



NASA LCLUC Meeting, Hanoi
January 31, 2024

OPEN-SOURCE POPULATION AND DEMOGRAPHIC DATA:

Considerations for Advanced Integration with Remotely Sensed Data

Forrest R. Stevens, Andrea E. Gaughan

Department of Geographic and Environmental Sciences



Son Nghiem
NASA Jet Propulsion Laboratory

Deborah Balk

Marxe School of Public and International Affairs

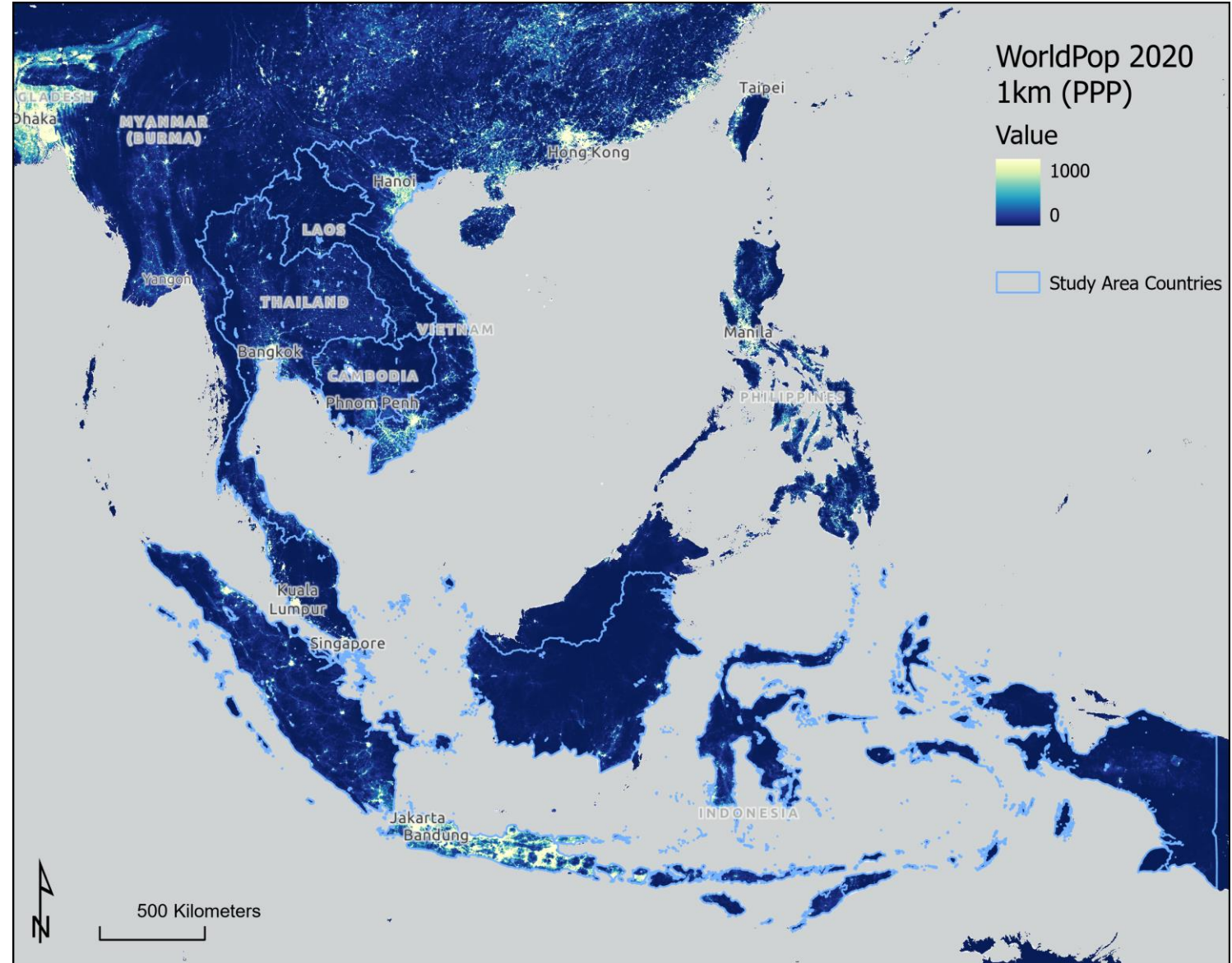


NASA LCLUC SARI Synthesis

SE Asia has seen extensive change in environmental and socio-demographic patterns from 2000-2020.

These complex dynamics transition differently along the rural-urban continuum.

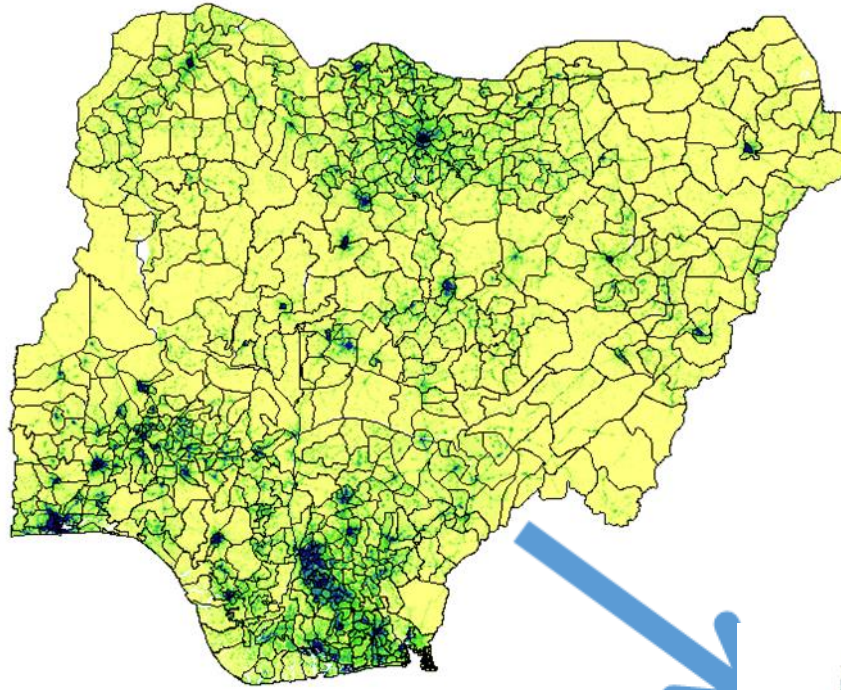
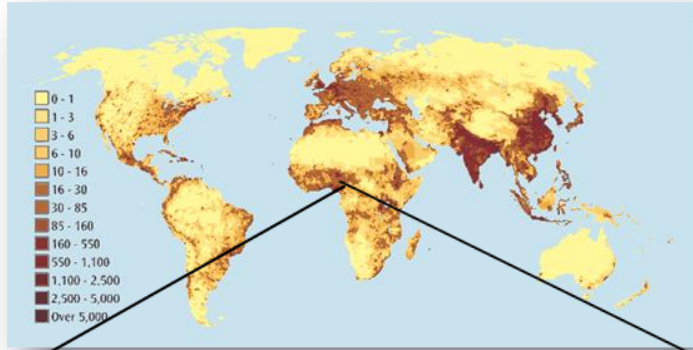
Essential to LCLUC synthesis research is the inclusion of the human element.



Introduction to gridded population data

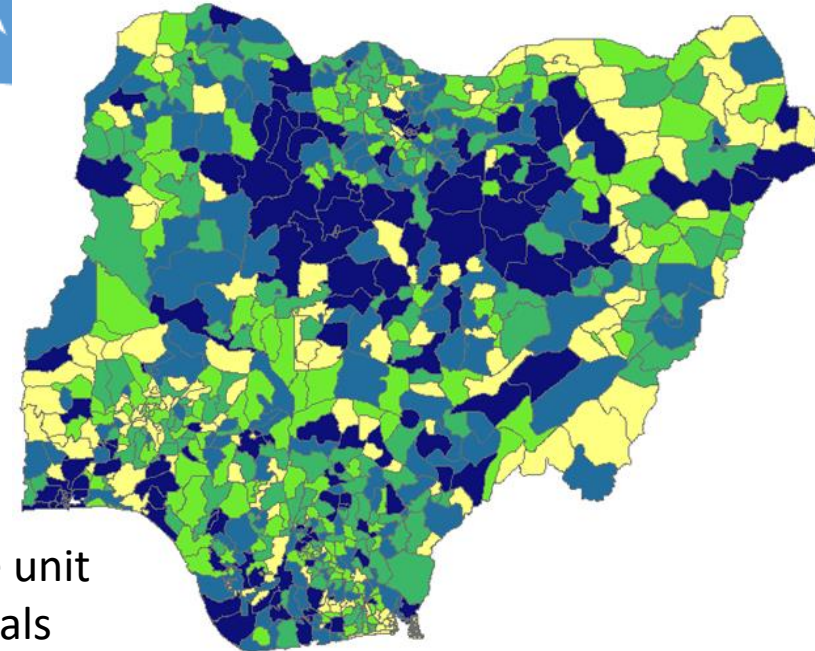
- Population data are usually collected from censuses of population and housing undertaken by national statistical offices. Additional data sometimes come from surveys.
- These data are almost always collected for administrative units that are useful for statistical units and are in vector-format. The boundaries of such units are intrinsically
 - irregular
 - change over time
- Transforming such inputs to a grid standardizes the spatial units and thus increases their usability in a wide variety of environmental usages
 - A limitation is that most global grids include information on population counts (Worldpop, by age and sex) but not on other socio-demographic or housing information

Gridded Data



Grids: flexibility in summarization to any administrative unit level

100 x 100m gridded estimates of population density



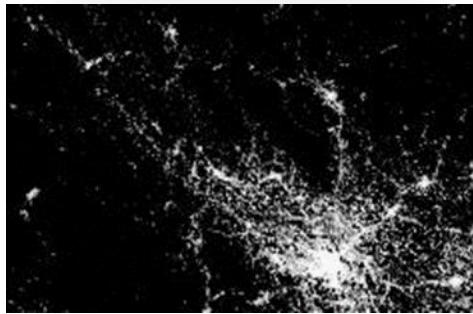
Administrative unit population totals

Grids: consistent and comparable format and a framework for integrating differing data types

