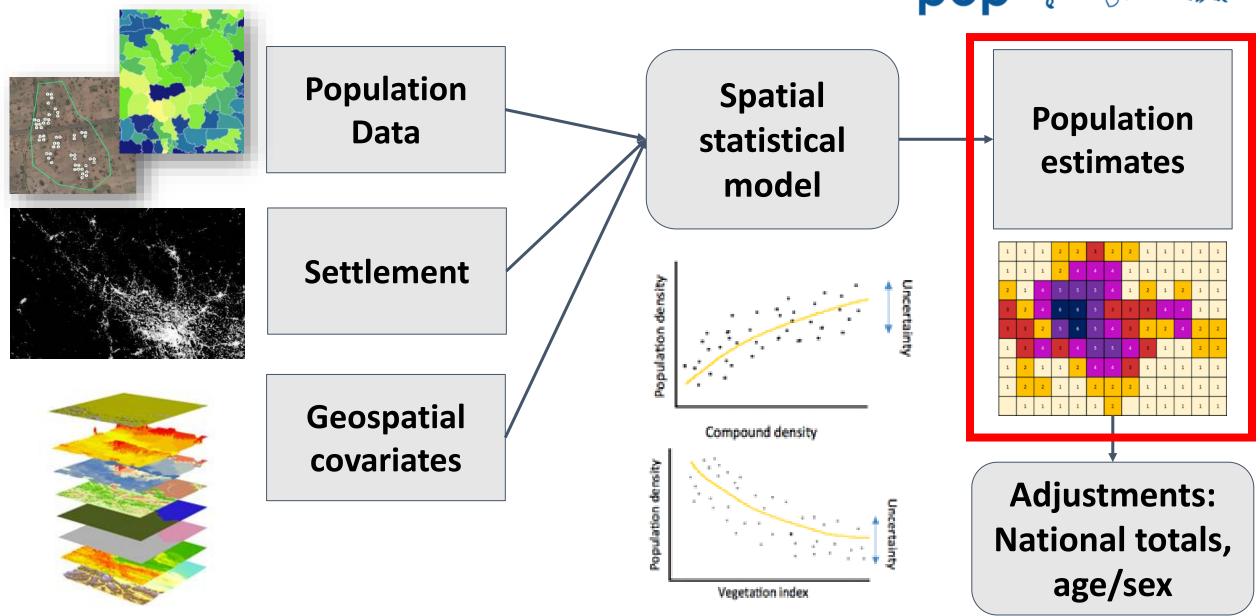
'Bottom up' bespoke country estimates Population counts per pixel



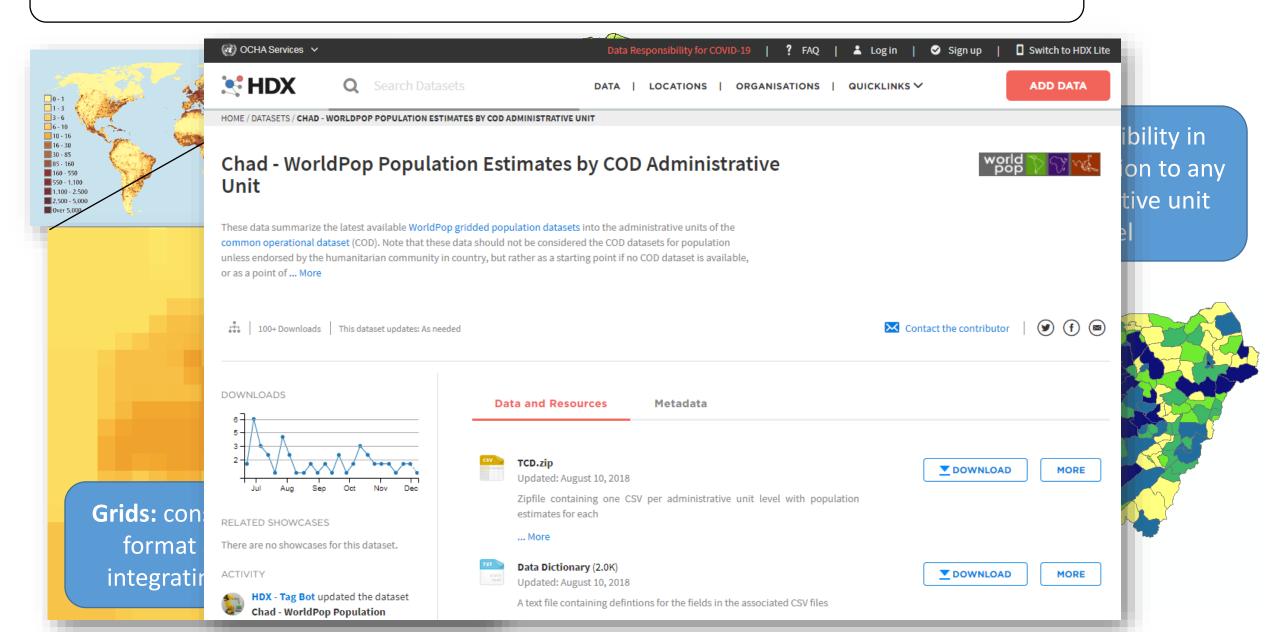


Population distribution modelling

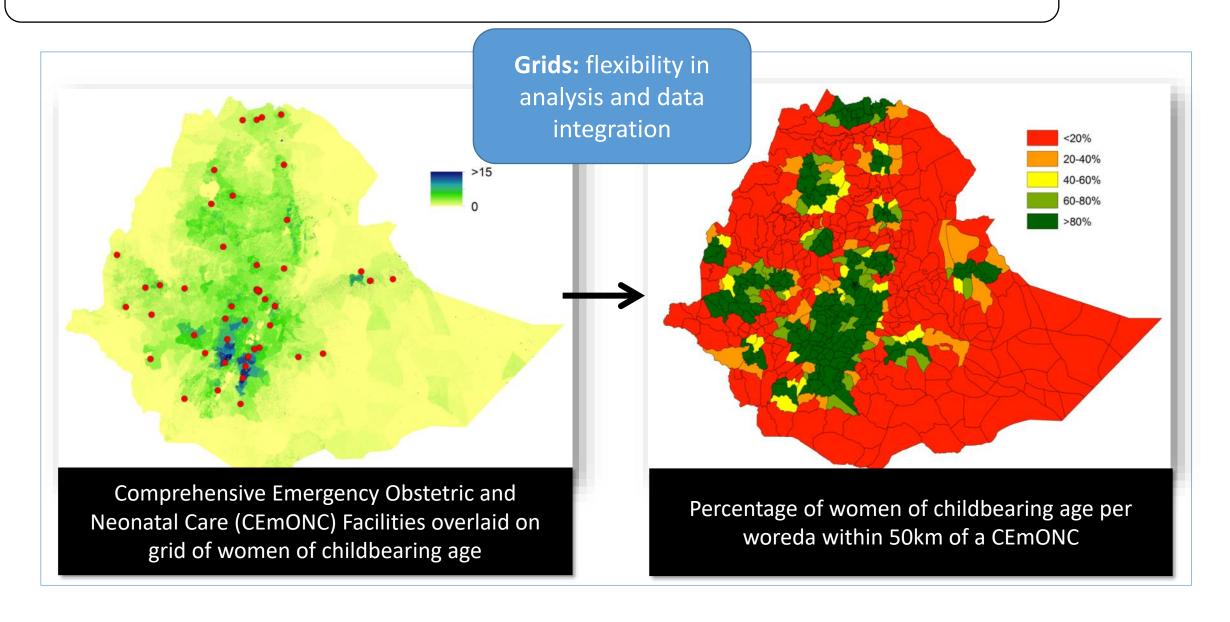




Benefits of 'gridded' demographic data

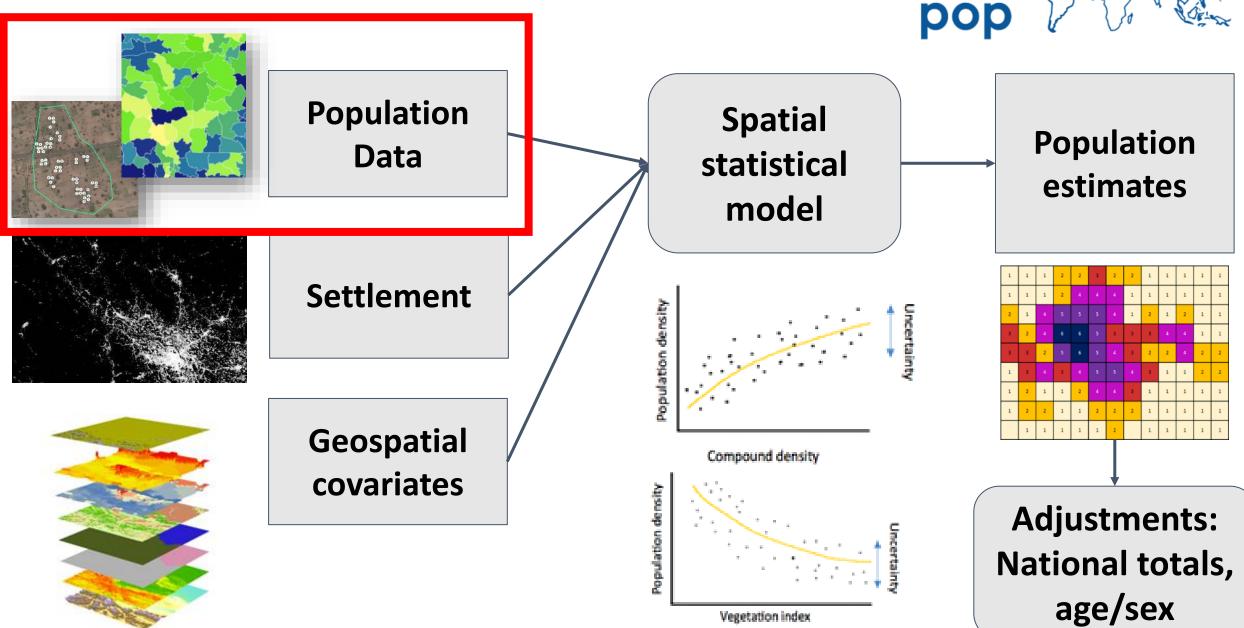


Benefits of 'gridded' demographic data

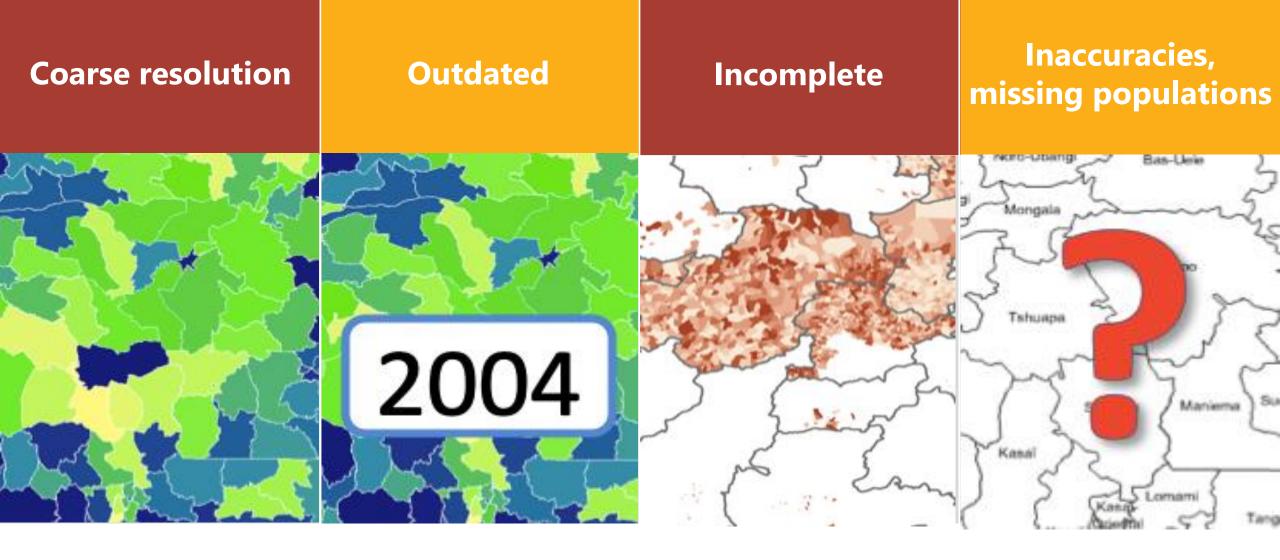


Population distribution modelling





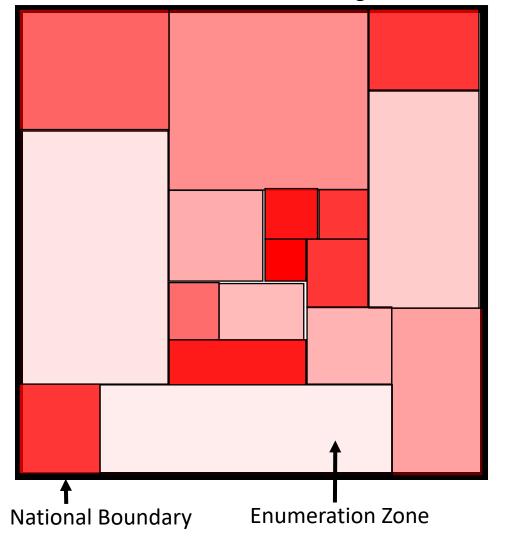
Vegetation index



Demographic data challenges

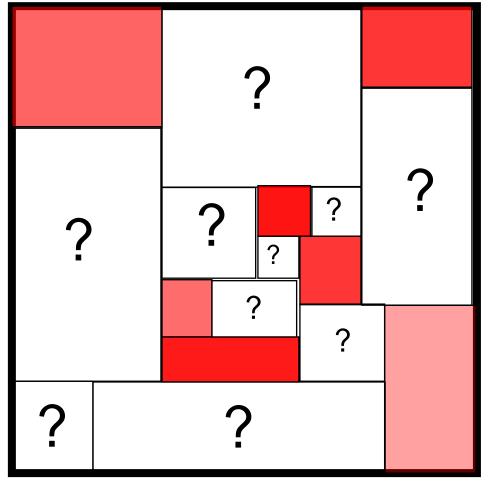
Complete: census or projections

Population totals for enumeration zones with full national coverage.



Incomplete: partial census or enumeration surveys

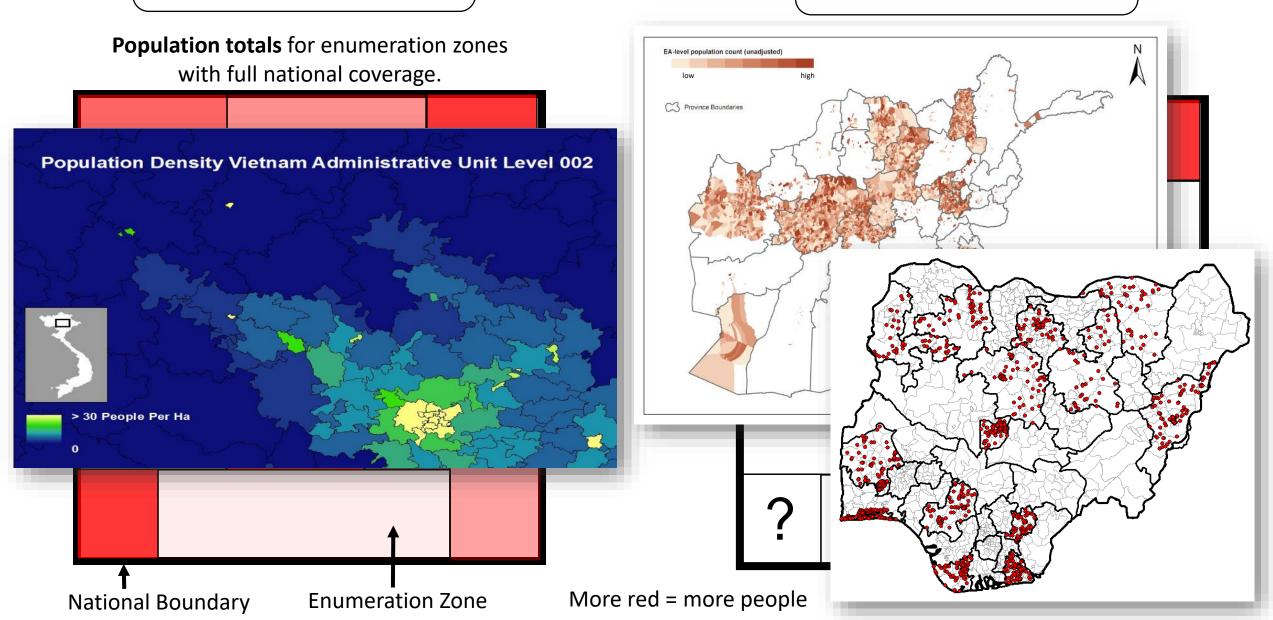
Population totals for a sample of enumeration zones.



