Review

Small area population denominators for improved disease surveillance and response

A.J. Tatem

WorldPop, School of Geography and Environmental Science, University of Southampton, UK

ARTICLE INFO

ABSTRACT

The Covid-19 pandemic has highlighted the value of strong surveillance systems in supporting our abilities to respond rapidly and effectively in mitigating the impacts of infectious diseases. A cornerstone of such systems is basic subnational scale data on populations and their demographics, which enable the scale of outbreaks to be assessed, risk to specific groups to be determined and appropriate interventions to be designed. Ongoing weaknesses and gaps in such data have however been highlighted by the pandemic. These can include outdated or inaccurate census data and a lack of administrative and registry systems to update numbers, particularly in low and middle income settings. Efforts to design and implement globally consistent geospatial modelling methods for the production of small area demographic data that can be flexibly integrated into health-focussed surveillance and information systems have been made, but these often remain based on outdated population data or uncertain projections. In recent years, efforts have been made to capitalise on advances in computing power, satellite imagery and new forms of digital data to construct methods for estimating small area population distributions across national and regional scales in the absence of full enumeration. These are starting to be used to complement more traditional data collection approaches, especially in the delivery of health interventions, but barriers remain to their widespread adoption and use in disease surveillance and response. Here an overview of these approaches is presented, together with discussion of future directions and needs.

1. The value of small area demographic data for effective disease surveillance

The Covid-19 pandemic has shone a spotlight on the importance of comprehensive, timely and accurate health surveillance and information systems for mitigating the impact of infectious diseases. Integrated systems that brought together timely data on cases, healthcare utilisation and deaths at small area scales, together with reliable data on underlying demographics, enabled rapid and accurate identification of outbreaks, key risk groups, scales of transmission and spread routes. These all facilitated the appropriate design of interventions and mitigation of impacts. Building such systems in the midst of a pandemic is challenging though and requires building upon existing routine systems and data collection. The Covid-19 pandemic exposed how many countries did not have such strong systems and underlying data in place (Aborode et al., 2021; Ibrahim, 2020).

Key challenges in disease surveillance and the achievement of 'precision' in public health, include the registration of births and deaths and tracking disease (Dowell et al., 2016). These are impacted in multiple ways by weaknesses in underlying demographic data. In terms of birth and death registration, it is hard to know whether a national deworming programme for children in one country or a vaccination programme for

pertussis in another is reducing mortality when less than 5 % of deaths are registered. However, even if 100 % of deaths are registered, it remains challenging to both implement the programmes and place the numbers of deaths in context without reliable multi-temporal, disaggregated data on population numbers and distributions, particularly when seasonal dynamics are strong and highly mobile population groups exist (Bharti et al., 2011; Buckee et al., 2017; Wesolowski et al., 2017). Careful surveillance can guide public health in a country and track disease outbreaks that could spread beyond borders (Dowell et al., 2016). Improving detection and measurement of the numerator without attention to the denominator however risks providing an inaccurate picture. Analyses in Namibia showed that improved quantification of denominator populations changed malaria incidence measures by more than 30 % (Zu Erbach-Schoenberg et al., 2016). Moreover, by pairing just a small number of physical autopsies with verbal autopsies on the same deaths, the much larger number of verbal autopsies can be calibrated - but the verbal autopsy data are often drawn from surveys built on static and outdated sample frames (Carr-Hill, 2013; Thomson et al., 2020) and again it remains challenging to place outputs in context in settings where denominators are uncertain and populations are mobile. The reliance on static and aging figures for denominators leads to the common occurrence of 200 % vaccination rates, or incidence measures

fluctuating by season where population mobility is high (Cutts et al., 2021).

The collection and maintenance of timely and accurate small area data on population distributions, demographics and dynamics represents a challenge across the World. In many countries, Covid has exposed a lack of registry systems for recording cases and deaths, but also a lack of reliable data on denominators. Moreover, Covid itself has disrupted the improvement of this situation. Seventy-three percent of NSOs had a Population and Housing Census planned in 2020 or 2021 before the pandemic hit (Contreras-Gonzalez et al., 2020; UNFPA, 2020a). In the low and lower-middle income group, 68 % of the NSOs that were planning a census had to postpone it (United Nations and World Bank, 2020).

The value of small area demographic data remain clear and the Covid-19 pandemic demonstrated key applications of modelled geospatial datasets. They formed the demographic basis for some of the most high profile and influential Covid transmission models used to guide policy (Imperial College, 2022; Institute for Health Metrics and Evaluation, 2022), and are now built into health information system software (University of Oslo, 2022). They were also used in the construction of microplans for the delivery of covid vaccines and in reacting and assessing responses to outbreak and mitigation measures (e.g. GRID3, 2021a; Shepherd et al., 2021). These applications demonstrate a growing acceptance and use of modelled population datasets built upon geospatial data from satellite imagery, GPS mapping and mobile data to fill gaps in small area demographic data availability. Nevertheless, the construction of such datasets remains an area of research, with substantial uncertainties remaining and differing inputs and approaches leading to large variations in output estimates.

