

Remote Sensing of Rural Populations: Lessons Learned

Hannah Rosenblum

Population Division

U.S. Census Bureau

Esri User Conference – July 2019

This presentation is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on technical issues are those of the author and not necessarily those of the U.S. Census Bureau.

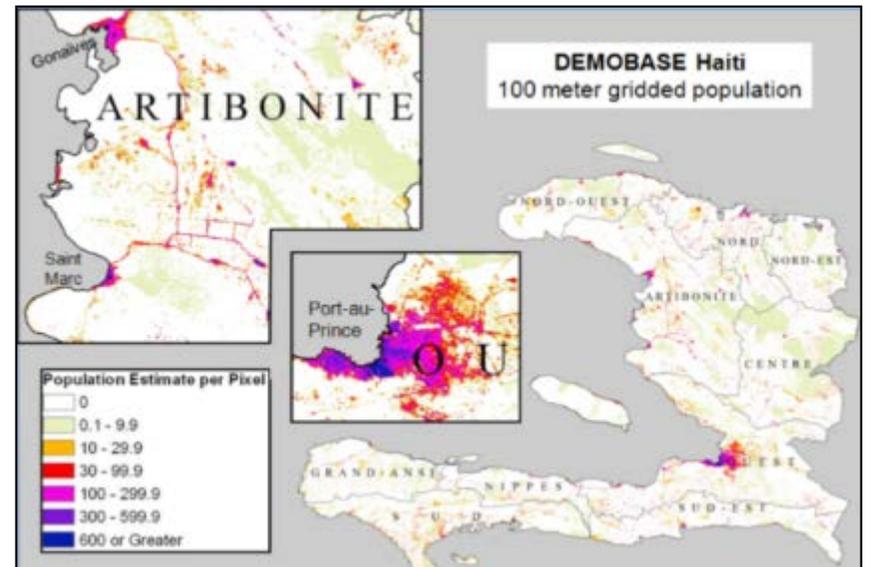
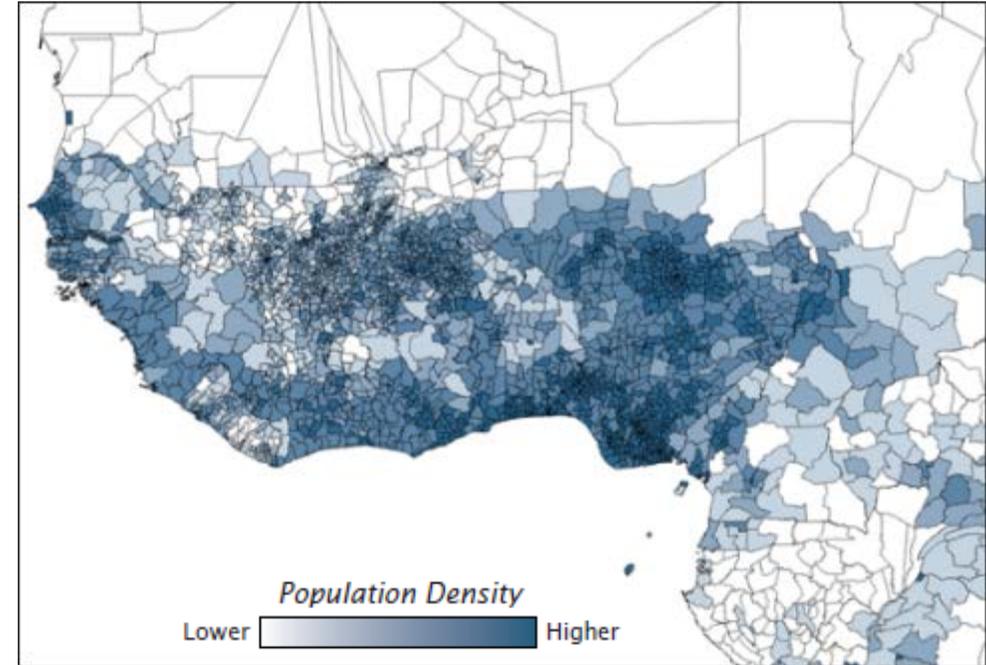
International Programs

- **Applied research and analysis:**

- International population estimates and projections.
- HIV/AIDS research tools.
- Global aging issues.

- **Geospatial products:**

- Subnational population data.
- Demobase.



Different Time Series

Censuses are held on different dates and at varying frequency

Longest period since last census:

- 1) Lebanon (1932)
- 2) Western Sahara (1970)
- 3) Afghanistan (1979)
- 4) D.R. Congo (1984)
- 5) Eritrea (1984)

Census rounds completed* (1990, 2000, 2010):



* At least one census held during the respective census round. Census rounds correspond to the years 1985-1994 (1990 round); 1995-2004 (2000 round); and 2005-2014 (2010 round).

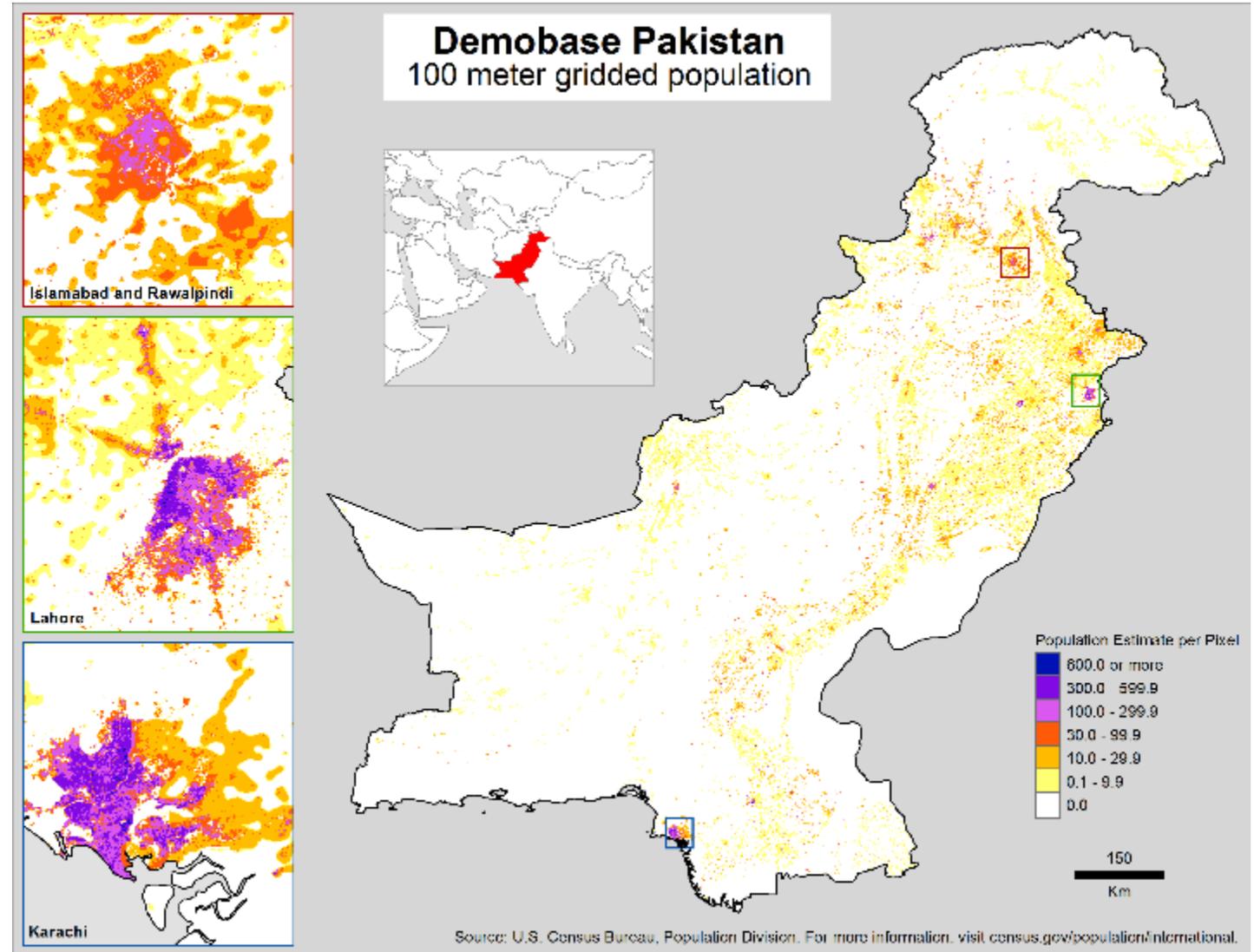
Source: United Nations 2020 World Population and Housing Census Program <unstats.un.org/unsd/demographic/sources/census/censusdates.htm>