More red = more people

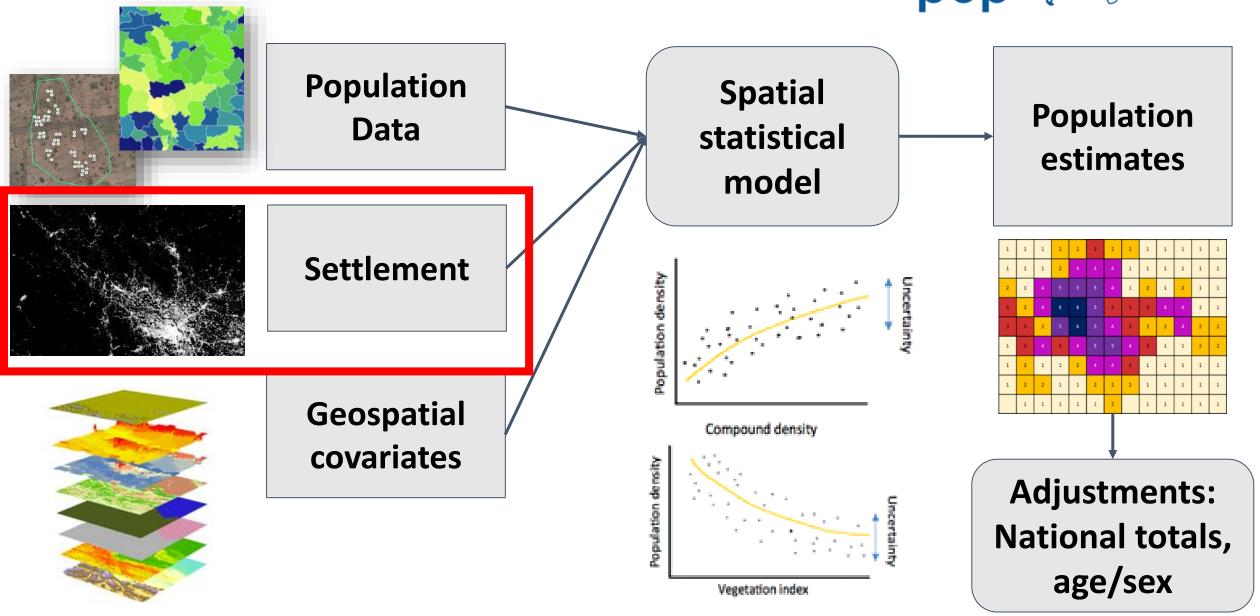
Complete: census or projections

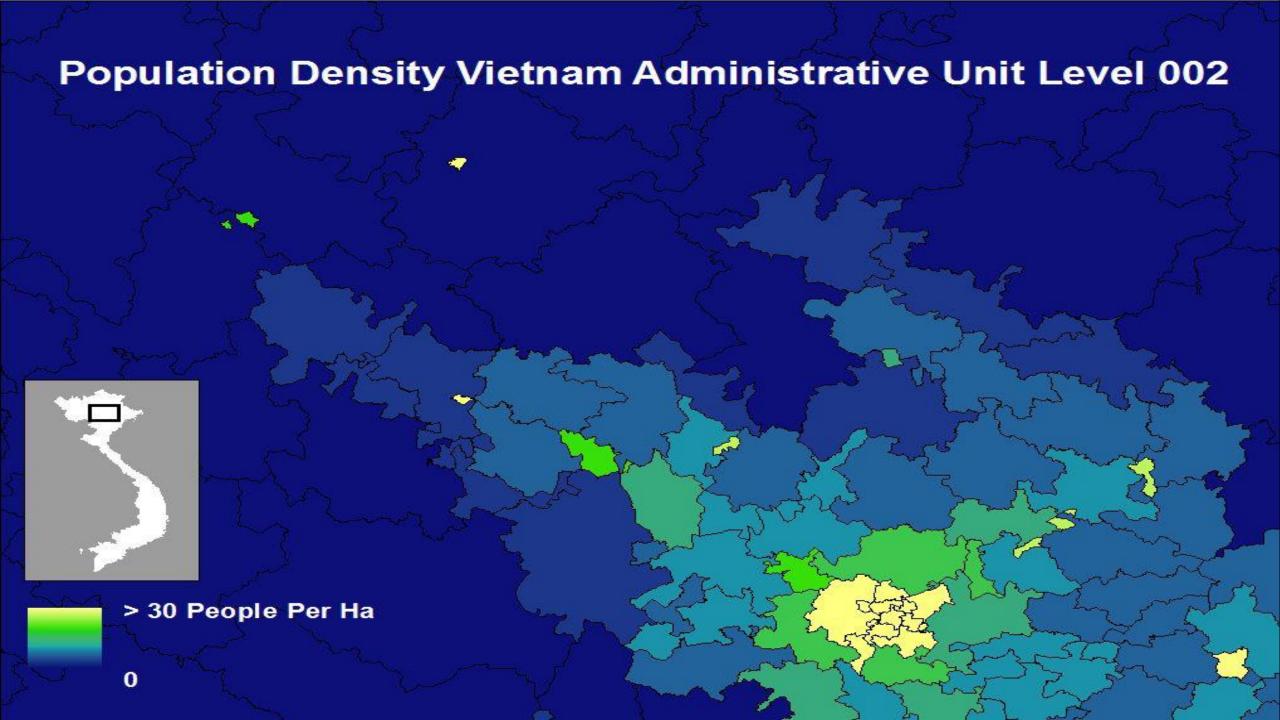
Incomplete: partial census or enumeration surveys