2. Modelled small area population estimates

Ideally, every country would have systems such as those in Nordic countries, where integrated registries and administrative data collection systems enable the production of timely small area data on population distributions and demographics, without the need for costly decennial national population and housing censuses (UNECE, 2018). While more countries move towards developing such systems, many are far behind, with the implementation of a national census every ten years still remaining a challenge. Even where these are implemented in a robust and rigorous way, demographic changes during intercensal periods can make the data rapidly outdated, particularly at small area scales where changes are harder to forecast. In some countries, registry and administrative systems can fill these gaps, but these can often be incomplete and inaccurate, especially in many low income settings. Spatial modelling approaches that aim to address some of these challenges through use of satellite imagery and other geospatial data to capture small area changes occurring over relatively short timescales compares to the decennial census have therefore become more prominent in recent decades (UNFPA, 2020b).

In the absence of publicly available small area data from census where it may be too sensitive to share, or in the absence of consistent data between countries and across continents, or in the absence of any recent and reliable data at all, spatial modelling approaches aim to fill gaps. Since the 1990 s (Tobler et al., 1997), a major focus has been on so-called 'top-down' disaggregation, whereby large area census data, or projections matched to relevant administrative/enumeration boundaries are disaggregated to grid squares, maintaining counts at original units and estimating distributions within these units (Leyk et al., 2019). The ongoing global assembly of such unit-based count data (Center for International Earth Science Information Network - CIESIN - Columbia University, 2018a) has meant that different approaches to disaggregation have been explored and applied over recent years. These have been driven by the availability of settlement maps, relevant geospatial covariates and computing power among other factors, as well as intended user needs. Some have maintained the simplicity of simply spreading available population data equally over a grid, some have allocated population counts to mapped settlements, while others have built more complex models that use a set of geospatial covariates to try and capture variations within settlements and across countries (Leyk et al., 2019).

The variety of 'top down' modelling approaches results in differing disaggregations of the same aggregate population count data. Using the example of the five most westerly provinces of the Democratic Republic of the Congo (DRC) that encompass Kinshasa and its surrounding region, Fig. 1 highlights how different a selection of commonly used open model estimates can be at the scale of health zones. These will in turn result in differing surveillance indicators, denominators for health metrics (Nilsen et al., 2021), and target populations for interventions. Fig. 2a-d shows the relative patterns of estimated population distributions in each dataset at the grid square scale for Kinshasa and its immediate surrounding area with health area boundaries overlaid. While DRC represents an extreme example, having not conducted a national census since 1984, and the population estimates input to the models are for large administrative units, the figures highlight how variations in top-down modelling methods can result in substantial differences in predictions. An obvious question when presented with such differing estimates of population distribution is 'which is right?'. This is often a challenge to assess, as where detailed, recent population data exist, often it is the input to the models, leaving a lack of independent data to compare against. Studies using detailed census data (Bustos et al., 2020; Chen et al., 2020; Fries et al., 2021; Yin et al., 2021), as well as cross-validation (Reed et al., 2018; Stevens et al., 2015, 2020) have tried to assess how well different models replicate population numbers and distributions at the scale of available data, and unsurprisingly those more complex models using detailed settlement mapping and range of covariates tend to do best. However, multiple trade-offs exist in the production of such datasets that depend on aspects such as input data availability, geographical extent, temporal range, spatial resolution, intended use and user needs. For instance, a more complex model may produce more accurate outputs, but the production process can be more challenging to communicate to users. A dataset that estimates population distributions over multiple years will necessarily have to compromise on the quality and amount of data for older time periods, and a high spatial resolution dataset can be more difficult to process and incorporate into surveillance systems than a coarser scale one.

A bigger issue than whether one model or another more accurately distributes the population counts from administrative unit to grid square scale, is the fact that the population counts themselves can be inaccurate. The age, scale, type and quality of input population count data being disaggregated within top-down models is information that has often been poorly communicated, understood and addressed. While gridded outputs therefore tend to look similar between countries, there can be substantial variations in the accuracy of population estimates. The differences between the DRC and nearby Malawi are illustrative, with input population count data for DRC coming from uncertain projections from the 1984 census and an average unit size of 12,476 km2, while the data for Malawi come from their 2018 census, with an average unit size of just 9.4 km2. In these two settings, the choice of modelling approach has a significantly larger impact for the DRC where population counts for massively larger units are being disaggregated to the same size grid squares. Nevertheless, as highlighted above, even in settings with high quality regular censuses with data available mapped to small units, the processes of migration, displacement, urbanization and heterogeneous fertility and mortality can make these data quickly outdated and are hard to accurately forecast at small area scales (Wilson et al., 2021), resulting in potentially major impacts on reliable surveillance and health metrics (Tatem, 2014).

