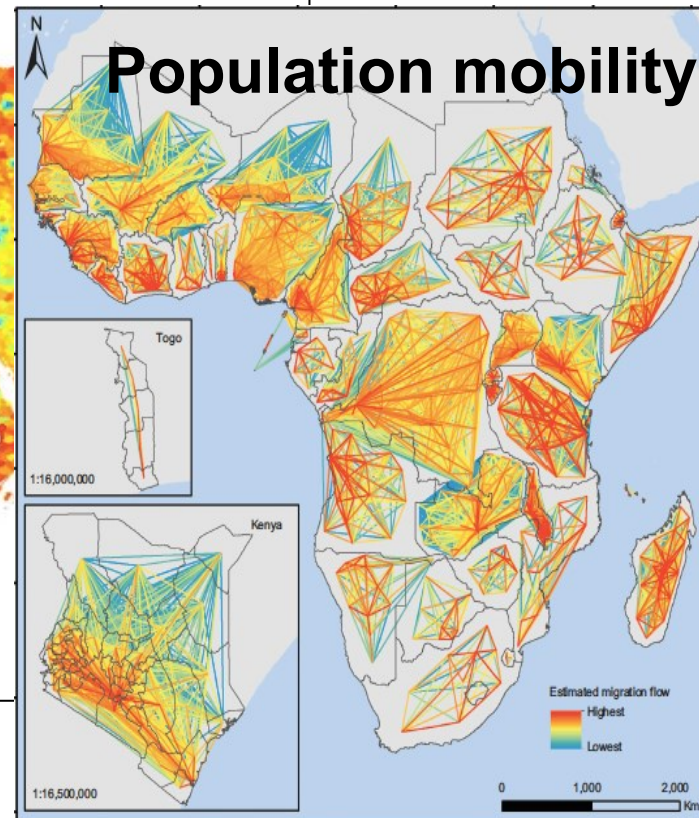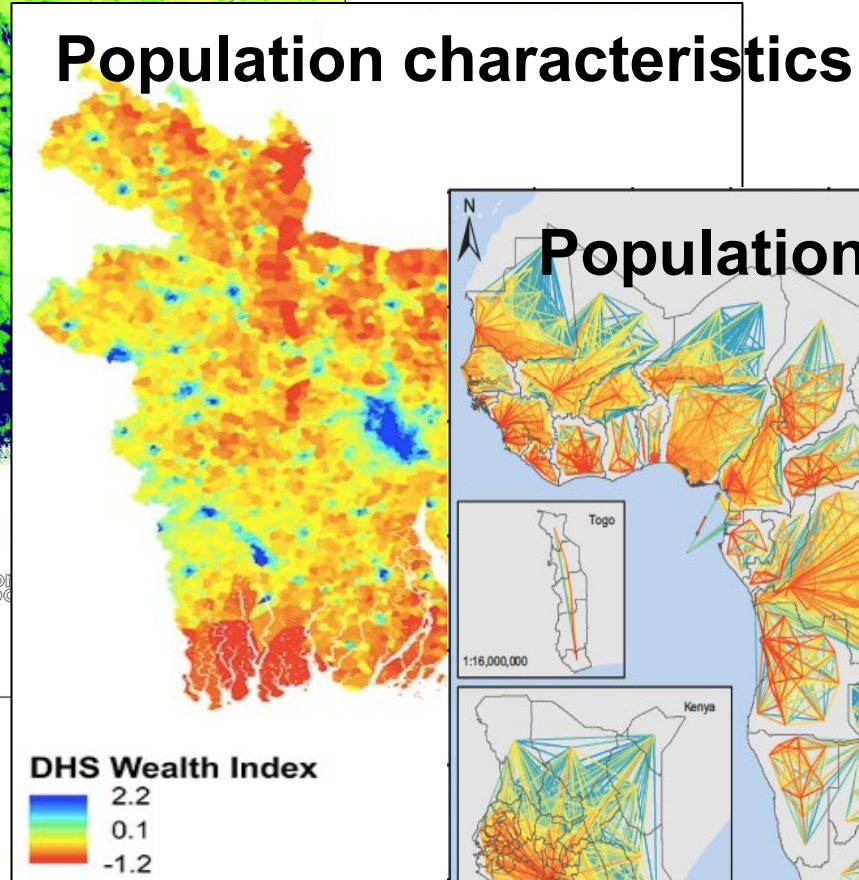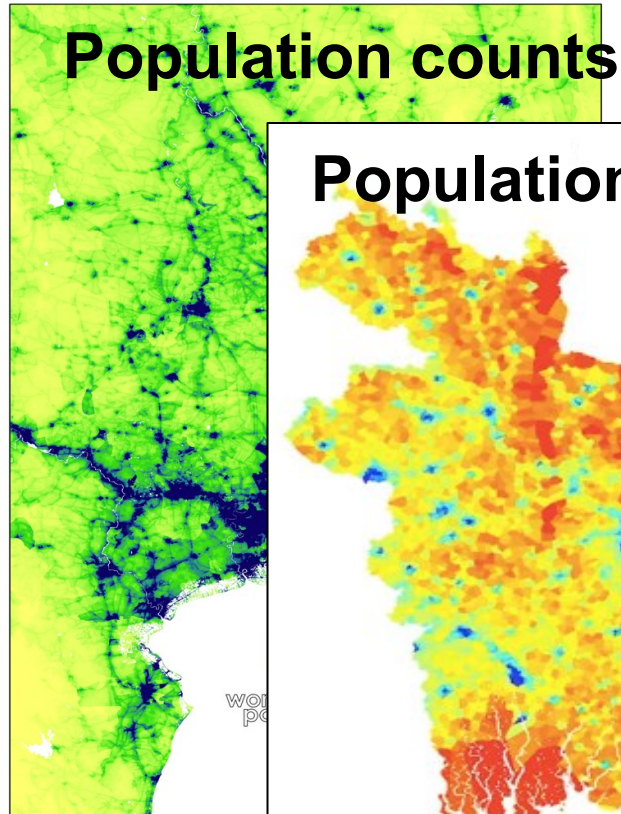# Mapping forcibly displaced people at high-resolution using machine learning and satellite-derivative datasets

Sarchil Qader[1], Alexey Noskov[1], Attila Lazar[1], Edith Darin[2], Ahmadou Dicko[3], Hisham Galal[3], Hyunju Park[3], Rebeca Moreno Jimenez[3], Andrew J Tatem[1]

[1]**WorldPop**; [2]University of Oxford; [3]UNHCR

Southampton Geospatial Summer School
12th September 2024

WorldPop

Population counts

Population characteristics

Population mobility

DHS Wealth Index
2.2
0.1
-1.2

Togo
1:16,000,000

Kenya

Estimated migration flow
Highest
Lowest

1:16,500,000

0   1,000   2,000
Km

www.worldpop.org

Applied research and implementation group

Open data, open peer-reviewed methods, co-development, capacity strengthening

Mapping small area demographics and dynamics for low and middle income countries

Gridded population datasets used by UN agencies, Governments, DHIS2

Major application areas in epidemiology, RMNCAH, childhood vaccination

But _why_ do we need to measure and map populations?

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

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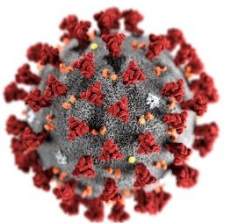
Background

☐ At the end of 2023, an estimated 117.3 million people worldwide were forcibly displaced due to persecution, conflict, violence, etc

☐ According to UNHCR report, forced displacement has continued to increase in the first four months of 2024 and by the end of April 2024 is likely to have exceeded 120 million.

Forced displacement data challenge and intervention

❑ High-quality baseline population data disaggregated down to local levels are fundamental for many applications, including needs assessment, planning and delivery of public services and response to disasters.

❑ Most of the short-term population change in many current and ongoing crises is caused by forced displacement. Forced displacement data, however, are complicated and often difficult to align with other sources of population, demographic, and humanitarian data.

❑ While datasets from Government and UN agencies such as UNHCR, OCHA, IDMC and IOM provide valuable insights, their granularity is often limited to administrative levels 1 or 2.

❑ This lack of detailed data poses significant challenges for policymakers, urban planners, national survey and researchers.

❑ Failure to reach displaced people will impede the Sustainable Development Goal's (SDGs) advancement and have negative effects on the peace and security of nations

WorldPop

Overview of approaches for creating gridded population datasets

"Top down"

"Bottom-up"

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

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

"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"

"We have conducted our census, but could not access certain areas and need estimates for these"

Top-Down

Bottom-up

Estimated of people per

www.worldpop.org/methods/populations

# Example of forced displacement database

UNHCR proGres Database:

❏ proGres database is developed by the United Nations High Commissioner for Refugees (UNHCR), which records all refugees defined as "individuals who are outside their country of origin and who are unable or unwilling to return there owing to serious threats to life, physical integrity or freedom resulting from generalized violence or events seriously disturbing public order"

❏ The granularity of the proGres data varied from one country to another and it is not available for all countries.

❏ At the end of March 2023, Cameroon had over 480,000 refugees and asylum seekers

❏ Inside the refugee camps in Cameroon, the data was comprehensive and detailed, but it was only available at the adm3 level for the rest of the country.

❏ With this level of the data, policy making decision, intervention and household survey will face critical challenges to reach these vulnerable population subgroups.
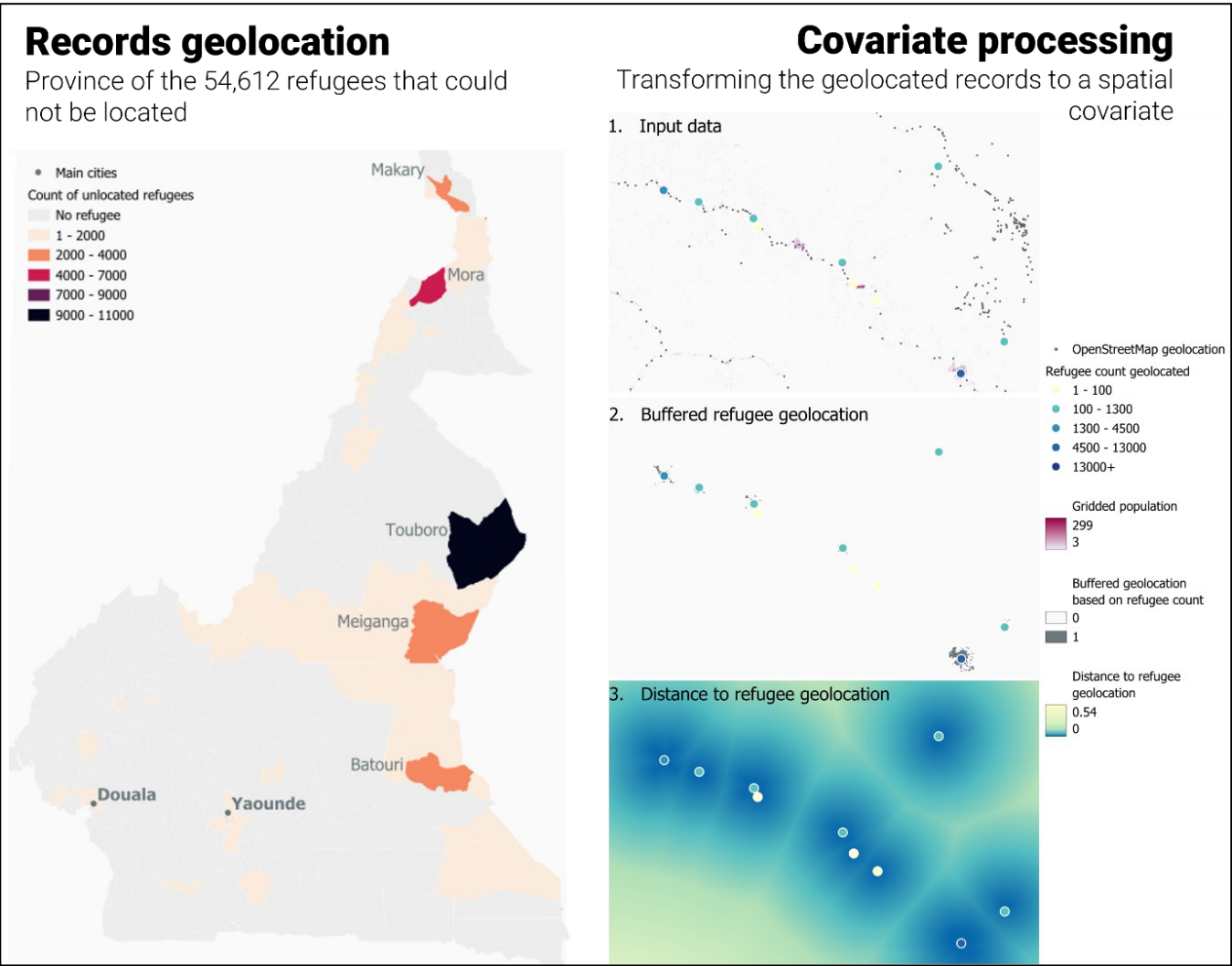