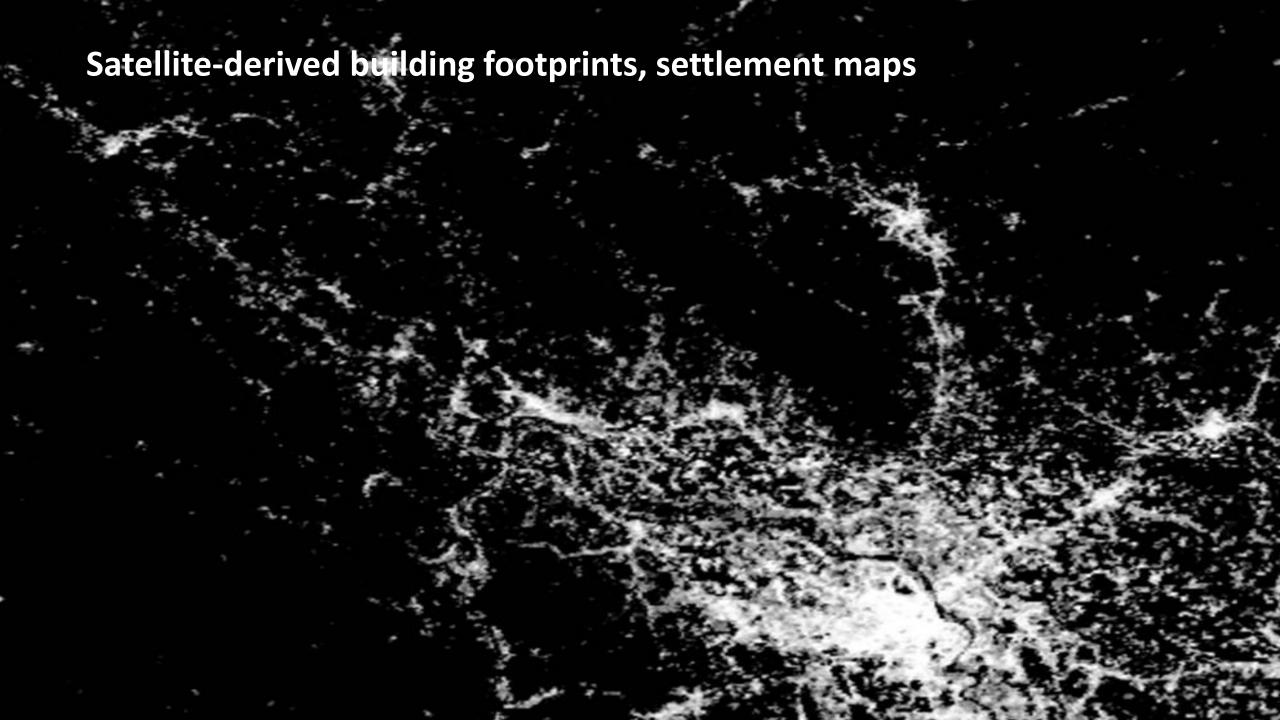


Population distribution modelling

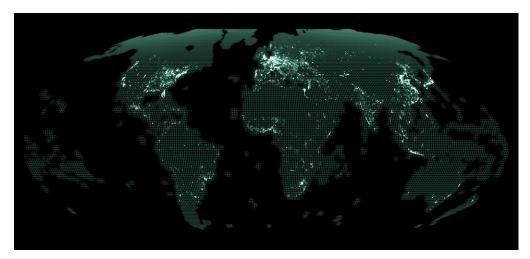




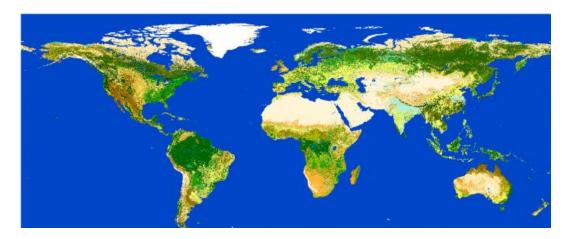




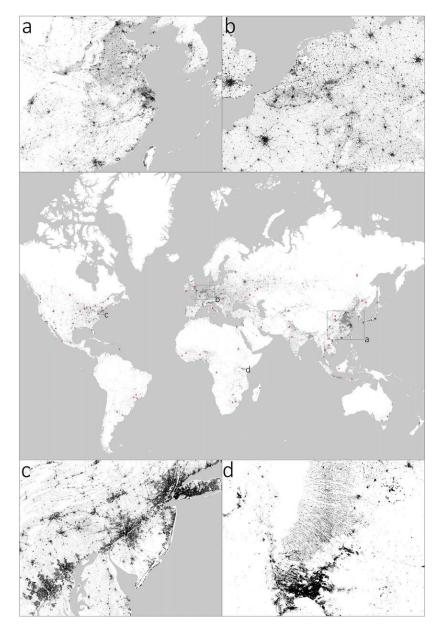
Global settlement datasets



Global Human Settlement Layer (GHSL) 1975, 1990, 2000, 2014

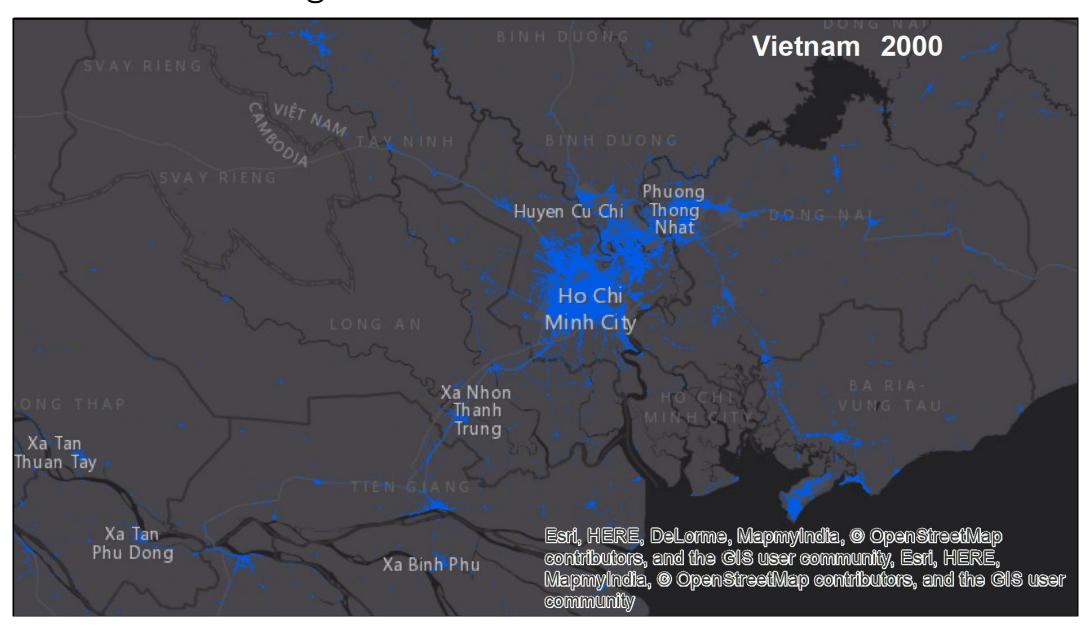


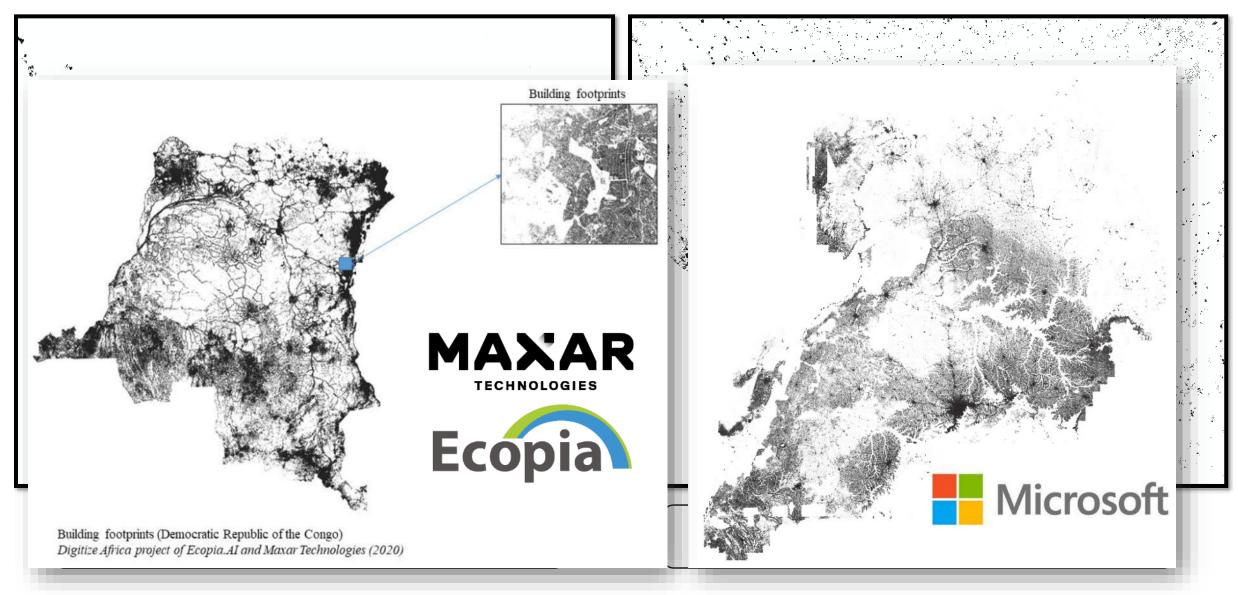
ESA CCI Land cover: 1992-2015



World Settlement Footprint (WSF): 2015

Built settlement growth model: 2000-2020





Nieves et al (2020) Computers, Environment and Urban Systems



Global + multitemporal vs Regional + recent

Built settlement growth model annual estimates 2000-2020

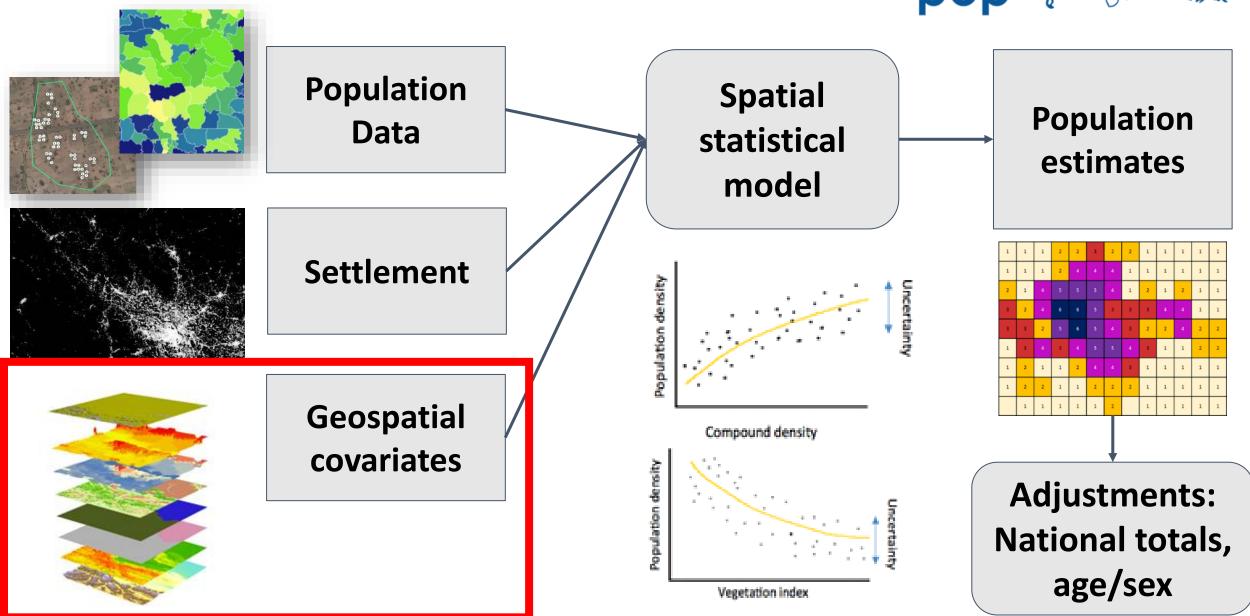


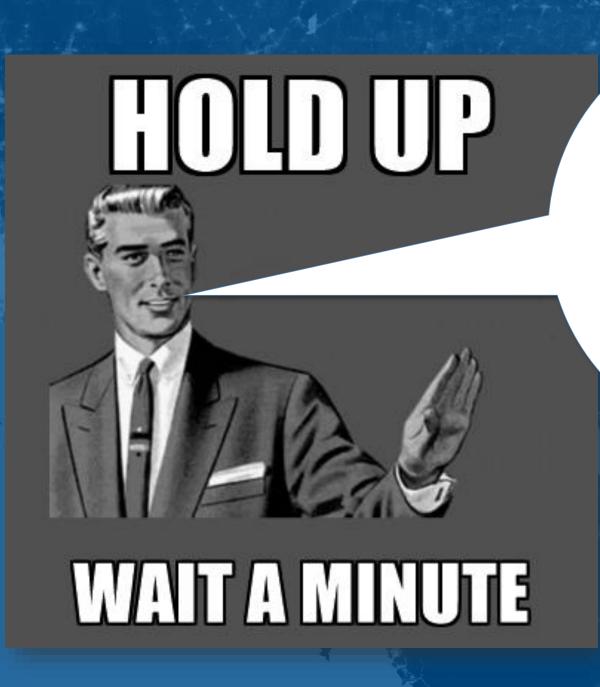
Building footprints circa 2018-20



Population distribution modelling



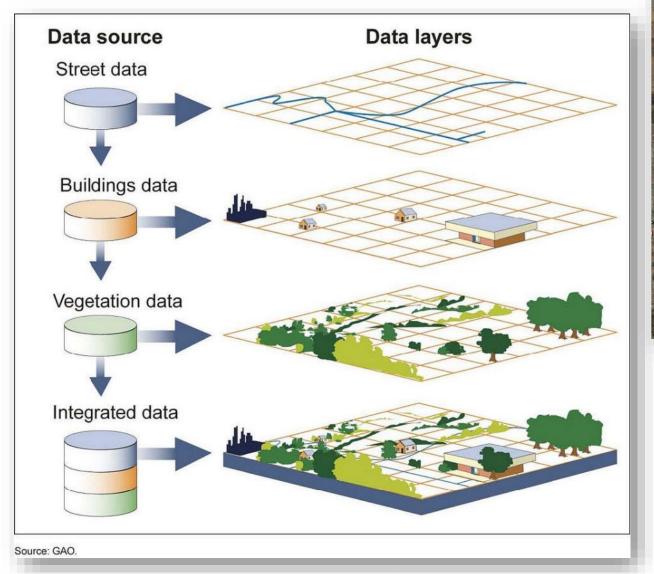


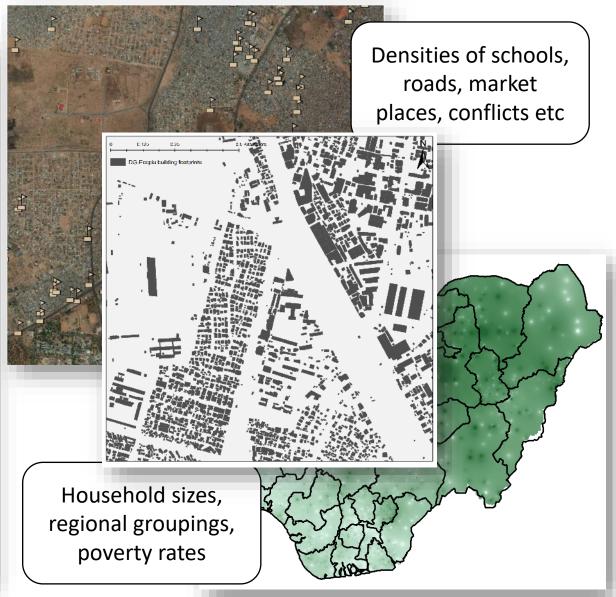


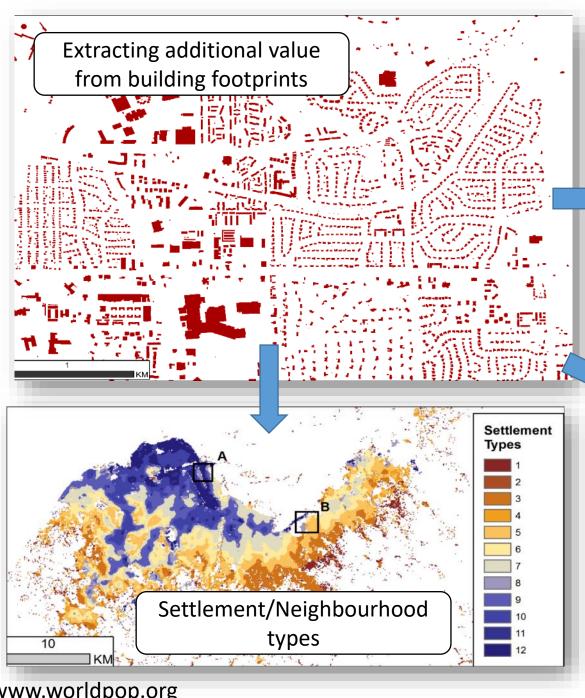
Why do we need these extra layers if we have buildings and people?

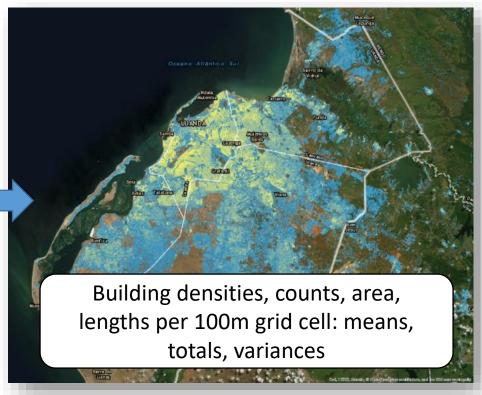


Capturing characteristics that determine variations in population density







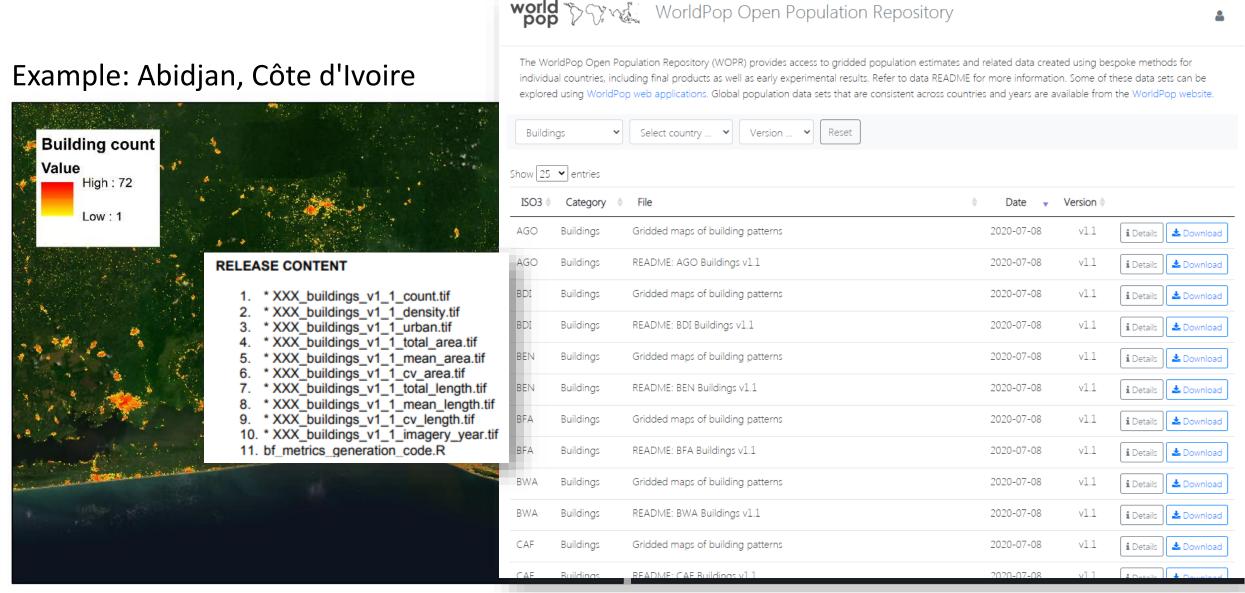


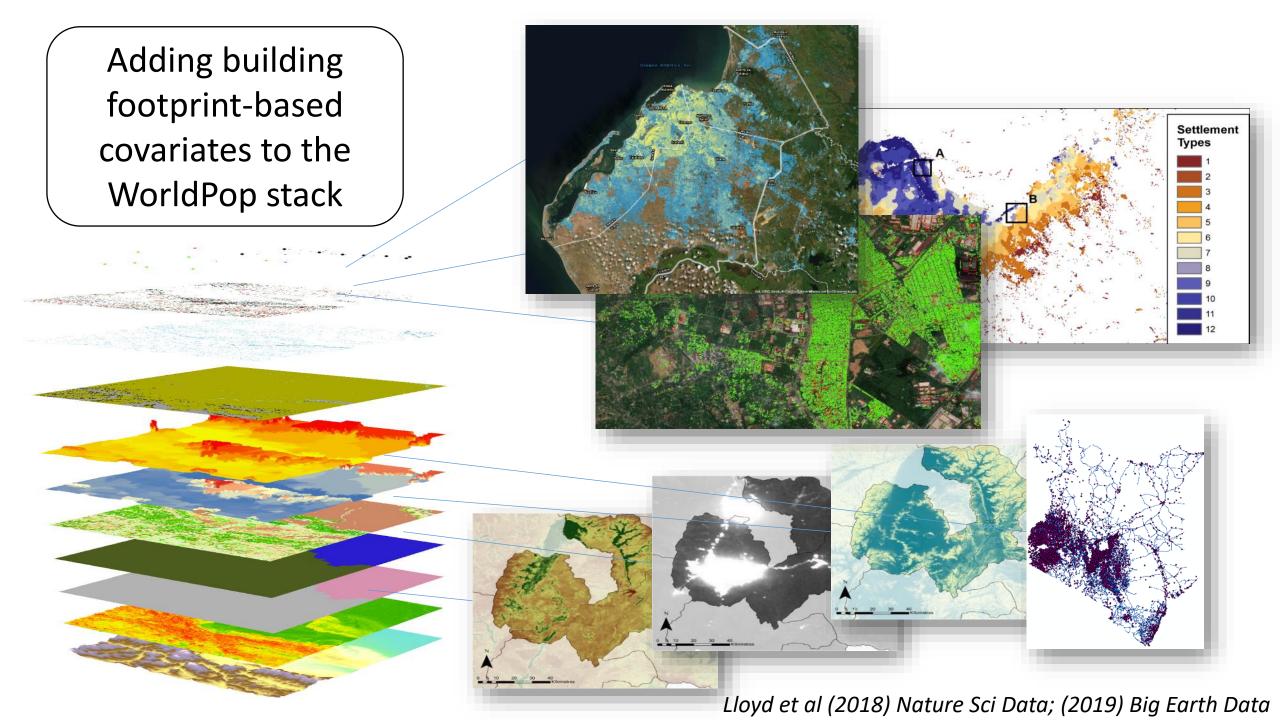


Jochem et al (2019, 2020), Dooley et al (2020), Lloyd et al (2020) in prep

Gridded maps of building patterns throughout sub-Saharan Africa

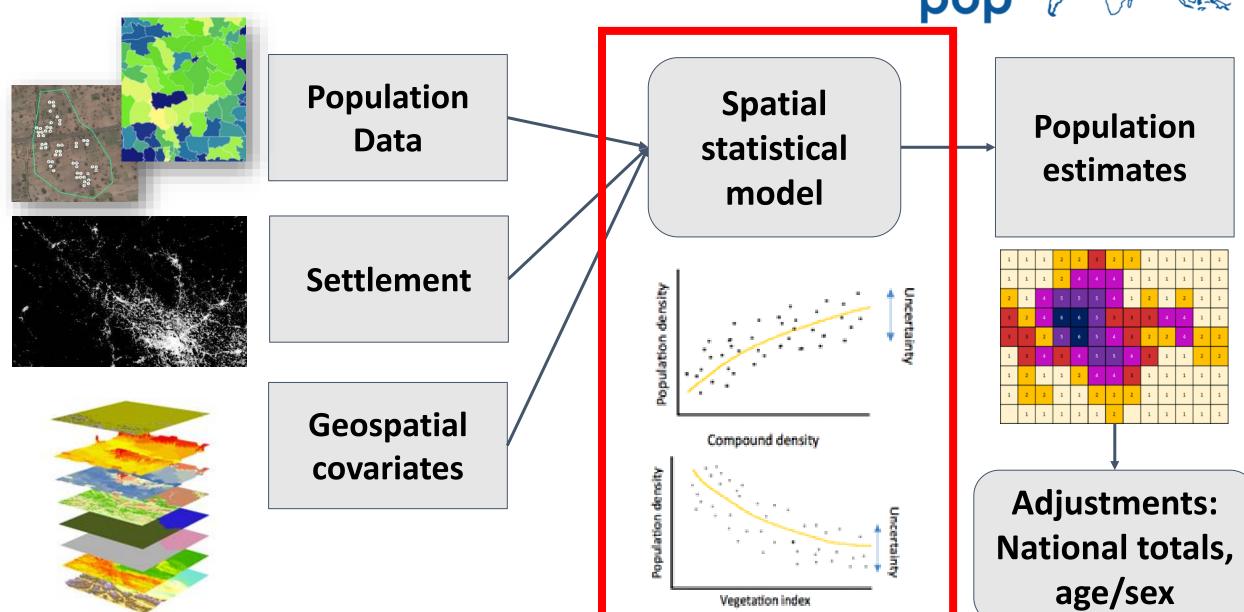
Service Layer Credits: Source: Esri,
DigitalGlobe, GeoEye, Earthstar Geographics,
CNES/Airbus DS, USDA, USGS, AeroGRID,



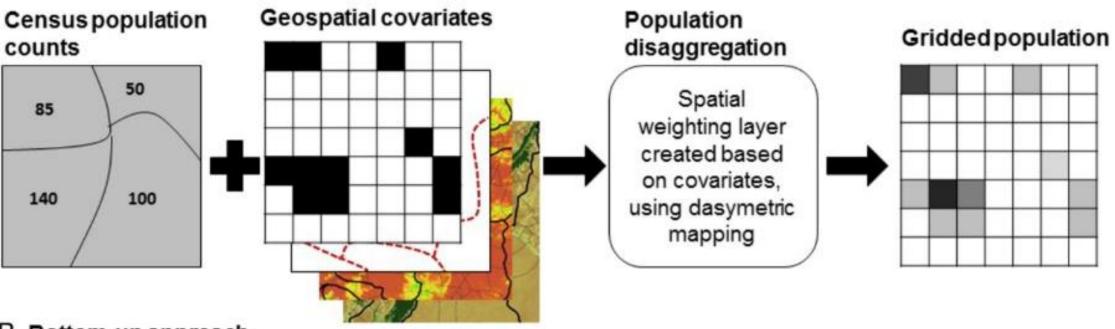


Population distribution modelling

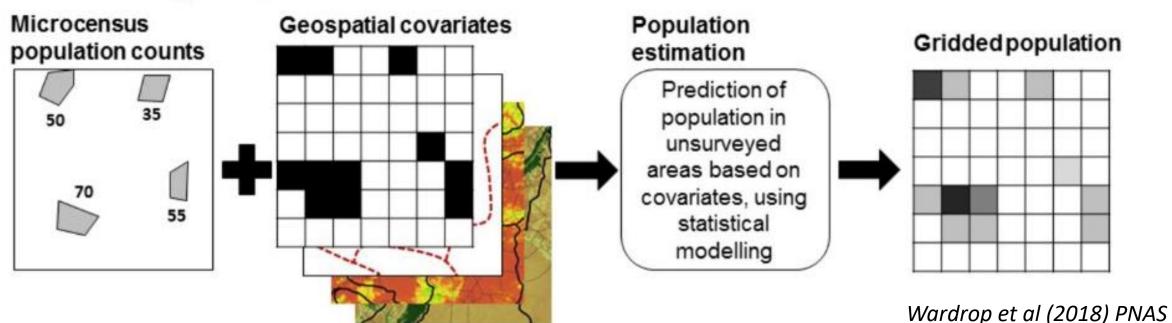


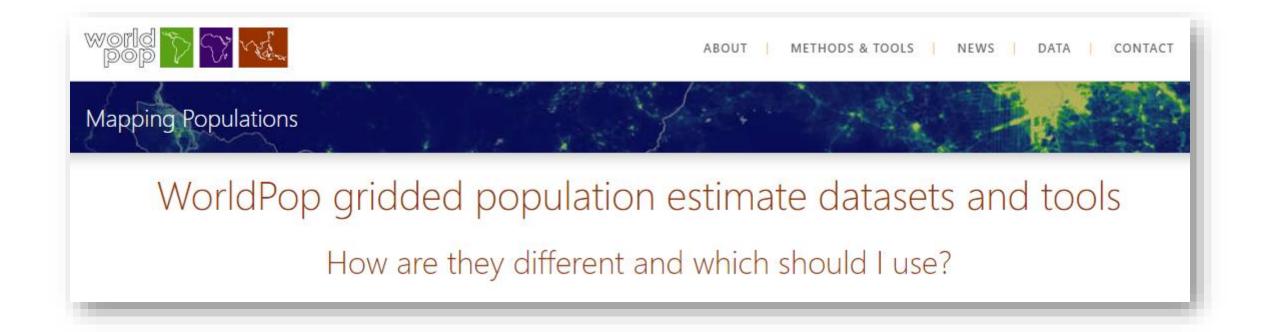


A Topdown approach



B Bottom-up approach

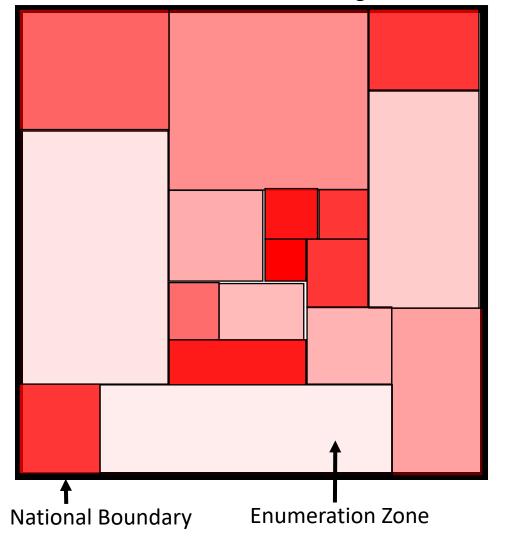




www.worldpop.org/methods/populations

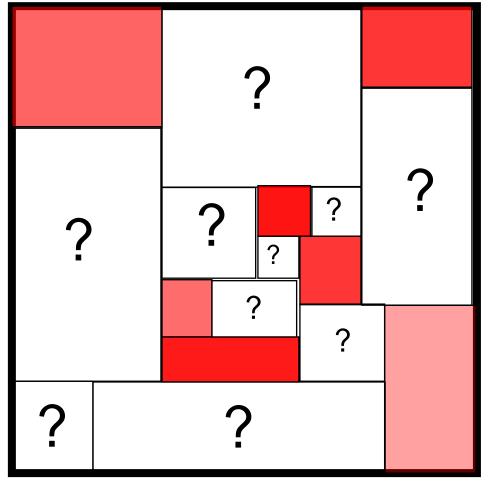
Complete: census or projections

Population totals for enumeration zones with full national coverage.