The issues outlined above have lead to the rise of census-independent small area population estimation methods (Francoise, 1990; Harvey, 2002; Hillson et al., 2015; Leasure et al., 2020; Wardrop et al., 2018; Weber et al., 2018). These 'bottom-up' modelling methods typically rely
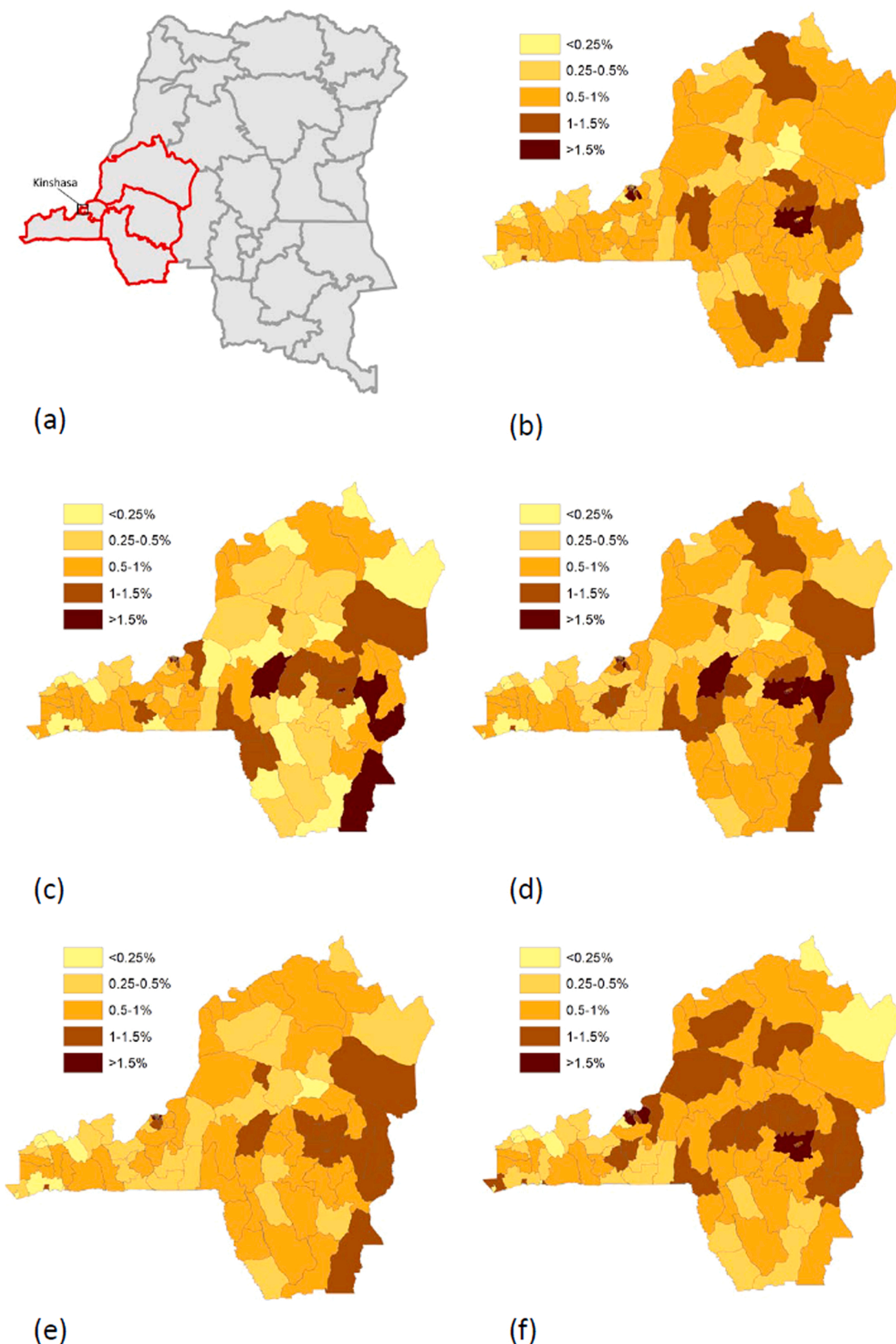
Fig. 1. Maps for the region encompassing the five most westerly provinces of the Democratic Republic of the Congo (Kinshasa, Kongo-Central, Kwango, Kwilu, and Mai-Ndombe, as shown in (a)), showing proportions of the total population of the region estimated to be in each health zone for a set of commonly used open gridded population datasets: (b) Gridded Population of the World version 4 (GPWv4) (Center for International Earth Science Information Network - CIESIN - Columbia University, 2018b), (c) GHS Population grid (Florczyk et al., 2019), (d) Meta Data for Good High resolution population density maps (Meta Data for Good, 2022), (e) WorldPop global constrained top down estimates (Bondarenko et al., 2020), (f) WorldPop/GRID3 bottom up population estimates (Boo et al., 2020).
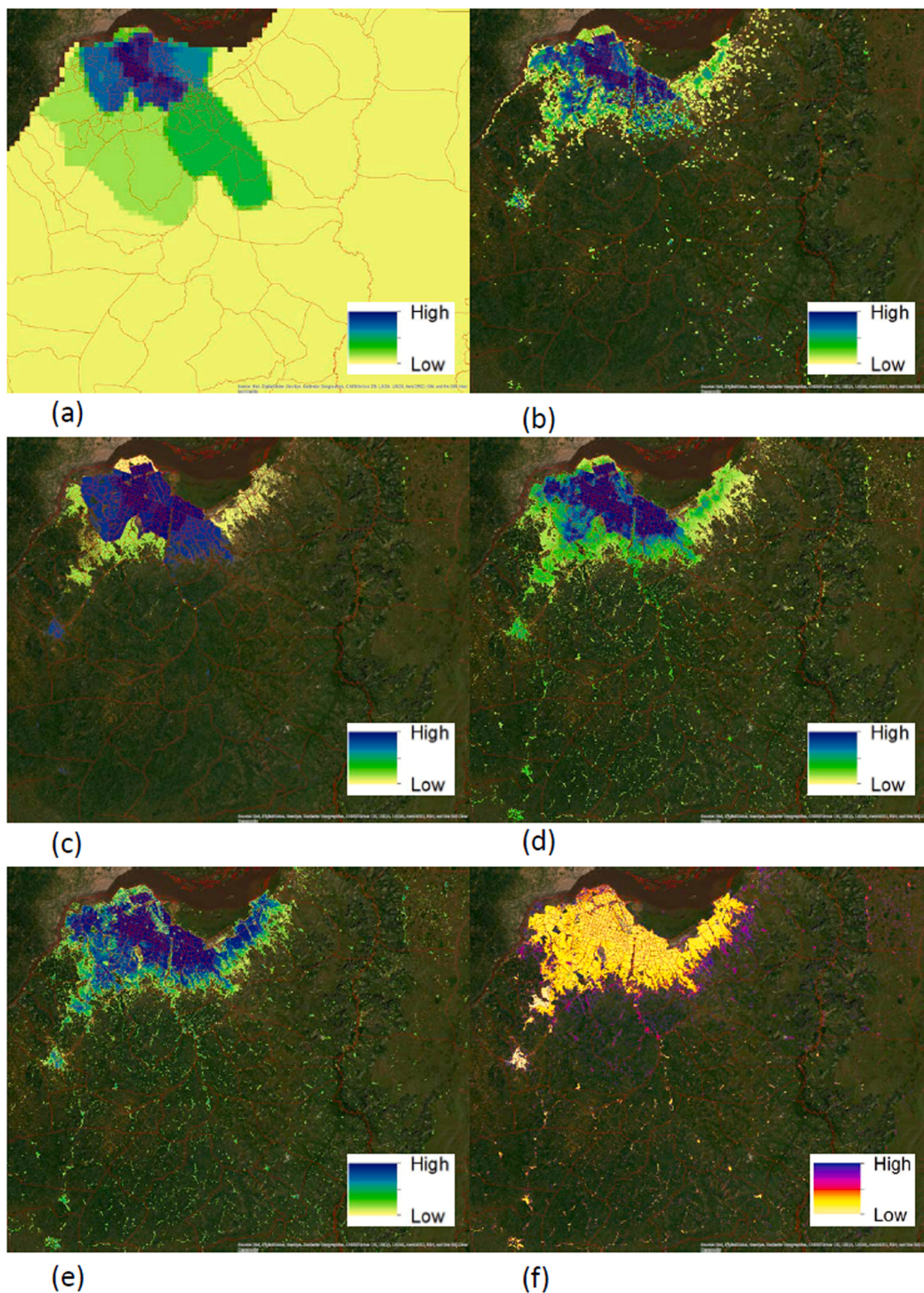
Fig. 2. Gridded population estimates from the same datasets as Fig. 1 for Kinshasa and surrounding area in the Democratic Republic of the Congo, with health areas overlaid. The extent of the area shown is highlighted in the black box in Fig. 1(a). Each dataset has been displayed using a scale of 20 quantiles within the area shown to highlight the inherent population distribution patterns. (a) Gridded Population of the World version 4 (GPWv4) (Center for International Earth Science Information Network - CIESIN - Columbia University, 2018b), (b) GHS Population grid (Florczyk et al., 2019), (c) Meta Data for Good High resolution population density maps (Meta Data for Good, 2022), (d) WorldPop global constrained top down estimates (Bondarenko et al., 2020), (e) WorldPop/GRID3 bottom up population estimates (Boo et al., 2020), (f) uncertainty of WorldPop/GRID3 bottom up population estimates, measured as the difference between the upper and lower 95 % credible intervals of the posterior prediction divided by the mean of the posterior prediction.