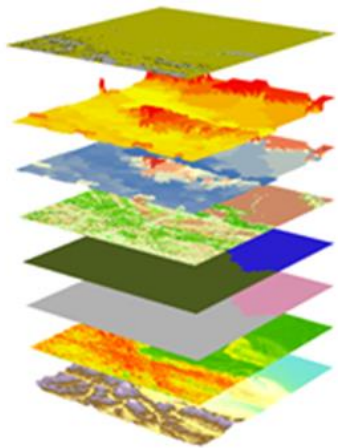
Spatial demographic modelling



Population Data

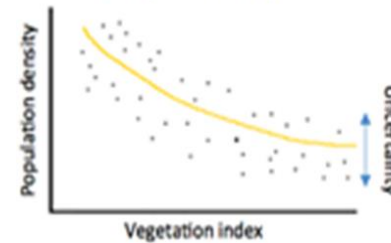
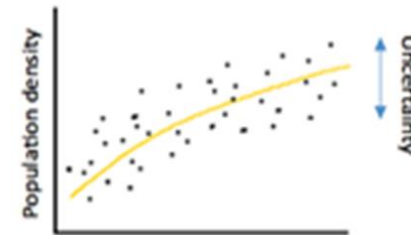


Settlement

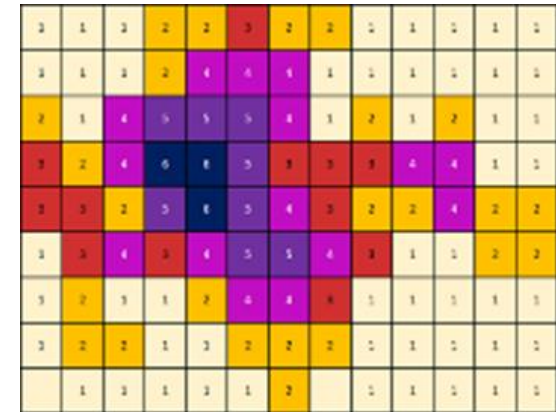



Geospatial covariates

Spatial statistical model



Population estimates
(ideally with uncertainty)

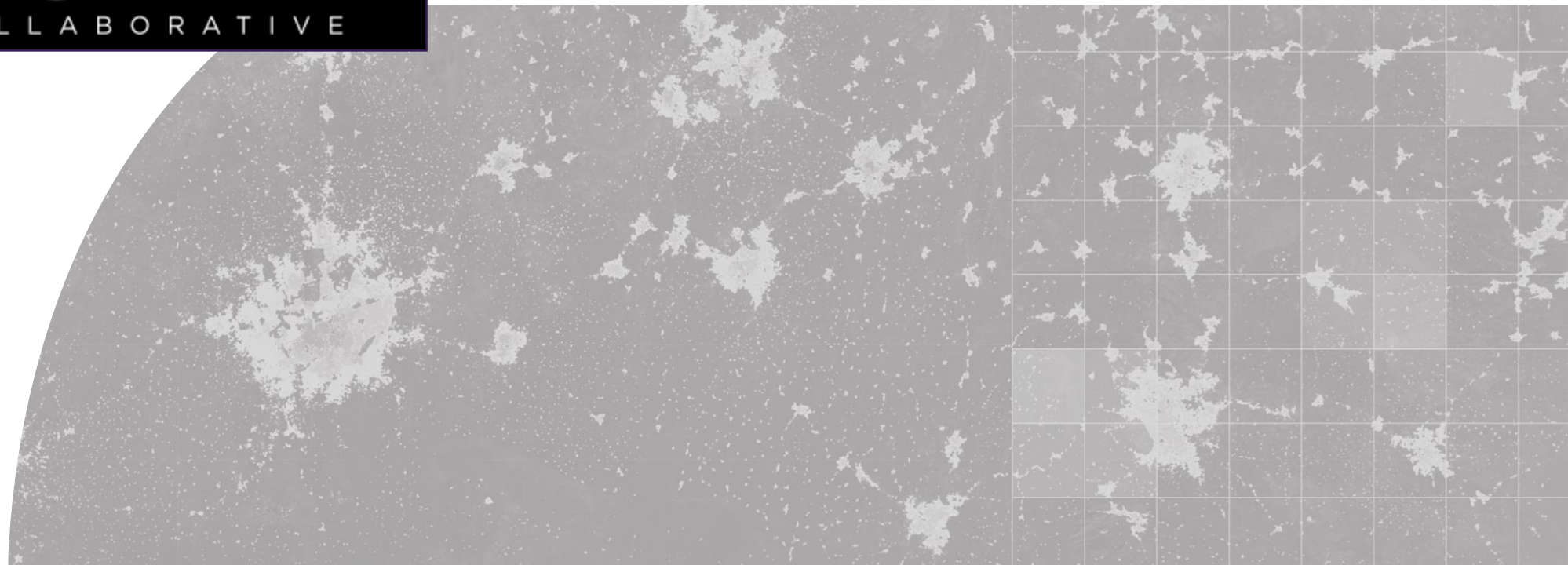




A Data Collaborative for
Settlement, Infrastructure, and Population Data



POPGRID
DATA COLLABORATIVE





Gridded Products

WorldPop

 COLUMBIA CLIMATE SCHOOL
CENTER FOR INTERNATIONAL EARTH SCIENCE
INFORMATION NETWORK

 European
Commission
Joint Research Centre

 ImageCat

 **esri**
THE SCIENCE OF WHERE™

 OAK RIDGE
National Laboratory

 CUNY INSTITUTE FOR
DEMOGRAPHIC RESEARCH

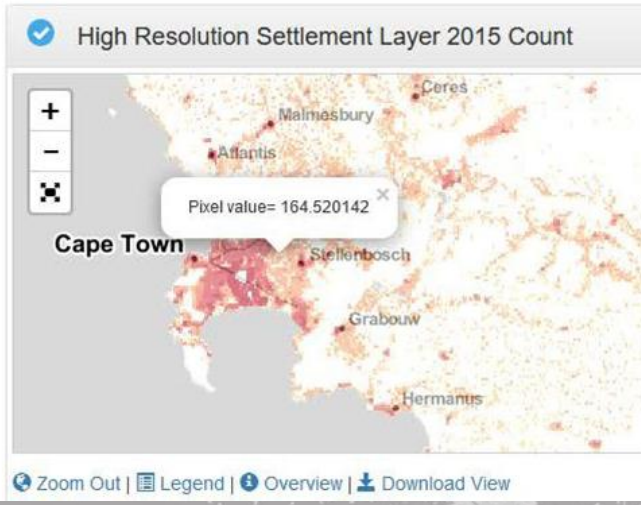
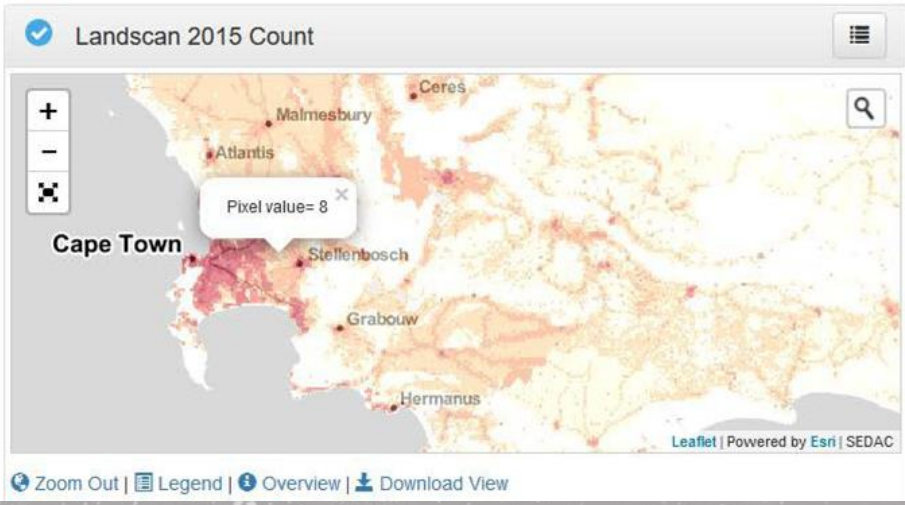
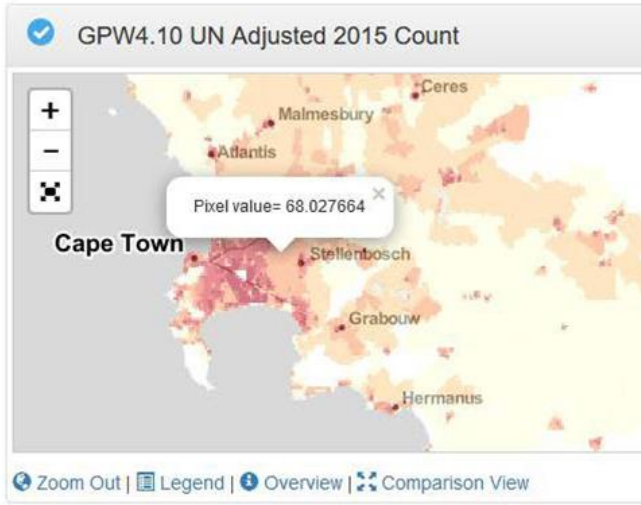
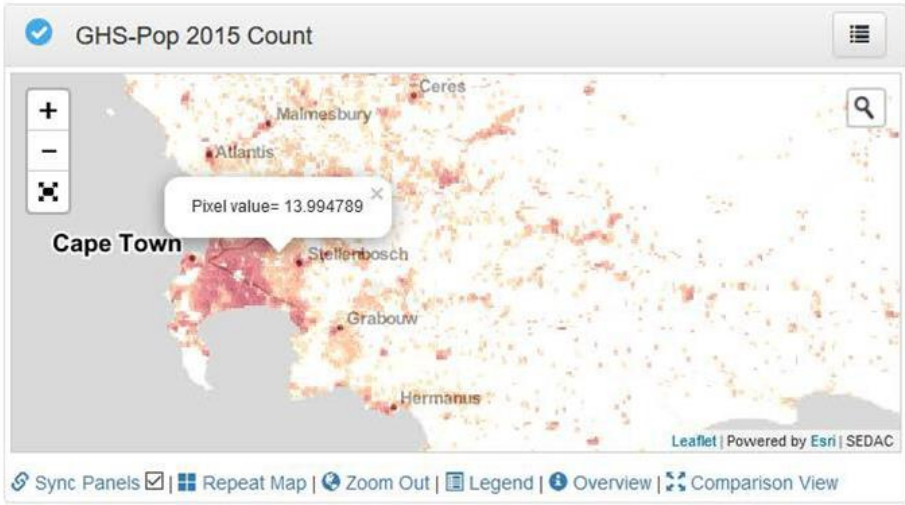
United States
Census
Bureau

 internet.org
by facebook

 Deutsches Zentrum
DLR für Luft- und Raumfahrt

Data Comparison Tools

NASA SEDAC POPGRID Viewer



Results

Population Estimates by Data Set

POPGRID

Source	Population Estimates
GHSL	1,570,029
GPWv4	1,479,828
Landscan	507,188
Worldpop	1,472,674
ESRI WPE	1,005,505
HRSL	1,560,679