Incomplete: partial census or enumeration surveys

Population totals for a sample of enumeration zones.



More red = more people

Inputs

Top-down

Population totals for enumeration zones with full national coverage.

"I have recent and reliable census counts matched to boundaries and want gridded estimates"

"I trust these subnational province-level projections, but need finer scale estimates"

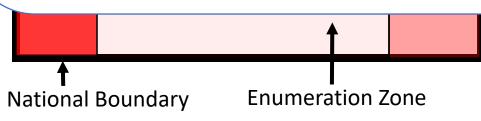
"I need gridded outputs that match my district unit totals"

Bottom-up

Population totals for a sample of enumeration zones.

"The last national census was 1984 so don't trust that data, but I need small area population data and do have some recent sample enumeration data"

"I have geolocated listings from a recent household survey and want to use these to build population estimates"

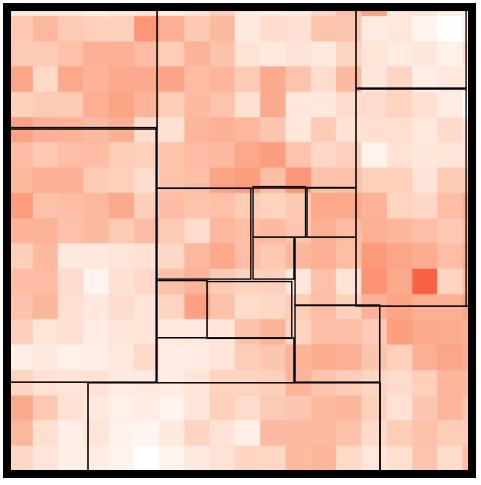


More red = more people

Inputs

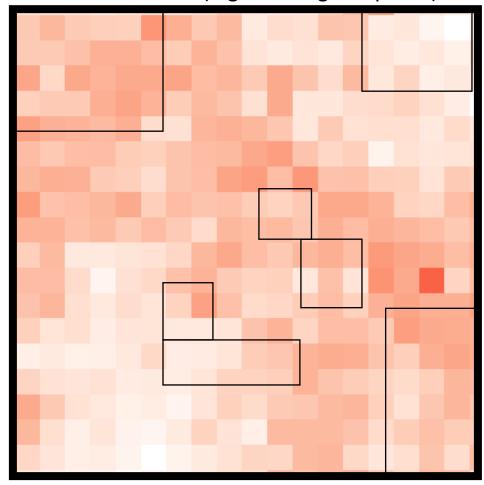
Top-down

Geospatial covariates and **settlement data** (e.g. building footprints)



Bottom-up

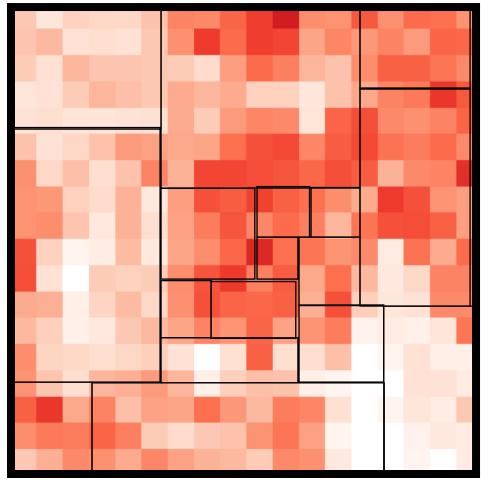
Geospatial covariates and **settlement data** (e.g. building footprints)



Outputs

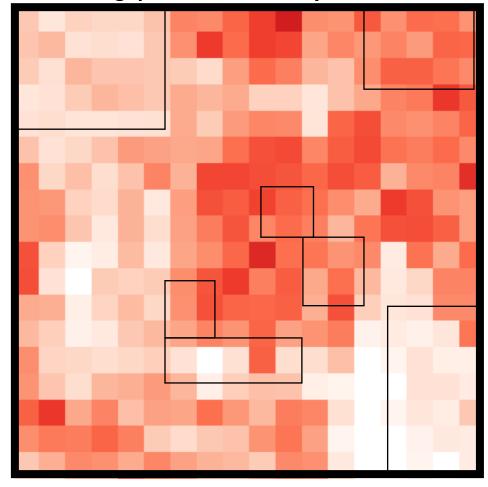
Top-down

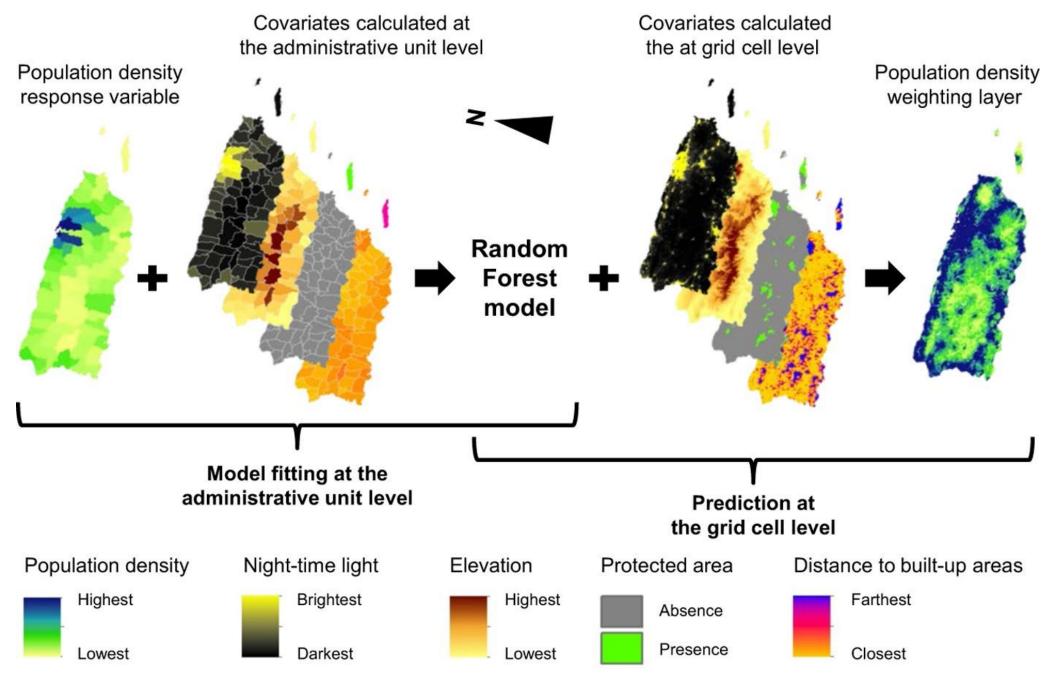
Gridded population estimates (100 m) that sum to pre-defined population totals.



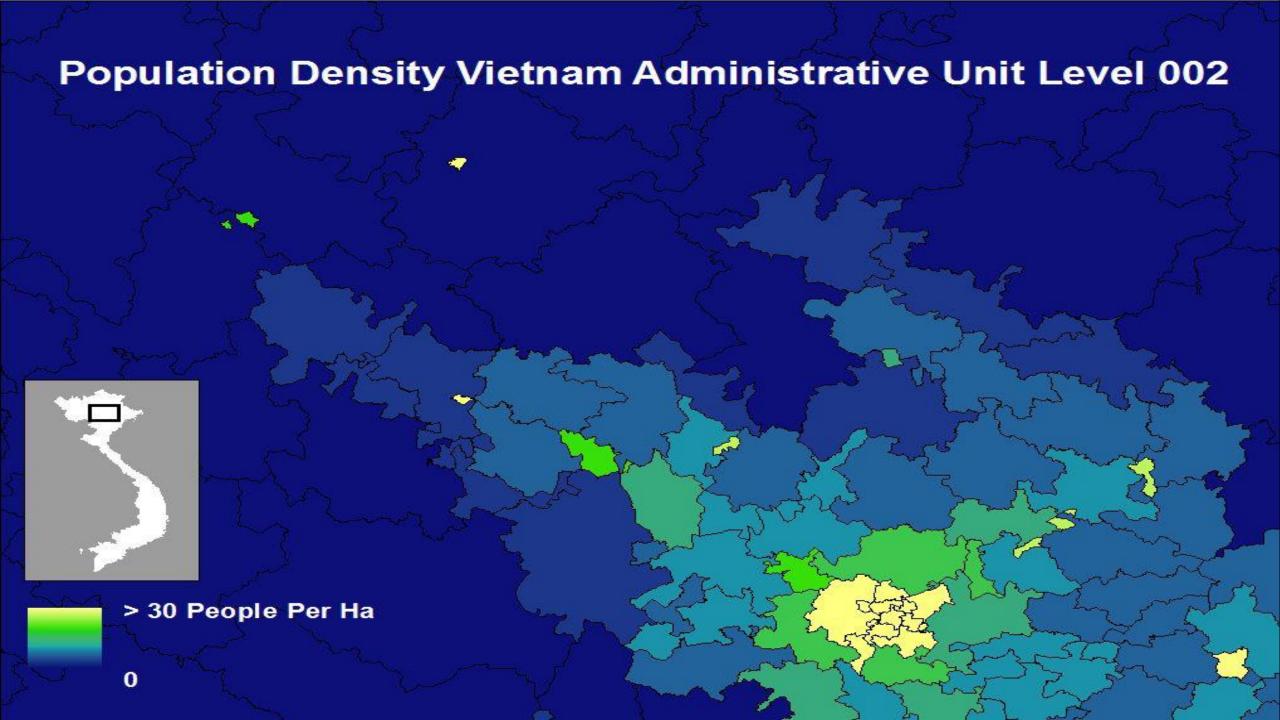
Bottom-up

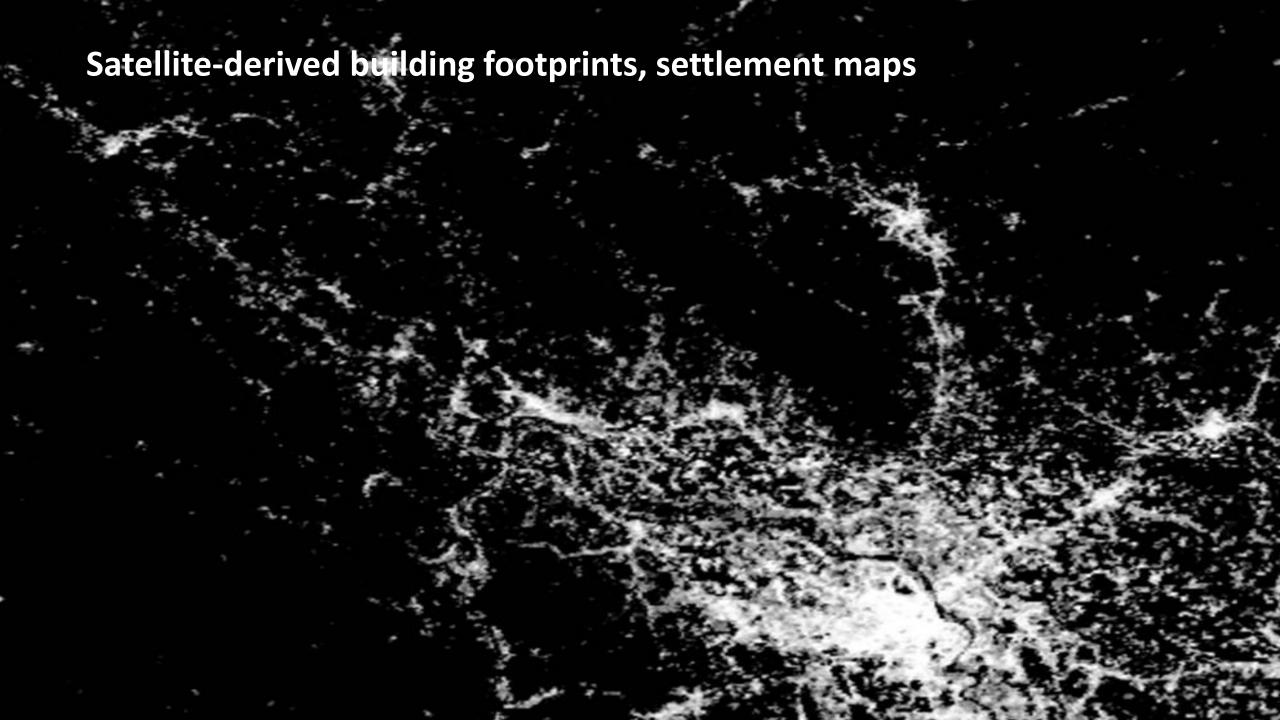
Gridded population estimates (100 m) that fill gaps between surveyed areas.

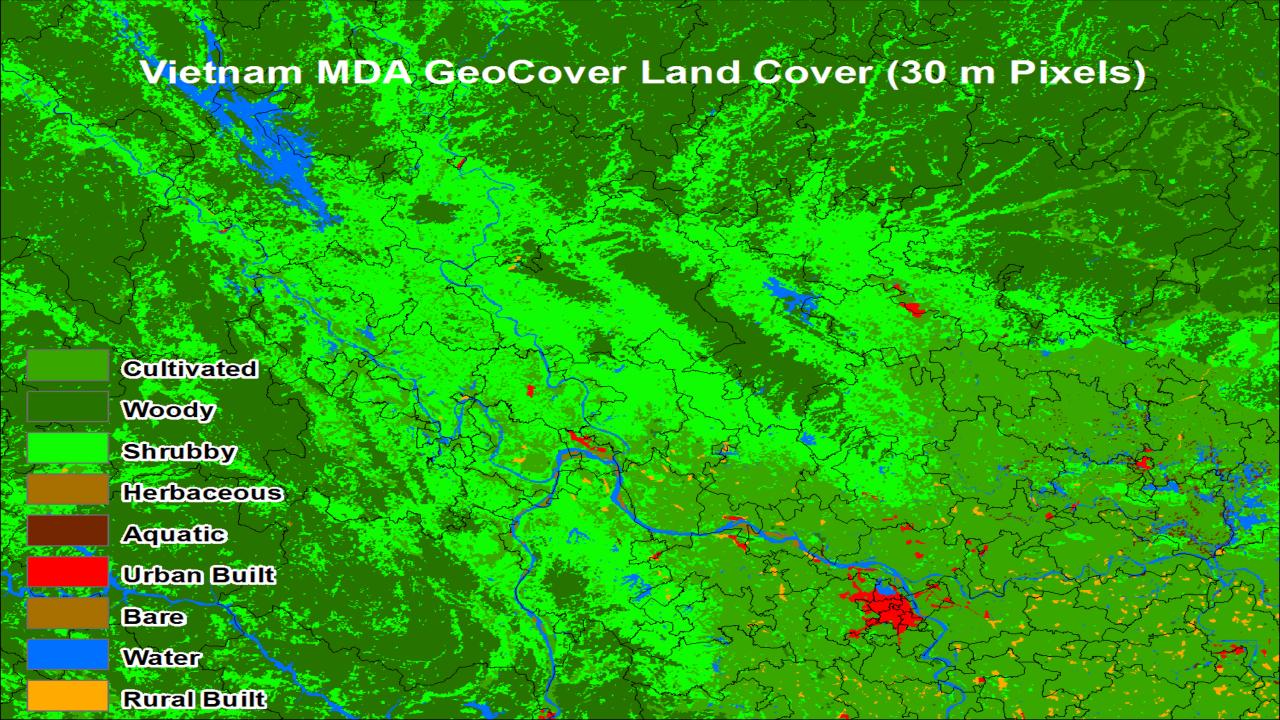




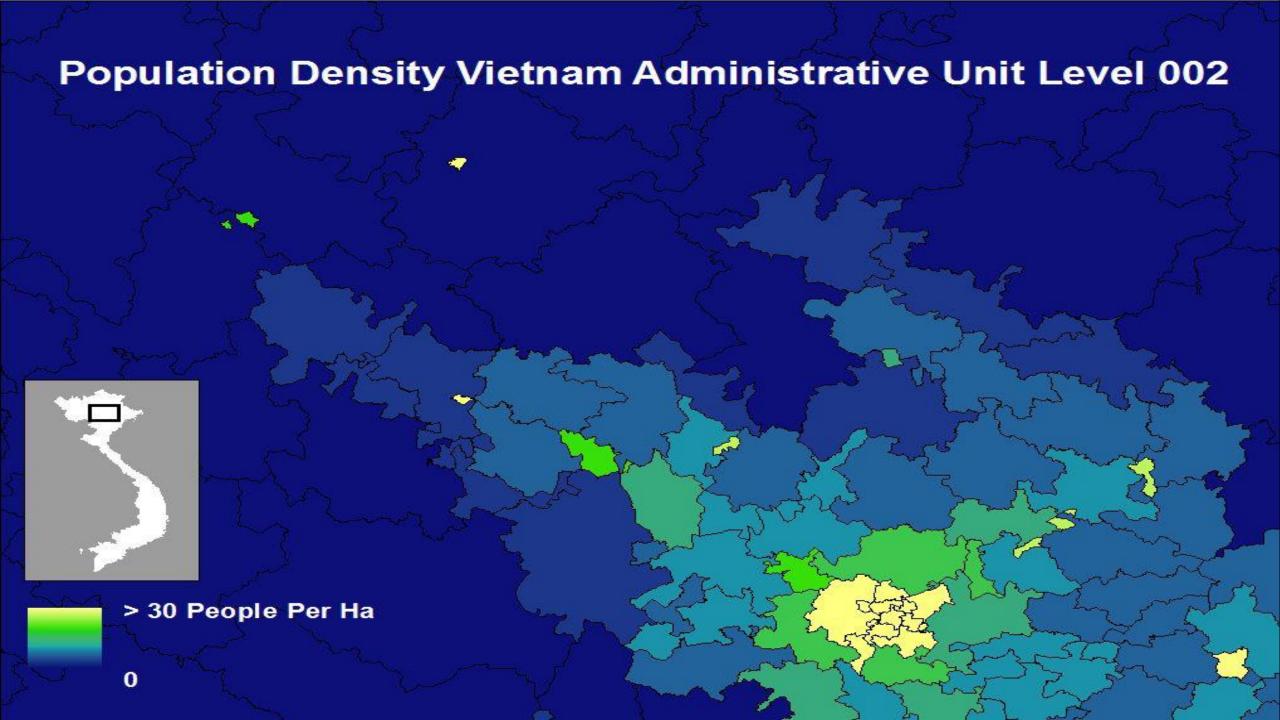
Stevens et al (2014) PLoS ONE; Sorichetta et al (2015) Nature Sci Data