on complete counts of population within small defined areas that can come from bespoke field surveys, listings from household surveys, or the results of rolling or partial censuses. Statistical models are then used to link these enumeration data to spatial covariate data, with full coverage over the regions of interest to predict population numbers in unsampled locations (UNFPA, 2020b; Wardrop et al., 2018). Often these are Bayesian models that produce full posterior prediction outputs for each grid square, meaning they can be summarised to produce 'best' estimates, but also measures of uncertainty at differing levels and spatial scales. Recent examples include the development of models to estimate population counts in areas that could not be enumerated in national census efforts in countries such as Burkina Faso, Afghanistan and Colombia (Darin et al., 2022; Sanchez-Cespedes et al., 2021; WorldPop, 2021a). In Zambia, models were developed to make use of household listings from recent survey data as training data to construct national small area estimates that formed the basis for census planning and health intervention delivery (Dooley et al., 2021; GRID3, 2021b; WorldPop, 2021b). Moreover, in settings where existing census data are outdated, bespoke field surveys were undertaken to obtain recent sample enumeration data for constructing modelled estimates in Nigeria, South Sudan and DRC (Leasure et al., 2020; Boo et al., 2022; UNFPA, 2021). The outputs of the bottom-up modelling efforts in DRC are presented in the same way as those from the multiple top-down models in Figs. 1f and 2e, but with the addition of a measure of prediction uncertainty mapped in Fig. 2f (Boo et al., 2022). Given the significantly different input data and approach to production of the estimates, it is no surprise to see more differences in estimated patterns of population distributions in the figures.