Gridded Population Mapping

- Estimate population for individual **grid cells**.
 - Advantage: construct custom geography.
 - Applications: **disaster response** and **scientific research**.
- Data providers:
 - Demobase (U.S. Census Bureau).
 - Gridded Population of the World (CIESIN).
 - Global Human Settlement Layer (Joint Research Centre).
 - WorldPop (University of Southampton).
 - LandScan (Oak Ridge National Laboratory).
 - Global Urban Footprint (German Aerospace Center).
- **Challenging to estimate population.**

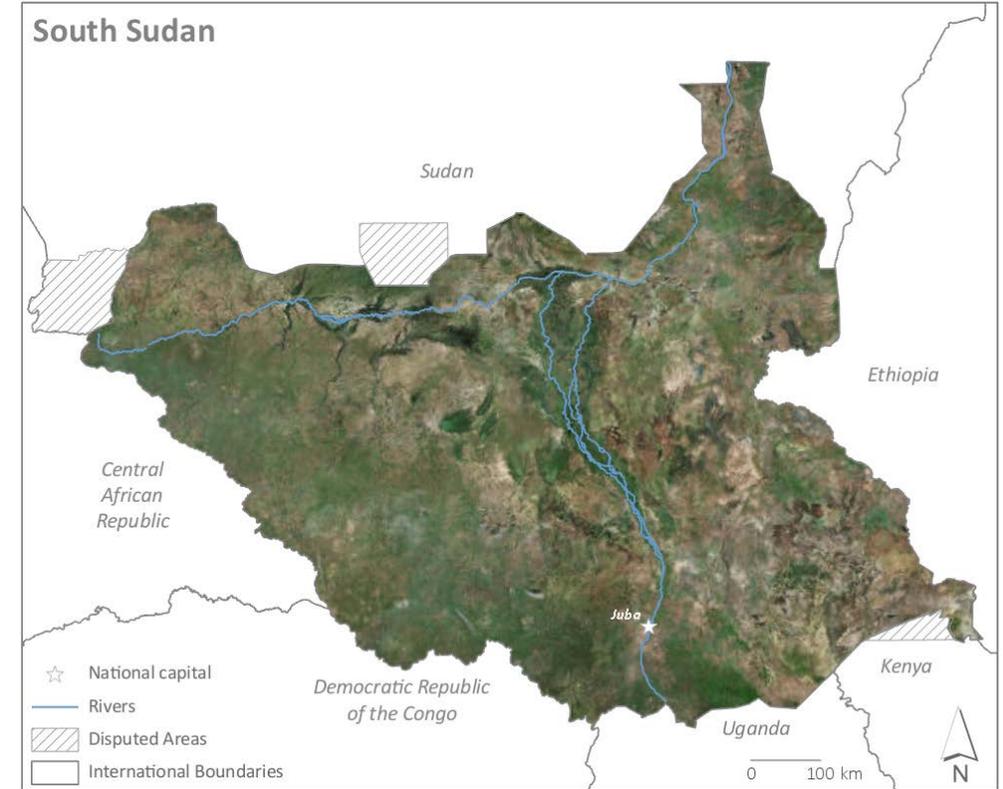
Demobase

- National Research Council (NRC) Report: “Tools and Methods for Estimating Populations at Risk”.
- Previous projects: Haiti, Pakistan, and Rwanda.
- Previously established the relationship between built-up areas and population density.



South Sudan

- Became an independent country in 2011.
- Has not had a census since **2008**, when it was still part of Sudan.
- Human activity is not uniformly associated with impervious surfaces.
- As much as **83 percent** of the population live in rural areas (South Sudan Bureau of Statistics).
- Violence has affected the spatial distribution of the population.



South Sudan Landscape

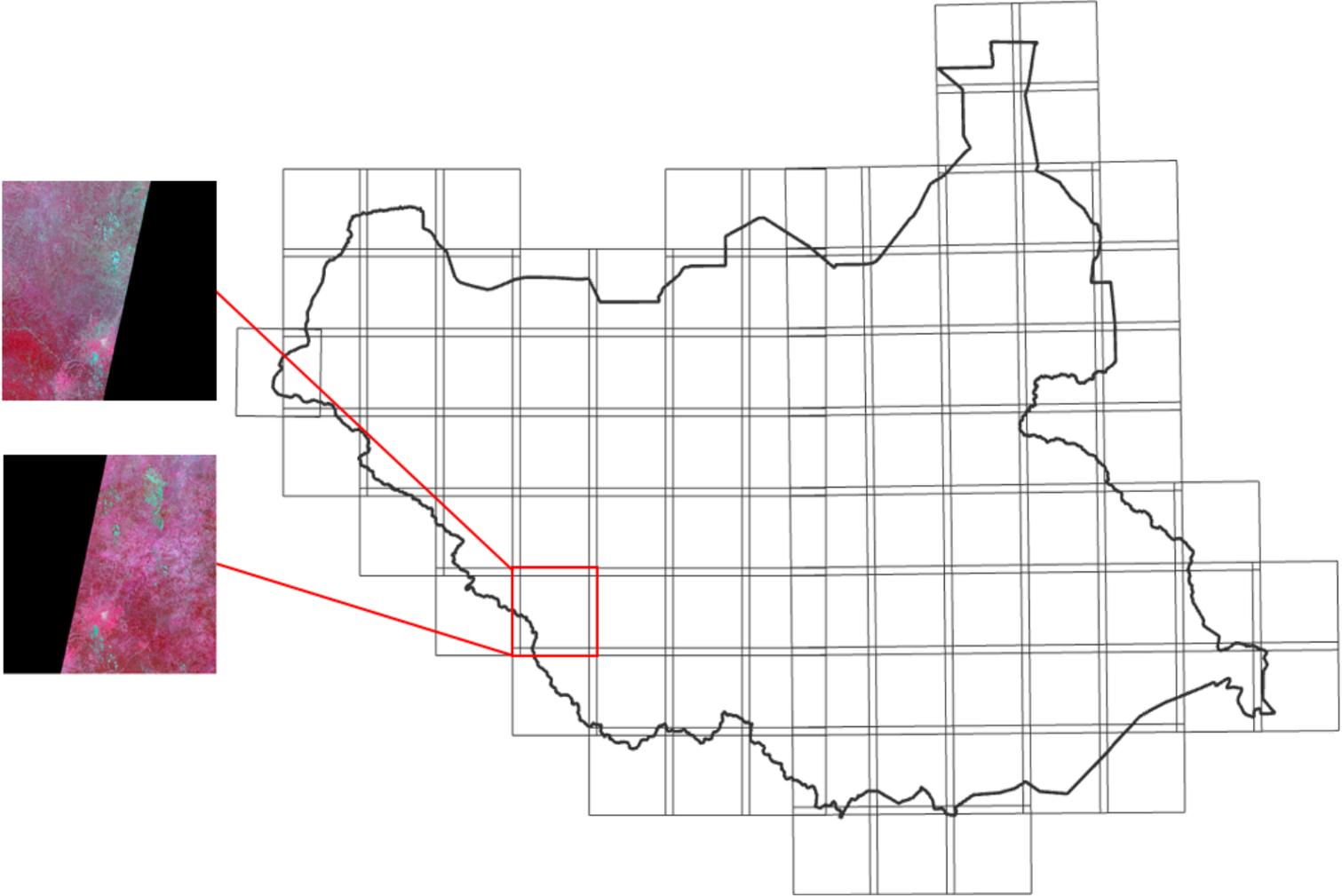


Source: DigitalGlobe

Methodology for South Sudan

- Pre-processing: automated thresholding into vegetation/non-vegetation images.
- Built-up area classification, using both **unsupervised** and **supervised classification** methods.
- **Population disaggregation/distribution**, using the built-up area classification along with ancillary datasets as inputs for the random forests algorithm.
- Compare output population datasets to other gridded population datasets.

Data Collection



Built-Up Area Classification

Iterative process of unsupervised Isodata clustering, followed by supervised training, until classification is satisfactory.