Population Estimates by Data Set

Source	Population Estimates
GHSL	1570029
GPWv4	1479828
Landscan	507188
Worldpop	1472674
ESRI WPE	1005505
HRSL	1560679

Data Quality Message(s)

- The average national WPE reliability ranking is 2
- The average size of national input units in GPW is 1588 square kilometers
- HRSL has coverage

Leaflet | Powered by Esri | SEDAC

Fitness for Use

Concept of “**relative data quality**” (Tayi & Ballou 1998)

Assess the **appropriateness** of a given dataset for an **intended purpose**

Guide user community in making **informed decisions** by better understanding:

Spatial, thematic and temporal **accuracy** in relation to the intended use, driven by...

- (1) **Input population data** properties
- (2) **Modeling assumptions** behind products
- (3) **Ancillary** data

There are 4 major global gridded data sets (and other national or regional grids). Figuring out which is best for your intended use can take some time:

See this paper for guidance:

Leyk et al. 2019

["The Spatial Allocation of population: a review of large-scale gridded population data products and their fitness for use"](https://doi.org/10.5194/essd-11-1385-2019)

<https://doi.org/10.5194/essd-11-1385-2019>



Gridded Products

WorldPop

 COLUMBIA CLIMATE SCHOOL
CENTER FOR INTERNATIONAL EARTH SCIENCE
INFORMATION NETWORK

 European
Commission
Joint Research Centre

 ImageCat

 **esri**
THE SCIENCE OF WHERE™

 OAK RIDGE
National Laboratory

 CUNY INSTITUTE FOR
DEMOGRAPHIC RESEARCH

United States
Census
Bureau

 internet.org
by facebook

 Deutsches Zentrum
DLR für Luft- und Raumfahrt

A world map showing population density. Landmasses are colored in shades of yellow and orange, indicating higher population density, while oceans are dark blue. The map is centered on the Atlantic Ocean, showing North and South America on the left, Europe and Africa in the center, and Asia and Australia on the right.

WorldPop 'Global 1' Project

- Bill and Melinda Gates Foundation funding 2016-18
- Top-down disaggregation of GPWv4 2000-2020 database using Random Forests and global 100m geospatial covariate library
 - Assembly of subnational age/sex structure data and development of interpolation/projection methods

Map 2.2 Urban population by country, 2000 and 2010

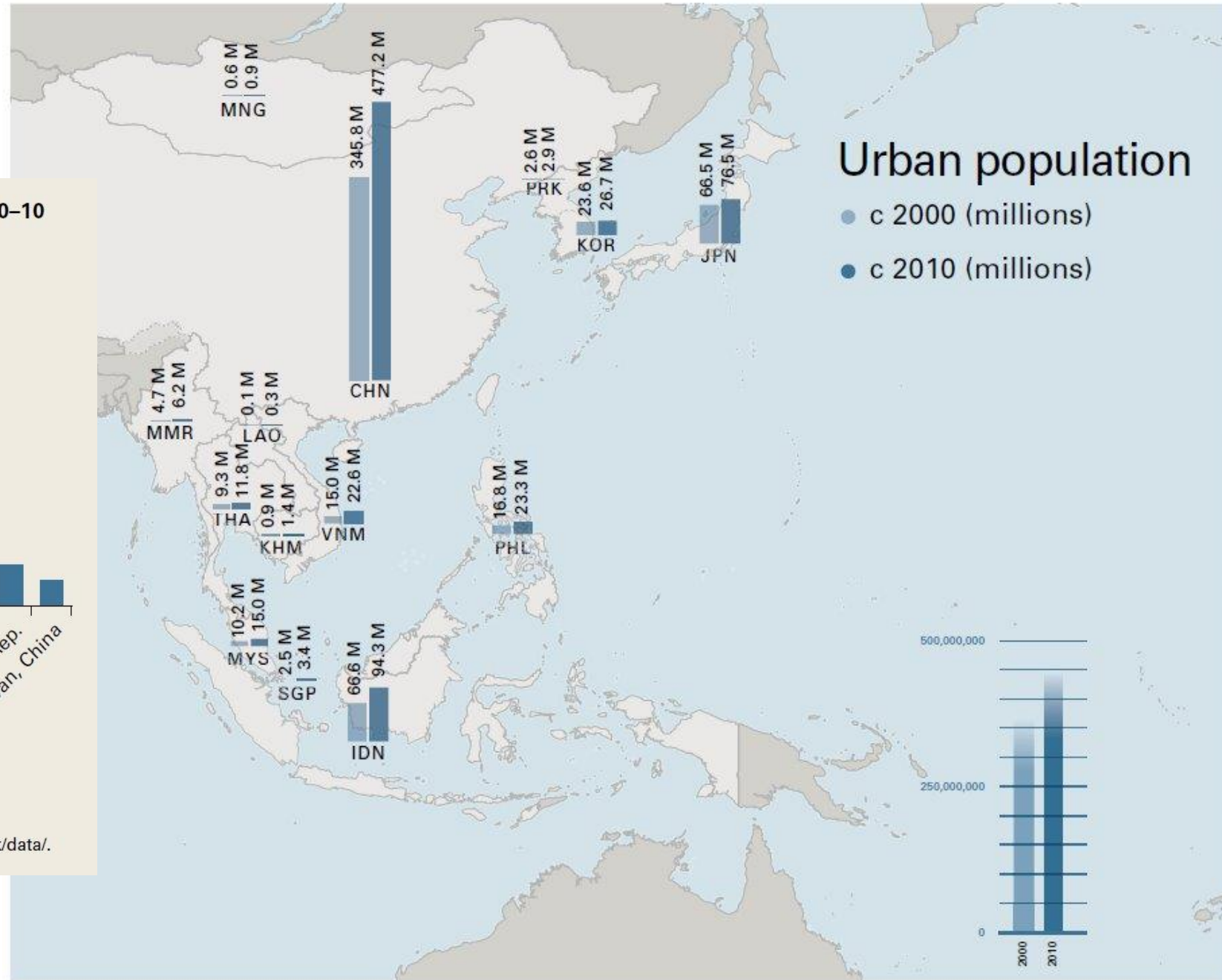
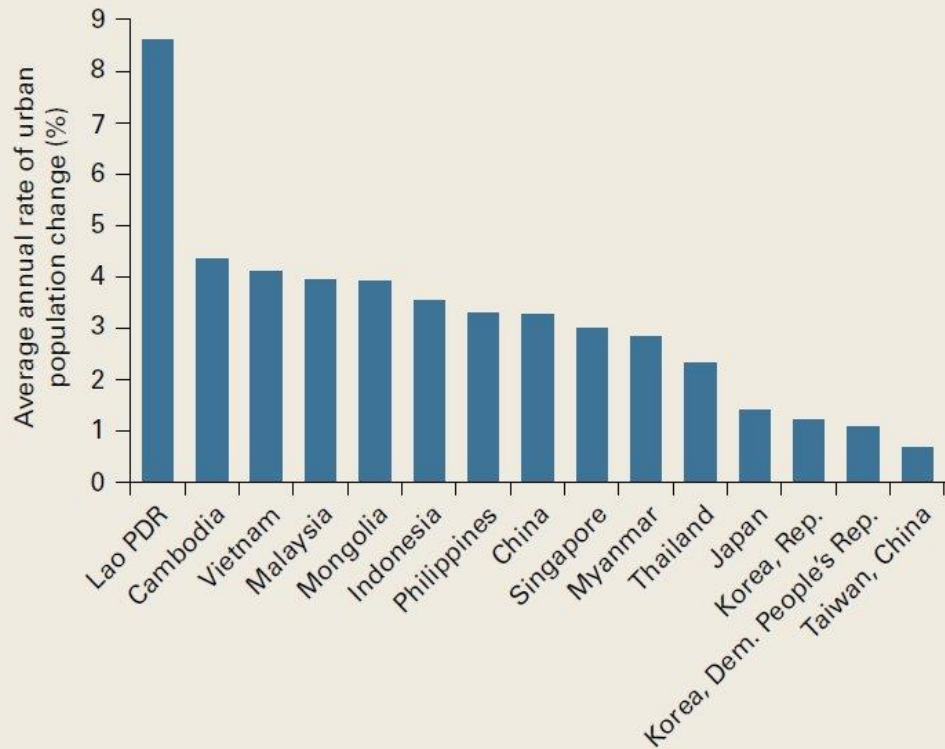


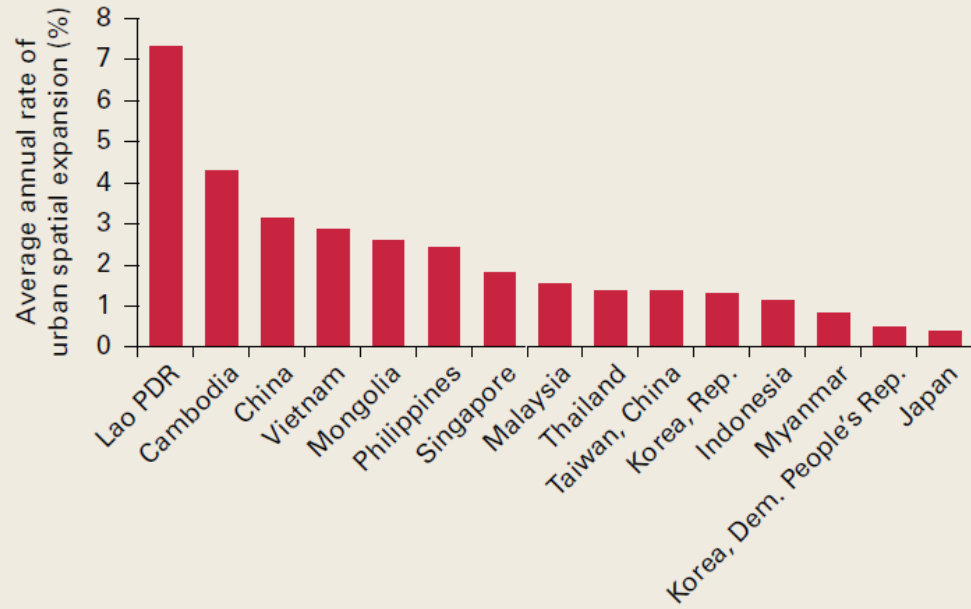
Figure 2.4 Rate of urban population growth by country, 2000–10