Are the bespoke bottom-up population estimation models producing more accurate estimates of small area population numbers than the top-down approaches? This is difficult to assess and remains context and location-specific, but there is growing evidence and many reasons to believe that the estimates are more reliable. While sample sizes are often small, building models upon recent enumeration data, rather than linear projections from census baselines many decades ago in settings where massive changes have occurred provides more confidence in outputs (Wardrop et al., 2018). A growing amount of anecdotal and quantitative feedback from field teams and national statistical offices on the accuracy of estimates adds to statistical evidence from model cross-validation, as well as assessments on the use of data in deriving metrics or in health delivery campaigns adds to this (Nilsen et al., 2021; Leasure et al., 2020; GRID3, 2021b, 2020; Boo et al., 2022; Thomson et al., 2021; Ali et al., 2020). Moreover, the explicit measurement and communication of uncertainty provides users with quantitative insights on where confidence in estimates is higher or lower, taking small area population estimates a step forward beyond the opacity of many top-down model outputs (Leasure et al., 2020).

3. The future for small area population data

Geospatial modelling approaches have made great advances in supporting the estimation of population numbers at small area scales. They should not be seen however as a replacement for censuses, surveys and systems of enumeration. These ensure that people are counted and are the source of a wealth of additional data that cannot be accurately estimated from models based on satellite imagery and other forms of geospatial data. Nevertheless, the challenges that physical enumeration of populations poses should be recognised (UNFPA, 2020b). National population and housing censuses are typically the largest peace-time operations that countries undertake, and the expense, complex logistics and disrupting factors, such as conflict or Covid, mean that full enumeration is not always possible. This in turn results in a lack of reliable and timely data that form a key component of disease surveillance systems. Recent years have seen a growing use of modelled population data by ministries of health, national statistics offices and international agencies, in particular estimates produced from bottom-up

models (Wardrop et al., 2018). As well as being used for health-related applications such as vaccination or bednet distribution campaigns (GRID3, 2021b; Ali et al., 2020), the value of these estimates in supporting the census process (UNFPA, 2020b) or collection of survey data (Thomson et al., 2020, 2017; Qader et al., 2021, 2020) has been shown.

Many challenges remain to be addressed in capturing accurate population numbers at small area scales through both enumeration and modelling approaches, but ongoing research points towards potential solutions. Multiple innovations in data collection and sample design are pushing forward the ability to directly enumerate populations and capture data from those that can be hard to reach (Hoogeveen and Pape, 2020; Tomaselli et al., 2021). These can provide valuable data to complement and improve upon traditional approaches to enumeration, as well as form the basis for the geospatial modelling efforts that are the focus here. Substantial variations in population densities and land uses over small spatial scales make accurately estimating and mapping populations within urban areas a difficult exercise, but the automated ability to accurately map building footprints from recent satellite imagery is helping to quantify some of this variability (Boo et al., 2022). The further processing of these building datasets to map neighbourhood types (Jochem et al., 2020) and categorize residential status (Sturrock et al., 2018) is supporting refinements to urban population modelling (WorldPop and National Population Commission of Nigeria, 2021). Moreover, the development of approaches for estimating building heights and volumes from satellites (e.g. Esch et al., 2022) presents opportunities to account for high rise residential or commercial buildings in estimation modelling. Geostatistical modelling from GPS-located survey data also offer solutions for the small area mapping of population demographics to move beyond large area summaries or outdated census data (e.g. Alegana et al., 2015).