WorldPop
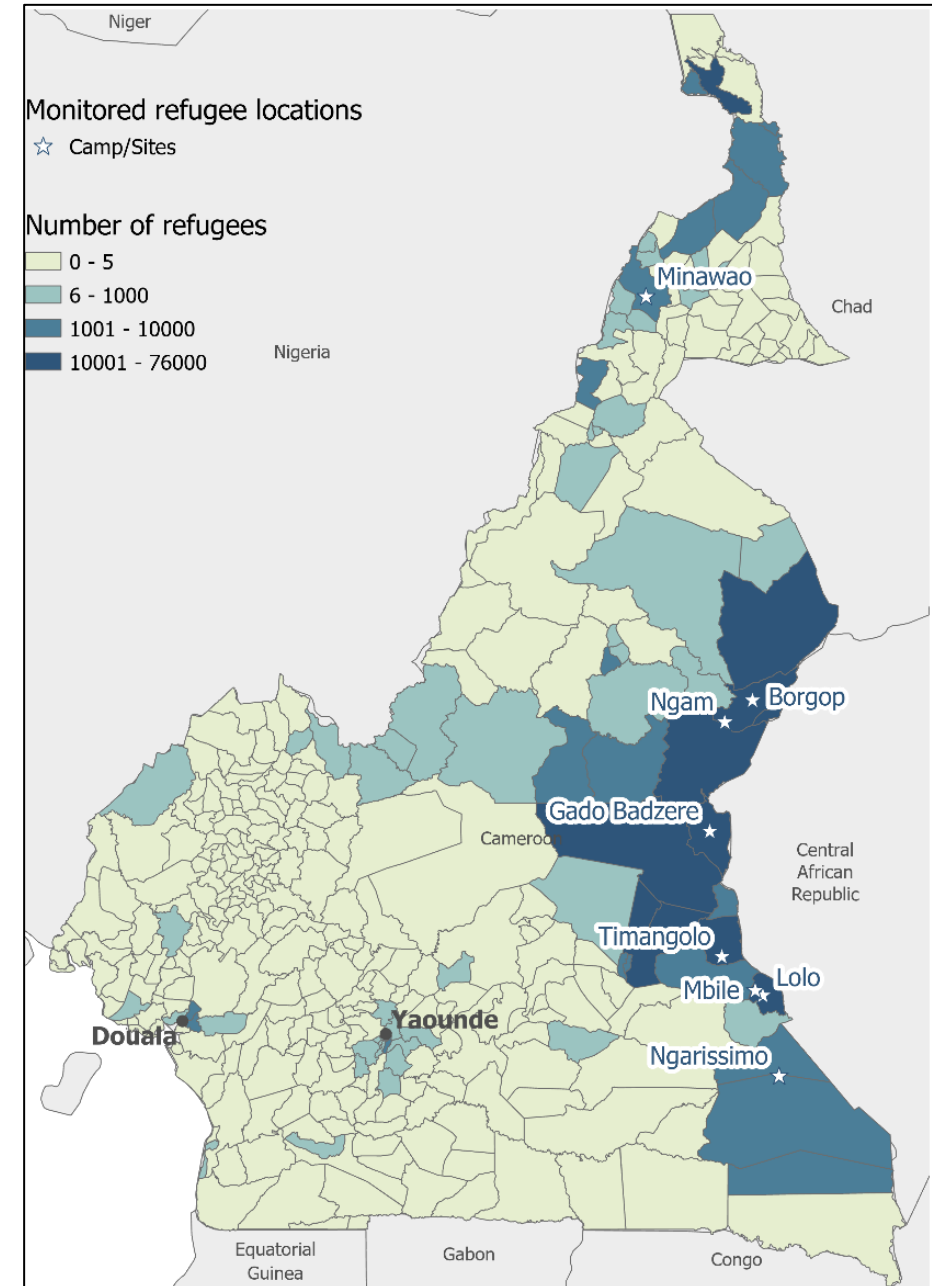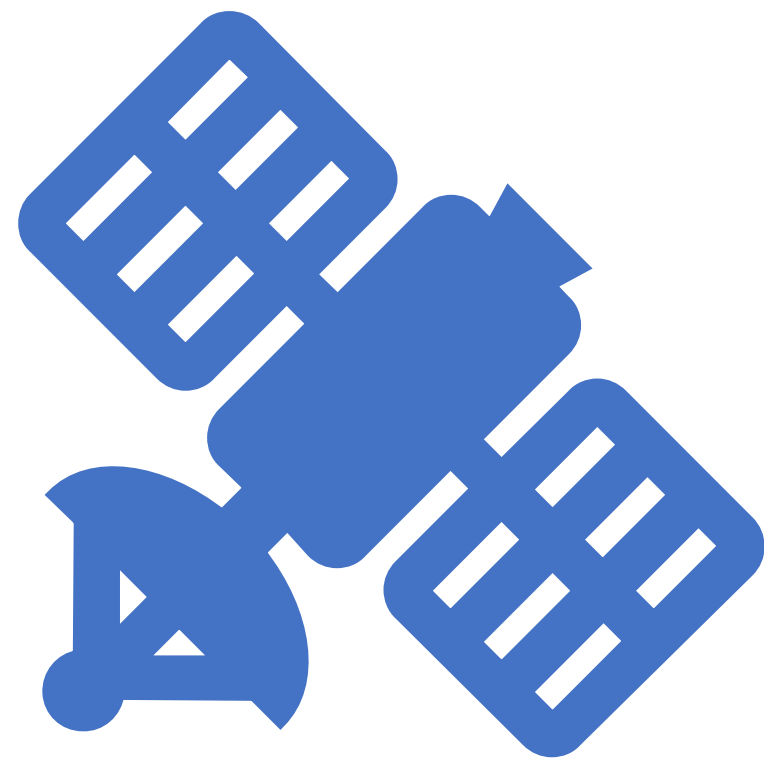
Preparation for the geospatial modelling
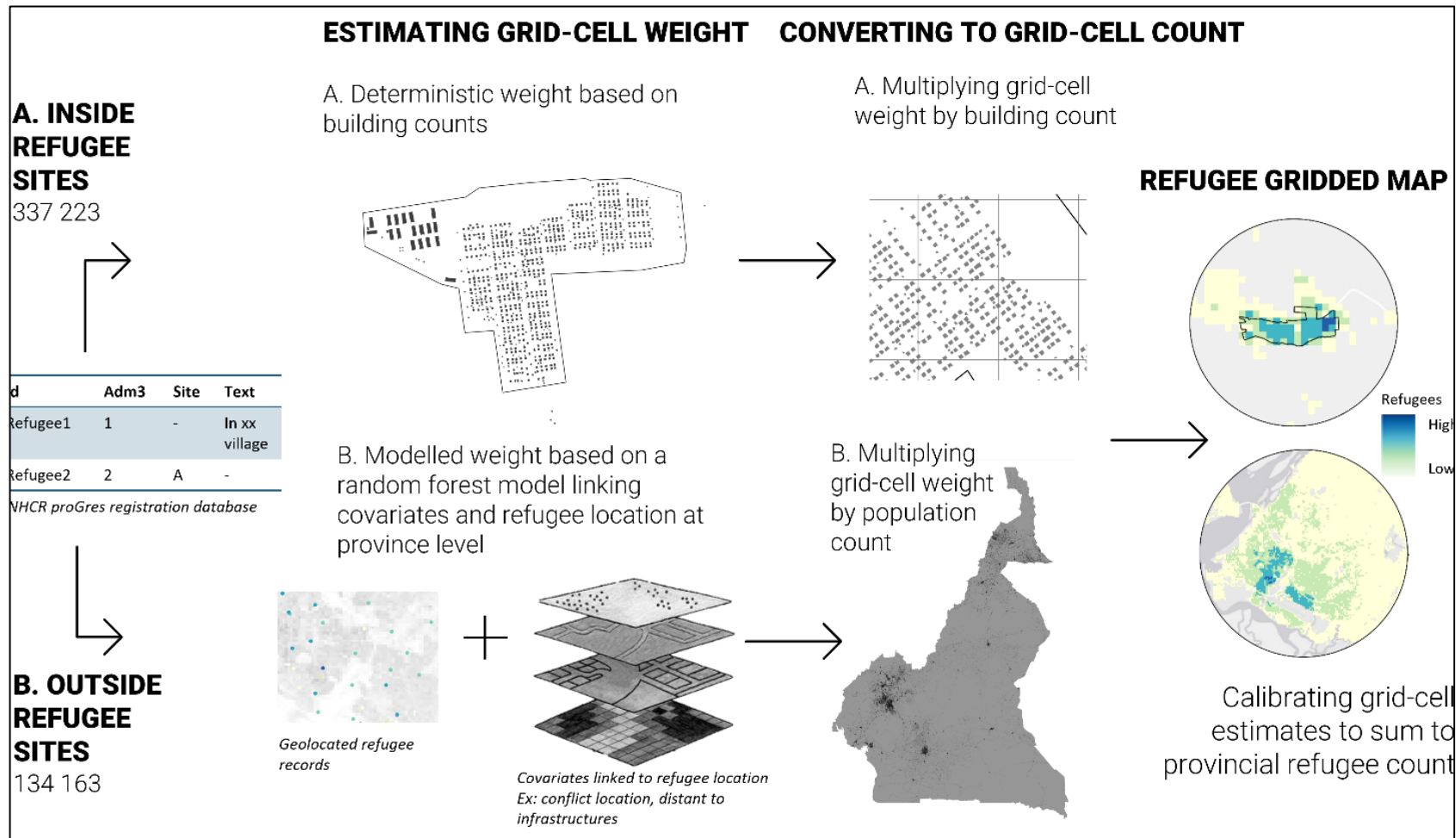
WorldPop

