Supporting Countries in Estimating Population and Demographic Dynamics

Dr Natalia Tejedor Garavito, 20 August 2025

Transcript

Thank you for the invitation. I'm going to talk a little about the projects we work on and about what we do in the WorldPop team and about the projects we carry out supporting countries in population estimation and demographic dynamics.

At WorldPop we built an applied research and implementation team. Normally, as Andrea mentioned, we have open data, and we try to publish all our methodologies. We do a lot of co-development and capacity building in the countries we work with. We produce dynamic demographic maps of small areas in low- and middle-income countries, and our data are always used by agencies like the UN and many others. Our projects are also based on census support, reproductive, maternal, neonatal, child and adolescent health, as well as child vaccinations and epidemiology.

The basis of all our projects is demographic data. However, demographic data are very challenging to identify many are low-resolution, outdated, incomplete, imprecise, or missing. Today, however, we have many high-resolution geospatial datasets, such as building footprints, nighttime lights, mobility data from Facebook or Google, and also urban expansion data like this one here from GHSL. All these spatial data help us infer information for models to estimate population.

What we do is take all these demographic covariates and put them together, so they help us infer data in places with no information. For our spatial models, we use some population data - later I'll explain which types we use. Some of these data are censuses, surveys, micro censuses, and settlement maps where we expect people to live, combined with geospatial covariates. All of these are put together in statistical models to estimate population, ideally, always with some degree of uncertainty at the grid level.

We produce grids, normally 100m x 100m, where an image of an area is divided into grids representing the estimated number of people living there. You'll notice that the numbers are decimal. We are working on maps with integers, but so far we can only redistribute the total population across the grids as probabilities and estimates.

Why are grids useful? Because they allow us to later summarize this information using different polygons, for example, administrative areas, health zones, enumeration areas, settlement extents, or circles around health centres or conflict sites. This flexibility makes grids very important.

At WorldPop, we have two types of models: a "top-down" model and a "bottom-up" model. The top-down model is used when we have complete national data. The bottom-up model is used in countries without recent censuses, for example since 1984, where only small-scale surveys exist, or where conflict prevented surveys. We base estimates on reliable survey data and project to areas without data. For the top-down model, we take administrative areas, use covariates, feed them into a Random Forest model, and disaggregate data across 100m x 100m grids.

We are about to release new data covering 2015 to 2030, globally, by country, and aggregated worldwide, at 1 km resolution, disaggregated by age and sex, including population density. These data are based on subnational censuses structured by age and sex and on projections. Some countries have two census points—for example, 2010 and 2020. We also share 73 covariate datasets: topography, climate, night lights, land cover, inland water, infrastructure, protected areas, built environment, etc. All aligned to 100m grids, freely available. We also model urban expansion, for example in Iraq, looking at growth in surface, height, and volume. Combined with censuses and covariates, we produce global gridded datasets in five-year intervals, by age and sex. These will be on Living Atlas, Google Earth Engine, and HDX.

We are working on integer-based datasets, as previous models used decimals. Our data are used by many organizations and portals: DHS2, FAO, national portals like Sierra Leone's infrastructure planning, etc. We also provide tutorials and co-development examples, e.g., with Brazil.

Many countries have outdated censuses - Congo's last was in 1987, Afghanistan's in 1979. For these, we use bottom-up methods, combining surveys or micro censuses to estimate populations across grids. For example, in Afghanistan, local censuses exist in some areas, and we project from there.

In Colombia, we supported the hybrid census (2018–2019), helping with undercounting in conflict or remote areas using Bayesian hierarchical inference models, workshops with Indigenous communities, and satellite imagery of built-up areas.

These top-down and bottom-up datasets have many uses: census planning in Benin, official statistics in South Sudan, Nigeria, Congo, vaccination strategies in hard-to-reach places, school planning in Sierra Leone, exposure to air pollution, disaster preparedness, etc. They have also been used in humanitarian crises: recalculating Ukraine's population after the war began, measuring exposure to conflict via ACLED data, and estimating populations within 1–5 km of conflict sites.

We study mobile populations: refugees, migrants, seasonal migration, using phone data, Google locations, and night lights. For example, Namibia (mobile phone data), Nigeria (border lights), and nomadic groups crossing artificial borders. These data inform vaccination and health services.

We also analyse disaster response movements: Nepal 2015 earthquake, Hurricane Matthew 2016, Haiti earthquake, Cyclone Mahasen in Bangladesh - using mobile data to see where people moved, to direct aid efficiently. We map refugee and displacement populations, for example in Cameroon and Nigeria, using settlement data and Random Forest disaggregation to estimate populations in camps and track migration flows.

We provide open tools: QGIS plugins to make gridded population estimates, tutorials on YouTube, and sampling tools for surveys and censuses. These help plan enumeration, logistics, and microplanning using population, settlements, roads, slopes, and land use.

We also build health indicators from population data and surveys and focus on co-development and training, so countries feel ownership of the data.

I think I'm finishing now. Our data are open and will soon be updated on the Humanitarian Data Exchange and our GitHub. This is not just my work, but the work of the whole WorldPop team.

And of course - thank you.