The dynamics of urban populations present challenges, with constant changes in densities each day, week and season, and urbanization trends changing the shape and extent of settlements rapidly. Here again, new forms of geospatial digital datasets offer possibilities to capture and quantify such changes that would be costly to measure with surveys or full enumeration. These include the use of mobile phone call records (Bergroth et al., 2022; Deville et al., 2014), satellite-based measures (Bharti et al., 2016) and models that integrate multiple forms of spatially referenced data (Martin et al., 2015). Subnational changes in population distributions induced by migration and displacement have been shown to be reliably captured by models driven by mobile phone record data (Lai et al., 2019; Bengtsson et al., 2011), and approaches for incorporating such flow data into small area population estimation models continue to be explored (Dooley et al., 2020). Improved understanding of the processes and dynamics of population changes at small area scales in turn offer the potential for improved forecasting (Wilson et al., 2021).

Ultimately, national statistical offices hold the responsibility for the production and maintenance of the official population data that feed into disease surveillance and health information systems. Many have invested in geospatial, geostatistical and data science skills to capitalise on advances in small area population data production. Across low and middle income regions, resources for such investment are often limited, and therefore challenges exist in abilities to adopt, develop and integrate new geospatial methods to complement more traditional enumeration approaches. Official population data underlie governance and can be highly sensitive since they determine allocations of resources, representations in parliaments and delivery of services. It is generally insufficient therefore for models developed and implemented elsewhere to be handed over to statistical offices. Country ownership and the ability of statistical offices to explain and defend methods adopted in the production of small area population data becomes vital for acceptance and use. This will remain a substantial barrier to the use of geospatial modelling to complement more traditional data collection, but the examples highlighted above from countries such as Burkina Faso, Colombia, Nigeria, DRC, Afghanistan and Zambia show how co-development and country ownership are beginning to address these

challenges, providing data that can then be used by Ministries of Health as an important component of surveillance and healthcare delivery.

The Covid pandemic has necessitated the acceleration of vaccines and treatment development, of the science of disease modelling, and of the shape and scale of infectious disease surveillance. The disruption that the pandemic has caused to traditional methods of population enumeration and the reduced budgets available to undertake such efforts will likely also necessitate the accelerated development and adoption of modelling approaches to fill gaps. Moreover, increased demands for timely and reliable small area denominator data to support the needs of expanded and reshaped disease surveillance systems will bring extra demands. Geospatial modelling can become an additional tool in integrated approaches to the production of small area population data. For example, where a survey or new census is being planned, modelled estimates can provide a sample frame where previous census-based data are outdated, and in turn, the data collected using the new frame can be used to improve and update the modelled estimates for future use. These complimentary activities offer the possibility of moving towards a kind of 'living' census, something that is currently only a reality in countries with strong registry-based systems. Having access to timely, reliable population data at small area scales that is able to be regularly updated should ideally be a goal for the World to work towards to ensure appropriate, effective and efficient responses when the next outbreak or pandemic arrives.

Declaration of Competing Interest

None.

Acknowledgements

References

Aborode, A.T., et al., 2021. Impact of poor disease surveillance system on COVID-19 response in africa: time to rethink and rebuilt. Clin. Epidemiol. Glob. Health 12, 100841.

Alegana, V.A., et al., 2015. Fine resolution mapping of population age-structures for health and development applications. J. R. Soc. Interface 12, 105.

Ali, D., et al., 2020. A cost-effectiveness analysis of traditional and geographic information system-supported microplanning approaches for routine immunization program management in northern Nigeria. Vaccine 38 (6), 1408–1415.

Bengtsson, L., et al., 2011. Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti. PLoS Med. 8 (8), e1001083.

Bergroth, C., et al., 2022. A 24-hour population distribution dataset based on mobile phone data from Helsinki Metropolitan Area, Finland. Sci. Data 9 (1), 39.

Bharti, N., 2011. Explaining seasonal fluctuations of measles in Niger using nighttime lights imagery. Science 334 (6061), 1424–1427.

Bharti, N., et al., 2016. Measuring populations to improve vaccination coverage. Sci. Rep. 5, 34541.

Bondarenko, M., et al., Census/projection-disaggregated gridded population datasets, adjusted to match the corresponding UNPD 2020 estimates, for 51 countries across sub-Saharan Africa using building footprints, U.o.S. WorldPop, Editor. 2020.

Boo, G., et al., Modelled gridded population estimates for the Kinshasa, Kongo-Central, Kwango, Kwilu, and Mai-Ndombe provinces in the Democratic Republic of the Congo, version 2.0., U.o.S. WorldPop, Editor. 2020.