Data input: settlement maps
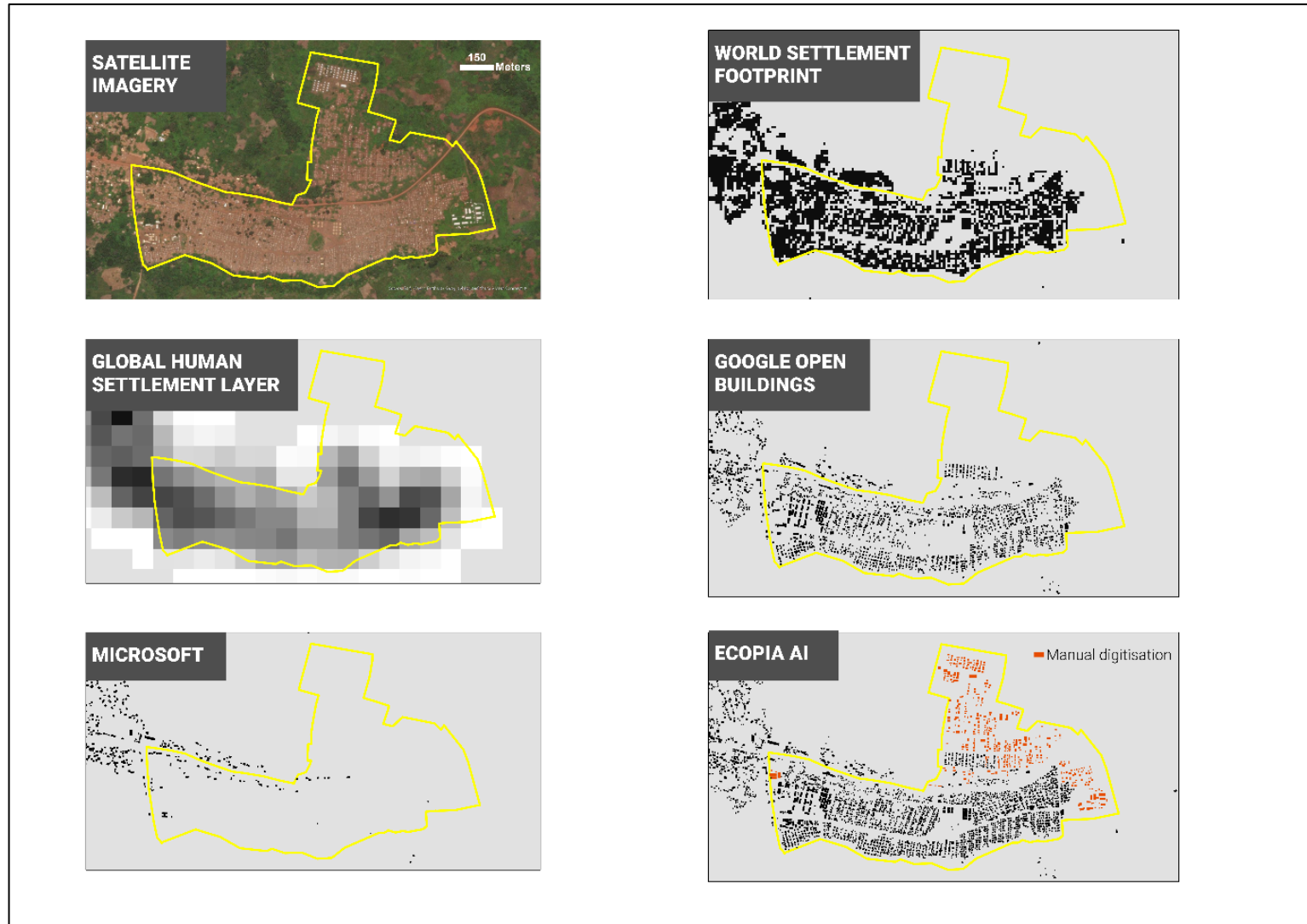
EXAMPLE OF A SETTLEMENT MAP ASSESSMENT IN A UNHCR-LED SITE.

WorldPop

Data input: Covariate preparation

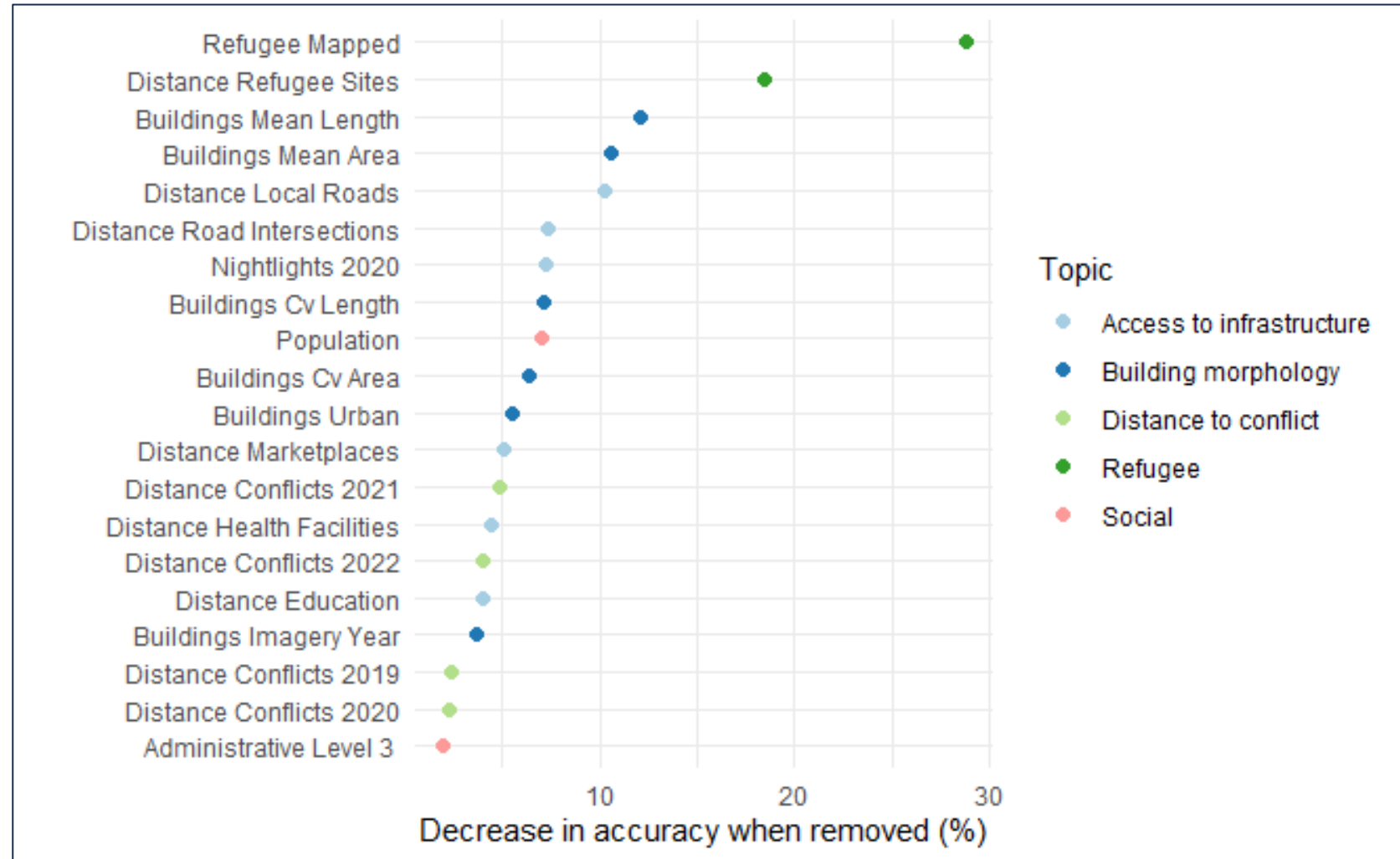GEOLOCATING PROGRES REFUGEE RECORDS

WorldPop

Modelling pipeline

GRID-BASED MAPPING OF REFUGEES: A TWO-STAGE WORKFLOW

WorldPop

Darin et al 2024

Outcome: covariates importance

COVARIATE IMPORTANCE IN MODELLING REFUGEES. CV STANDS FOR COEFFICIENT OF VARIATION.