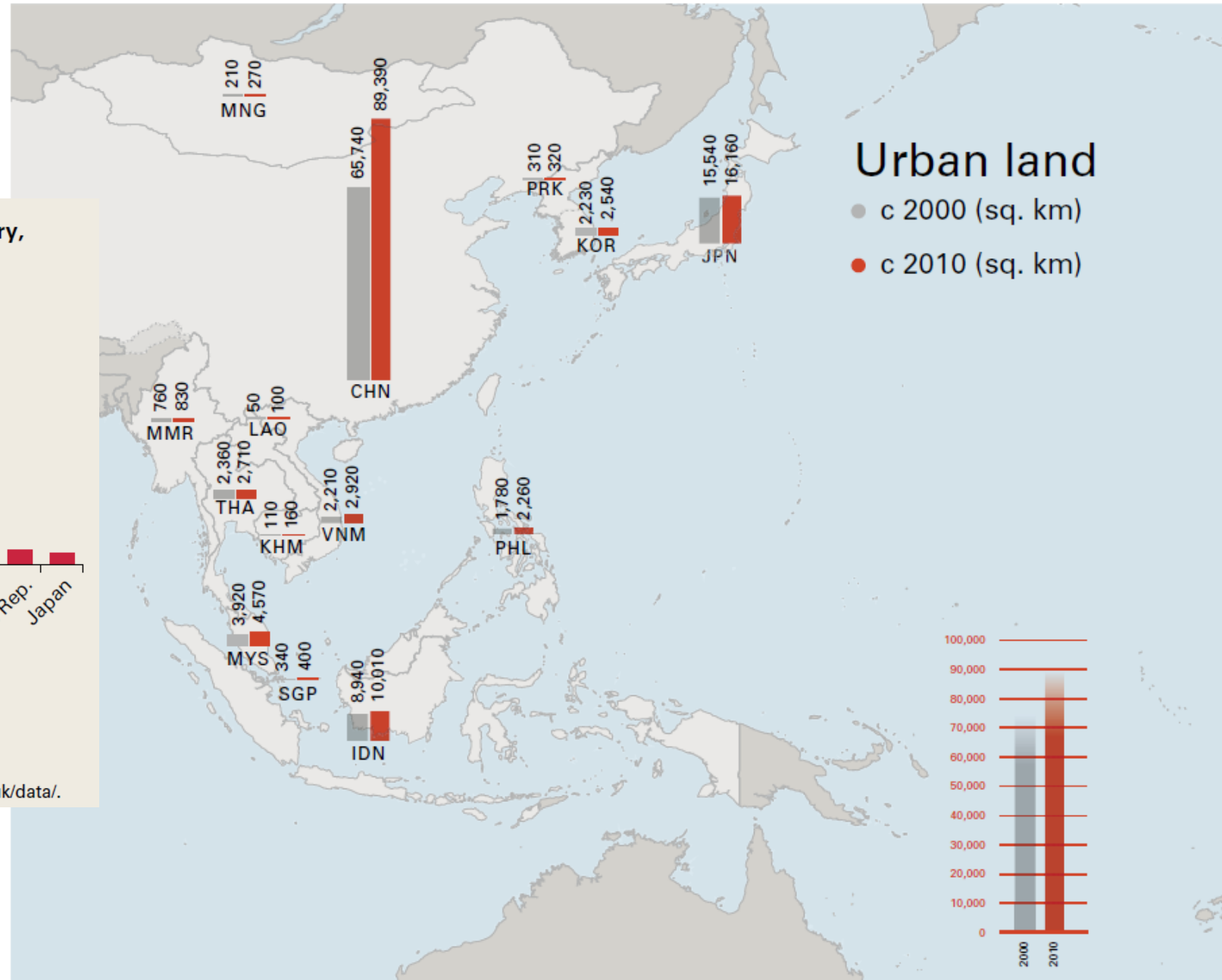
Source: Study team, incorporating WorldPop data, <http://www.worldpop.org.uk/data/>.

Map 2.1 Urban land by country, 2000 and 2010

Figure 2.2 Annual rate of urban spatial expansion by country, 2000–10

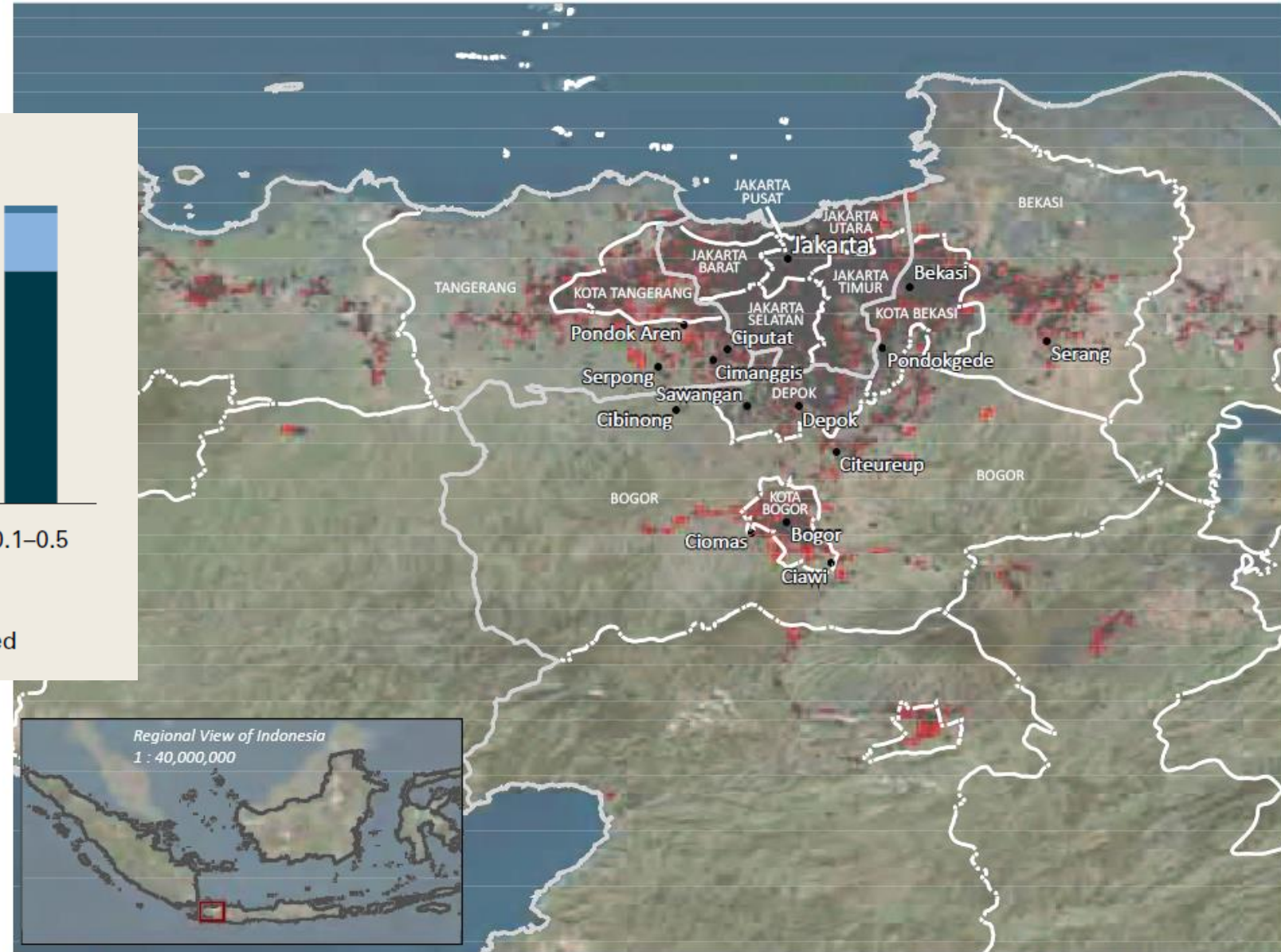


Source: Study team, incorporating WorldPop data, <http://www.worldpop.org.uk/data/>.



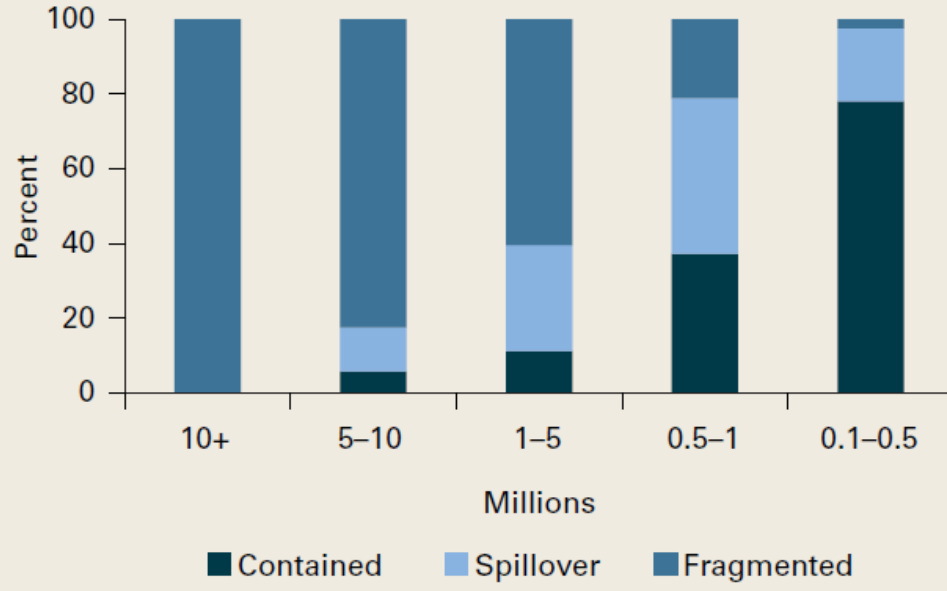
Source: Study team, incorporating WorldPop data, <http://www.worldpop.org.uk/data/>.