Boo, G., et al., 2022. High-resolution population estimation using household survey data and building footprints. Nat. Commun. (in press).

Buckee, C.O., Tatem, A.J., Metcalf, C.J.E., 2017. Seasonal population movements and the surveillance and control of infectious diseases. Trends Parasitol. 33 (1), 10–20.

Bustos, M.F.A., 2020. A pixel level evaluation of five multitemporal global gridded population datasets: a case study in Sweden, 1990–2015. Popul. Environ. 42, 255–277.

Carr-Hill, R., 2013. Missing millions and measuring development progress. World Dev. 46, 30–44.

Center for International Earth Science Information Network - CIESIN - Columbia University, 2018a. Gridded Population of the World, Version 4 (GPWv4):

Administrative Unit Center Points with Population Estimates, Revision 11. NASA Socioeconomic Data and Applications Center (SEDAC),, Palisades NY, USA.

Center for International Earth Science Information Network - CIESIN - Columbia University, 2018b. Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades NY, USA.

Chen, R., 2020. Multiple global population datasets: differences and spatial distribution characteristics. ISPRS Int. J. Geo-Inf. 9, 637.

Contreras-Gonzalez, I.M., 2020. Dealing With Disruption: National Statistical Offices are Adjusting to Continued Challenges as COVID-19 Keeps Affecting Their Operations. World Bank.

Cutts, F.T., Danovaro-Holliday, M.C., Rhoda, D.A., 2021. Challenges in measuring supplemental immunization activity coverage among measles zero-dose children. Vaccine 39 (9), 1359–1363.

Darin, E., et al., *La population vue du ciel: Quand l'imagerie satellite vient au secours du recensement* Population, 2022. in press.

Deville, P., 2014. Dynamic population mapping using mobile phone data. Proc. Natl. Acad. Sci. U.S.A. 111 (45), 15888–15893.

Dooley, C.A., 2020. Description of Methods for South Sudan 2020 Gridded Population Estimates from Census Projections Adjusted for Displacement, Version 2.0. WorldPop, University of Southampton, Southampton, UK.

Dooley, C.A., 2021. Description of methods for the Zambia modelled population estimates from multiple routinely collected and geolocated survey data. WorldPop, University of Southampton,, Southampton, UK.

Dowell, S.F., Blazes, D., Desmond-Hellmann, S., 2016. Four steps to precision public health. Nature 540, 189–191.

Esch, T., 2022. World settlement footprint 3D - a first three-dimensional survey of the global building stock. Remote Sens. Environ. 270, 112877.

Florczyk, A.J., et al., *GHSL Data Package 2019*, in *Publications Office of the European Union, Luxembourg,* J.R.C.o.t.E. Commission, Editor. 2019.

Francoise, D., 1990. Utiliser la télédétection spatiale pour observer les populations citadines. Villes Et. Citad. du Tiers-Monde 4, 23–55.

Fries, B., et al., 2021. Measuring the accuracy of gridded human population density surfaces: a case study in Bioko Island, Equatorial Guinea. PLoS One 16 (9), e0248646.

GRID3 GRID3 data as a building block for COVAX interventions: spotlight on Nigeria microplanning. 2021a.

GRID3 *Data-driven interventions bring Zambia closer to controlling malaria.* 2021b.

GRID3 *Outside the box: how Nigeria won the fight against polio.* 2020.

Harvey, J.T., 2002. Estimating census district populations from satellite imagery: some approaches and limitations. Int. J. Remote Sens. 23, 2071–2095.

Hillson, R., et al., 2015. Stratified sampling of neighborhood sections for population estimation: a case study of Bo city, Sierra Leone. PLoS One 10 (7), e0132850.

Hoogeveen, J. and U. Pape, eds. *Data collection in fragile states.* 2020, CHAM.

Ibrahim, N.K., 2020. Epidemiologic surveillance for controlling Covid-19 pandemic: types, challenges and implications. J. Infect. Public Health 13 (11), 1630–1638.

Imperial College, 2022. COVID-19 CovidSim Model. Imperial College, London, UK.

Institute for Health Metrics and Evaluation, 2022. COVID-19 model FAQs. IHME, University of Washington, Seattle, USA.

Jochem, W.C., et al., 2020. Classifying settlement types from multi-scale spatial patterns of building footprints. Environ. Plan. B: Urban Anal. City Sci. 48 (5), 1161–1179.

Lai, S., 2019. Exploring the use of mobile phone data for national migration statistics. Palgrave Commun. 5.

Leasure, D.R., 2020. National population mapping from sparse survey data: a hierarchical Bayesian modeling framework to account for uncertainty. Proc. Natl. Acad. Sci. U.S.A. 117 (39), 24173–24179.

Leyk, S., 2019. The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use. Earth Syst. Sci. Data 11 (3), 1385–1409.

