A pixel level evaluation of five multitemporal global gridded population datasets: a case study in Sweden, 1990–2015

Maria Francisca Archila Bustos¹ · Ola Hall¹ · Thomas Niedomysl^{2,3} · Ulf Ernstson⁴

Published online: 1 September 2020
© The Author(s) 2020

Abstract

Human activity is a major driver of change and has contributed to many of the challenges we face today. Detailed information about human population distribution is fundamental and use of freely available, high-resolution, gridded datasets on global population as a source of such information is increasing. However, there is little research to guide users in dataset choice. This study evaluates five of the most commonly used global gridded population datasets against a high-resolution Swedish population dataset on a pixel level. We show that datasets which employ more complex modeling techniques exhibit lower errors overall but no one dataset performs best under all situations. Furthermore, differences exist in how unpopulated areas are identified and changes in algorithms over time affect accuracy. Our results provide guidance in navigating the differences between the most commonly used gridded population datasets and will help researchers and policy makers identify the most suitable datasets under varying conditions.

Keywords Gridded population · Population estimation · Human population distribution · Dasytetric mapping

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11111-020-00360-8>) contains supplementary material, which is available to authorized users.

✉ Maria Francisca Archila Bustos
maria.archila@keg.lu.se

✉ Ola Hall
ola.hall@keg.lu.se

Extended author information available on the last page of the article

Introduction

Knowledge of the spatial distribution of human activity is imperative in order to understand and address many of the global challenges that we face today, challenges that have been generated and are driven, in large part, by such human activity. Detailed information about human population distribution is a fundamental component of such knowledge. However, population distribution information is most often aggregated to administrative areas, as is the case with census data. These aggregations assume evenly distributed populations and can be difficult to use when analyzing irregular phenomena that do not follow the same administrative boundaries, which both human and natural phenomenon often do not follow.

Gridded population datasets disaggregate census data to a more useable scale and format by modeling the distribution of the human population on Earth at the grid or pixel level. This reduces issues related to statistical fallacies, such as the modifiable areal unit problem (MAUP) (Openshaw 1984), and increases possibilities related to modeling which often requires disaggregated data as input. There are various methods for generating gridded population datasets which vary in the type of input data used, the extent of the study area, and the assumptions made (Leyk et al. 2019; Wu et al. 2005).

The use of gridded population datasets in research has increased over the last 25 years and most dramatically in the last 10 years, as seen in Fig. 1. Gridded population data can be useful in order to identify populated and unpopulated places as, for example, in disaster management to determine where aid should be sent. It can also be used to estimate the number of people located in certain areas, as is often done when estimating the effects of environmental catastrophes or the spread of disease. Finally, it can be used in modeling and projections such as those done for varying climate scenarios. Gridded datasets on global human population distribution are currently applied in a wide variety of research and decision-making, including recently within environmental degradation and monitoring (Hjort et al. 2018; Sordo-Ward et al. 2019), development and inequalities (Li and Liu 2018; Melchiorri et al. 2019), disaster management (Dou et al. 2018), accessibility (Cheng et al. 2019; Weiss et al. 2018), and epidemiology (Haw et al. 2019; Siraj et al. 2018).

Specifically, there are five main datasets that have emerged as the leading datasets for use in research and decision-making, likely as a result of both their global nature and because they are freely available online.¹ These include the Gridded Population of the World (GPW), the Global Rural Urban Mapping Project (GRUMP), the Global Human Settlement Layer-Population (GHS-POP), the LandScan Population database, and the WorldPop database.

However, most studies which employ these datasets do not discuss their dataset choice and it has been shown that the choice of dataset can have a significant effect on results (Mondal and Tatem 2012; Tatem et al. 2011). While use of gridded population data is widespread, there have been few studies which compare existing datasets systematically. While individual comparisons to known population datasets have been carried out (e.g., Calka and Bielecka 2019), comprehensive evaluations of multiple datasets are largely missing in the literature. One exception is a recently published paper by Leyk et al. (2019) which provides an in-depth comparison of several gridded

¹ Landscan is only freely available for the educational community

Fig. 1 “Gridded population” in research output between 1993 and 2018. Number of publications with the term “gridded population” found within the text at least once. Counts taken from Digital Science’s Dimensions database

population datasets by comparing their data inputs and processing methodologies and providing commentary on their fitness for use. Additionally, the POPGRID Data Collaborative, which is “an international community of data providers, users, and sponsors concerned with georeferenced data on population, human settlements and infrastructure” has recently emerged as an effort to improve access to, simplify interpretation of, reduce confusion over, and encourage innovative use of the diverse gridded population datasets that have become available (POPGRID Data Collaborative 2020).