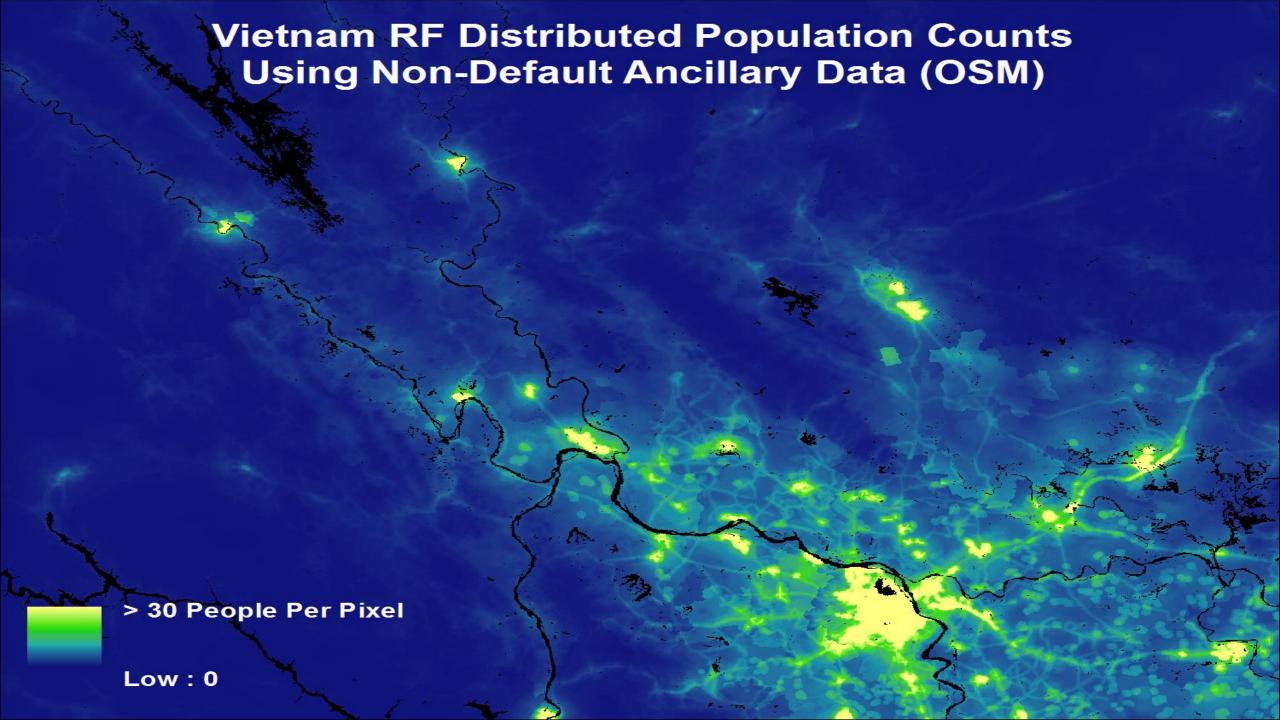




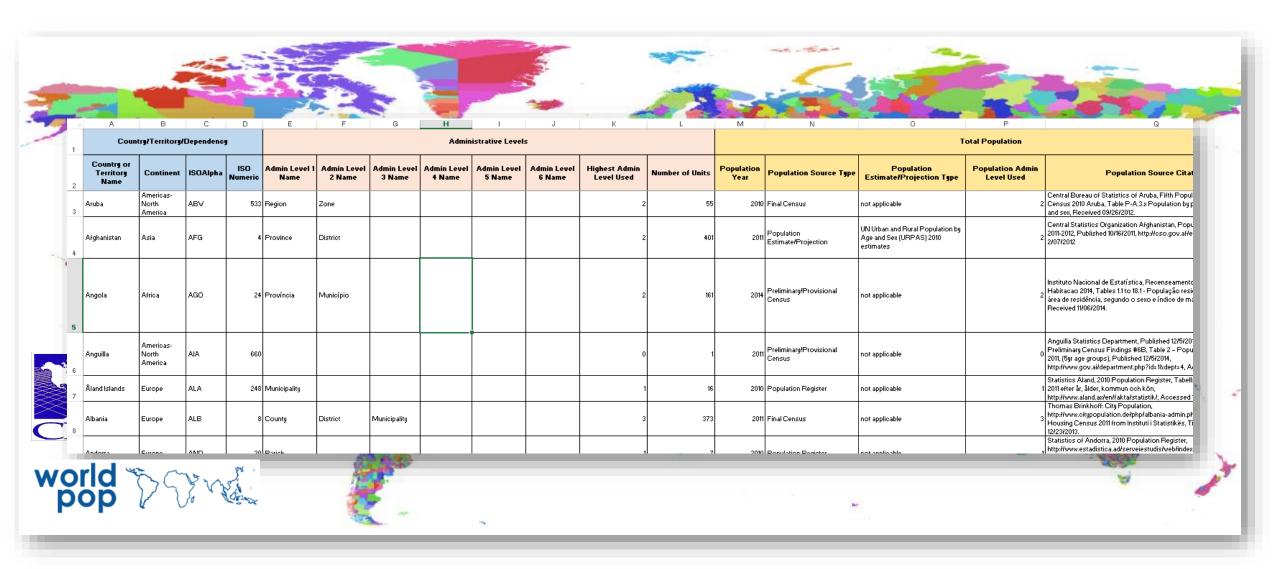


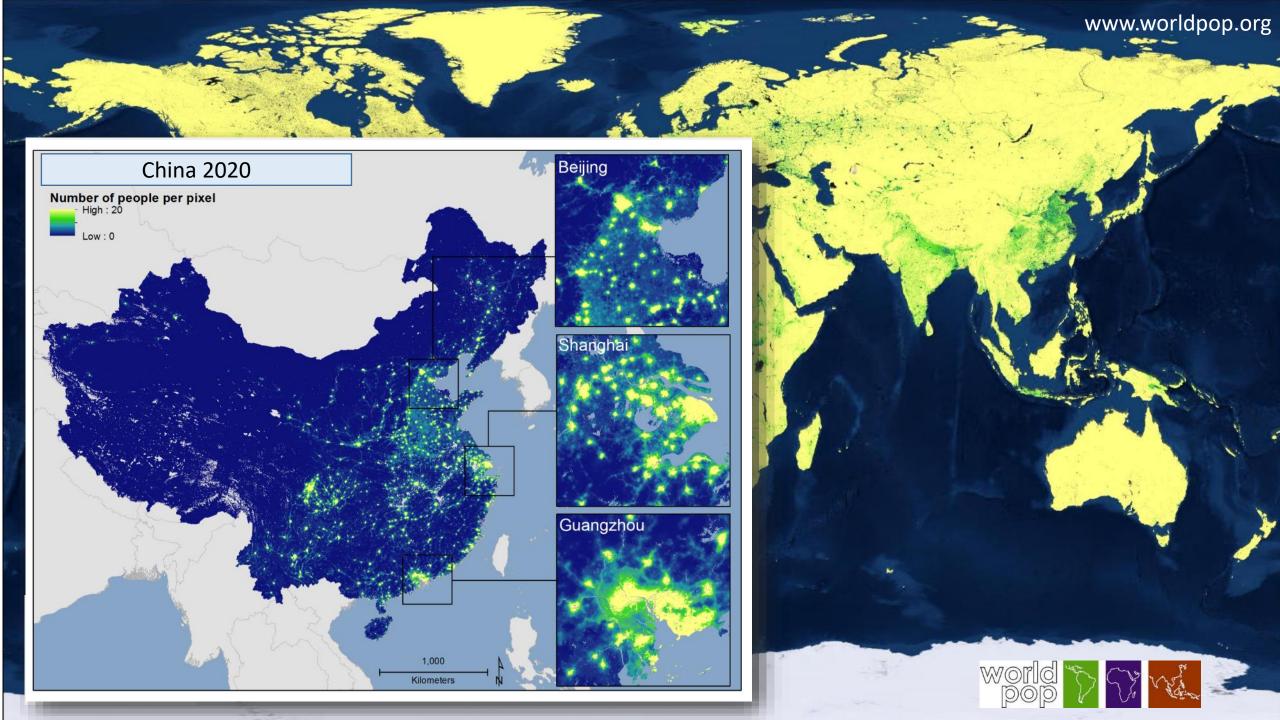
NOAA Suomi VIIRS-derived Lights at Night 2012 for Vietnam





2000-2020 subnational census/projections



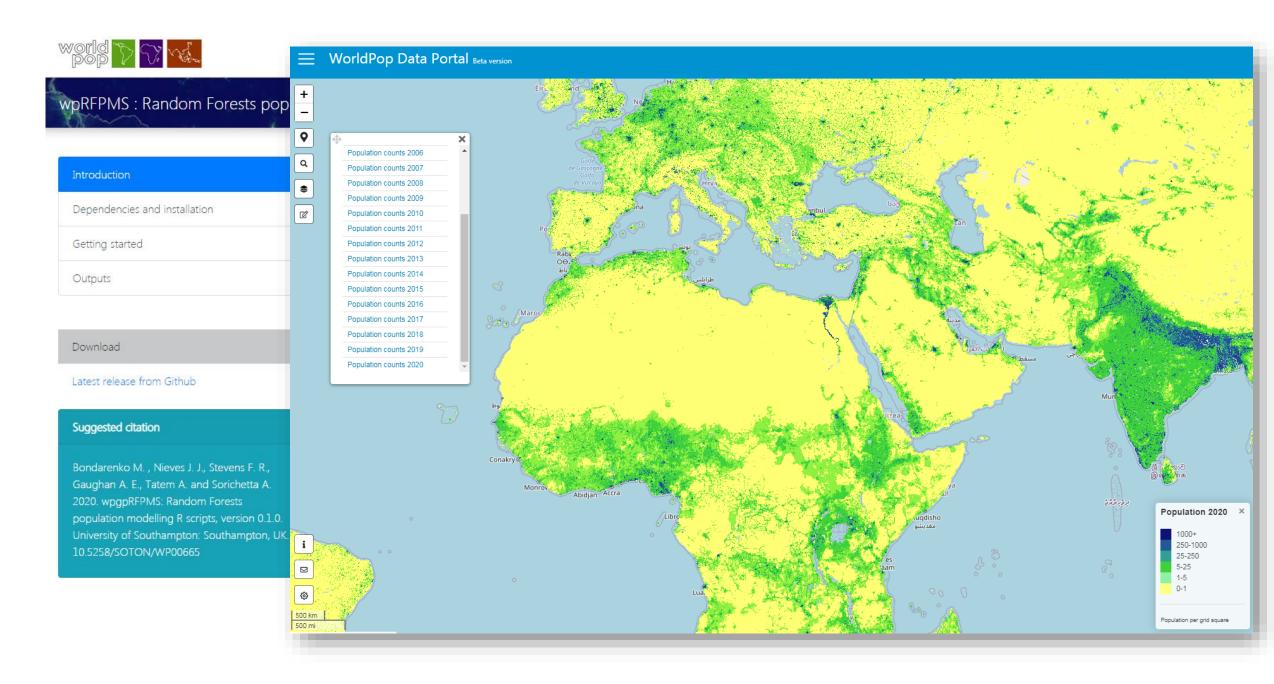


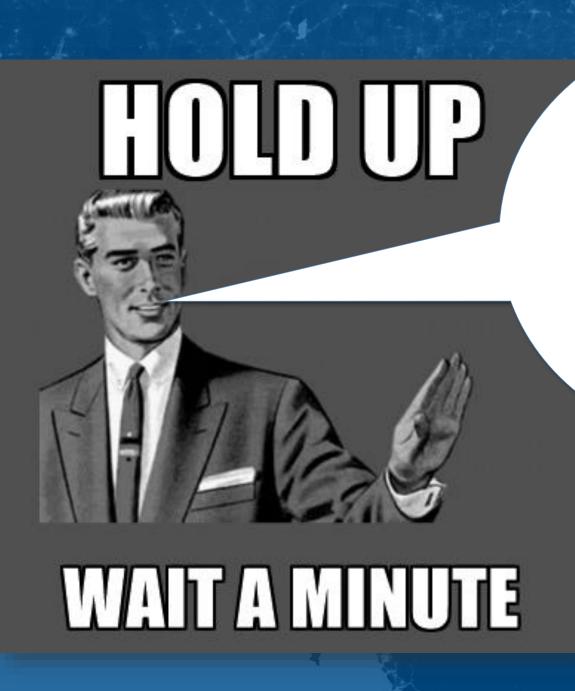
Algeria Population Map Metadata Report

Covariate importance for DZA RF Model proximately 100 m at the equator) obtained using the Random Forest (RF)-Elevation used to dasymetrically disaggregate population counts from administrative eing adjusted to match the most recent UNPD estimates# available at the al distribution of each covariate, is also provided in this metadata report. Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. Random Forest Diagnostics Covariate importance for DZA RF Model 2000 Prediction Error (MSE) Stability for DZA RF Model 1500 Covariate Metadata 1000 - 500 Distance to cultivated areas 2006 Distance to woody areas 2006 Distance to cultivated areas 2006 Distance to herbaceous areas 2006 Distance to enarce vegetation areas 2006

Торо

Data Source: Viewfinder Panoramas (http://viewfinderpanoramas.org/), Digital Elevation Data (http://viewfinderpanoramas.org/dem3.html) Download :ftp://ftp.worldpop.org.uk/WP515640_Global/time_invariant





Didn't you mention 'unconstrained' and 'constrained' estimates?