Wet season

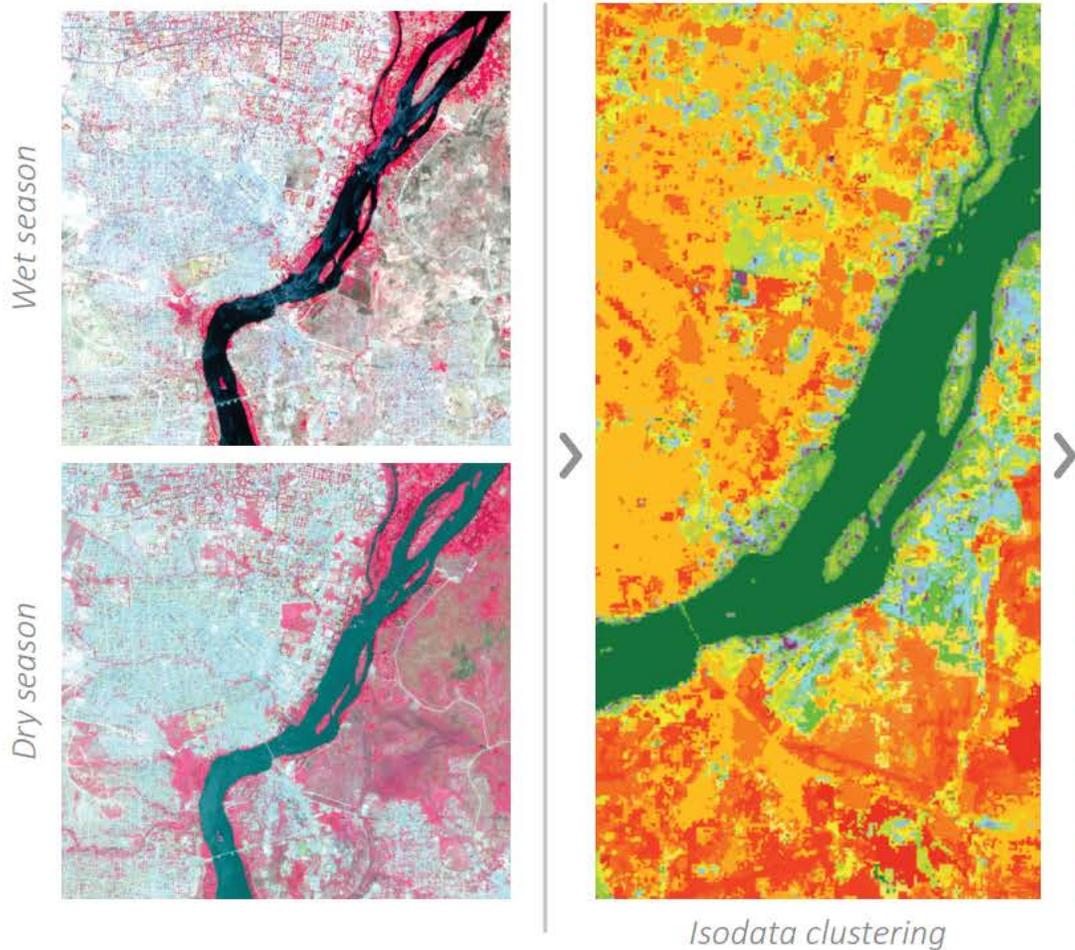


Dry season



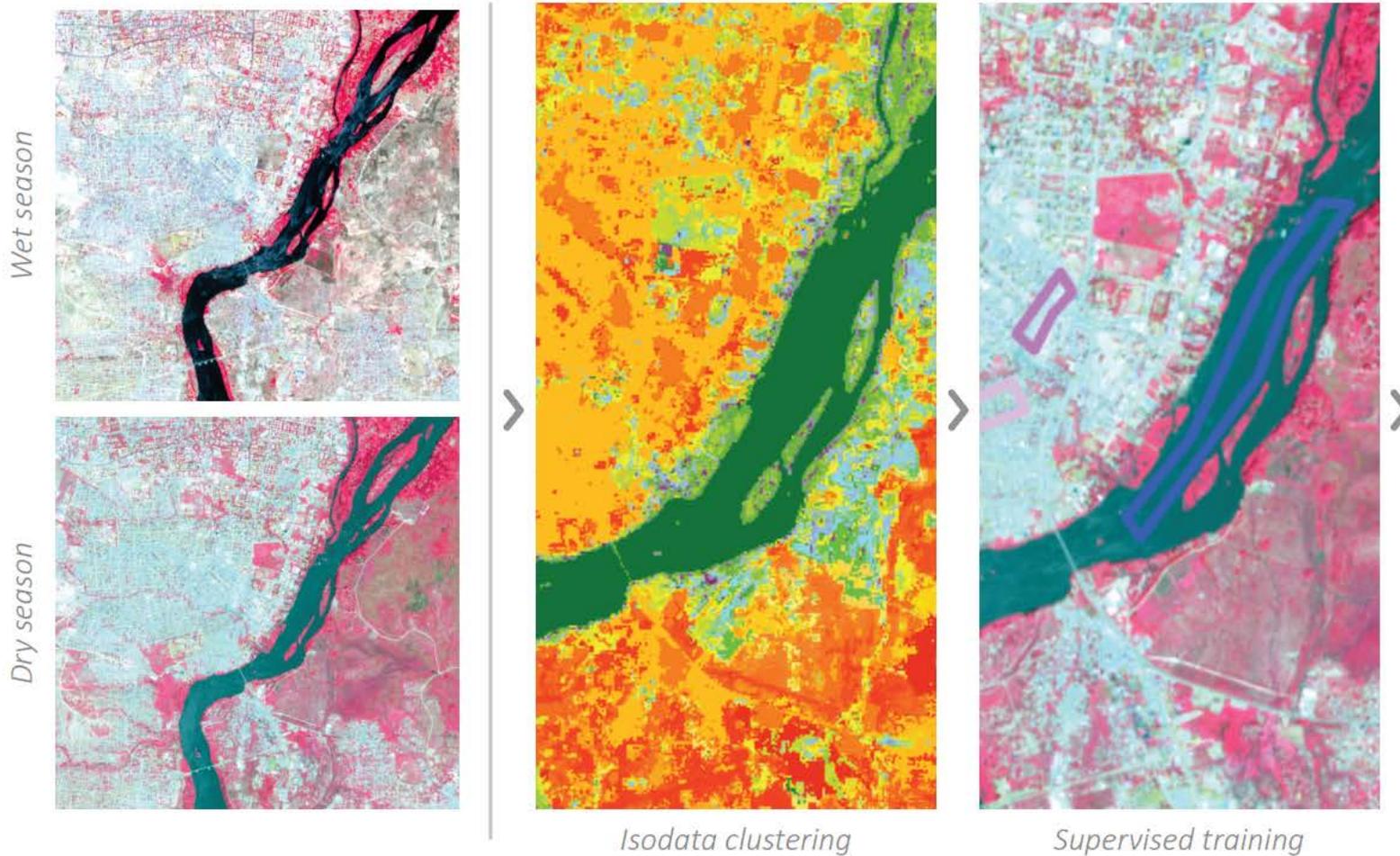
Built-Up Area Classification

Iterative process of unsupervised Isodata clustering, followed by supervised training, until classification is satisfactory.