WorldPop

Darin et al 2024

Outcome: fine-resolution refugee map

COMPARISON OF THE HIGH-RESOLUTION MAPPING WITH UNHCR CONVENTIONAL REFUGEE MAPPING REPORT AND UNHCR INTERACTIVE MAPPING
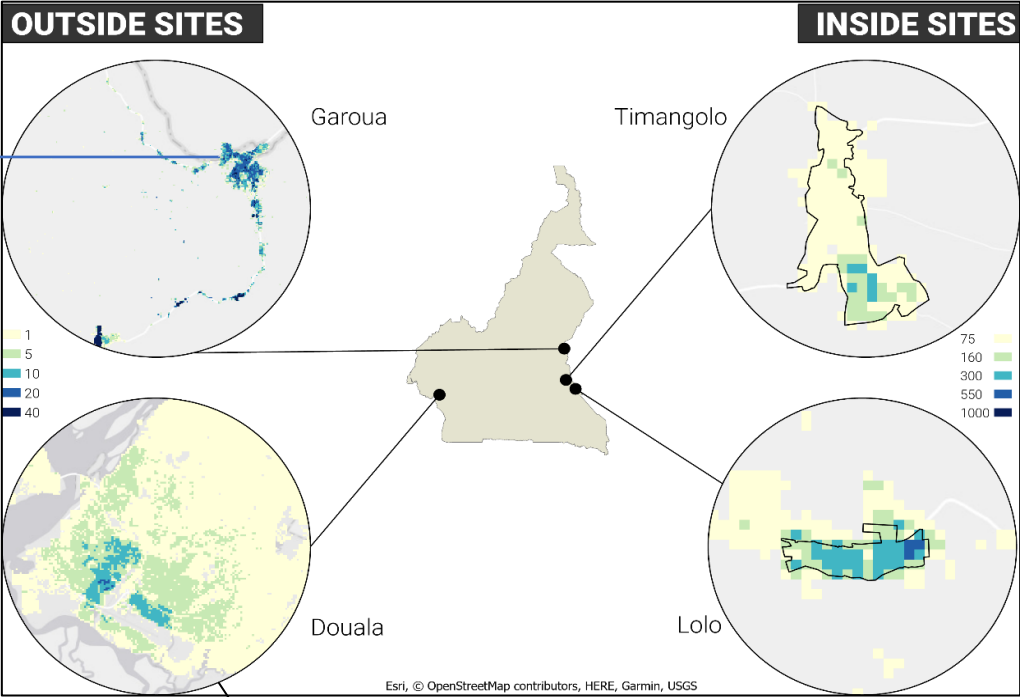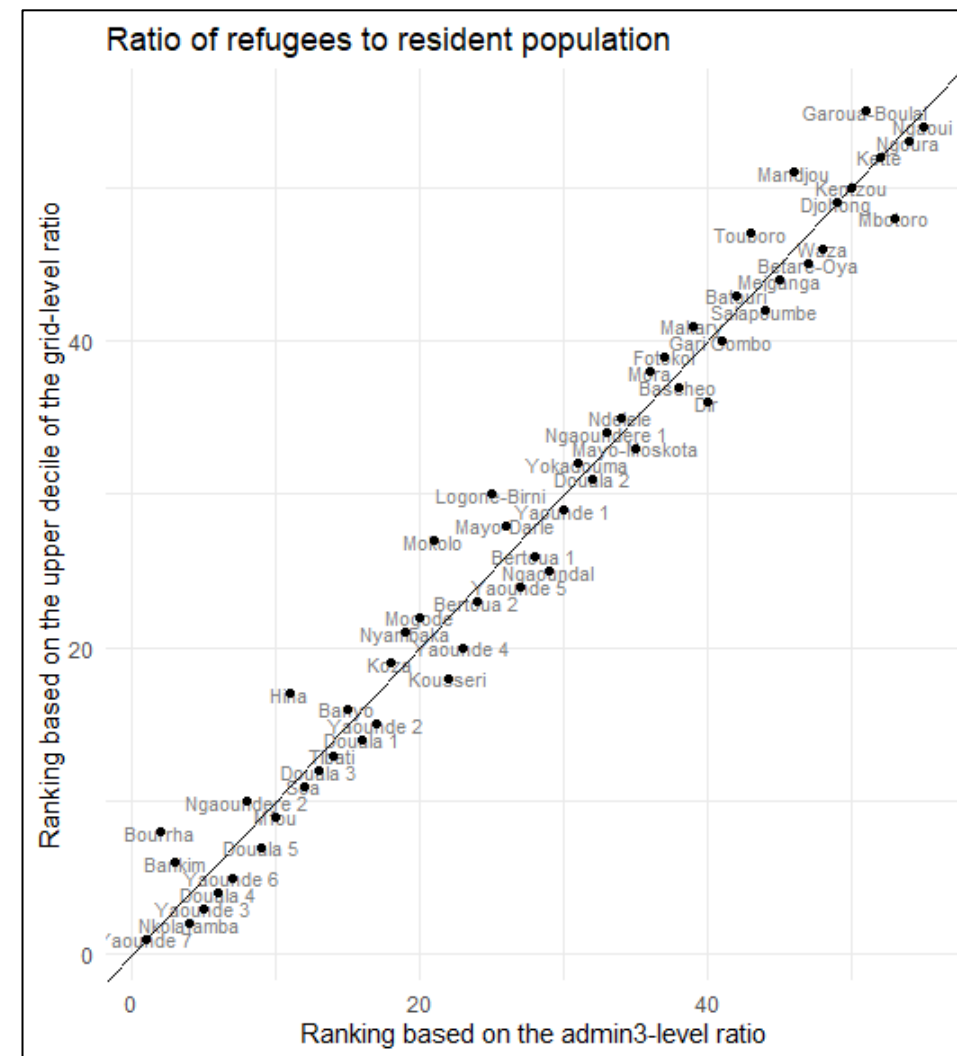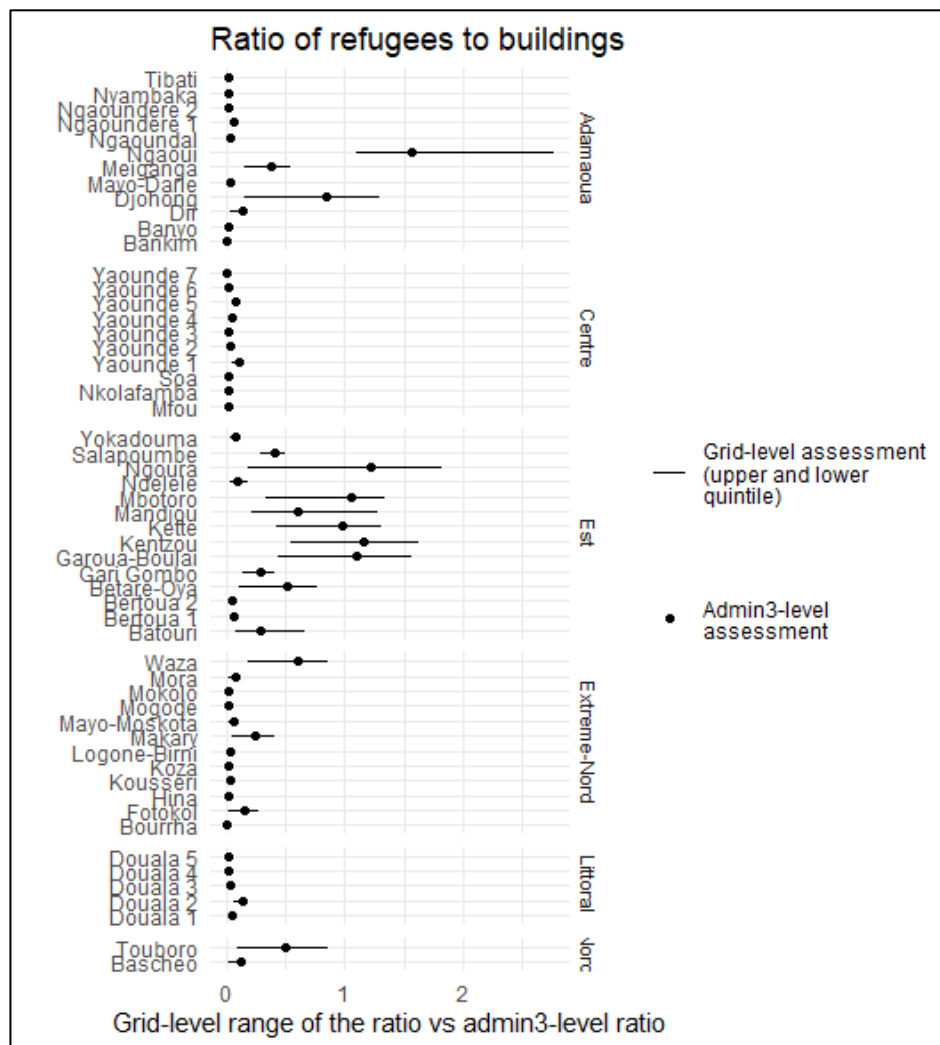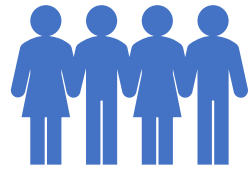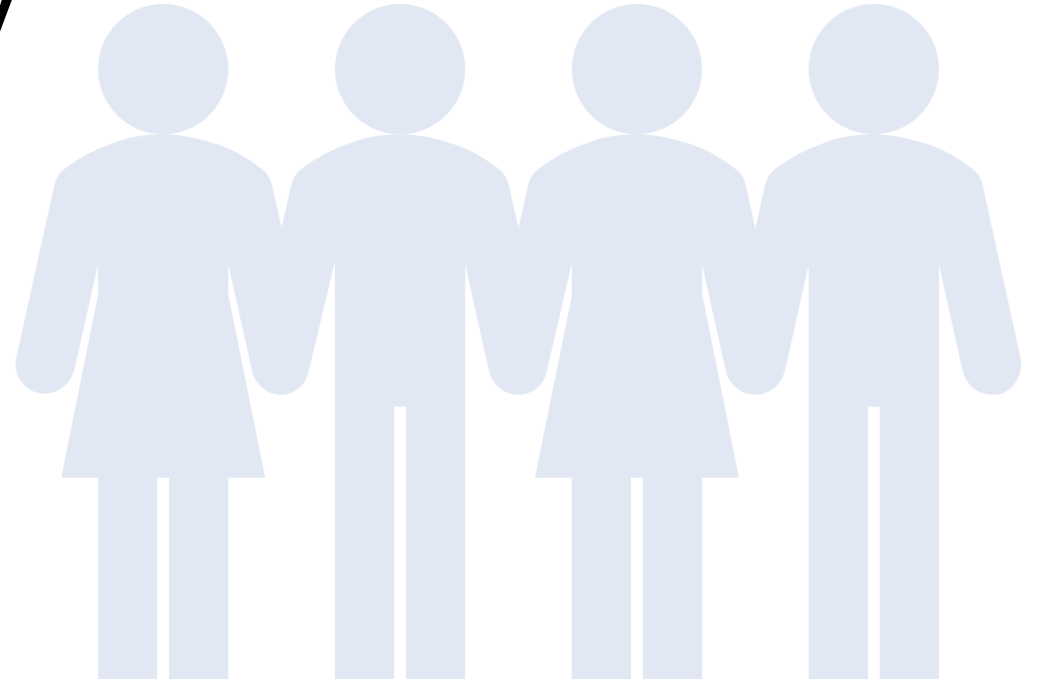

Dasymetric Mapping Result

Darin et al 2024
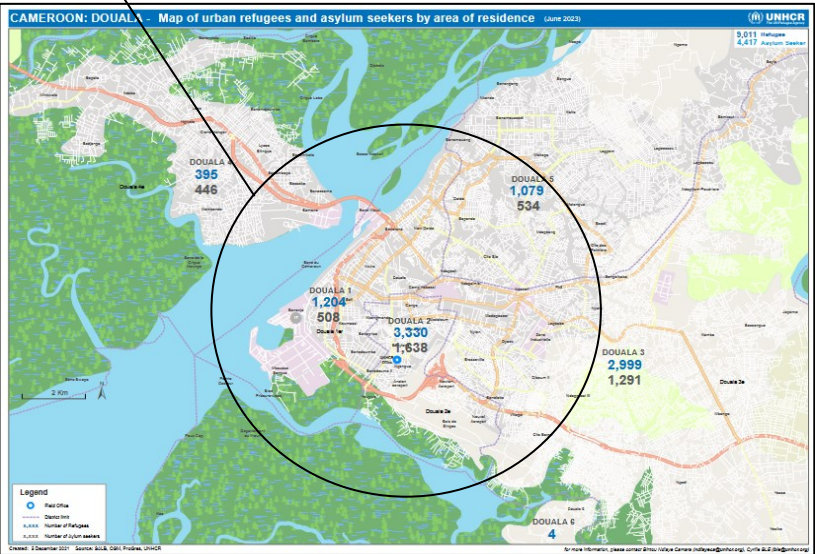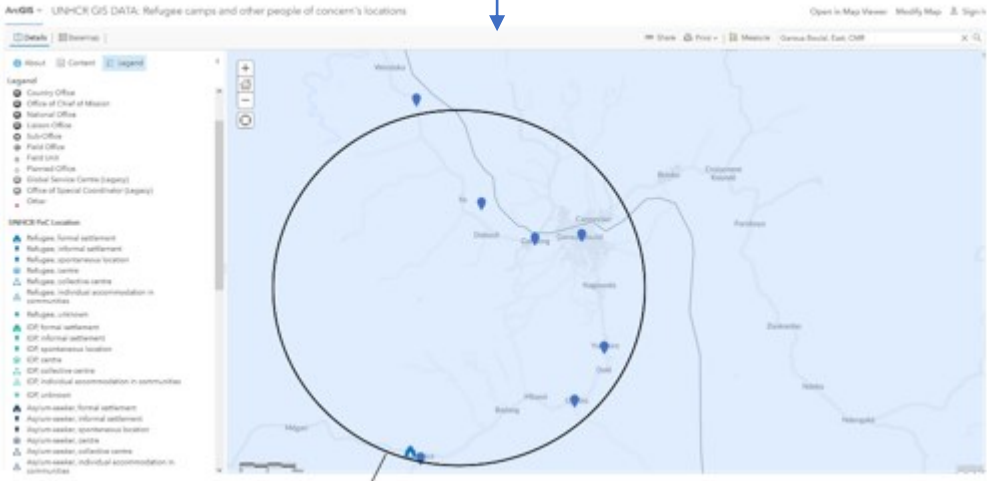
UNHCR Standard Report

WorldPop

What for?

