

Spatial approach to probability sampling: Lessons from PMA Kaduna State GPS mapping exercise

Elizabeth Omoluabi, Samson Adamu, Julien Nobili, Blake Zachary, Temitope Erinfolami

elizomoluabi@gmail.com



Goal

Compare micro-censuses conducted by PMA Kaduna State to gridded population dataset estimates of the same polygons





Potential benefits of these validation tests?

If the comparison gives satisfactory results:

- Gridded population datasets can replace the outdated census population estimates
- The cluster selection process can easily be automated, reducing error and saving cost and time
- Allows autonomy in the selection process and weighting of samples
- Clusters drawn from a grid dataset will be more accurate and have geo-referenced boundaries • It will enable excellent QA of cluster identification, listing, and confirmation of sample randomness

Estimation Method

1. Digitize the census 3. Extract the boundaries population obtained from estimate from the gridded population paper maps. 22 datasets. 2- Listing of households and PM 2. Collect GPS coordinates of the 4. Compare the survey EA 01 results with the boundary and 22 modeled estimates. each residential

- Extraction of the population count in each EA from th Pop = 290 4 - Population could

What are the potential risks?

• PMA results are no longer comparable to those of previous census based rounds We won't know until we try....

Gridded Population Datasets

Estimated population density applied to a map grid

The map is independent of administrative boundaries

Limitations: Sources: • Need for updating Worldpop • Some outliers with • GeoPoDe large variance • Landscan



Cluster Mapping and Listing Method: Example

Round 4 A boundary and listing geo-points

structure.



Verified Round 4 A cluster and Round 4B clusters.



Before: The Resident Enumerator did not list all the (structures) within the EA.



After: All structures and households previously missed are now listed.







Red polygon: round 4 a (initial boundary) cluster Blue polygon: round 4 b (verified boundary) cluster **Green pins**: boundary verification geo-points recorded by resident enumerator (re)

polygon: round 4 cluster. e pins: household listing geo-points

Yellow polygon: round 5 cluster and EAs **black** oval polygons: results before and after re-listing

Population Count Extraction Method: GeoPoDe

GeoPoDe - an open-source, public data repository for geospatial reference datasets funded through the Bill and Melinda Gates Foundation.

In ArcGIS, export the cluster polygons to KML format



Import the KML file to GeoPoDe, select and save the



Results: Total Population

PMA Kaduna micro-census compared to Landscan, WorldPop and GeoPoDe

PMA	Landscan	WorldPop	Geopode
Round 4 Total pop	+14%	-226%	+35%
Round 4 Rural	+15%	-808%	+33%
Round 4 Urban	+13%	-146%	+37%
Round 5 Total pop	-32%	-260%	+1%
Round 5 Rural	-30%	-1685%	-2%
Round 5 Urban	-34%	-165%	+2%



Results: Summary

Round 4

- PMA population data for Round 4 showed gaps in listing (about 20% estimated omission in the 60 clusters Using gridded population maps for population estimates is feasible and has been validated for Kaduna State in our study. assessed)
- Worldpop's population estimates: **226%** lower than PMA
- LandScan's population estimates: 14% higher than PMA
- Geopode's population estimates: **34%** higher than PMA

Round 5

- PMA population data for Round 5 was carefully collected, listing and randomness verified using GIS
- Worldpop's population estimates: **1685%** lower than PMA
- LandScan's population estimates: **32%** lower than the PMA
- Geopode's population estimates: 1% percent higher than PMA





Conclusion

- This has provided a critical mass of data to enable PMA Nigeria to make an evidence-based decision on a sampling strategy for the future.

What next?

• We have used gridded dataset for probability sampling in more than 3 other surveys