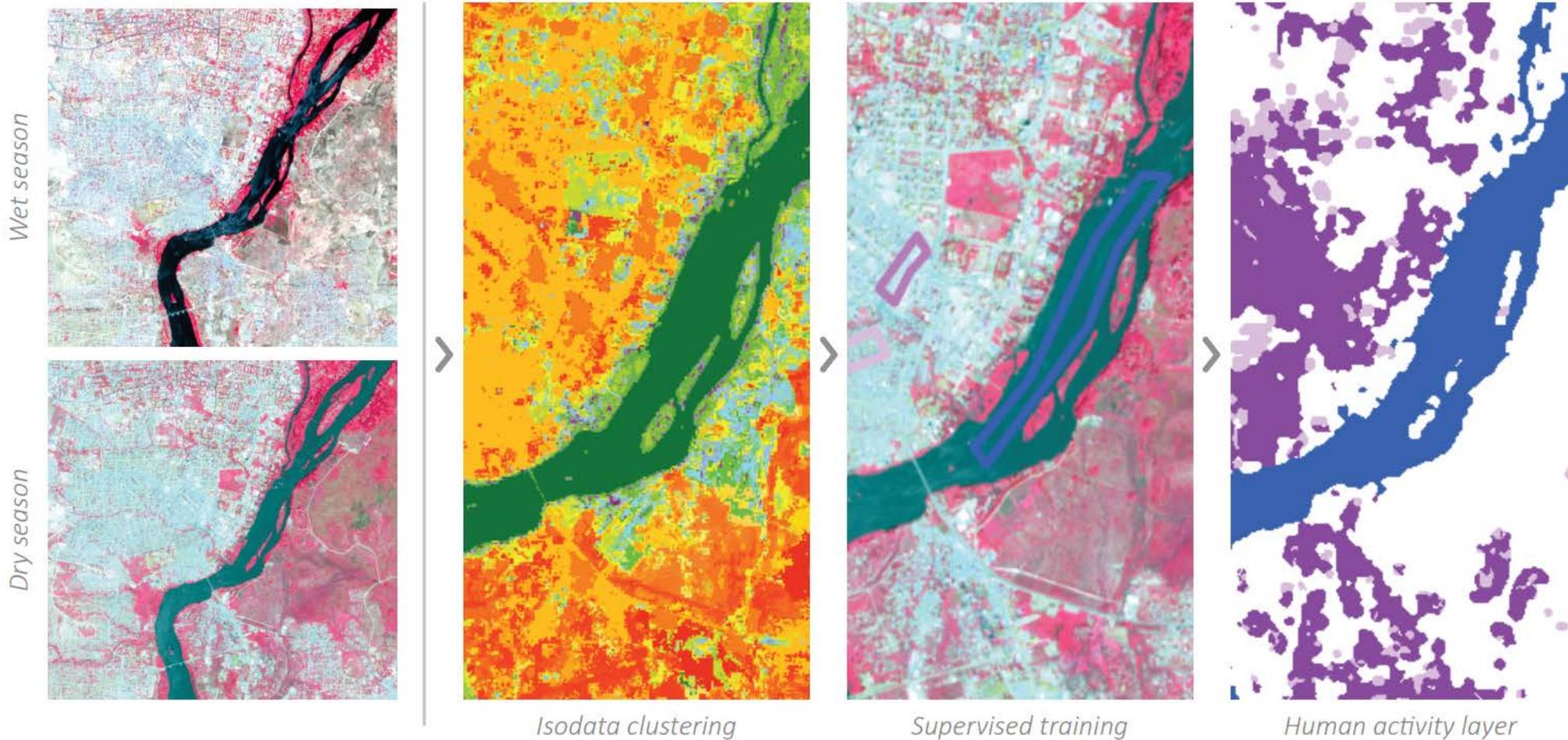
Built-Up Area Classification

Iterative process of unsupervised Isodata clustering, followed by supervised training, until classification is satisfactory.



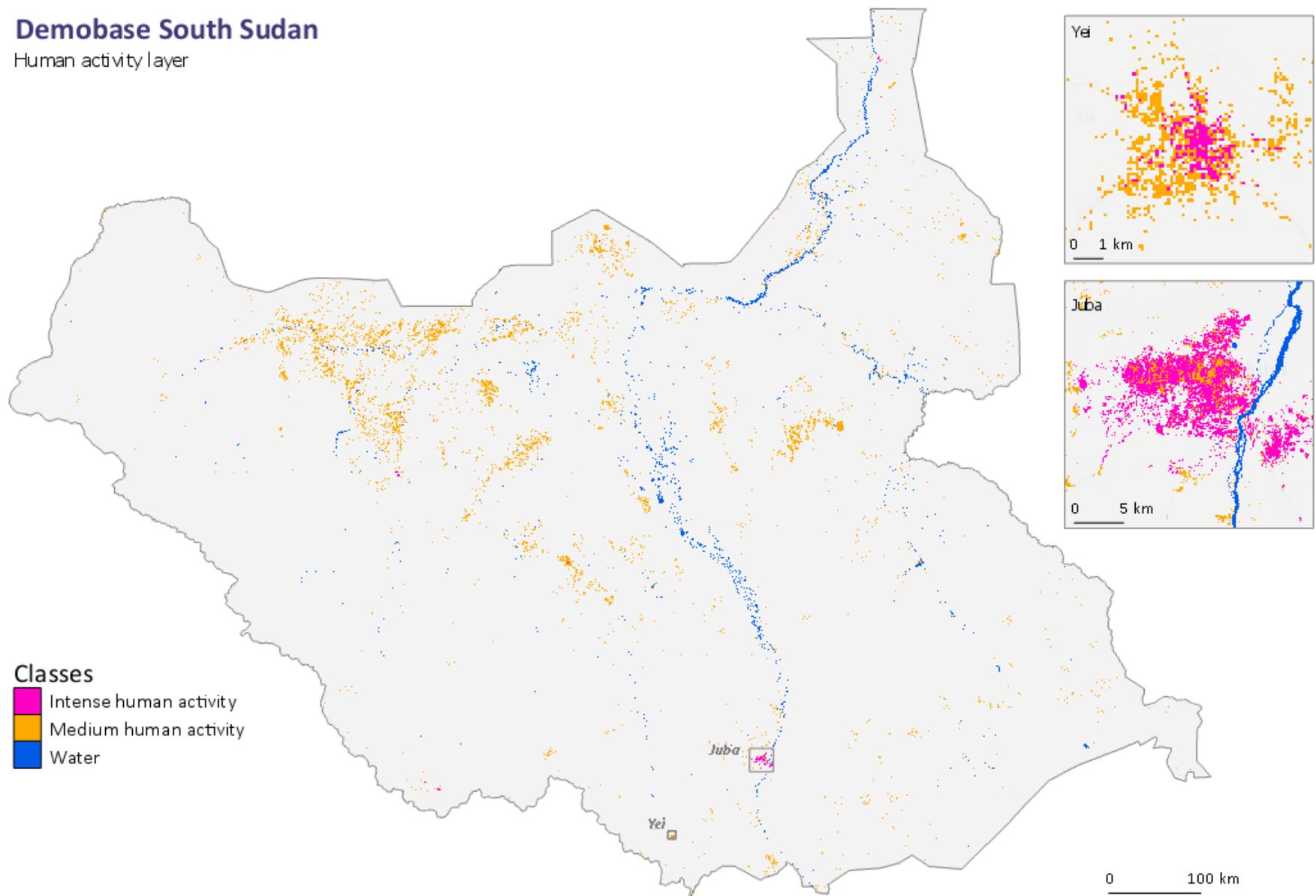
Built-Up Area Classification

Iterative process of unsupervised Isodata clustering, followed by supervised training, until classification is satisfactory.



Demobase South Sudan

Human activity layer



Population Census (2008)

- A census is a **complete count** of the population.
- The most recent census, in 2008, was conducted prior to South Sudan's independence from Sudan.
- Results were available for as low as the ADM4 level.

Subnational Administrative Boundaries

- To geo-enable the population data, subnational administrative boundaries were needed for matching.
- Despite tabulated census data being available at ADM4 (boma), only ADM3 (payam) and higher level boundaries were available.

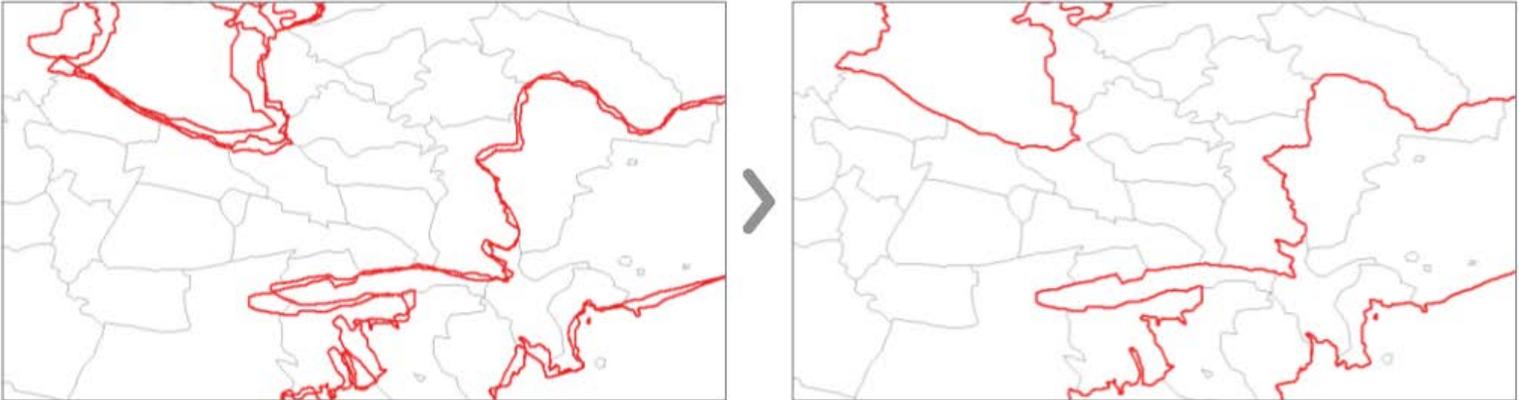
Population Estimates (2017)

- Estimates were produced by projecting 2008 census data forward to 2017 using assumptions about births, deaths, and migration (these were generated by the South Sudan International Food Security Phase Classification Technical Working Group).
- Due to the complexity of producing estimates and limitations in the underlying assumptions, these data were only available at ADM2.