GAIN FOR CRISIS ANALYTICS OF SPATIALLY DISAGGREGATING THE REFUGEE POPULATION

WorldPop

Darin et al 2024

# Mapping Internally Displaced People (IDP) at high resolution

Example of IDP datasets in Nigeria

| No	Internal Displacement Monitoring Centre (IDMC)	International Organisation for Migration-Displacement Tracking System (IOM-DTM)
1	GeoJSON	XLSX
2	142 records	1637 records
3	44 attributes (columns)	109 attributes (columns)
4	Years: 2023, only	Years: 2014-2020, 2021, 2022,2023
5	No sex/age data	No sex/age data
6	Reasons: Flood, Mixed disasters, Non-International armed conflict (NIAC), Other situations of violence (OSV), Rogue Wave, Storm	Reasons: Banditry and Kidnapping, Communal clashes, Herdsmen attack, Insurgency, Natural disaster
7	Origin: file only https://www.internal-displacement.org/database/displacement-data/	Origin: file and API https://dtm.iom.int/datasets

WorldPop

IDMC data visualisation
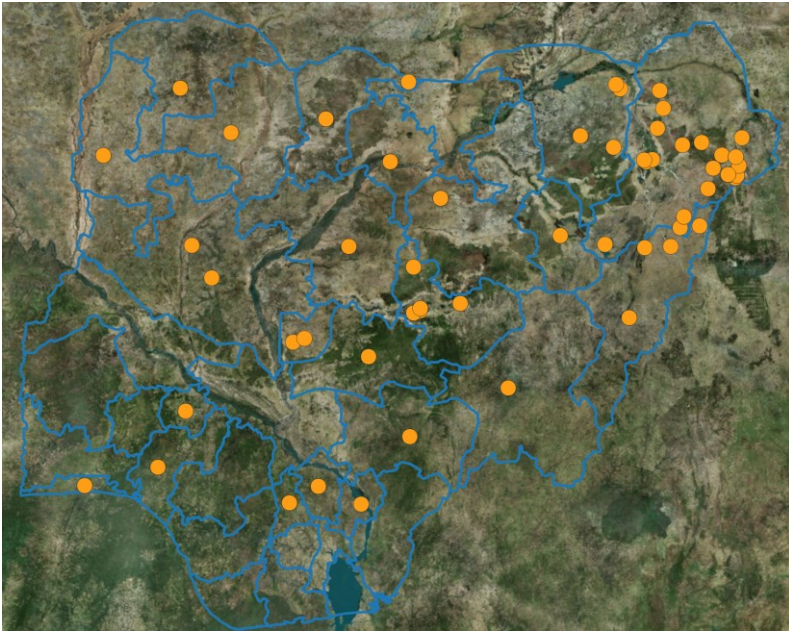
IDMC Raw Data Visualization (point layer)
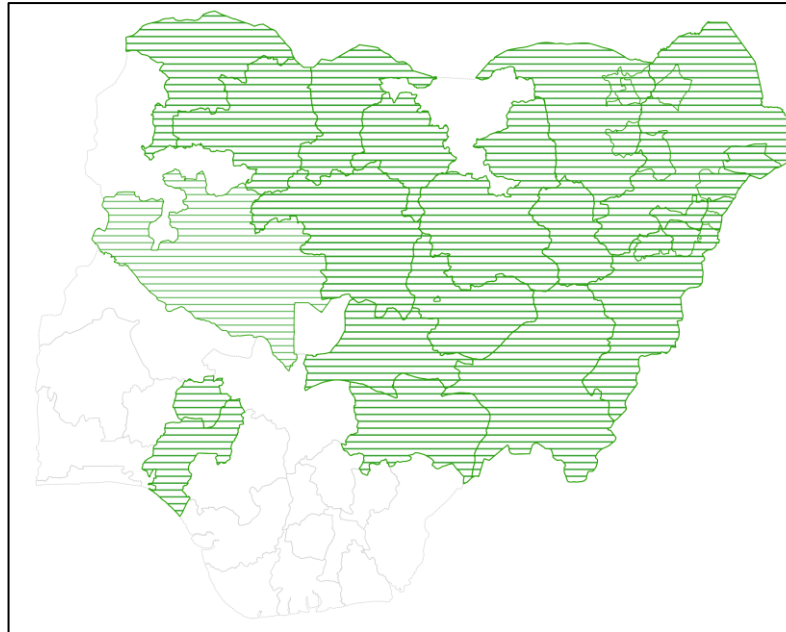
IDMC – Linking to Administrative Boundaries (Origin)

IDMC – Linking to Administrative Boundaries (Destination)

IDMC Orig
IDMC Dest
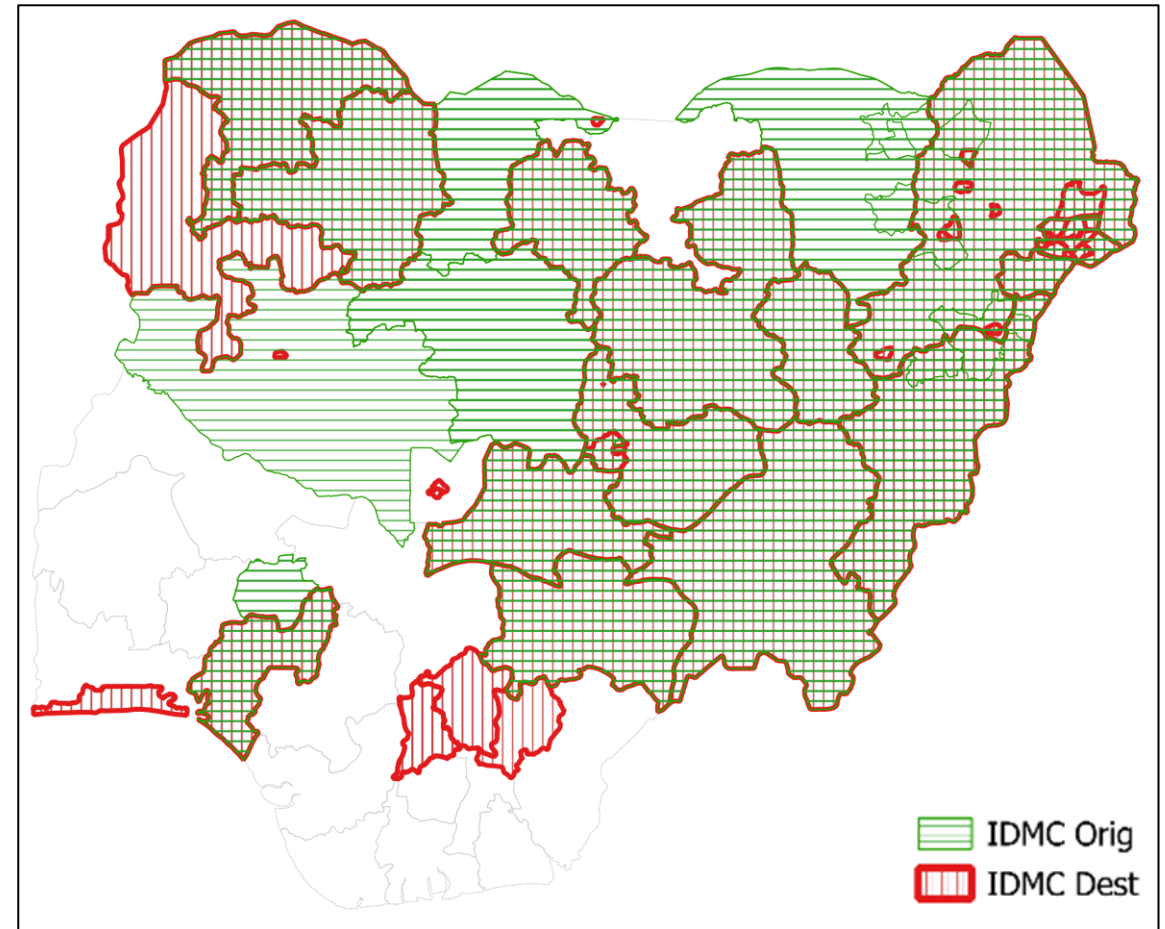
WorldPop
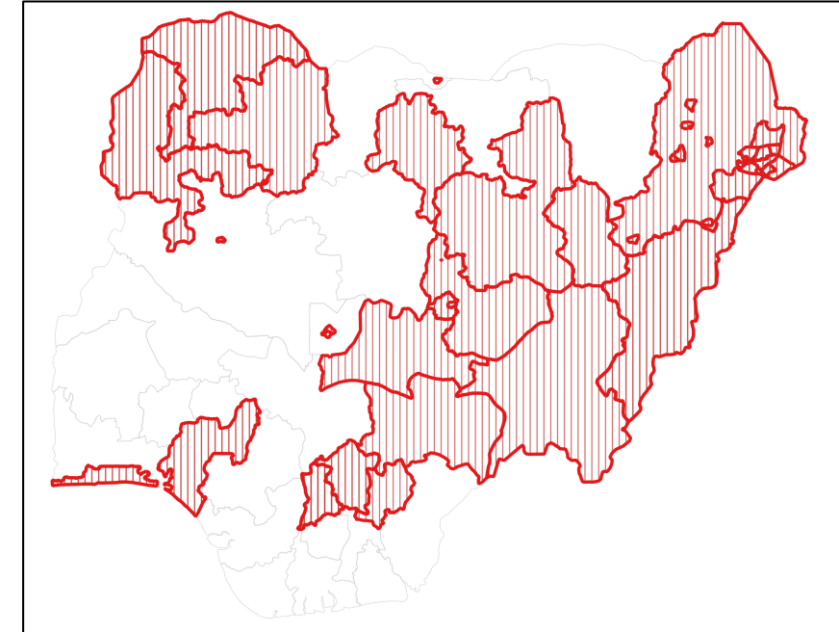
IDMC data Pros & Cons