Martin, D., Cockings, S., Leung, S., 2015. Developing a flexible framework for spatiotemporal population modeling. Ann. Assoc. Am. Geogr. 105, 754–772.

Meta Data for Good, Democratic Republic of the Congo: High Resolution Population Density Maps and Demographic Estimates, Meta, Editor. 2022.

Nilsen, K., 2021. A review of geospatial methods for population estimation and their use in constructing reproductive, maternal, newborn, child and adolescent health service indicators. BMC Health Serv. Res 21 (Suppl. 1), 370.

Qader, S.H., 2020. Using gridded population and quadtree sampling units to support survey sample design in low-income settings. Int J. Health Geogr. 19 (1), 10.

Qader, S.H., 2021. Semi-automatic mapping of pre-census enumeration areas and population sampling frames. Humanit. Soc. Sci. Commun. 8 (3).

Reed, F.J., 2018. Gridded population maps informed by different built settlement products. Data (Basel) 3, 33.

Sanchez-Cespedes, L.M., et al., *Harnessing the predictive power of community workshops, geospatial data, and Bayesian statistics to address census omission in remote areas of Colombia.* 2021 International Population Conference, 2021.

Shepherd, H.E.R., 2021. Domestic and international mobility trends in the United Kingdom during the COVID-19 pandemic: an analysis of facebook data. Int J. Health Geogr. 20 (1), 46.

Stevens, F.R., 2015. Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. PLoS One 10 (2), e0107042.

Stevens, F.R., 2020. Comparisons of two global built area land cover datasets in methods to disaggregate human population in eleven countries from the global South. Int. J. Digit. Earth 13, 78–100.

Sturrock, H.J.W., et al., 2018. Predicting residential structures from open source remotely enumerated data using machine learning. PLoS One 13 (9), e0204399.

Tatem, A.J., 2014. Mapping the denominator: spatial demography in the measurement of progress. Int Health 6 (3), 153–155.

Thomson, D.R., 2017. GridSample: an R package to generate household survey primary sampling units (PSUs) from gridded population data. Int J. Health Geogr. 16 (1), 25.

Thomson, D.R., 2020. Gridded population survey sampling: a systematic scoping review of the field and strategic research agenda. Int J. Health Geogr. 19 (1), 34.

Thomson, D.R., 2021. Evaluating the accuracy of gridded population estimates in slums: a case study in Nigeria and Kenya. Urban Sci. 5, 48.

Tobler, W., 1997. World population in a grid of spherical quadrilaterals. Int. J. Popul. Geogr. 3, 203–225.

Tomaselli, V., 2021. Methods, developments, and technological innovations for population surveys. Soc. Sci. Comput. Rev. 1–20.

UNECE, Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses. 2018: Geneva.

UNFPA, 2020a. Technical Brief on the Implications of COVID-19 on Census. UNFPA, New York.

UNFPA, 2020b. The Value of Modelled Population Estimates for Census Planning and Preparation. UNFPA, New York, USA.

UNFPA South Sudan *Population Estimation Survey launched: Gov't, UN underscore importance for dev't planning, SDGs monitoring*. 2021.

United Nations and World Bank, Monitoring the State of Statistical Operations under the COVID-19 Pandemic. 2020: Washington DC, USA.

University of Oslo. DHIS2 Documentation: Using the Maps App. 2022; Available from: ⟨https://docs.dhis2.org/en/use/user-guides/dhis-core-version-236/analysing-data/maps.html⟩.

Wardrop, N.A., 2018. Spatially disaggregated population estimates in the absence of national population and housing census data. Proc. Natl. Acad. Sci. U.S.A. 115 (14), 3529–3537.

Weber, E.M., 2018. Census-independent population mapping in northern Nigeria. Remote Sens Environ. 204, 786–798.

Wesolowski, A., 2017. Multinational patterns of seasonal asymmetry in human movement influence infectious disease dynamics. Nat. Commun. 8 (1), 2069.

Wilson, T., 2021. Methods for small area population forecasts: state-of-the-art and research needs. Popul Res Policy Rev. 1–34.

WorldPop Mapping Afghanistan's population. 2021a.

WorldPop, 2021b. A simulation study exploring weighted Bayesian models to recover unbiased population estimates from weighted survey data. University of Southampton, Southampton, UK.

WorldPop and N.P.C.o. Nigeria;, Bottom-up gridded population estimates for Nigeria, version 2.0., U.o.S. WorldPop, Editor. 2021: Southampton, UK.

Yin, X., et al., 2021. Which gridded population data product is better? Evidences from mainland Southeast Asia (MSEA). ISPRS Int. J. Geo-Inf. 10, 681.

Zu Erbach-Schoenberg, E., et al., 2016. Dynamic denominators: the impact of seasonally varying population numbers on disease incidence estimates. Popul Health Metr. 14, 35.