Population

Geometry corrections

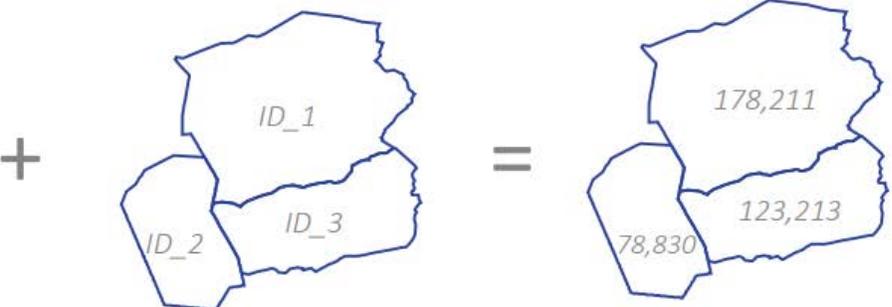
Repair subnational topology and align international boundaries to LSIB



Matching

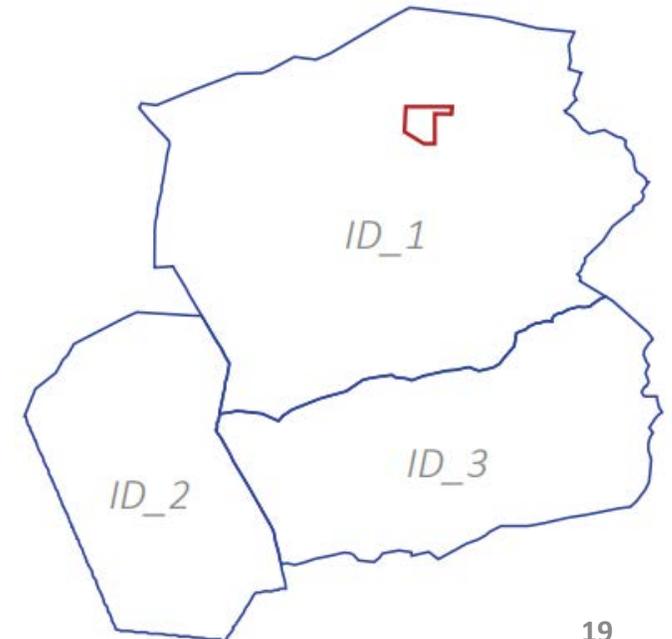
Match population estimates to appropriate geographies

ADM	POP
ID_1	178,211
ID_2	78,830
ID_3	123,213



Integrating IDP site estimates (2017 only)

- To enhance the accuracy of the population density model, the boundaries of major IDP sites were delineated and unioned with the ADM 2 boundaries.
- The July 2017 population estimates for both ADM2s and IDP sites were linked to the corresponding geographic areas.
- The population estimate for each IDP site was subtracted from its respective ADM2 .

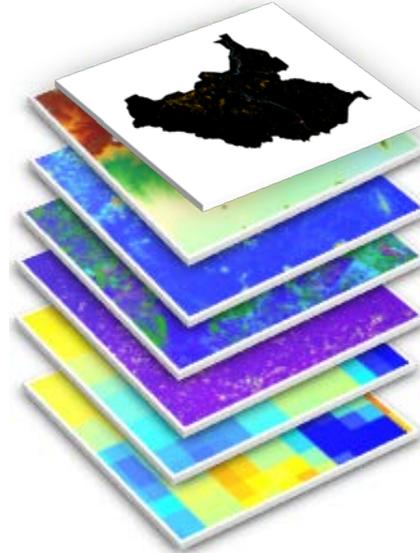


Random Forests

Census data (2008)/
Population estimates (2017)

+

Human activity layer,
ancillary datasets



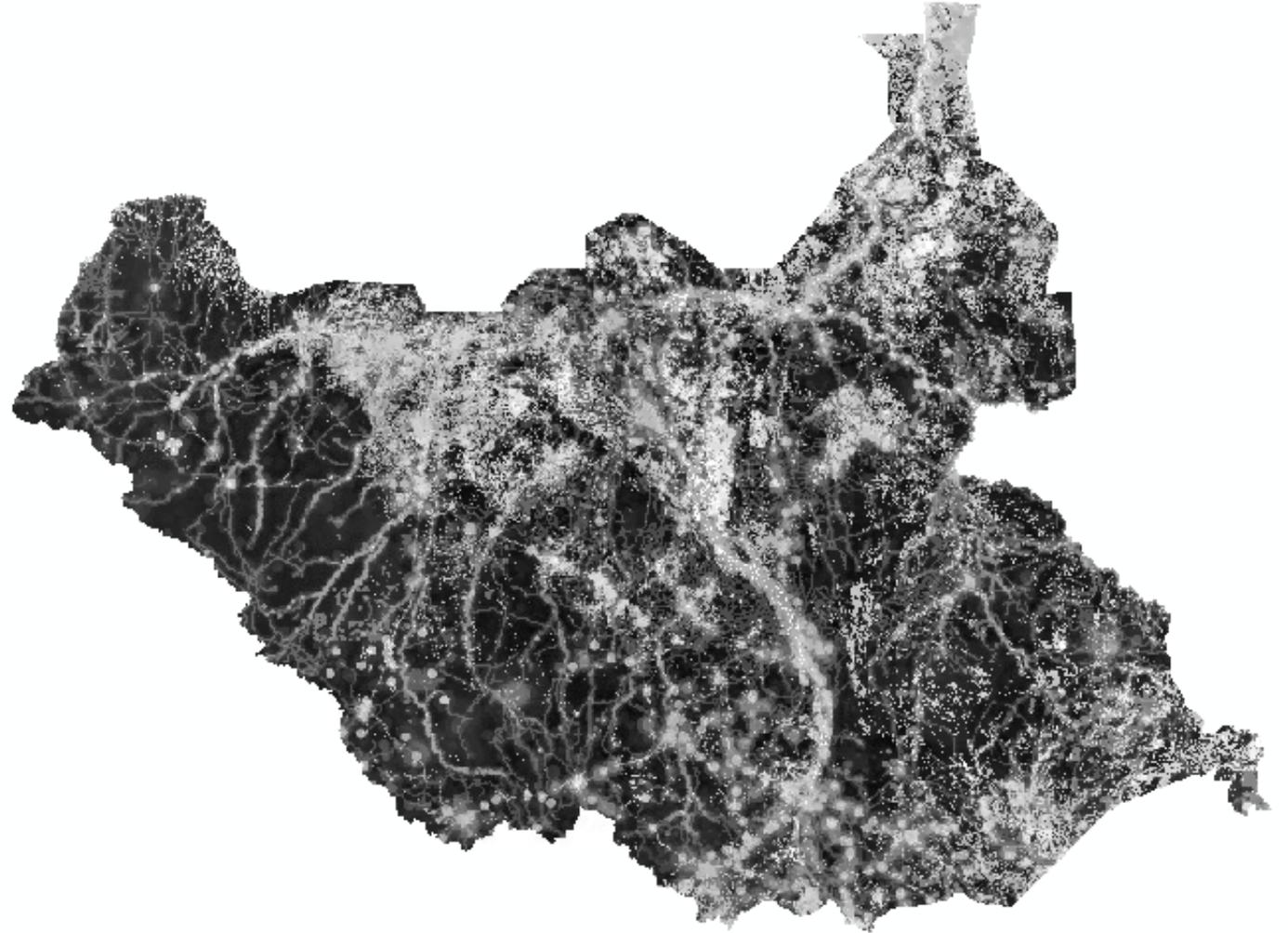
- Administrative level 1 capitals
- Administrative level 2 capitals
- Education facilities
- Elevation
- Health facilities
- Internally displaced persons camps
- Land cover
- Large towns
- Maximum precipitation
- Maximum temperature
- Minefields
- Minimum temperature
- Miscellaneous points of interest
- Night lights
- Net primary productivity
- Places of worship
- Rivers
- Roads
- Slope
- Small towns
- Town areas
- Village
- Water facilities