IDMC – Promising, but low granularity

Pros:
- ❑ Multiple Data Sources
- ❑ Flexible Data Model (Multiple Origins and Destinations)
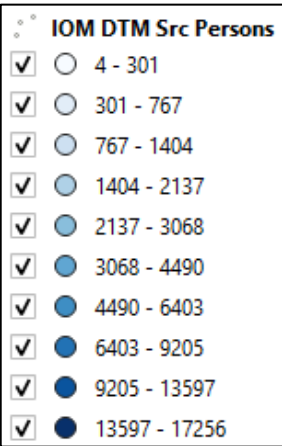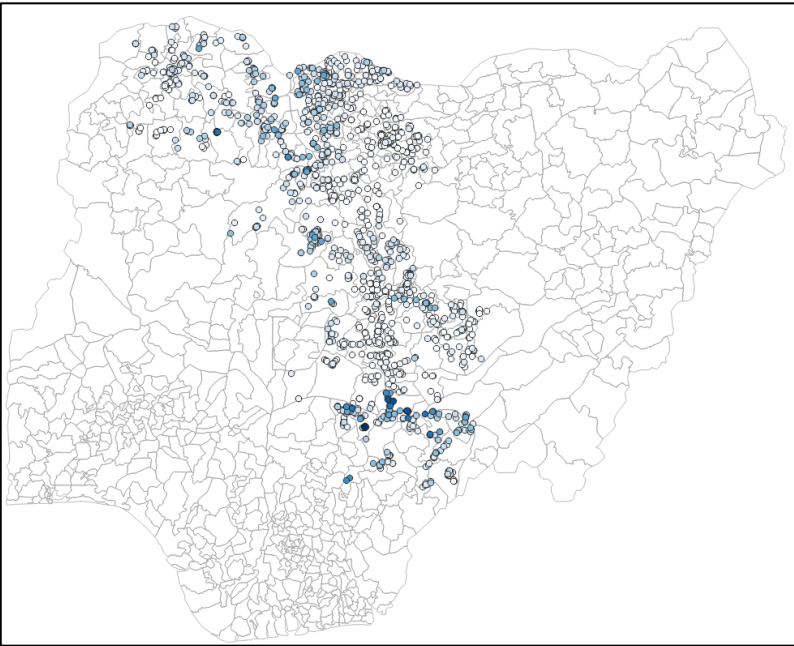- ❑ Origin/Destination Point Data
- ❑ Abundant Attributes

Cons:
- ❑ Low Granularity
- ❑ Limited Timeframe (currently, only 2023 for Nigeria)
- ❑ Inconsistency in Origin/Destination Data
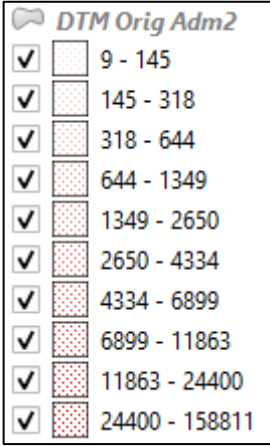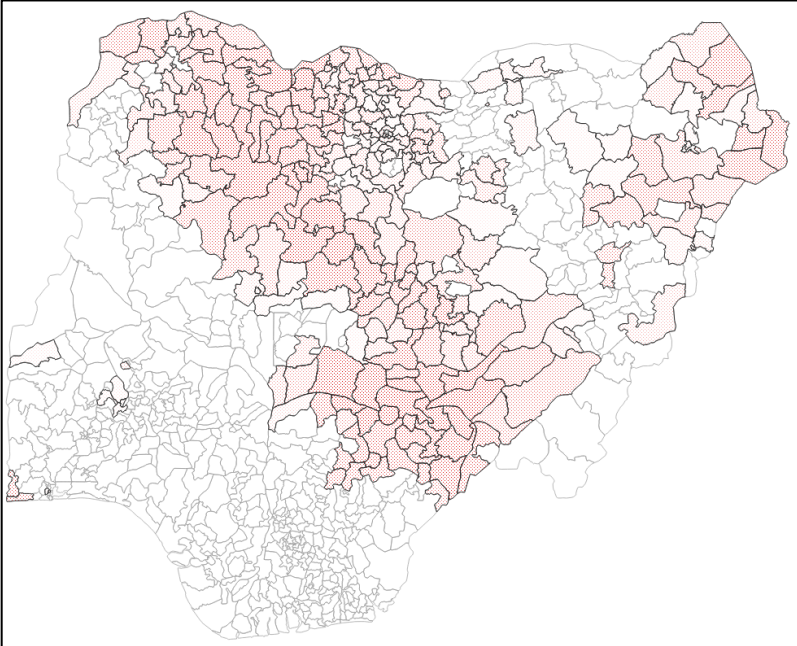- ❑ Complex Data Model

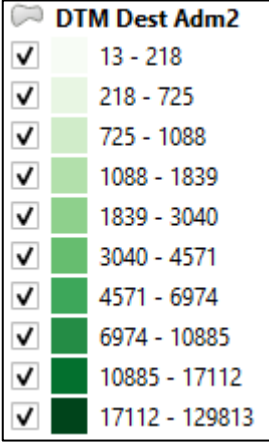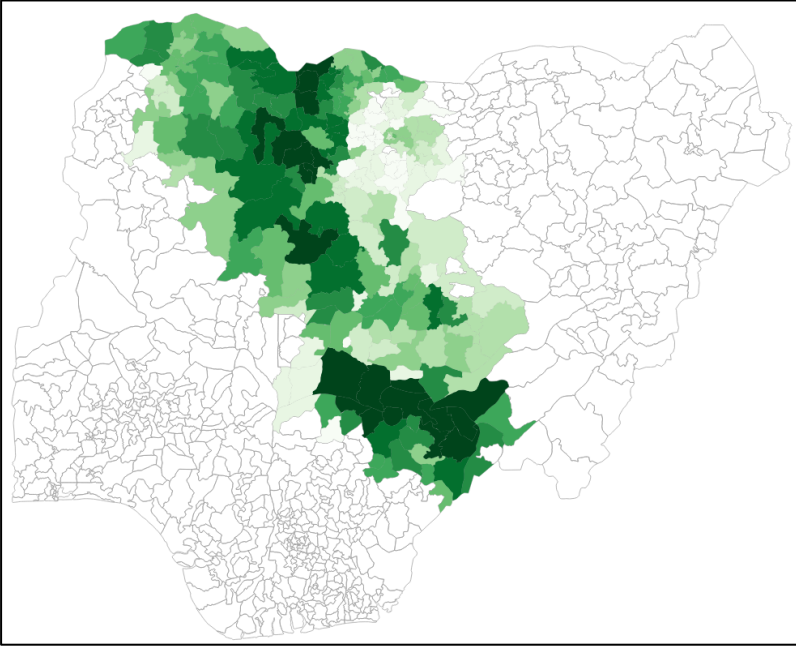WorldPop

IOM-DTM data visualization

DTM Raw Data Visualization (point layer)

DTM – Linking to Administrative Boundaries (Origin)

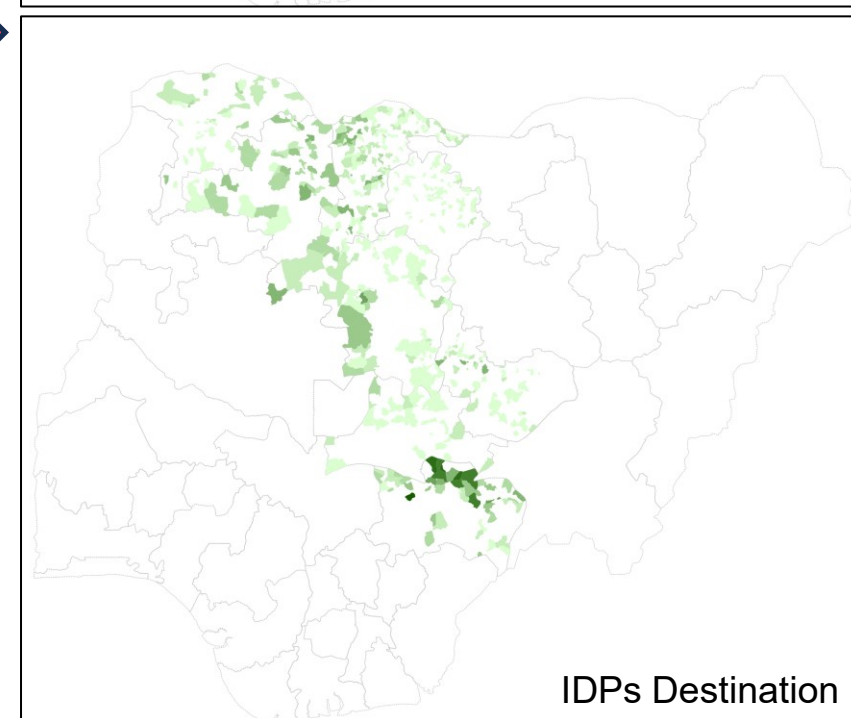DTM – Linking to Administrative Boundaries (Destination)

IOM DTM Src Persons
- 4 - 301
- 301 - 767
- 767 - 1404
- 1404 - 2137
- 2137 - 3068
- 3068 - 4490
- 4490 - 6403
- 6403 - 9205
- 9205 - 13597
- 13597 - 17256

DTM Orig Adm2
- 9 - 145
- 145 - 318
- 318 - 644
- 644 - 1349
- 1349 - 2650
- 2650 - 4334
- 4334 - 6899
- 6899 - 11863
- 11863 - 24400
- 24400 - 158811