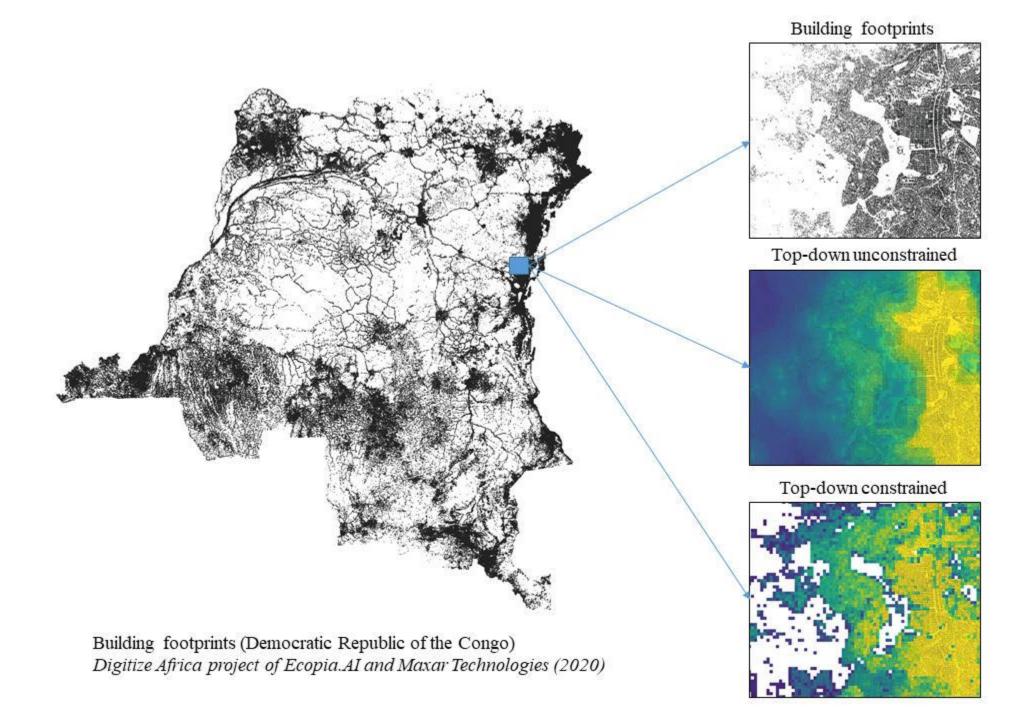
Global + multitemporal vs Regional + recent

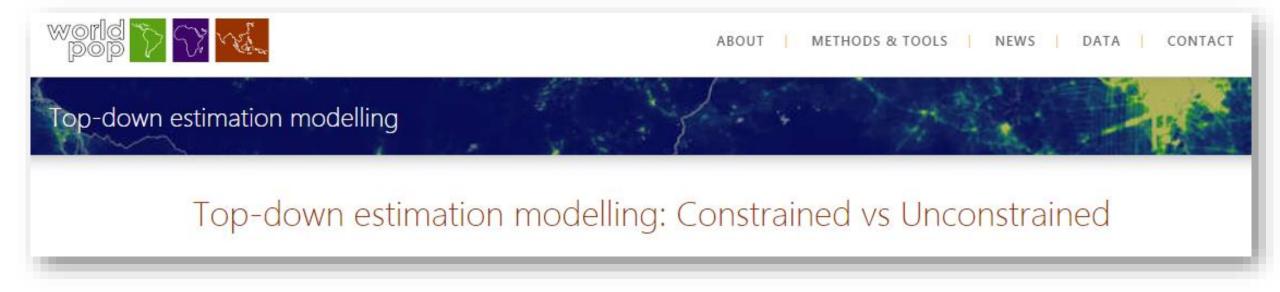
Built settlement growth model annual estimates 2000-2020



Building footprints circa 2018-20

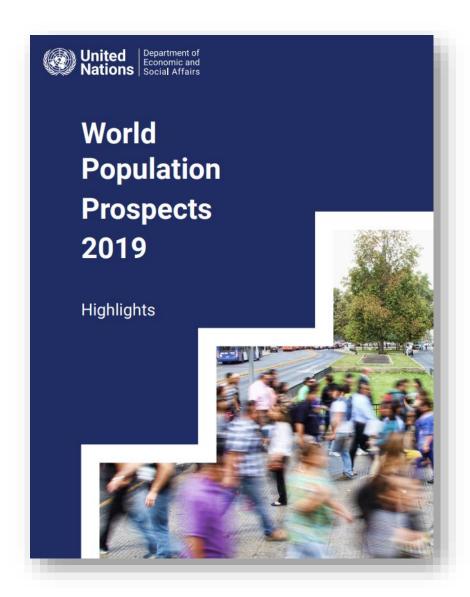






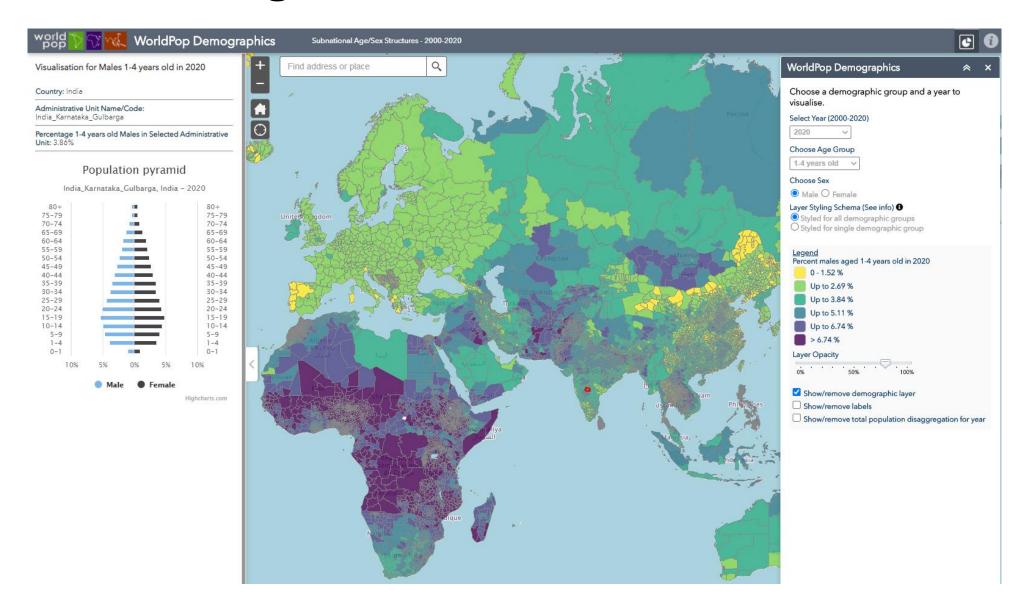
www.worldpop.org/methods/top_down_constrained_vs_unconstrained

Adjustments: National population totals



- non-adjusted = maintaining GPWv4
 census database estimates
- UN-adjusted = national population total adjusted to match World
 Population Prospects 2019 estimates

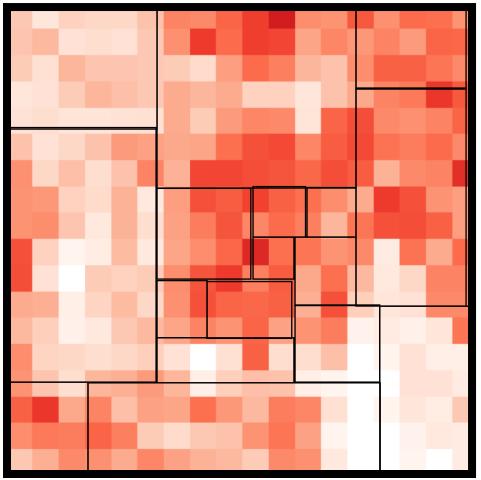
Adjustments: Age/sex structures



Outputs

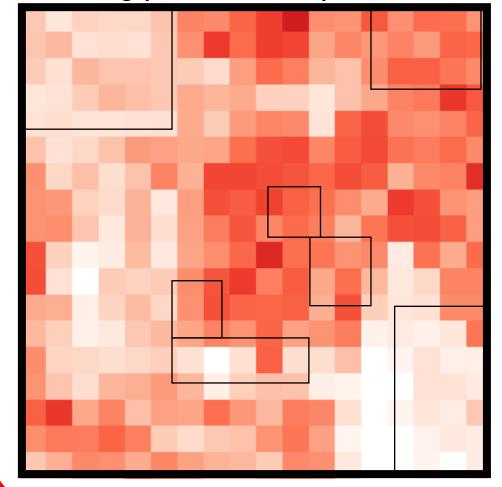
Top-down

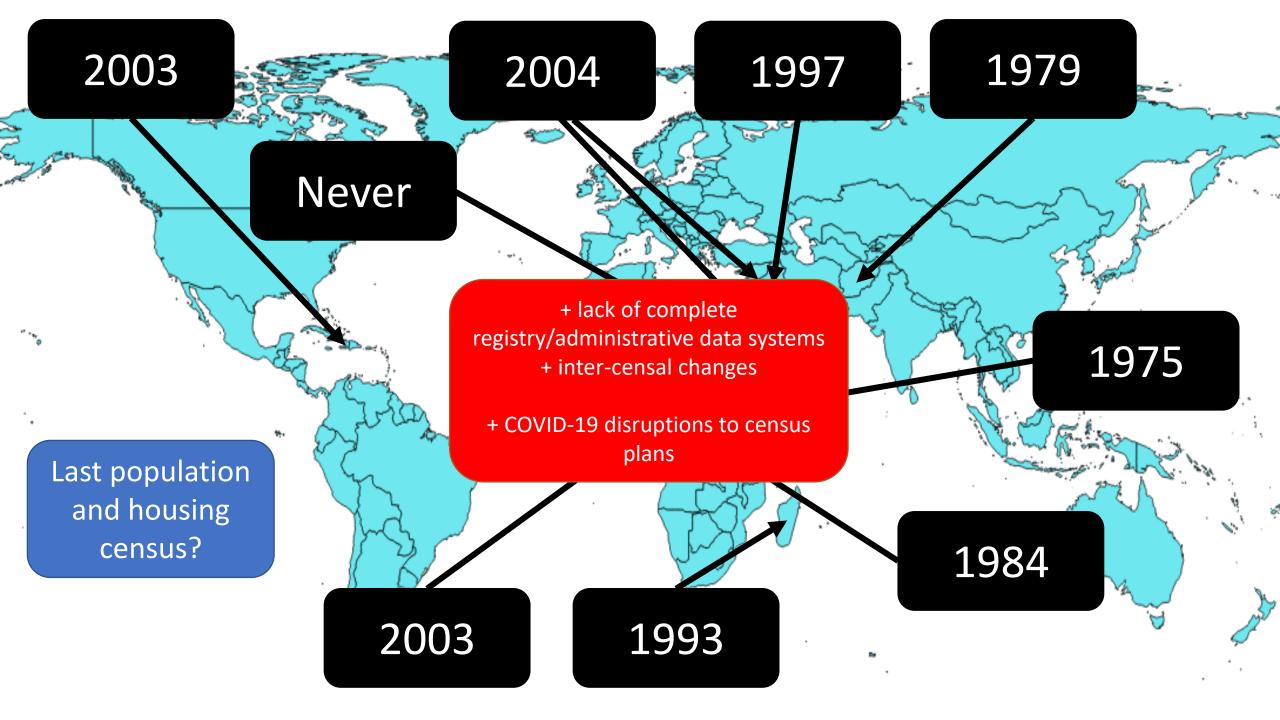
Gridded population estimates (100 m) that sum to pre-defined population totals.



Bottom-up

Gridded population estimates (100 m) that fill gaps between surveyed areas.





Afghanistan

- Last national population census in 1979
- Current estimates largely based on projections
- Significant uncertainties in national and subnational estimates
- One-third of country covered by a rolling census, but insecurity preventing additional data collection

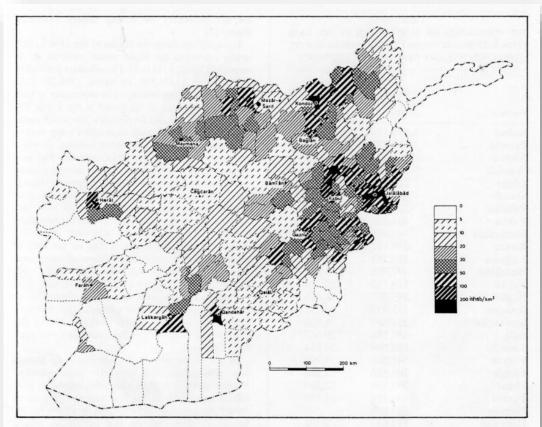
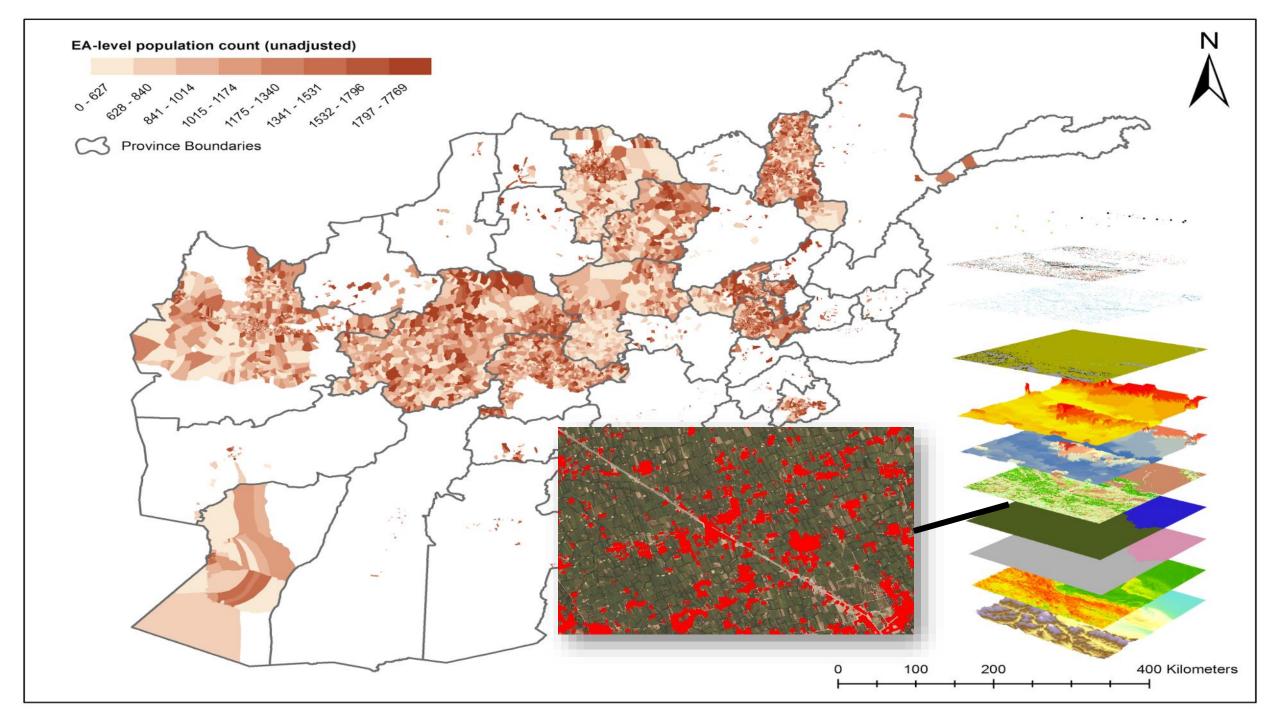
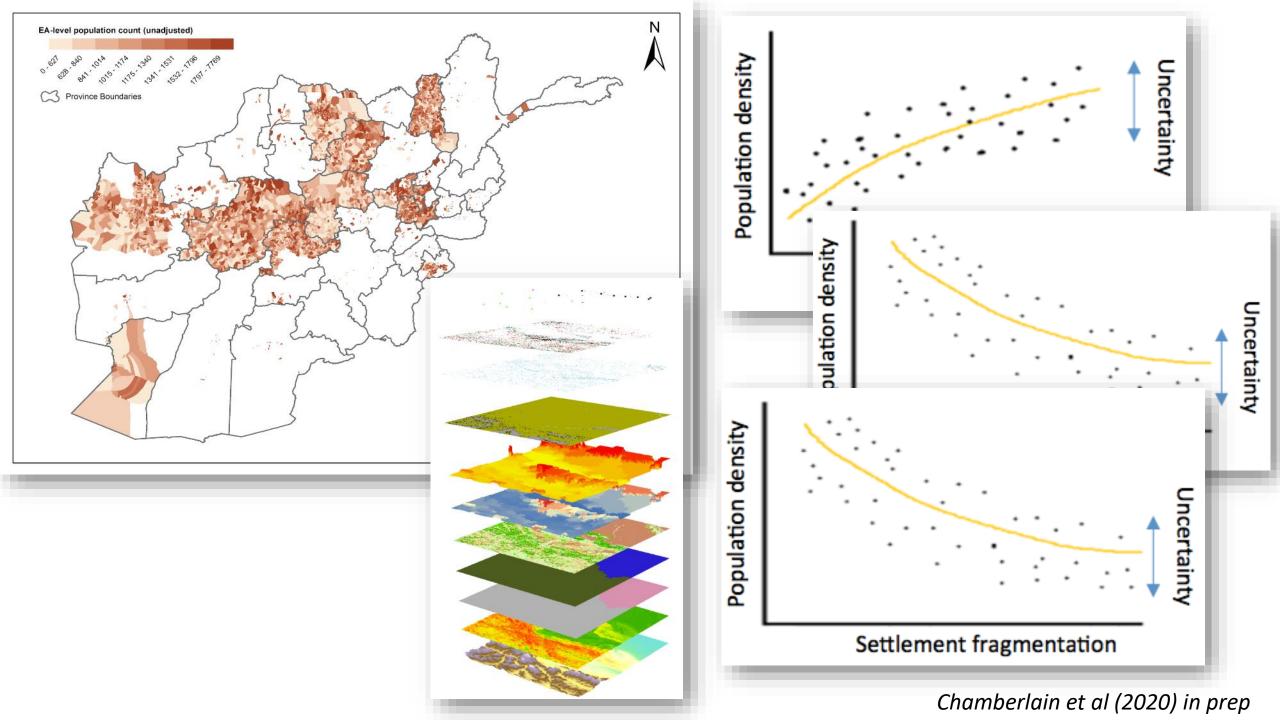
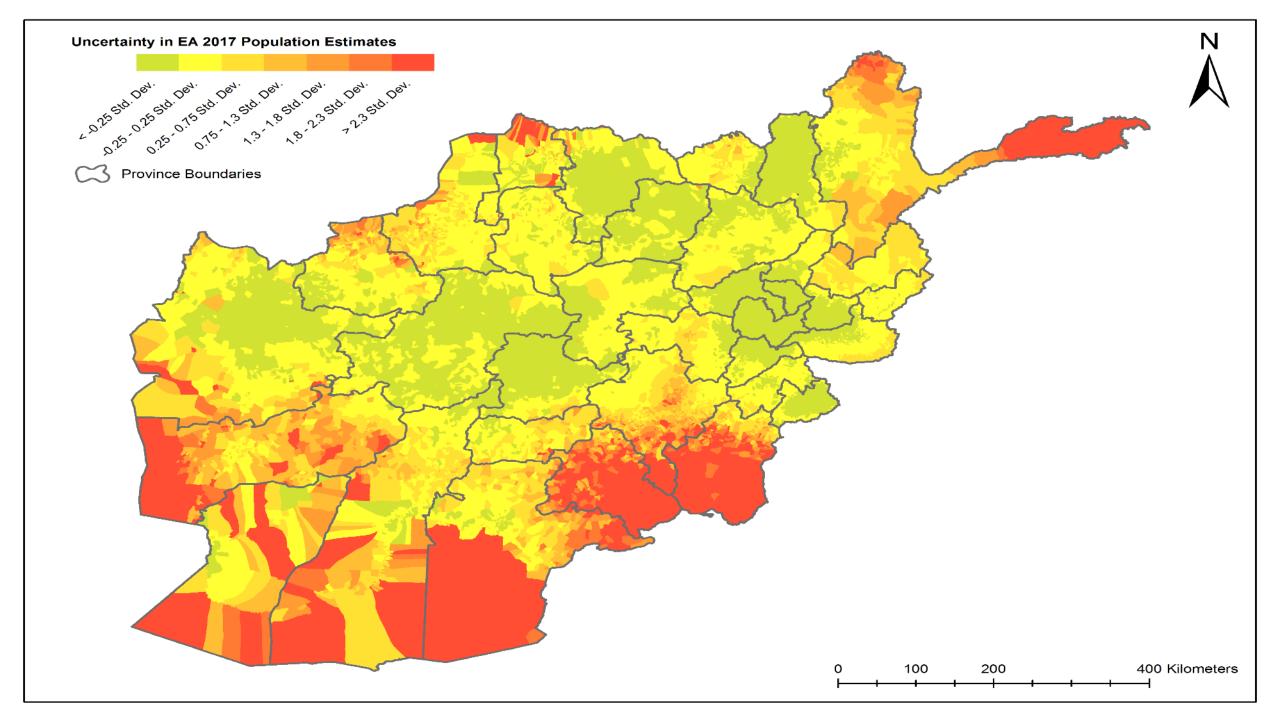


Figure 12. Geographical distribution of the settled population in Afghanistan, according to the 1358 Š./1979 census. Source: CSO, *Natāyej*, pp. 148ff.