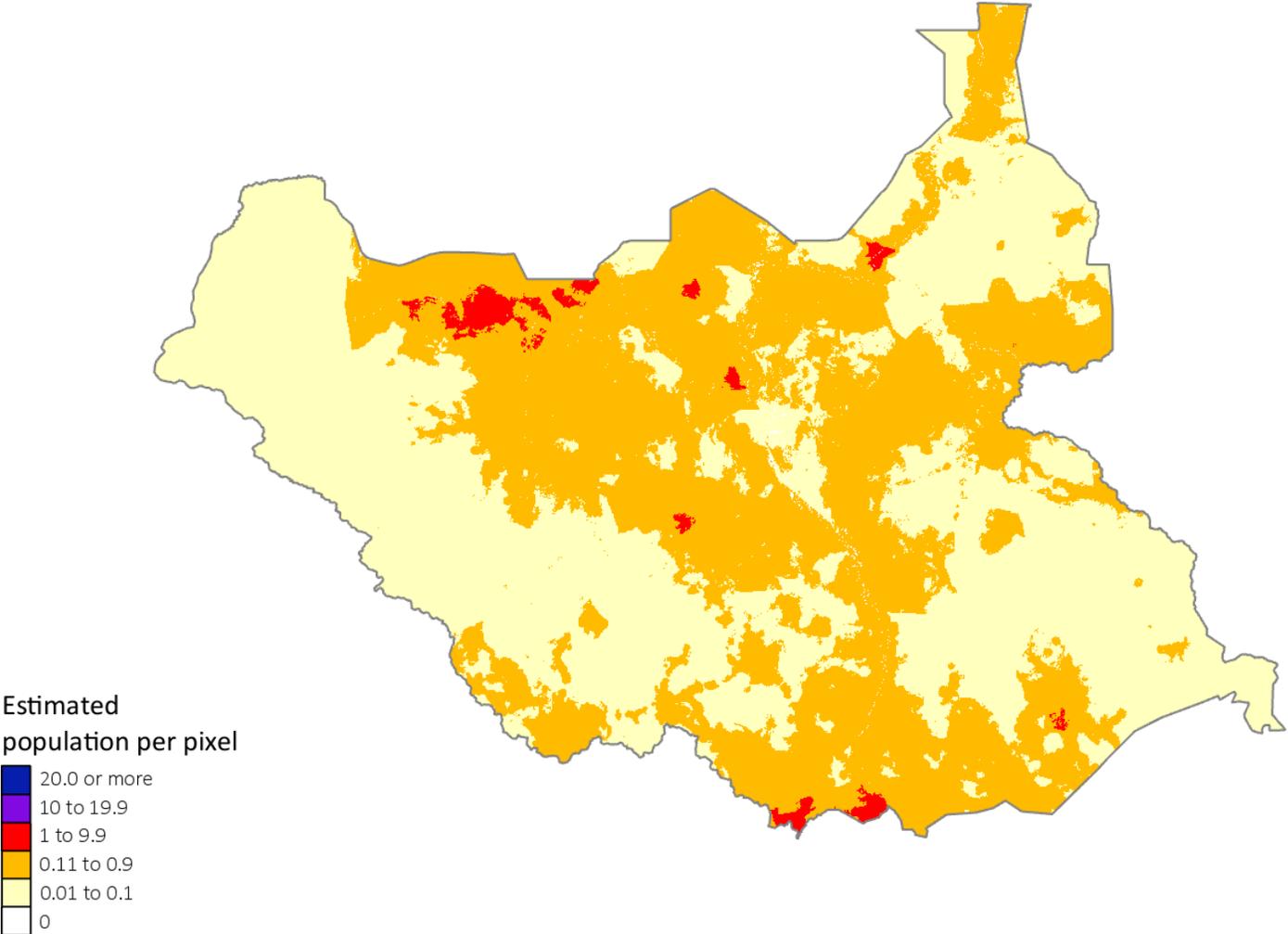
Random forests model

Population Disaggregation to Density Surface

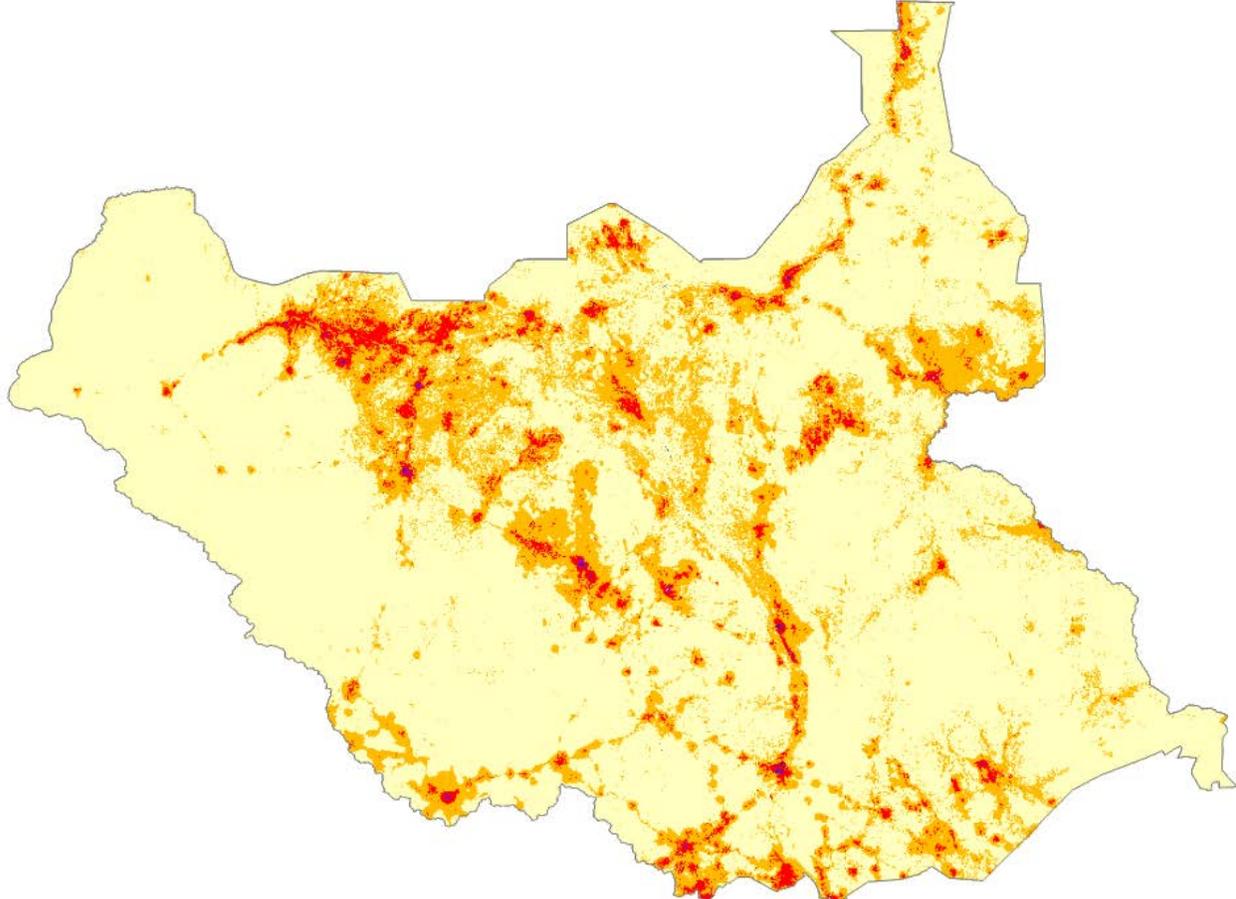
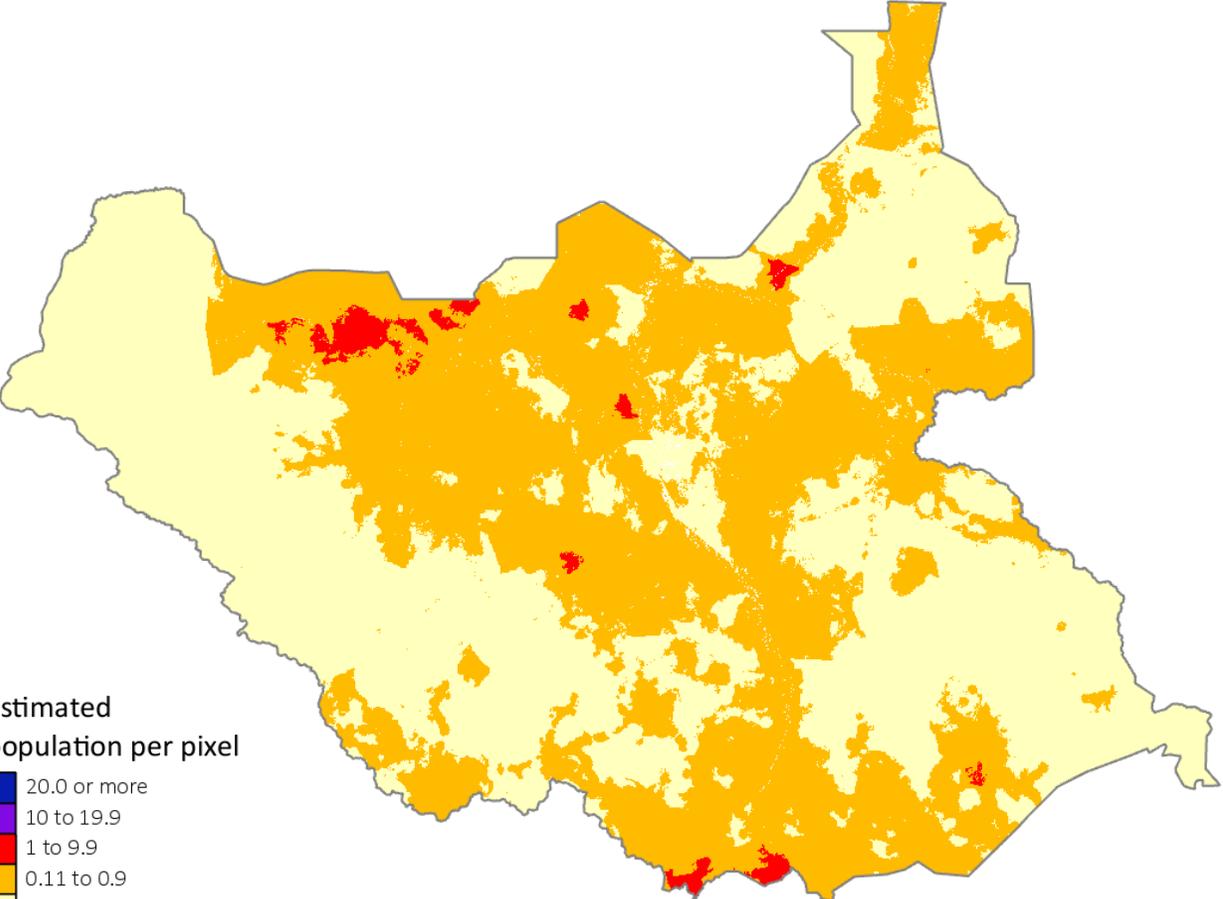
- Population counts for each administrative division were then disaggregated onto the density surface, resulting in gridded population datasets for 2008 and 2017.