Map 2.6 The Jakarta, Indonesia, urban area covers 1,600 square kilometers and 12 jurisdictions



Maps produced by University of Wisconsin-Madison, August 2013
 1:750,000
 Albers equal-area conic projection
 Administrative boundaries from GADM, levels 1 and 2

Figure 2.18 Fragmentation by size category



Fragmentation increases steadily with increasing population size

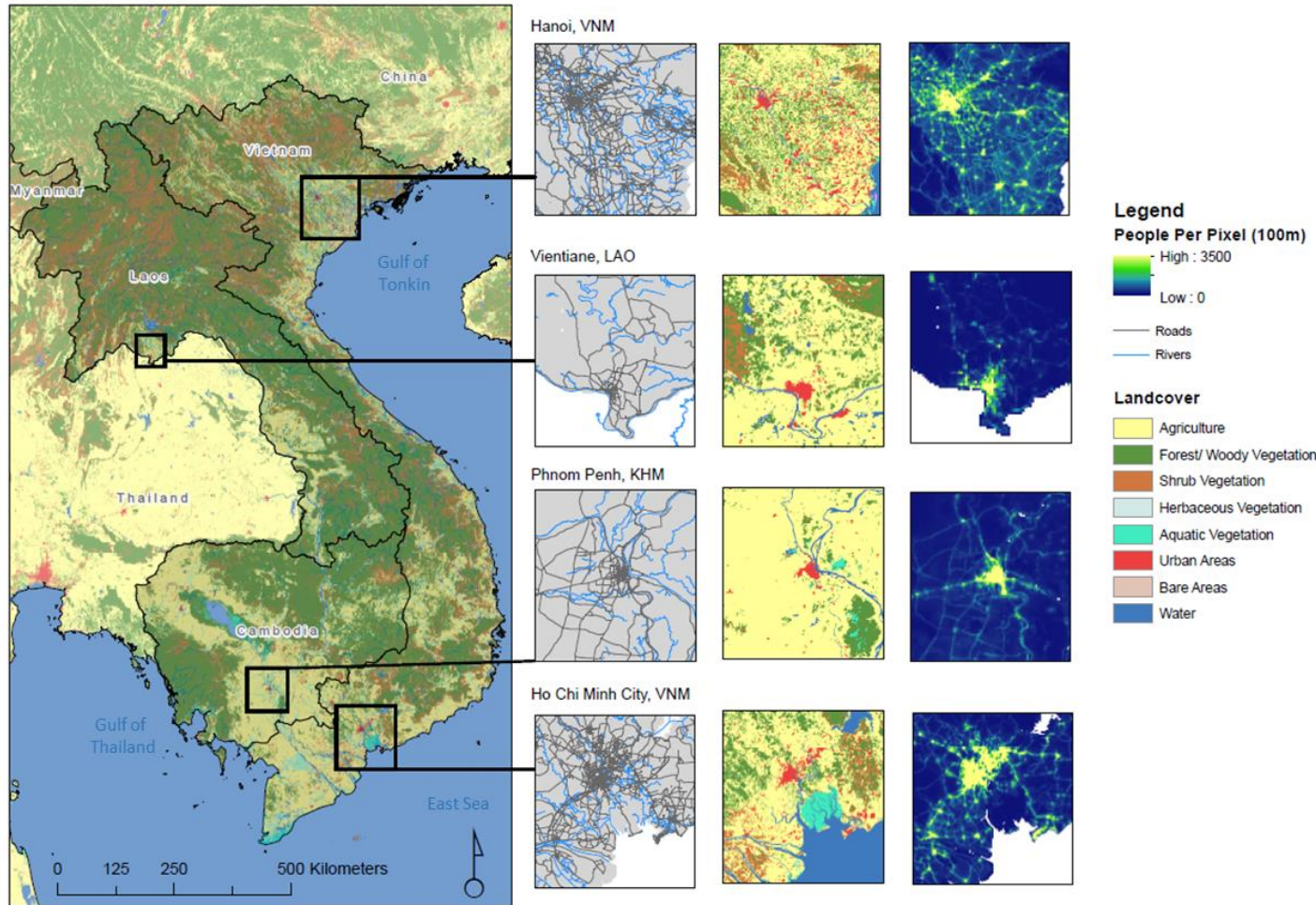
Majority of urban areas with <500,000 people historically within their boundaries but how changing?



PREVIOUS WORK AND SYNTHESIS AREAS

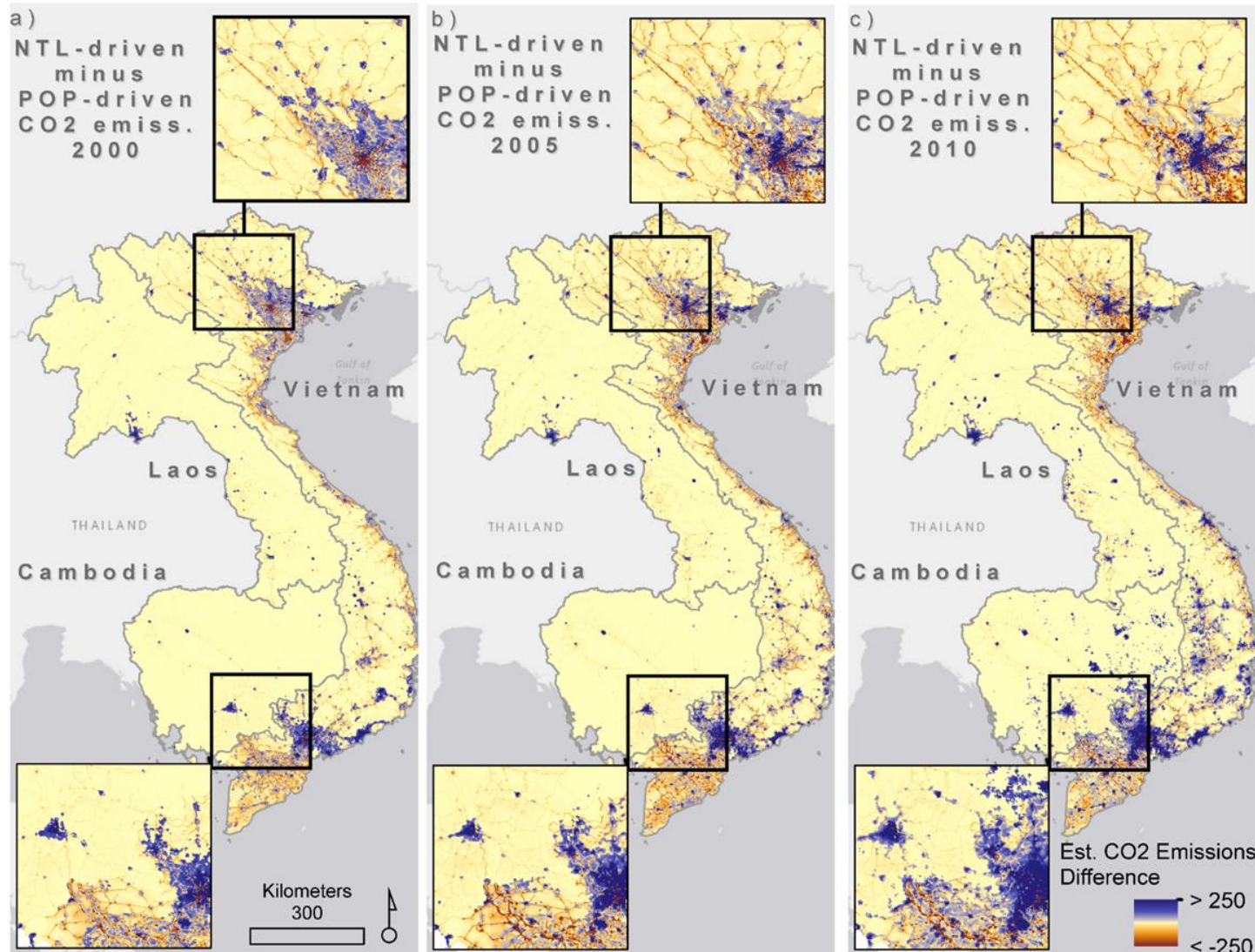
Evaluation of Gridded CO₂ Emissions from Night-Time Lights Compared with Geospatially-Derived Population Distributions for Vietnam, Cambodia, and Laos

Vietnam, Laos, and Cambodia 2010



Compare two gridded data products, modes of disaggregating regional/country level non-point source CO₂ emissions

Evaluation of Gridded CO₂ Emissions from Night-Time Lights Compared with Geospatially-Derived Population Distributions for Vietnam, Cambodia, and Laos