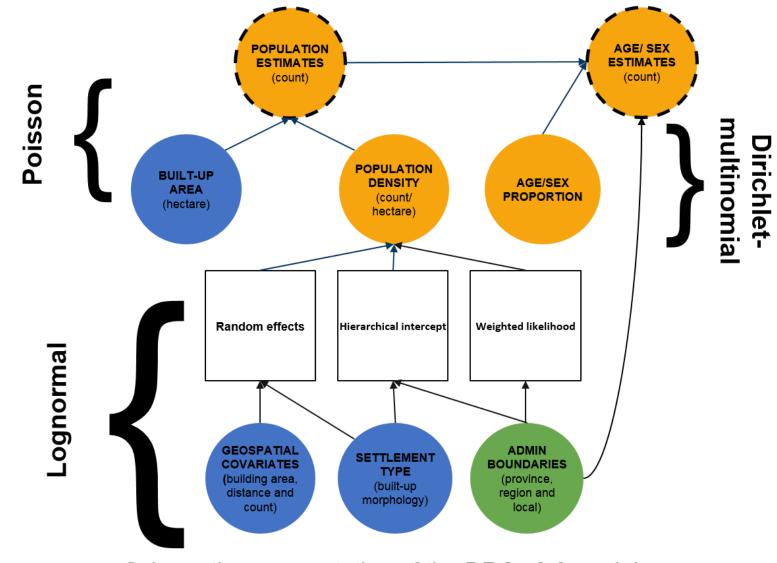


Bespoke models

Bayesian hierarchical model

Production of population estimates with uncertainty metrics broken down by age/sex

Variants developed and implemented for Zambia, Nigeria, DRC, Burkina Faso



Schematic representation of the DRC v2.0 model



Population

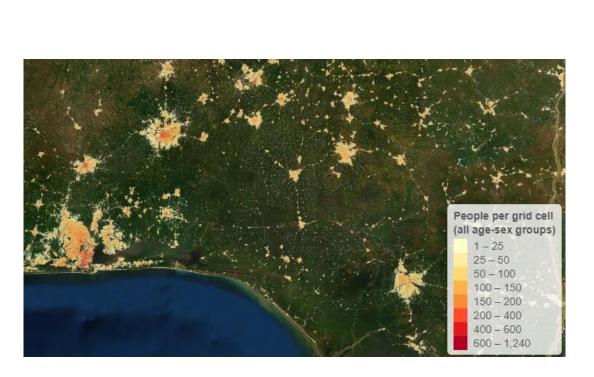
WorldPop Open Population Repository

Version ...



The WorldPop Open Population Repository (WOPR) provides access to gridded population estimates and related data created using bespoke methods for individual countries, including final products as well as early experimental results. Refer to data README for more information. Some of these data sets can be explored using WorldPop web applications. Global population data sets that are consistent across countries and years are available from the WorldPop website.

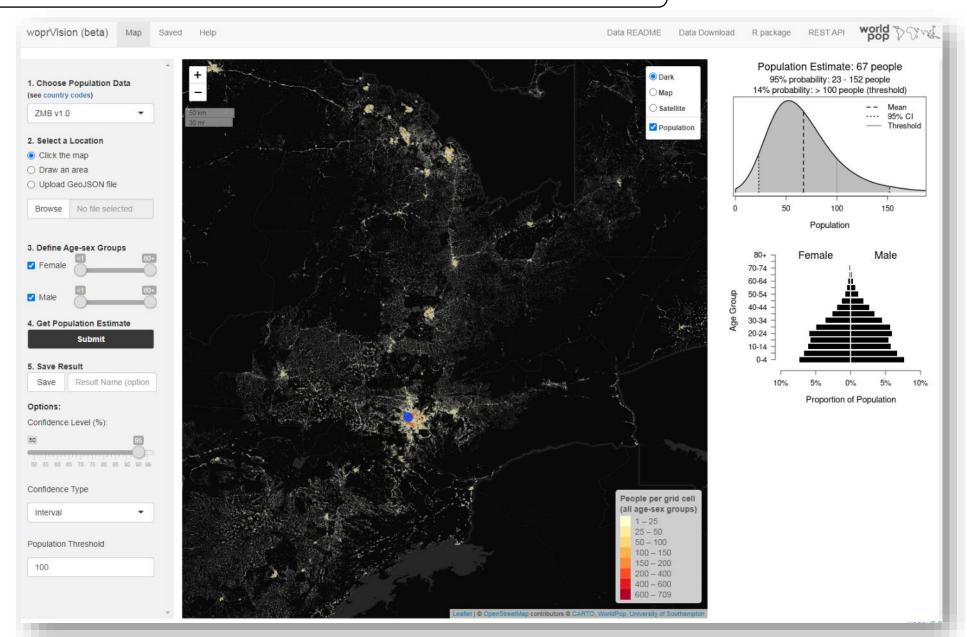
Reset



Select country ...

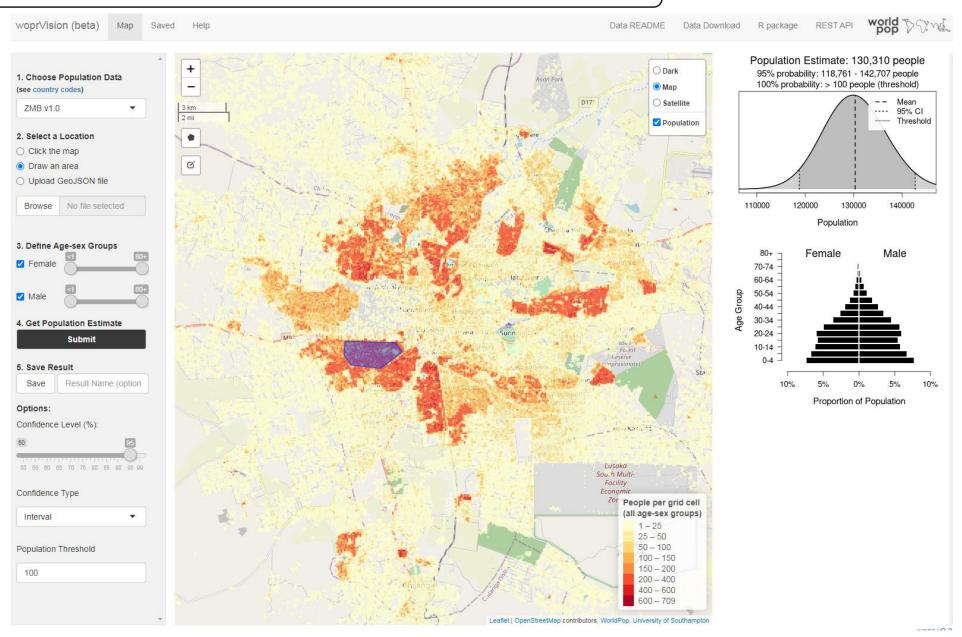
https://wopr.worldpop.org/

		,	•		
ZMB	Population	Bottom-up gridded population estimates (~100m) for specific age-sex groups	2020-04-07	v1.0	i Details
ZMB	Population	Bottom-up gridded population estimates (~100m) for Zambia	2020-04-07	v1.0	i Details
ZMB	Population	Mastergrid for gridded population estimates	2020-04-07	v1.0	i Details
ZMB	Population	README: ZMB Population v1.0	2020-04-07	v1.0	i Details
ZMB	Population	SQL database with Bayesian posterior population predictions	2020-04-07	v1.0	i Details
ZMB	Population	Image tiles for gridded population estimates	2020-04-07	v1.0	i Details
NGA	Population	Bottom-up gridded population estimates (~100m) for specific age-sex groups	2020-03-16	v1.2	i Details
SSD	Population	Top-down gridded population estimates (~100m) for South Sudan	2019-12-06	v1.0	i Details
SSD	Population	README: SSD Population v1.0	2019-12-06	v1.0	i Details
NGA	Population	Bottom-up population totals for administrative units in Nigeria	2019-07-10	v1.2	i Details
NGA	Population	Bottom-up gridded population estimates (~100m) for Nigeria	2019-07-10	v1.2	i Details
NGA	Population	Mastergrid for gridded population estimates	2019-07-10	v1.2	i Details
NGA	Population	README: NGA Population v1.2	2019-07-10	v1.2	i Details
NGA	Population	SQL database with Bayesian posterior population predictions	2019-07-10	v1.2	i Details
NGA	Population	Image tiles for gridded population estimates	2019-07-10	v1.2	i Details
COD	Population	Bottom-up gridded population estimates (100m) for Democratic Republic of the Congo	2019-05-20	v1.0	i Details
COD	Population	Mastergrid for gridded population estimates	2019-05-20	v1.0	i Details
COD	Population	README: COD Population v1.0	2019-05-20	v1.0	i Details
COD	Population	SQL database with Bayesian posterior population predictions	2019-05-20	v1.0	i Details
COD	Population	Image tiles for gridded population estimates	2019-05-20	v1.0	i Details

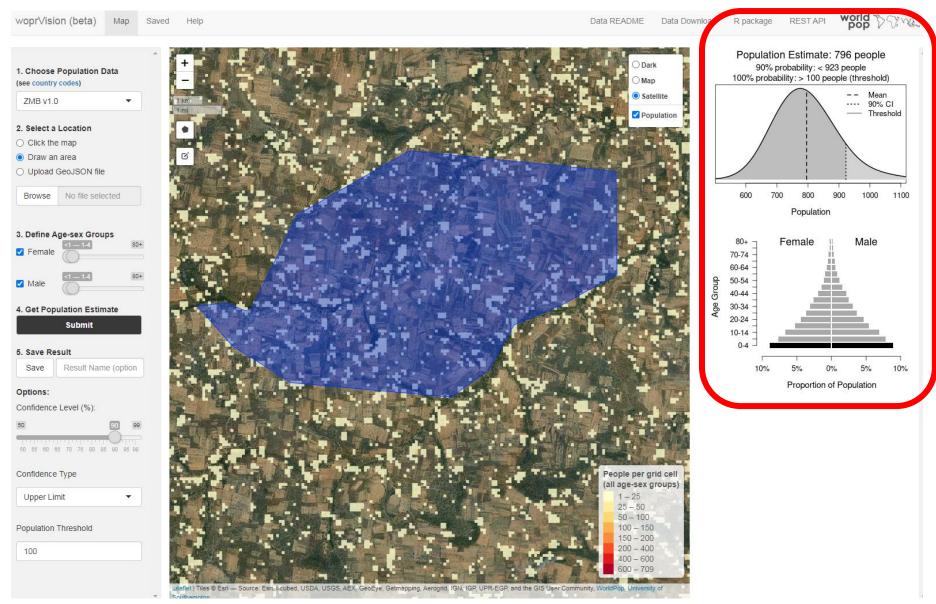


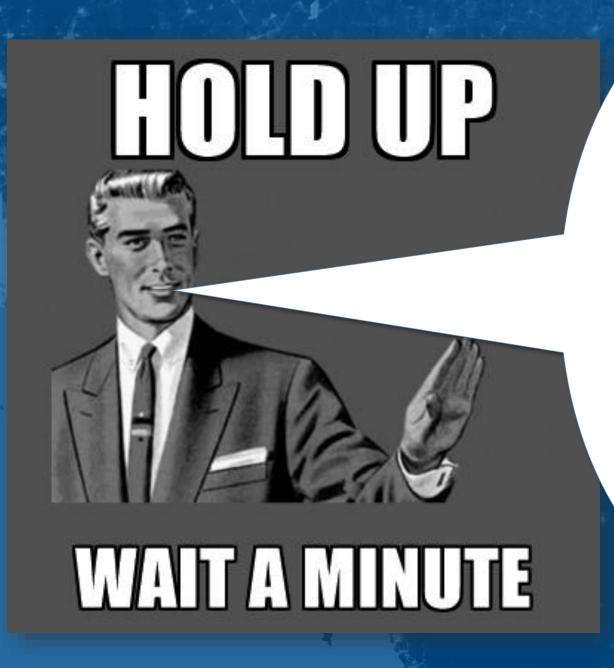
Interacting with modelled estimates: WoprVision

https://apps.worldpop.org/woprVision/

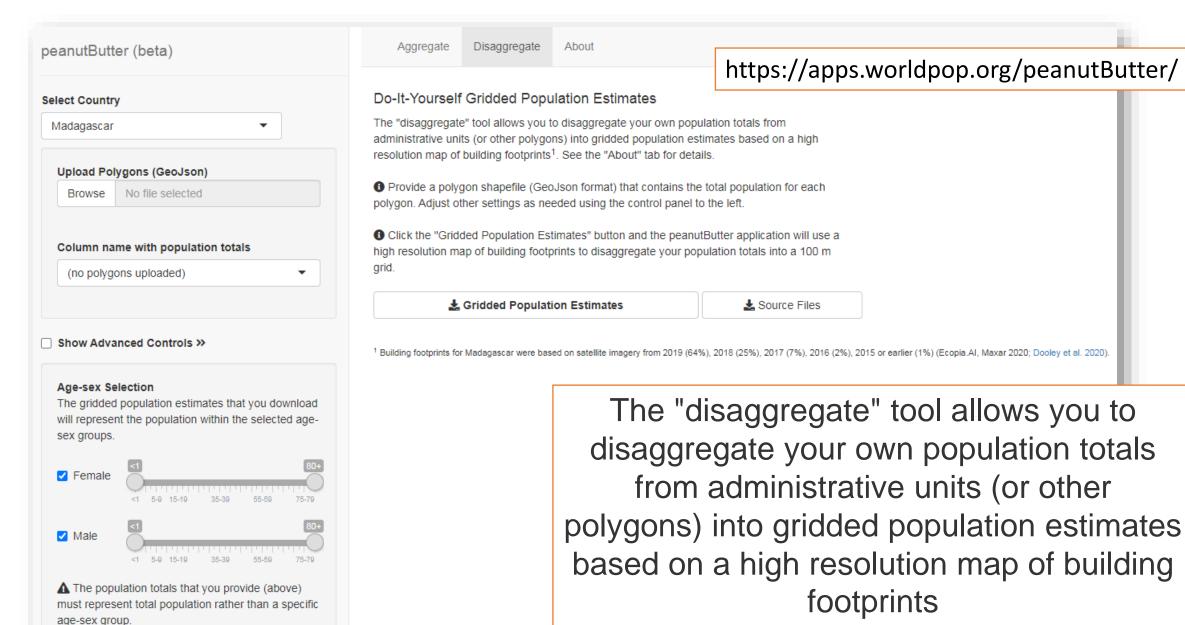


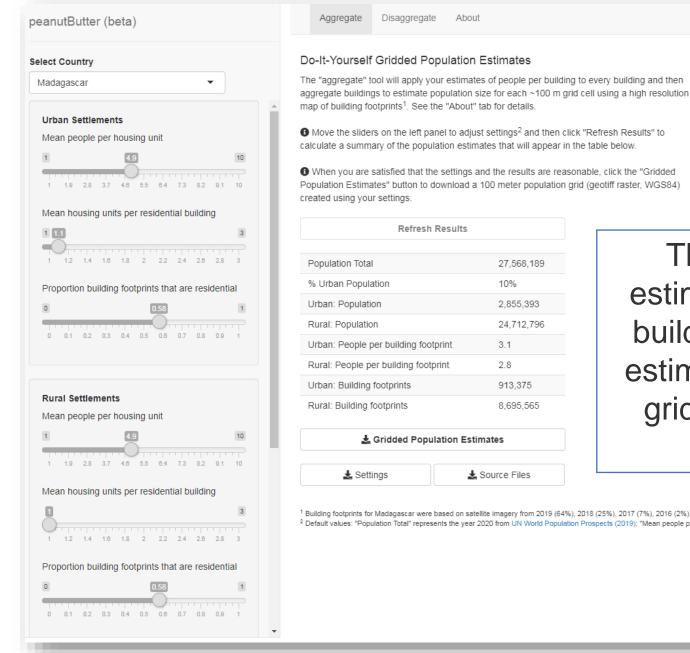
https://apps.worldpop.org/woprVision/





What if I have my own sample enumeration data or projections and want to produce a rapid gridded population map using building footprints?