DTM Dest Adm2
- 13 - 218
- 218 - 725
- 725 - 1088
- 1088 - 1839
- 1839 - 3040
- 3040 - 4571
- 4571 - 6974
- 6974 - 10885
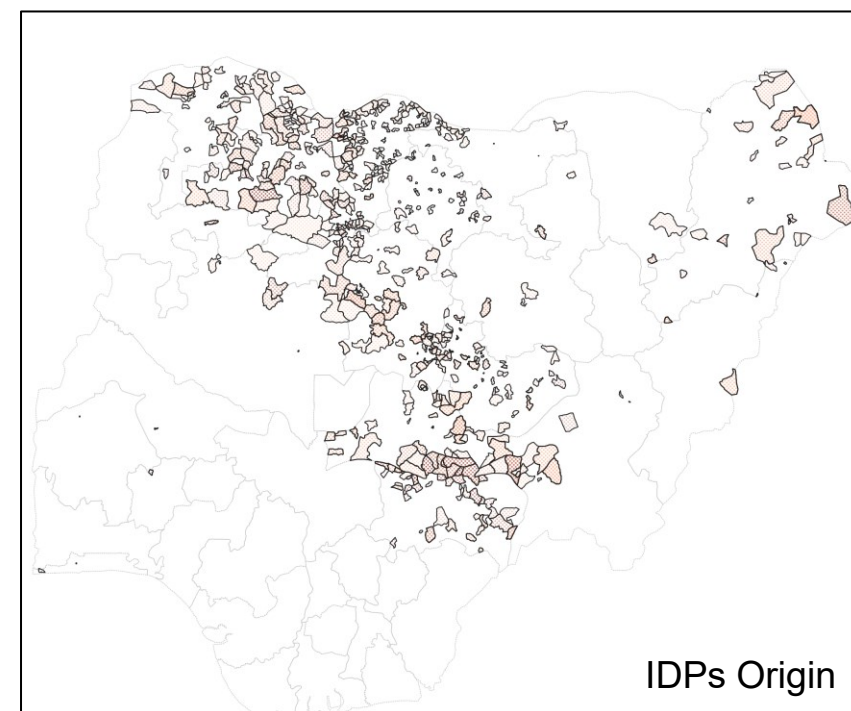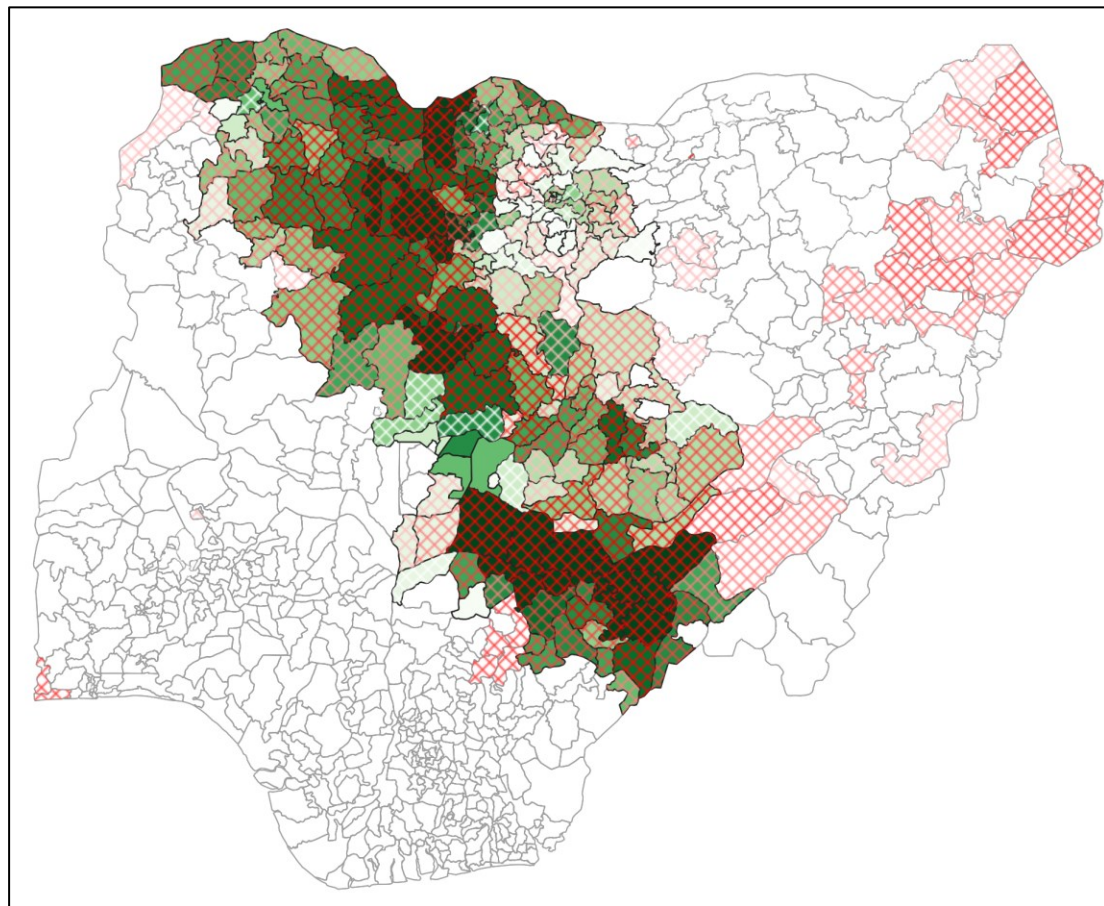- 10885 - 17112
- 17112 - 129813

WorldPop

Improve IOM-DTM granularity

IOM DTM State/LGA/Ward Names – GRID3 Admin 1/2/3 Polygons

DTM Roun	Date of Ass	KI's	Population Types	Region	State	S	LGA	I	Ward	\S	Site Name
R12	45215	3	IDPs dispersed in North cent	Benue	N	Logo	I	MBAGBER	I	EI	kyochi
R12	45210	2	IDPs dispersed in North cent	Benue	N	Makurdi	I	NORTH BANK I	I	E	Akuundu Ityough
R12	45213	3	IDPs dispersed in North cent	Benue	N	Kwande	I	YAAAV		I	ETuran Community Sec Sch
R12	45213	4	IDPs dispersed in North cent	Benue	N	Kwande	I	YAAAV	I	E	Enia Pri Sch
R12	45211	4	IDPs dispersed in North cent	Benue	N	Kwande	I	MBAIKYOR	I	E	Rcm Prim Sch Nyihemba

wardname	▲	rdco	lganame	acod	statename
1004 / Aboyade		L...	Eti Osa	2...	Lagos
A Ozizor		A...	Ogbaru	4...	Anambra
Aagba		O...	Boripe	3...	Osun

1. State DTM == State Adm1 -> define the state ***polygon A*** from Adm 1,

2. LGA == LGA Adm1 and LGA within the ***polygon A*** -> define the LGA ***polygon B*** from Adm 2,

3. Select all wards from Adm 3 within the ***polygon B***, find:

   o a ward with an equal name,

   o If no, remove all non-letter symbols and convert to the lower case,

   o check for equal names (lower case) or the names with the Levenshtein distance < 4 and <30% of the string length.
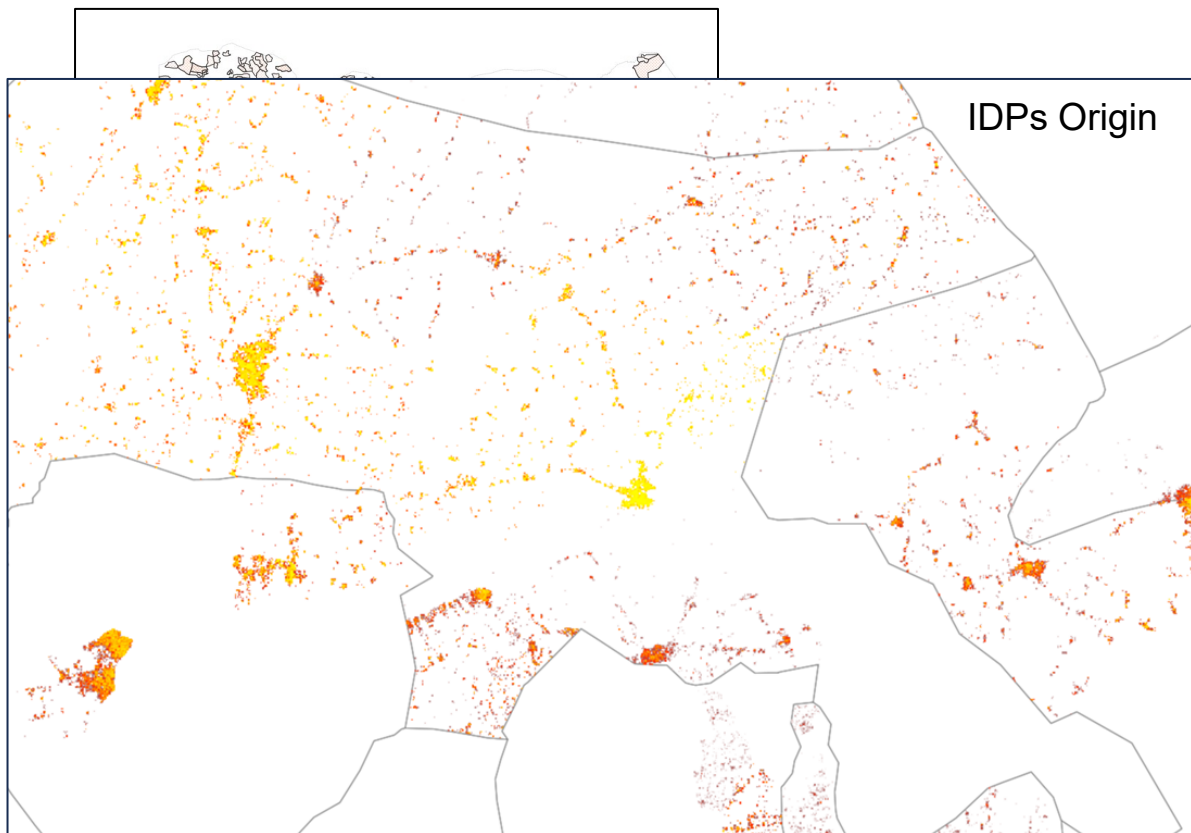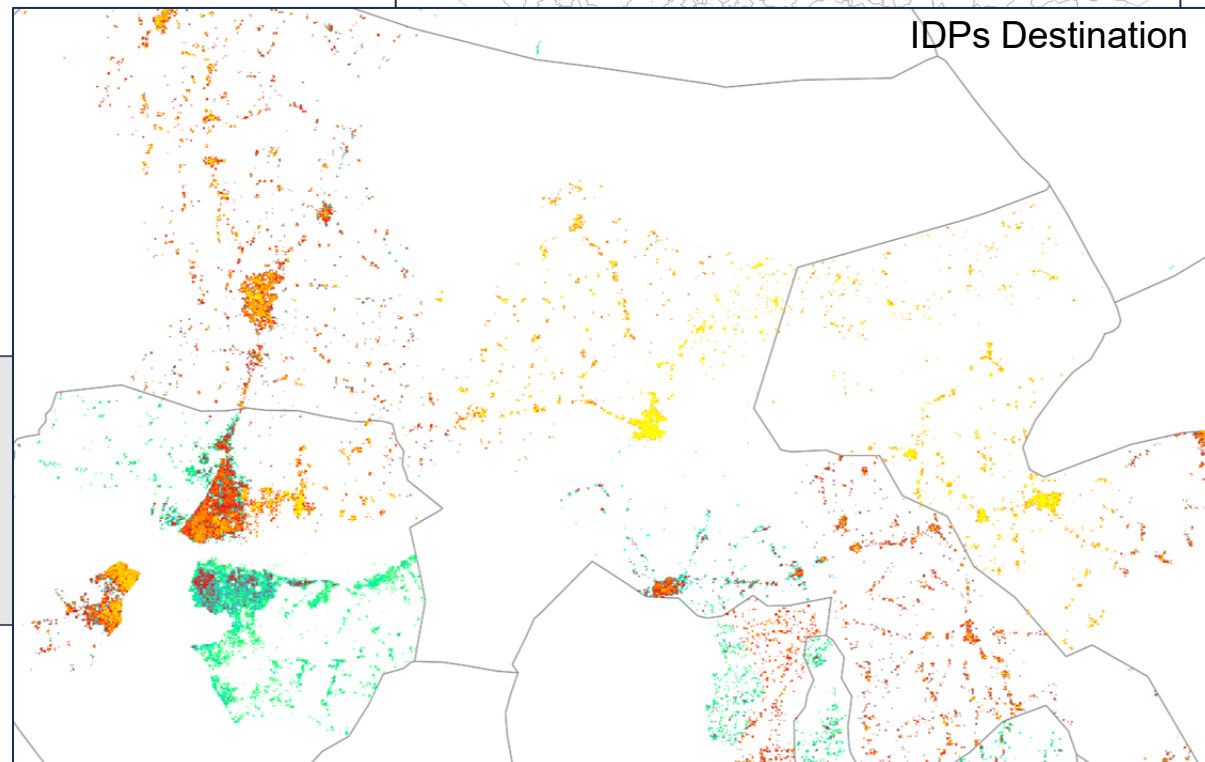
WorldPop

