



LEAVING NO ONE OFF THE MAP

A GUIDE FOR GRIDDED POPULATION
DATA FOR SUSTAINABLE DEVELOPMENT

EXECUTIVE SUMMARY





ACKNOWLEDGMENTS

This report was written by Hayden Dahmm (SDSN TReNDS) and Maryam Rabiee (SDSN TReNDS), with input from Jessica Espey (SDSN TReNDS), Susana Adamo (Columbia University), Alex de Sherbinin (Columbia University) and Robert Chen (Columbia University). We thank Alyson Marks (SDSN TReNDS) for her assistance with copy-editing and Luciana Berger for design.

We are grateful to the many gridded population data users and experts whose inputs have made this report possible. In particular, we would like to thank Kytt MacManus (Columbia University), Greg Yetman (Columbia University), Andy Tatem (University of Southampton), Stefan Leyk (University of Colorado Boulder), Forrest Stevens (University of Louisville), Andrea Gaughan (University of Louisville), Amy Rose (Oak Ridge National Laboratory), Marie Urban (Oak Ridge National Laboratory), Jacob Mckee (Oak Ridge National Laboratory), Kelly Sims (Oak Ridge National Laboratory), Thomas Kemper (European Commission Joint Research Center), Michele Melchiorri (European Commission Joint Research Center), Sergio Freire (European Commission Joint Research Center), Charlie Frye (Esri), Ashraf Abushady (World Food Programme), Dana Thompson (University of Southampton), Lorenz Noe (Open Data Watch), and Cascade Tuholske (University of California, Santa Barbara) for sharing their knowledge and perspectives in extended interviews. We would also like to recognize the valued contributions from our colleagues at the Global Partnership for Sustainable Development Data (GPSDD), Alonso Ortiz Galán and Charu Vijayakumar, who included questions on gridded population data during their survey of government officials from Belize, Colombia, Costa Rica, Ecuador, Ghana, Nepal, Paraguay, Sierra Leone, and Tanzania. Lastly, we are grateful to the Bill & Melinda Gates Foundation who helped make this report possible with their support of the POPGRID Data Collaborative.

EXECUTIVE SUMMARY

Each year, nearly 160 million people are impacted by natural disasters. During a natural disaster, every second counts to save lives and to ensure that critical supplies reach those in need. To respond quickly and confidently to identify impacted communities in the wake of natural disaster, emergency organizations, like the United Nations World Food Program (WFP), are using gridded population data to estimate near real-time impacts of earthquakes and tropical storms on people and infrastructure. Gridded population data have proven essential for informing emergency response efforts and minimizing suffering.

Having reliable and timely population data can make a life or death difference for individuals facing crises or living in conflict-ridden regions. These data are essential for addressing the above challenges and for critical decision-making and planning. We need to know where people are located, what conditions they are facing, what infrastructure is available, and what basic services they can access. When it comes to counting people in hard-to-reach or conflict-ridden environments, there are no silver bullets. However, gridded population data offer a promising option for delivering actionable data in difficult circumstances.

When 193 world leaders agreed upon the 17 Sustainable Development Goals (SDGs) in 2015, they promised to “leave no one behind.” But without reliable and timely population data linked to location, we cannot ensure that everyone is counted and that no one will be left behind. While, most governments and policymakers depend on traditional data sources, such as household surveys and population censuses, in order to develop the necessary policies and programs to eradicate poverty and improve health, education, and other basic services, these traditional data sources present a range of geographic, temporal, and logistical challenges. For example, at the time of writing, nearly 60 countries are facing confirmed or potential delays to their census operations because of the COVID-19 pandemic (UNFPA, n.d.). Fortunately, with recent advancements in Earth observations’ capabilities and statistical methods, it is now possible to obtain more frequent and more granular population estimates worldwide through the use of gridded population datasets.

Gridded (or raster) population maps represent the distribution of population in rows and columns of grid cells, typically defined by their latitude-longitude coordinates. An increasing number of data providers are combining information from censuses with satellite-derived geospatial features to redistribute populations and produce gridded population datasets. Despite this progress, there remains confusion or simply lack of awareness about gridded population data. The large number of different datasets now available can be overwhelming to users, particularly to those who lack the time and technical expertise to understand differences among the products and assess their strengths and weaknesses for potential applications.

The POPGRID Data Collaborative was established in 2018 to address many of these challenges by connecting the diverse data users, providers, and stakeholders from the public and private sectors working with georeferenced data on population, human settlements, and infrastructure.

In this report, *Leaving No One off the Map: A Guide for Gridded Population Data for Sustainable Development*, we aim to narrow this knowledge gap by helping to improve the accessibility and understanding of gridded population datasets for policymakers and other users. The report was written with two overarching questions in mind:

- ▶ How can gridded population data supplement current population data sources and support users from the sustainable development community to make timely, informed decisions?
- ▶ Which gridded population dataset is the most suitable for a user's intended use?

Drawing from an extensive literature review and interviews with key data providers and users in the POPGRID Data Collaborative, the report presents an overview, analysis, and recommendations for the use of gridded population datasets in a wide range of application areas, such as in disaster response, health interventions, and survey planning. Specifically, the report compares seven gridded population datasets from the POPGRID Data Collaborative, including an analysis of the underlying data, methods and basic assumptions, and the corresponding strengths and limitations of each dataset in simple terms. The report also presents an intercomparison assessment of the use of different datasets and their varying outputs, addresses many of the misconceptions around gridded population data, and concludes with nine guiding criteria to aid users in their selection process.



KEY MESSAGES FROM THE REPORT INCLUDE:

CENSUS DATA ARE STILL IMPORTANT; GRIDDED POPULATION DATA ARE NOT A SUBSTITUTE FOR CENSUS DATA.

The datasets featured in this report build off of national censuses as inputs to their population estimations, which can then be more regularly updated or calculated at a higher spatial frequency.

GRIDDED POPULATION DATA ARE NOT ERROR-FREE.

Although they address some of the limitations of traditional sources, the methods for producing estimates add their own sources of uncertainty. Users should be aware of and transparent about these potential uncertainties.

DATA USERS SHOULD CONSIDER A NUMBER OF FACTORS WHEN SELECTING AN APPROPRIATE DATASET FOR THEIR PARTICULAR NEEDS.

Each of the datasets are based on different underlying data and assumptions, so policymakers and researchers should examine the characteristics of available datasets carefully. These include: demographic characteristics; the spatial resolution required; the time periods of interest; data costs and rights to reuse the data; and more.

MORE VALIDATION WORK IS NEEDED TO COMPARE GRIDDED POPULATION DATA ESTIMATES AGAINST AUTHORITATIVE DATA ON POPULATION LOCATION.

There is a critical need for a more systematic analysis and objective validation of these products to further refine methods and improve their accuracy and utility; this work is underway through the POPGRID Data Collaborative.

With only ten years remaining to achieve the SDGs, we are at a crossroads. Gridded population data are already available to help fulfill these ambitious goals by improving the availability, consistency, and spatial disaggregation of SDG indicators, by helping national and international initiatives to better target their efforts to achieve the SDGs, and by identifying and locating those who might otherwise be left behind. However, these data are only as good as policymakers' understanding of their limitations, applications, and fitness for use. Furthermore, their full potential cannot be realized if they are not thoroughly validated against authoritative, real-world data. We must accelerate this important research and advancement of the application of gridded population data around the world to make sure that the SDGs are achieved and no one is left off the map.