2017 Initial Results



2017 Results – Using IDP Camps as Pseudo-administrative Boundaries



Estimated population per pixel

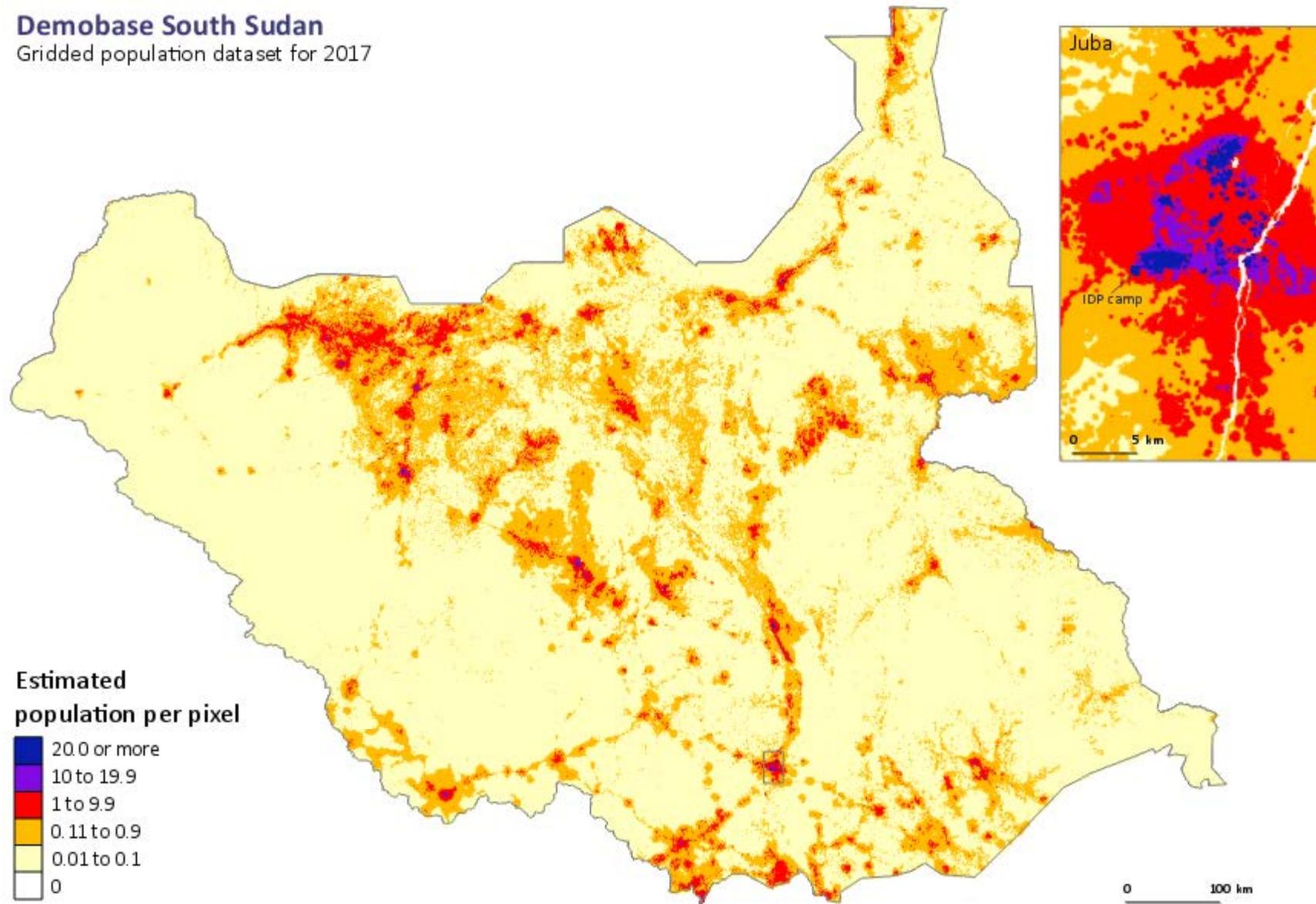
20.0 or more
10 to 19.9
1 to 9.9
0.11 to 0.9
0.01 to 0.1
0