Improve IOM-DTM granularity

Improve Granularity

IDPs Origin

IDPs Destination

WorldPop

Geospatial Modelling

WorldPop

Gridded IDP output

IDP Data
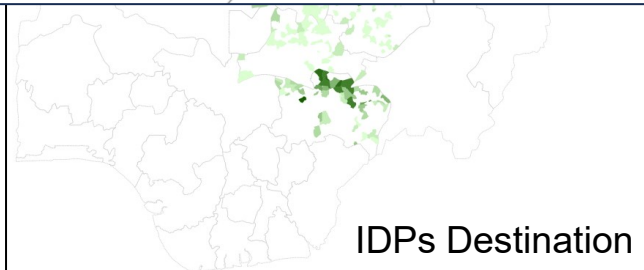
Disaggregation: IDPs (100m)

IDPs Origin
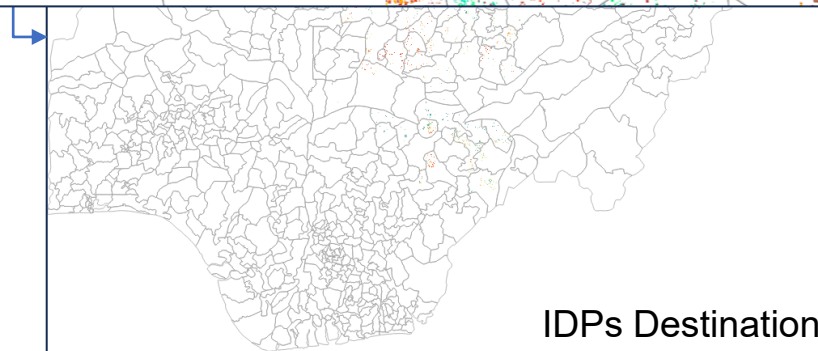
IDPs Destination

IDPs Destination

IDPs Destination

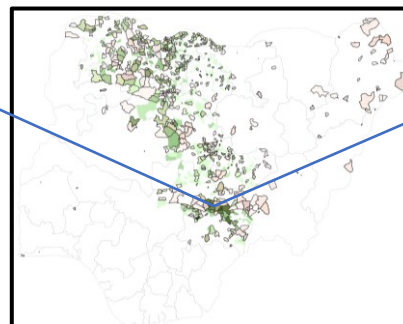**WorldPop**

Gridded IDP output

IDPs Origin
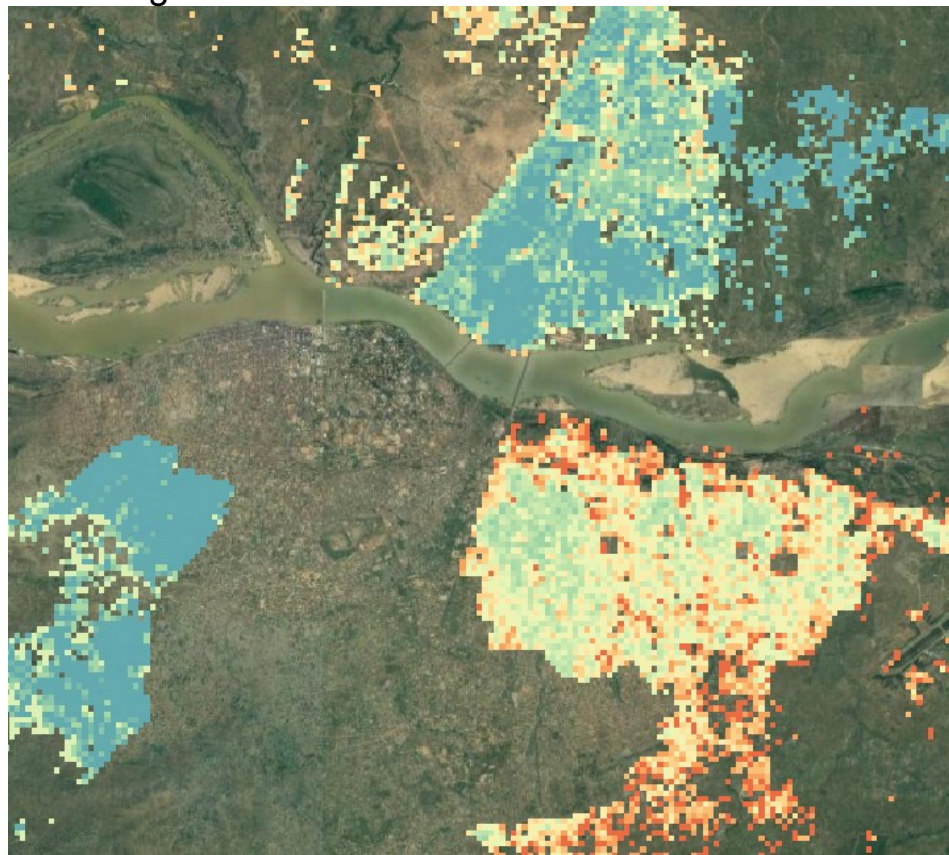
IDPs Destination

IDPs Origin and Destination

WorldPop

Why is an IDP map at high resolution necessary?

Our modelling output includes high-resolution IDP maps, which provide detailed insights into the distribution of displaced populations. For example, by overlaying satellite images, we can zoom in on specific areas to demonstrate the improved accuracy and granularity of our data.

High-resolution IDP maps offer several benefits:

❑ Flexibility to aggregate data to any geographic boundary.

❑ Easier identification of IDP locations.

❑ Enhanced utility for future national household surveys on IDPs.

❑ Improved resource allocation and policy planning.

❑ The IDP estimates can be break down by age and sex

❑ Enhances outreach to a particular age and gender group (Maximise inclusion)

WorldPop

Future Research
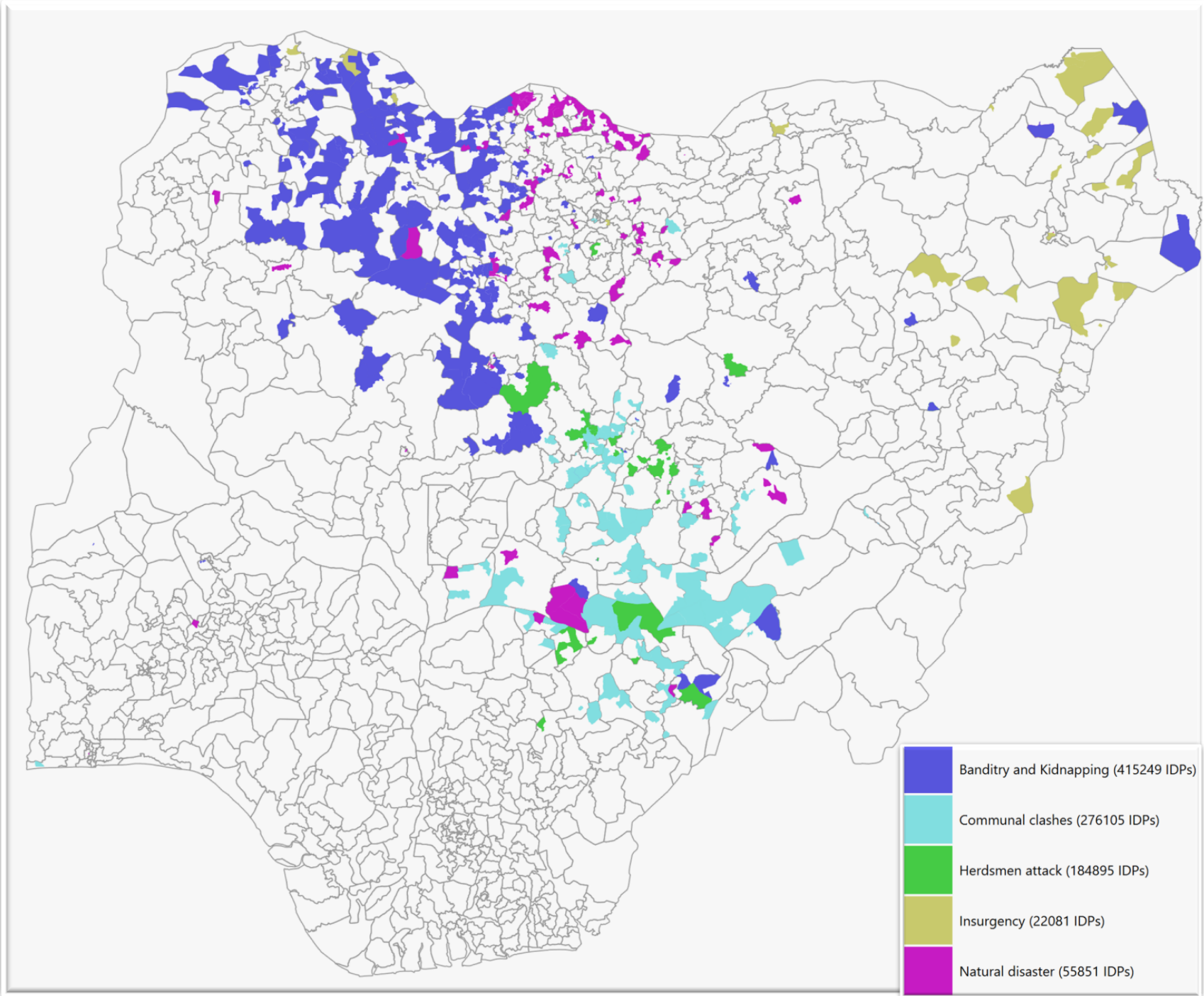
WorldPop

Automatic preEA boundary workshops and trainings
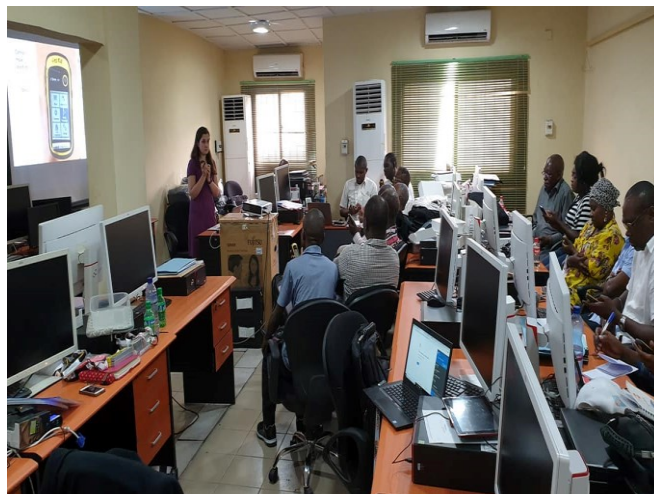
Thank You

Contact: Sarchil Qader

Email: S.Qader@soton.ac.uk

Twitter: @SarchilQ

WorldPop: @WorldPopProject

https://www.worldpop.org/

WorldPop