https://apps.worldpop.org/peanutButter/

The "aggregate" tool will apply your estimates of people per building to every building and then aggregate buildings to estimate population size for each ~100 m grid cell using a high resolution map of building footprints

1 Building footprints for Madagascar were based on satellite imagery from 2019 (64%), 2018 (25%), 2017 (7%), 2016 (2%), 2015 or earlier (1%) (Ecopia.AI, Maxar 2020; Dooley et al. 2020)

27,568,189

2,855,393

24,712,796

8,695,565

3.1

2.8 913,375

Source Files

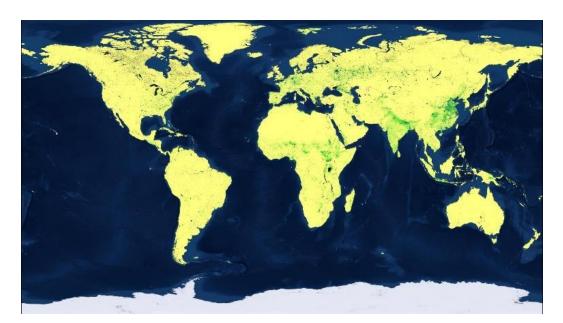
² Default values: "Population Total" represents the year 2020 from UN World Population Prospects (2019); "Mean people per housing unit" is from United Nations (2019) or United Nations (2017)



Gridded population datasets/methods

'Top down' global estimates

- 'Unconstrained' 2000-2020
- 'Constrained' 2020

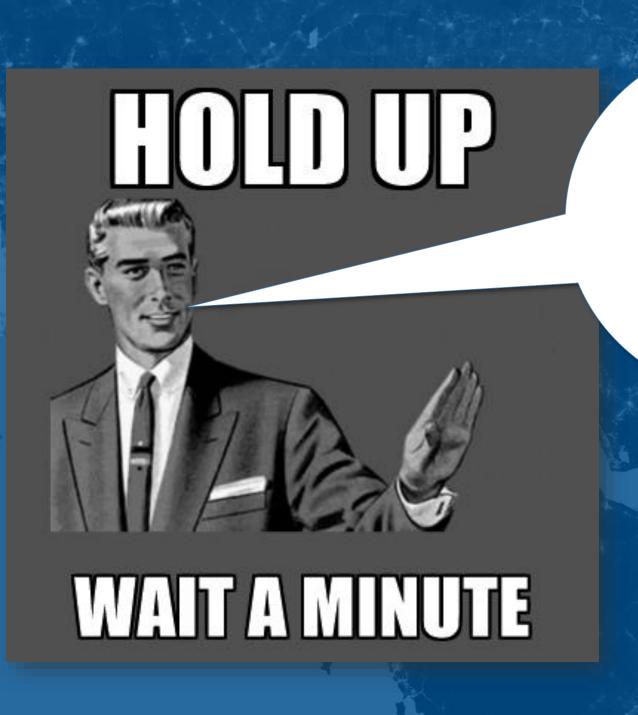




'Bottom up' bespoke country estimates Population counts per pixel







How are these estimates used?

https://www.unfpa.org/resources/value-modelled-population-estimates-census-planning-and-preparation

The Value of Modelled Population Estimates for Census Planning and Preparation





Publication Date: May 2020

Author: UNFPA

Download English

Spanish

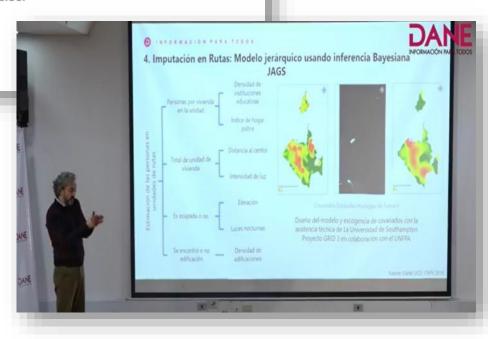
French

Arabic

Russian

This guidance note details how modelled population estimates can support national statistics offices in planning, conducting and performing quality assurance checks during the census exercise.





New EA boundaries

Old EA boundary





UNOSAT Tropical Cyclone ETA-20 Population Exposure Analysis in Honduras 5 November 2020



Honduras population 2020

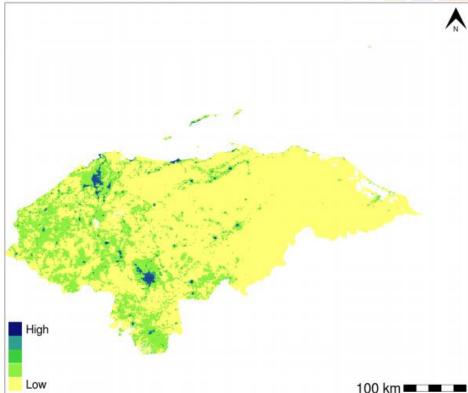
Estimated total number of people per grid-cell at a resolution of (3 arc seconds approximately 100m at the equator)





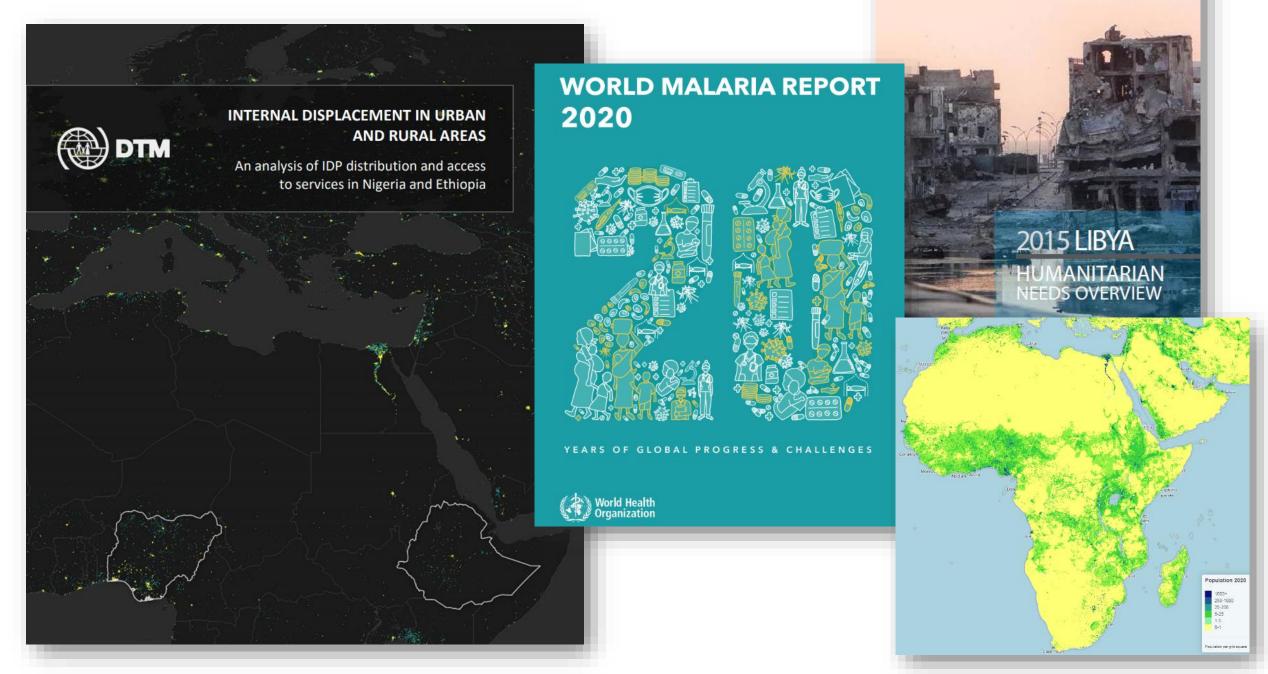




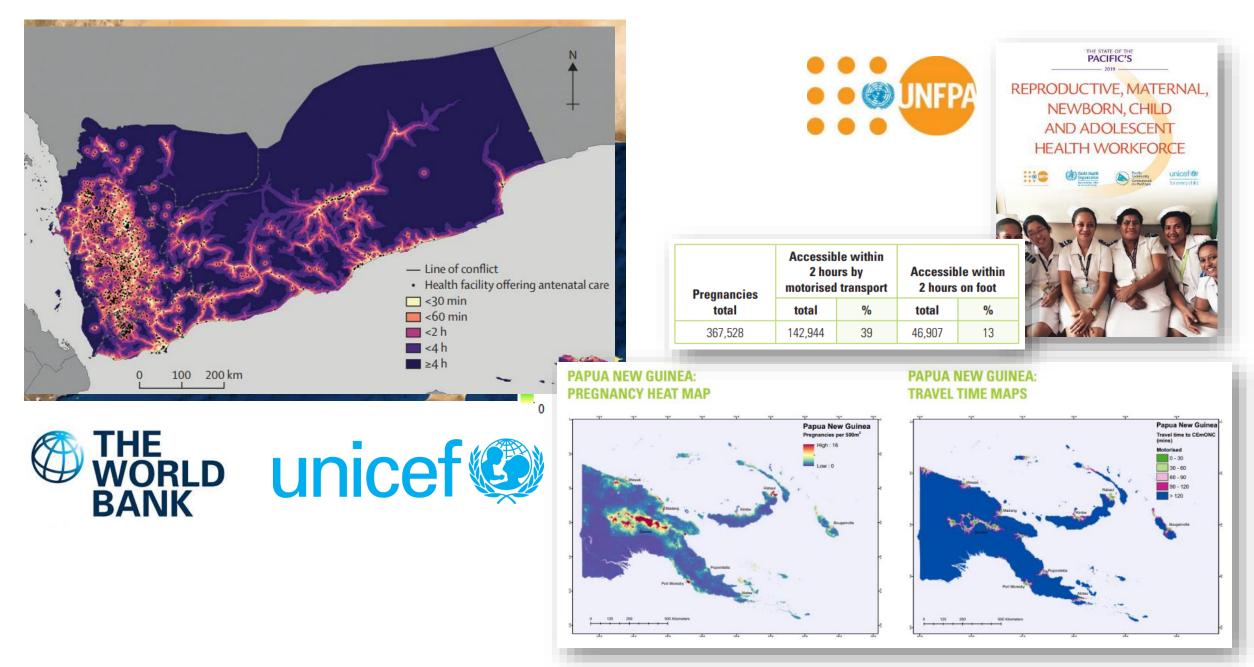


WorldPop (www.worldpop.org - School of Geography and Environmental Science, University of Southampton; Department of Geography and Geosciences, University of Louisville; Département de Géographie, Université de Namur) and Center for International Earth Science Information Network (CIESIN), Columbia University (2018). Global High Resolution Population Denominators Project - Funded by the Bill and Melinda Gates Foundation (OPP1134076), https://dx.doi.org/10.5258/SOTON/WP00645

©2018 This work is licensed under a Creative Commons Attribution 4.0 International License



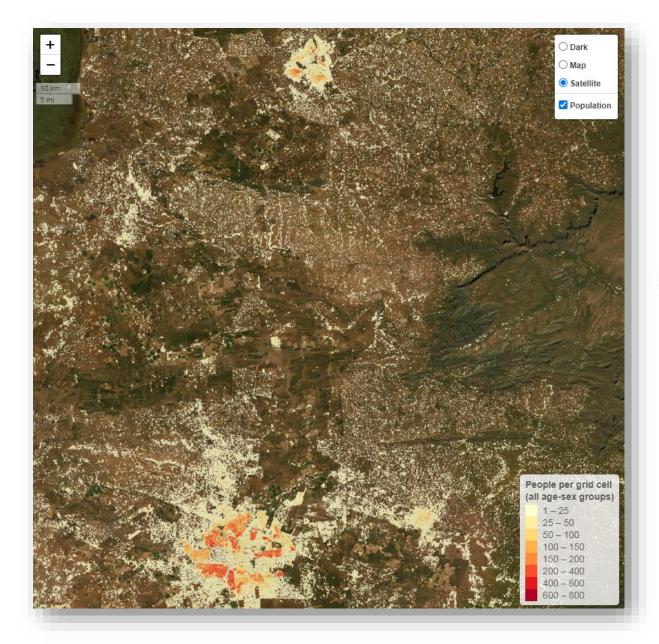
https://displacement.iom.int/system/tdf/reports/urban_displacement_21-05-19.pdf



https://pacific.unfpa.org/en/publications/state-pacifics-rmncah-workforce-2019-report



https://grid3.org/news/outside-the-box-how-nigeria-won-the-fight-against-polio



https://wopr.worldpop.org/?ZMB/Population

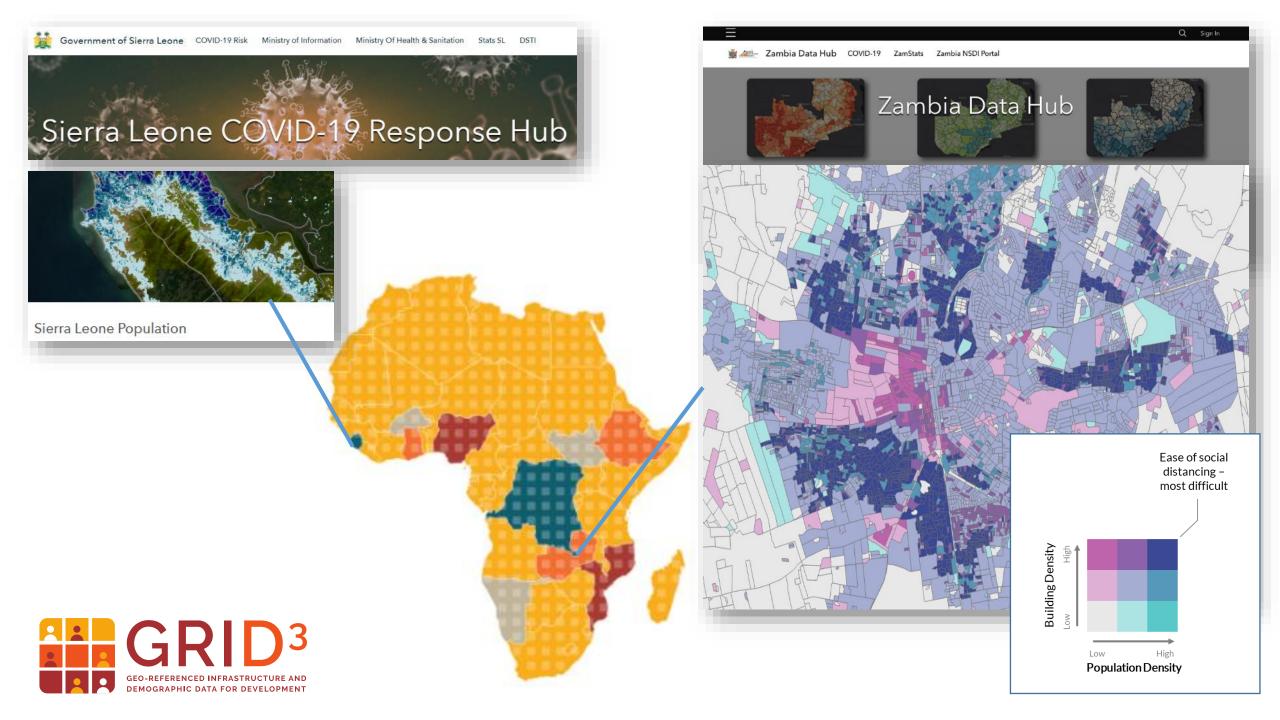




Akros @akros_global · Nov 26

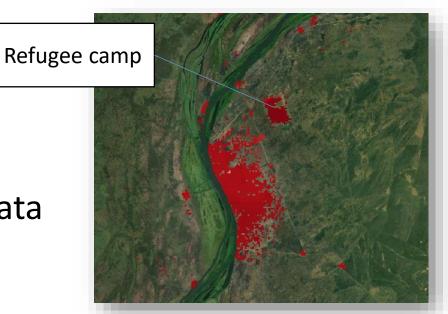
Last month, Akros joined Zambia's NMEP to assist w/ IRS & #ITN resource microplanning for vector control—together making ~120 district maps using enumeration & @grid3global data. We know these maps will prove invaluable to get ahead on vector control before the annual rains come!





Some next steps

- Small area projections: integrating displacement data
- Seasonal population mapping: geospatial data integration
- New covariates: neighbourhood mapping, residential/non-residential, building heights
- Methods development: developing and validating rapid and bespoke models
- Methods documentation and webinar development
- *GRID3*: supporting new countries



South Sudan: Dooley et al (2020) https://wopr.worldpop.org/

Key messages

- Small area population estimates form the basis of decision making across multiple fields
- In resource poor settings, population data can often be coarse, outdated and/or unreliable
- A range of geospatial modelling methods exist to complement traditional data collection methods to produce timely small area estimates
- Models are never perfect = importance of validation, and of measuring, communicating and using uncertainty
- End-user engagement and capacity strengthening vital to adoption and sustained use



Acknowledgements

These slides contain input from members of the WorldPop group: Maksym Bondarenko, Gianluca Boo, Heather Chamberlain, Donna Clarke, Claire Dooley, Chris Jochem, Doug Leasure, Chris Lloyd, Jeremiah Nieves, Sarchil Qader, Alessandro Sorichetta

Thanks to them and both the whole WorldPop group and GRID3 team for wider input and support



Further information



www.worldpop.org





www.grid3.org



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



Questions & Answers

Thank you!

centre.humdata.org

centre for humdata