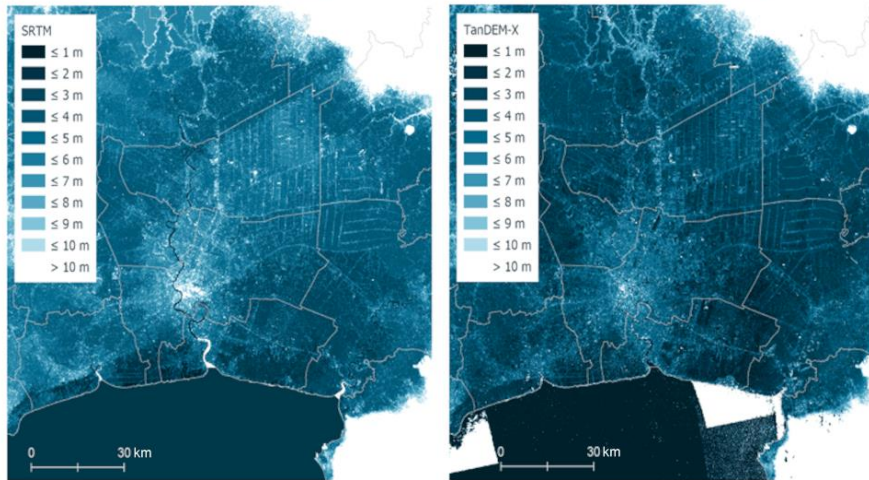
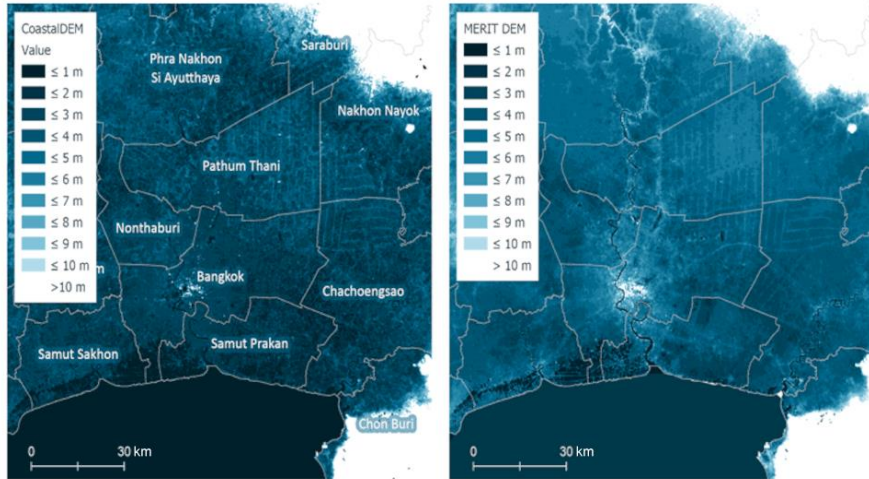
Per pixel differences in CO₂ emissions estimates produced using only nighttime light intensity, minus those produced using population estimates (per capita emissions).

Units are expressed in tonne carbon/year/grid cell

Results are separated by the years 2000 (a), 2005 (b), and 2010 (c)

Estimating population and urban areas at risk of coastal hazards, 1990–2015: how data choices matter

Kytt MacManus ✉, Deborah Balk, Hasim Engin, Gordon McGranahan, and Rya Inman



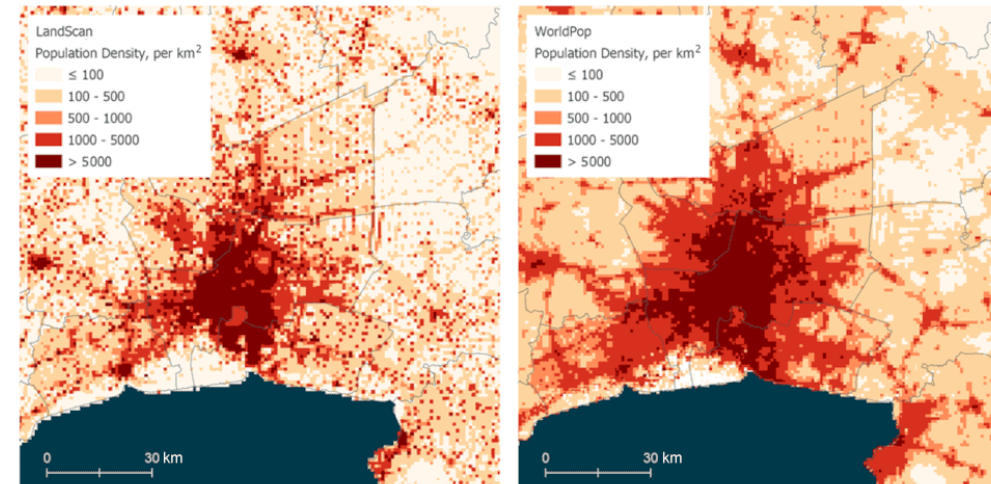
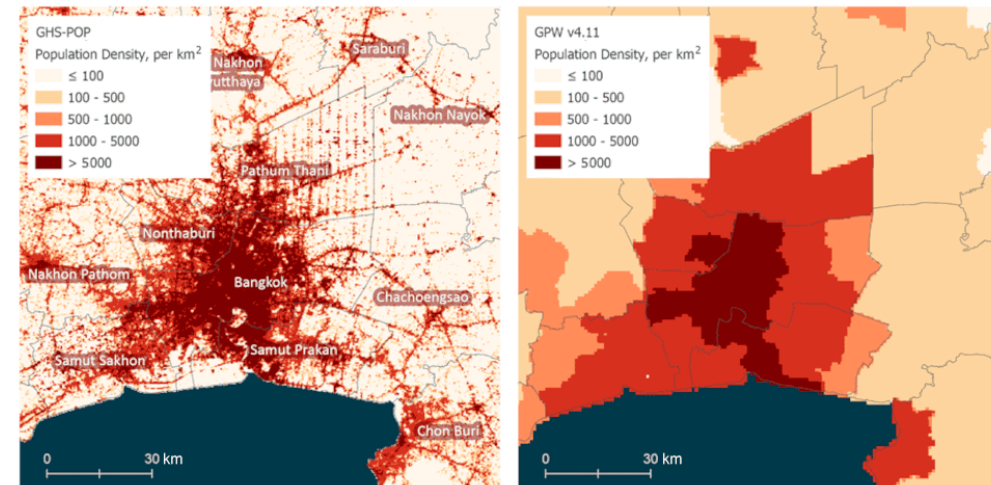
Bangkok

<< 4 Elevation data sets --> LECZ

4 gridded pop >> data sets

4 urban data sets (not shown)

Yield different estimates of population exposures to coastal flooding



Prior work of Balk & Nghiem

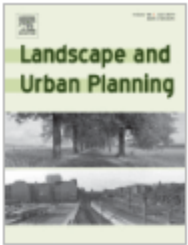
Leverages different remote-sensing data for the built-environment with census data to examine urban change

Some methods applicable to current study: see slide 29 for similar census or survey data that we can use









Landscape and Urban Planning

Volume 187, July 2019, Pages 199-209



Research Paper

Up and out: A multifaceted approach to characterizing urbanization in Greater Saigon, 2000–2009

[Deborah L. Balk](#)^a  , [Son V. Nghiem](#)^b , [Bryan R. Jones](#)^c , [Zhen Liu](#)^{d e} ,
[Gillian Dunn](#)^{f e} 

Agricultural Transitions, Dragon Fruit Cultivation, and Electrification in Southern Vietnam

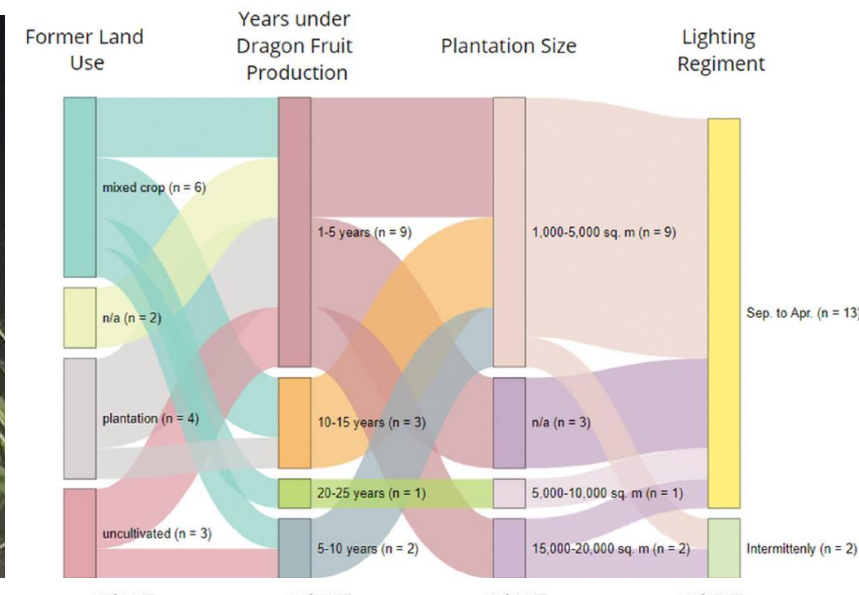
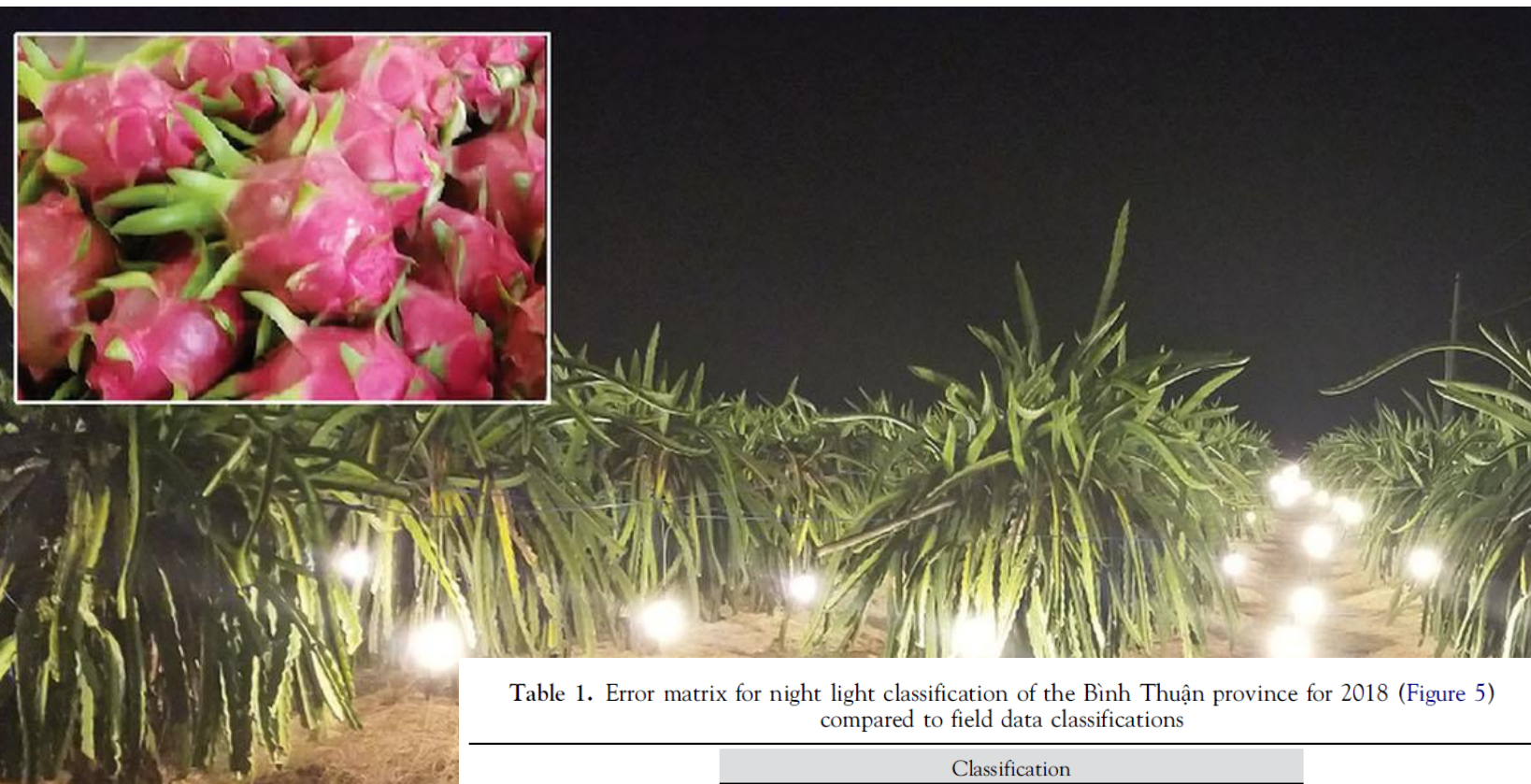
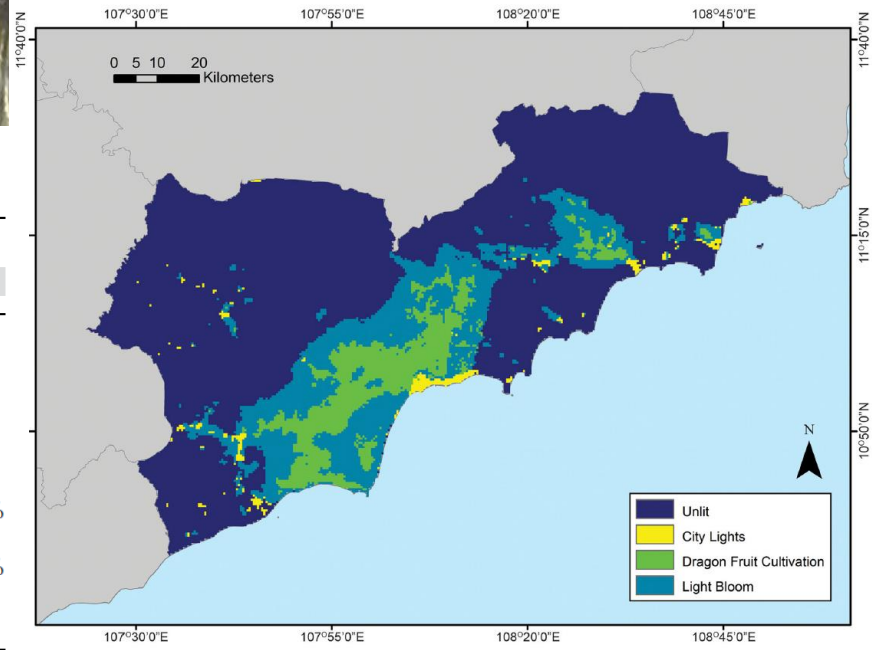