Before

After

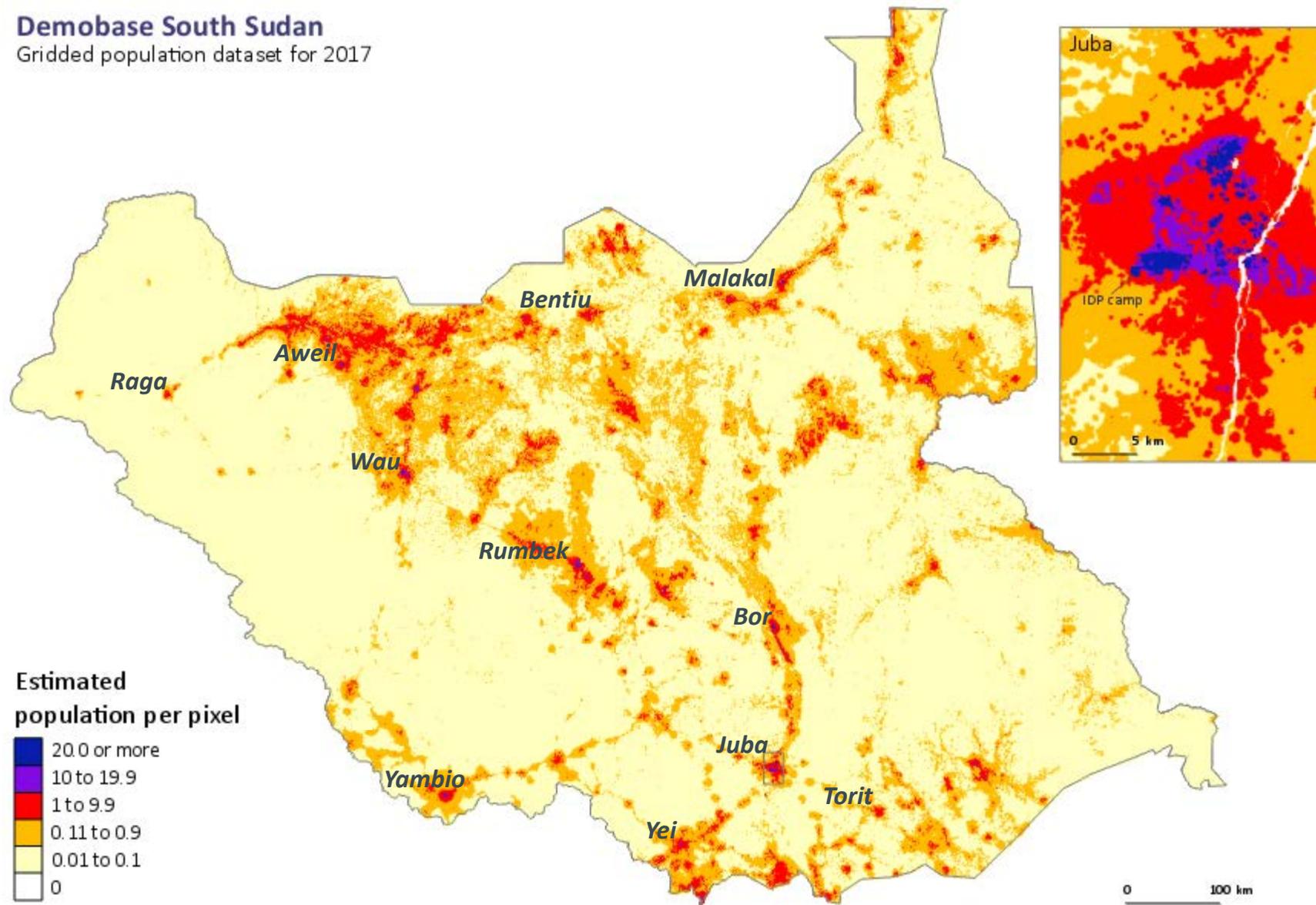
Demobase South Sudan

Gridded population dataset for 2017



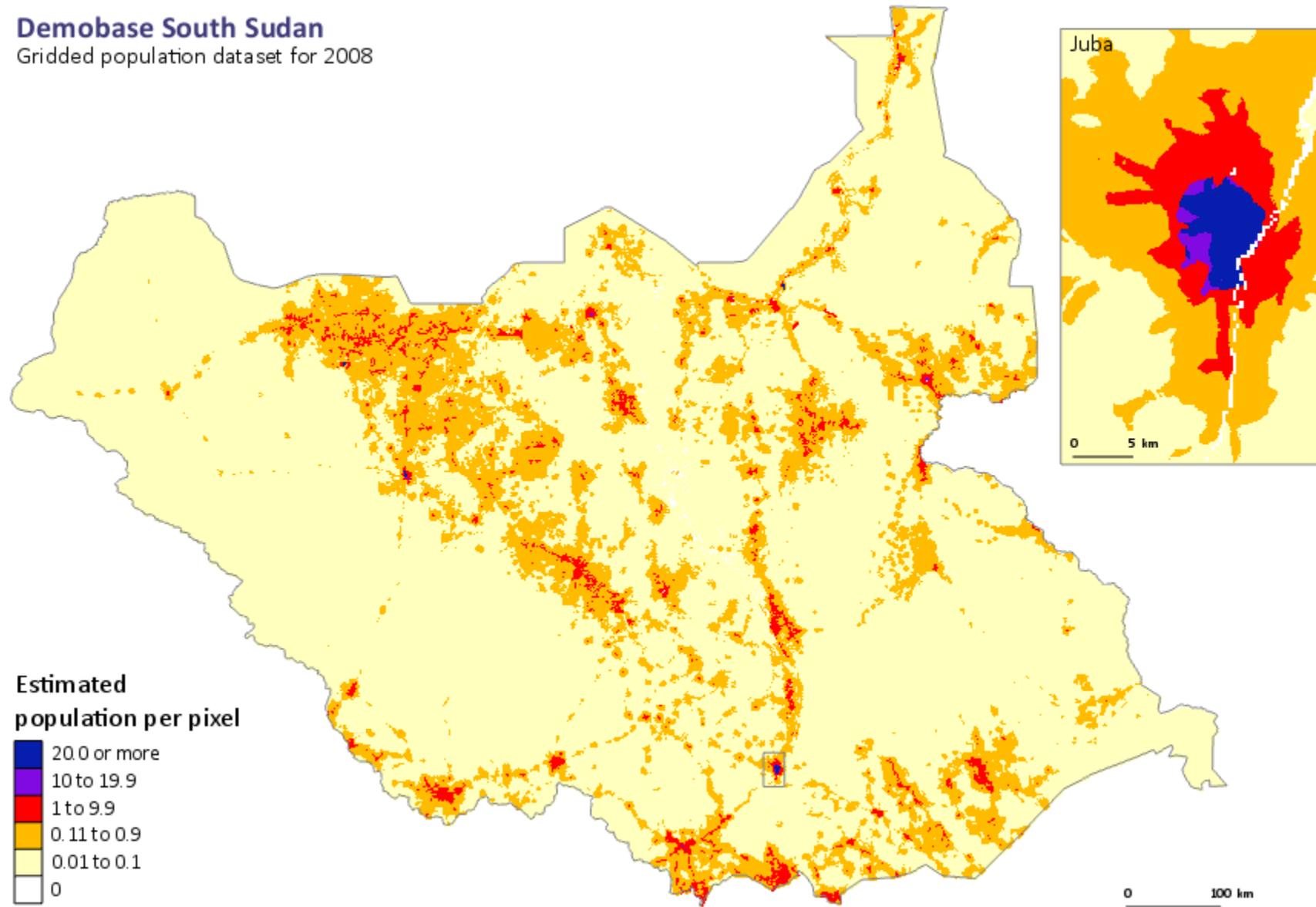
Demobase South Sudan

Gridded population dataset for 2017



Demobase South Sudan

Gridded population dataset for 2008

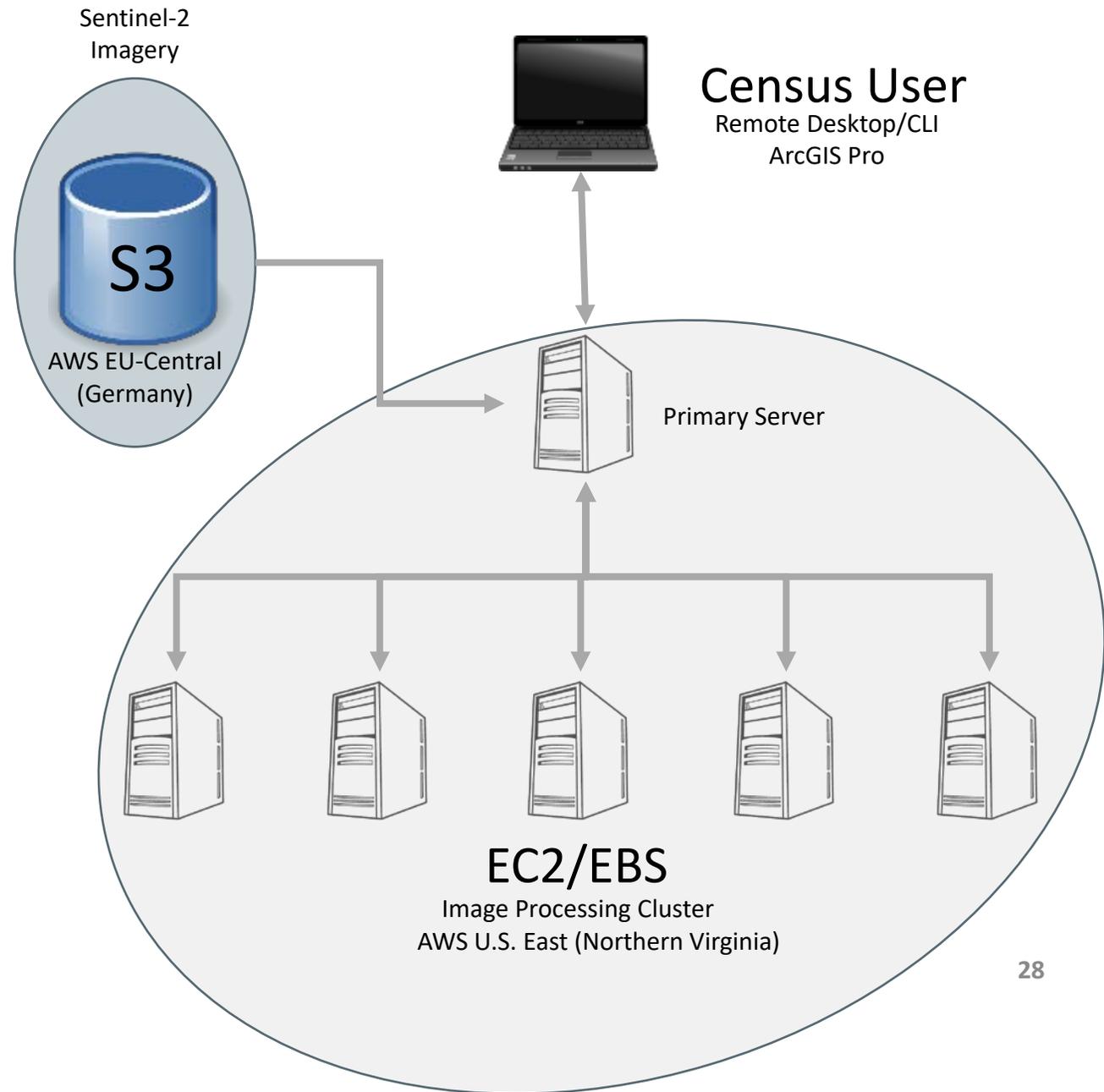


Challenges

- Some dispersed huts with little or no impervious surface around them may not be detected, difficult to distinguish between abandoned and populated huts.
- Most recent census is 2008.
- Making assumption that census is accurate.
- No independent, spatially explicit population data available for validation.
 - Instead, compared results to other gridded population datasets.

Next Steps

- Move to cloud environment for future projects.



Thank you!

Hannah Rosenblum

Hannah.Rosenblum@census.gov

Website:

[census.gov/programs-surveys/international-programs.html](https://www.census.gov/programs-surveys/international-programs.html)

2017 Results – Comparisons to other gridded datasets