Table 1. Error matrix for night light classification of the Bình Thuận province for 2018 (Figure 5) compared to field data classifications

Reference	Classification				Total
	Unlit	City lights	Dragon fruit	Light bloom	
Unlit	16	0	1	3	20
City lights	1	14	0	2	17
Dragon fruit	1	3	48	13	65
Light bloom	0	1	17	15	33
Total	18	18	66	33	135
User's accuracy	80.0%	82.4%	73.9%	45.5%	
Error of commission	20.0%	17.7%	26.2%	54.6%	Average = 28.8%
Producer's accuracy	88.9%	77.8%	72.7%	45.5%	
Error of omission	11.1%	22.2%	27.3%	54.5%	Average = 29.6%
Overall statistic	68.9%				
Kappa statistic	53.5%				



THE GREATER MEKONG REGION: HUMAN POPULATION AND HEAT EXPOSURE



The Greater Mekong Region: Human Population and Heat Exposure

-

The global rate of urban transition has been immense in the past half century.

-

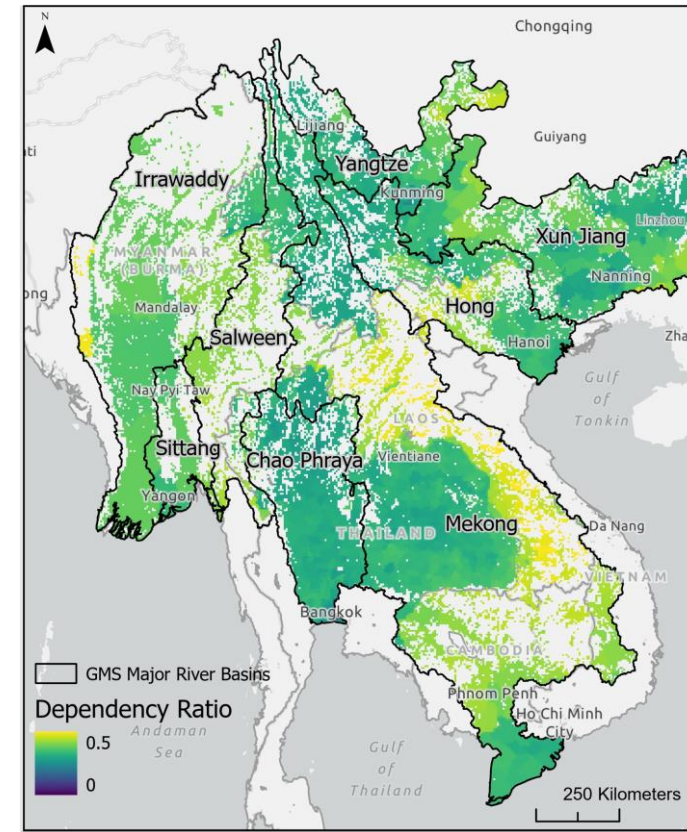
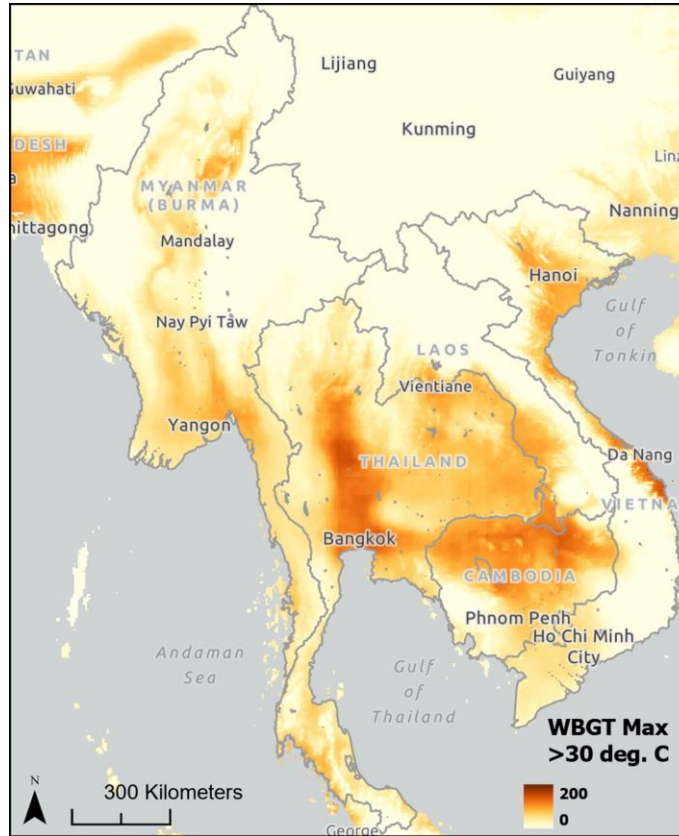
Much of that transition and associated population growth occurs across coastal and riverine parts of Asia.

-

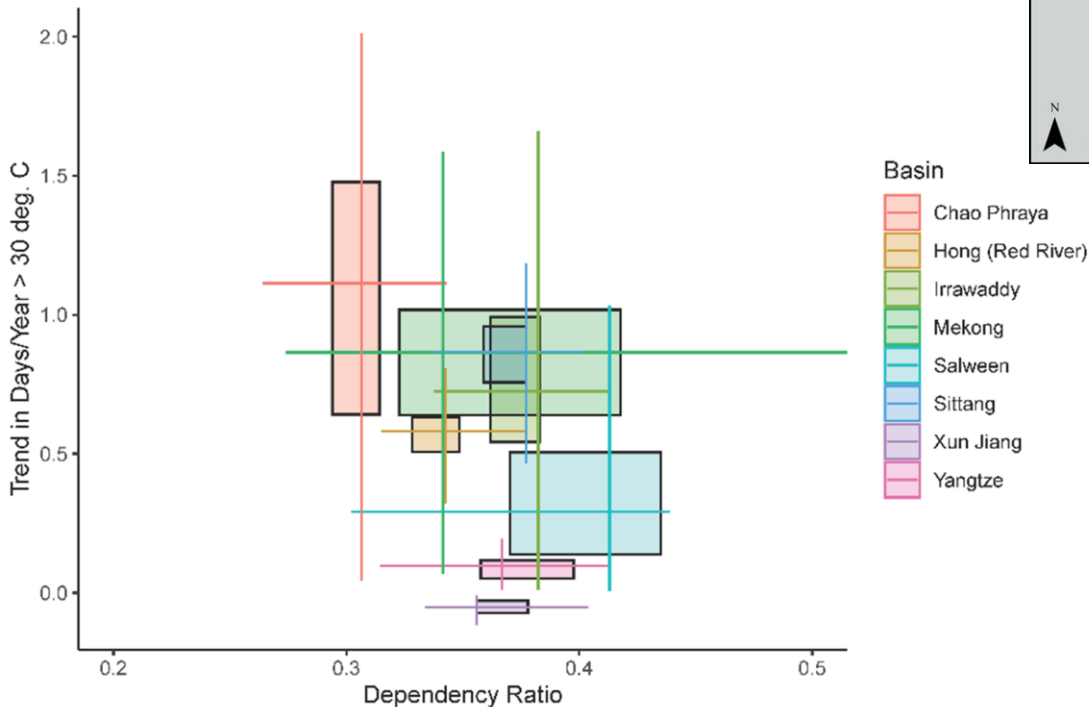
At the same time, changes in climate patterns have disproportionate impacts on populations at risk, the old, and young. Notably, increasing temperatures impact different populations in more or less extreme ways based on livelihoods, age, environment.

The Greater Mekong Region: Human Population and Heat Exposure

An application of leveraging Google Earth Engine and gridded data of demographic patterns to illustrate heat exposure for the Lower Mekong Region.



Bivariate Summary of Dependency Ratio and Trend in Climatic Extremes for the Mekong and Adjacent Basins



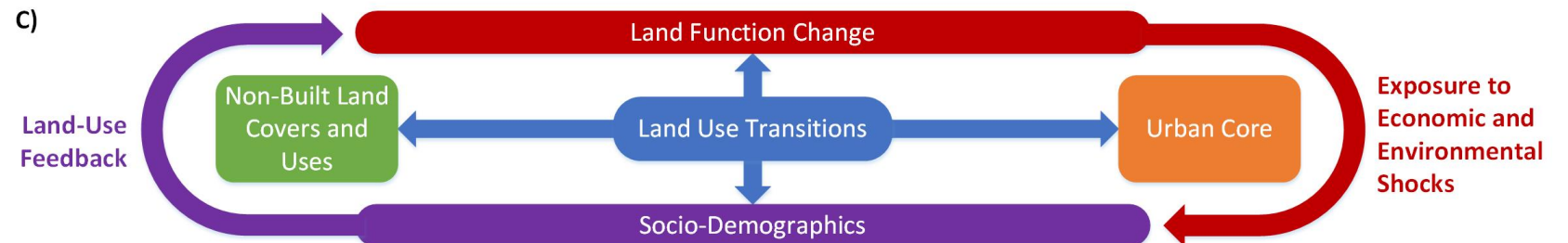
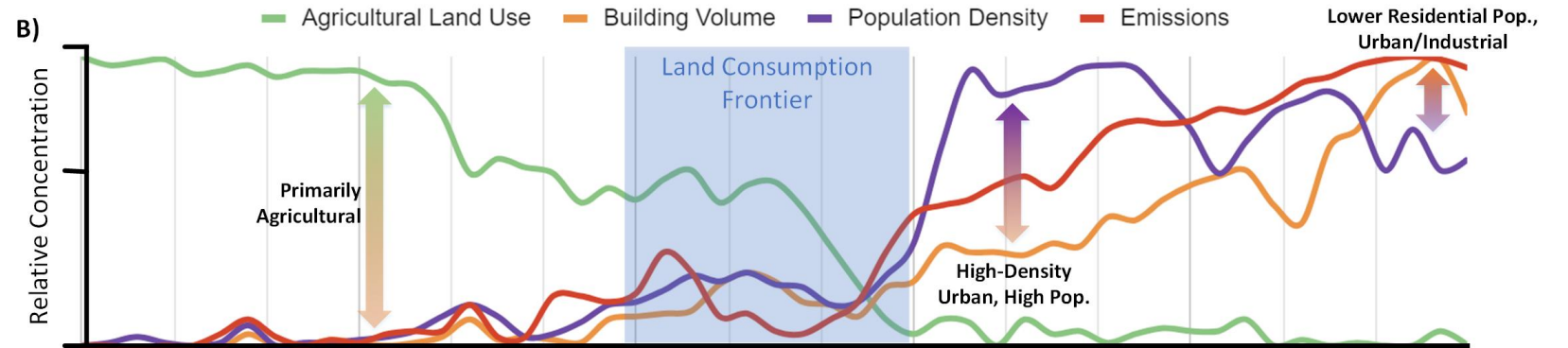
Gridded annual counts of the number of days where the maximum WBGTmax exceeded > 30°C for 2016.

Dependency ratio (2020) across the Greater Mekong Subregion river basins

The Mekong and Chao Phraya basins have the most notable impact of heat stress on key demographic populations of young and old, but also have the highest range of dependency ratios.

Synthesis Framework for Integrating Population Variation & Change with LCLUC

Rural to Urban Continuum (RUC)



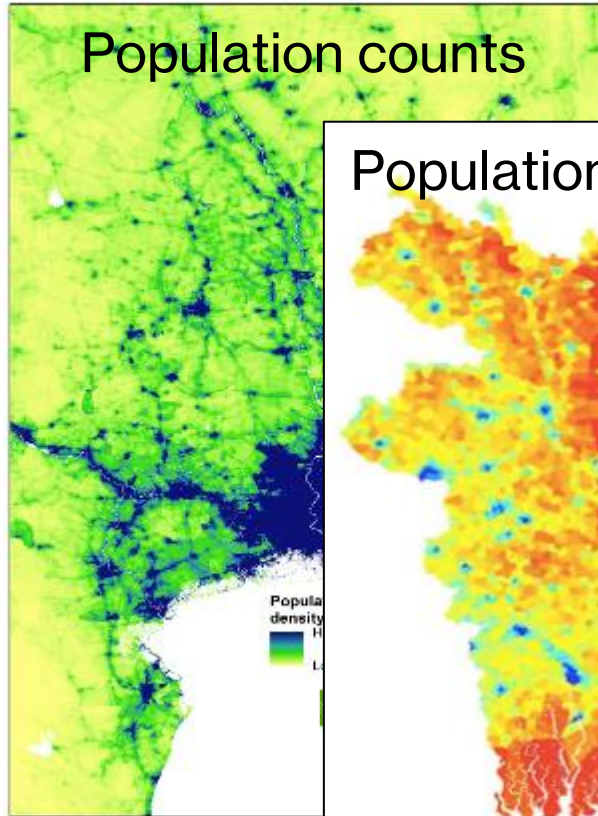
Further Information

E-mail: forrest.stevens@louisville.edu

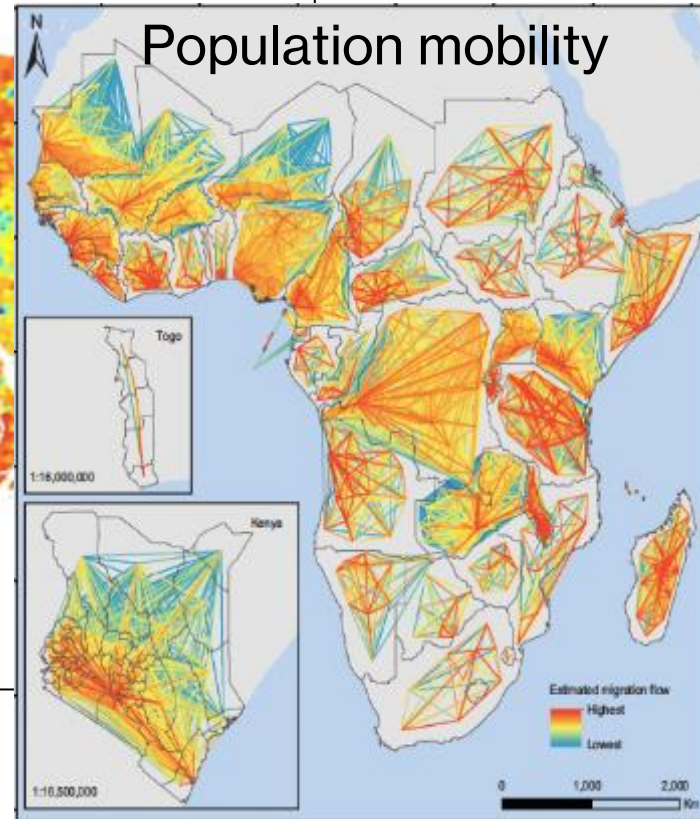
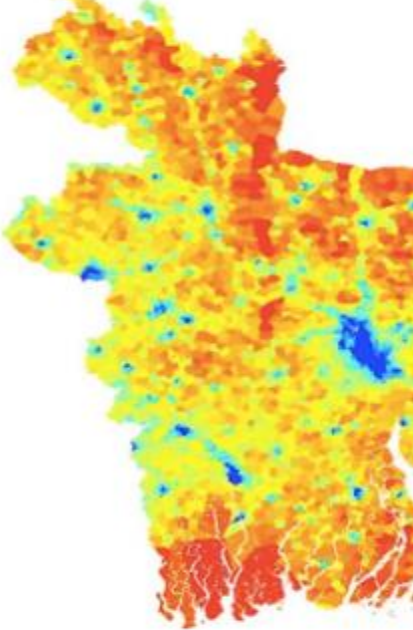
WorldPop

UNIVERSITY OF
LOUISVILLE[®]





Population characteristics



Applied research and implementation group with 25+ staff

Method development for mapping small area demographics and health/development metrics

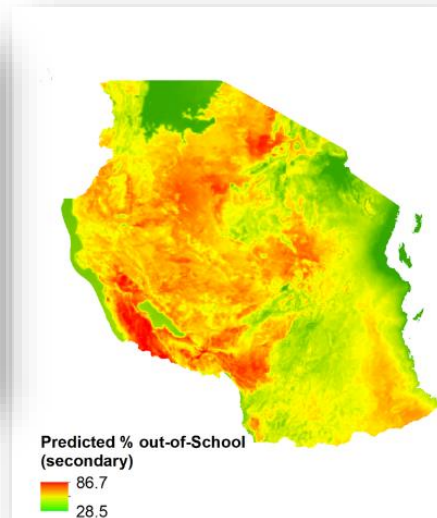
Analysis of spatial data to support decision making

Open data, data analysis tools and dashboards, co-development, capacity strengthening

WorldPop

Bespoke country analyses

Support to census processes; Vaccination microplans; Health system denominators; Education planning; Humanitarian response



Global modelling

Health metrics; Epidemiological modelling; Humanitarian response; Vaccination strategies; Scientific analyses