The aim of this paper is to contribute to this recently started discussion on accuracy and fitness for use of various gridded population datasets by evaluating the differences between five commonly used, freely available population datasets, and official population statistics from Sweden on a grid cell level. This contribution is twofold. First, this paper attempts to quantify the reliability of existing datasets by comparing them to a known population distribution on a pixel level. Most existing comparisons are done on an aggregated level because of lack of access to reliable high-resolution population data. Second, it includes a comparison over an extended time period. Most quantified comparisons of gridded population datasets to known population distributions evaluate only one or two datasets and are restricted temporally to 1 year. Comparisons of a greater number of datasets, such as that done by Leyk et al. (2019), are often not quantified. Additionally, the disaggregation of the analysis by population density and population change provides further detail through which the findings can be meaningful in other similarly developed regions outside of the Swedish context.

In this research, we carried out a grid cell by grid cell comparison of the estimated population value for each of the five most commonly used gridded population datasets against the known Swedish population distribution at their place of residence with 100-m resolution. We examined the latest version of each of the five datasets over 5-year periods between 1990 and 2015 (where 5-year periods were not available, we used all

of the years available). To quantify the differences between the estimated and known populations, we calculated percent root mean square error (%RMSE) and percent mean absolute error (%MAE) for the whole population and under different density and growth conditions. We also examined the linear association between the estimated and known populations using Pearson's r and each dataset's ability to correctly identify populated and unpopulated grid cells.

The structure of this paper is as follows: first, we introduce the study area, data, and methods. Next, the results are presented followed by a discussion and conclusions.

Study area, data, and methods

Study area

The population of Sweden was 10,319,601 at the end of October 2019 with an annual population growth rate of approximately 1% (Statistiska centralbyrån (SCB), 2019a). Thus, Sweden can be characterized as having a slow growing population. Similar rates of growth exist in countries such as Ireland, Vietnam, Argentina, Mexico, and Indonesia (based on 2018 annual growth rates) (World Bank 2018). The population density of Sweden was approximately 25 people/km² in 2018 (SCB, 2019b). However, this varies widely throughout the country. The population density in the most populated municipalities is approximately 5000 people/km² while the population in the most sparsely populated municipalities is lower than 1 person/km² (SCB, 2019b). Most of the population (87%) lives on approximately 630,000 ha of land, which equates to approximately 1.5% of Sweden's landmass. Approximately 32% of the Swedish population live in cities with a population over 100,000 (of which there are nine cities). Since 2015, 40% of Sweden's population increase has occurred in these cities and a further 37% has occurred in cities with populations between 50,000 and 99,999 (SCB, 2019c).

Data

Here, we describe the five gridded population datasets (hereafter referred to as “candidate population grids”) which were evaluated in this study as well as the Swedish reference dataset (hereafter referred to as the “known population grid”) which was used for comparison. We first provide a general background on existing methods for generating gridded population datasets and then provide details on each individual candidate population grid dataset.

Areal interpolation is a technique commonly used to disaggregate census population data gathered from various sources to the grid level (Wu et al. 2005). Areal interpolation refers to transforming statistical information from one set of geographical boundaries to another separate set of boundaries and can be accomplished using a number of different methods (Deichmann 1996; Fisher and Langford 1995, 1996; Mennis 2003). Most relevant to the global gridded population datasets used in this study are the areal interpolation methods of areal weighting and dasymetric mapping. Areal weighting makes the assumption that population is equally distributed across administrative areas and distributes census population counts by assigning a population value to each grid cell based on the percentage of the administrative unit covered by each cell (Bhaduri

et al. 2007; Lloyd et al. 2017; Mennis 2003). Of the five global gridded population datasets, GPW is the only one which uses areal weighting, while the others use varying interpretations of dasymetric mapping. Dasymetric mapping makes assumptions about the relationship between population density and various geographic and land characteristics and introduces ancillary data to help dictate where population should be assigned and how much should be assigned in each location (Bhaduri et al. 2007; Linard et al. 2011; Mennis 2003, 2009; Mennis and Hultgren 2006; Zandbergen and Ignizio 2010). GRUMP and GHS-POP use a binary reallocation that distinguishes rural and urban areas and built-up areas respectively, while LandScan and WorldPop use a number of different ancillary datasets to model population distribution. In the following subsections, we provide a brief summary of the characteristics of each of the datasets and the methods used to produce them. Table 1 provides an overview. A similar discussion can be found in (Leyk et al. 2019). For full details, please consult documentation related to each dataset.

GPWv4

GPWv4 models population distribution on a continuous and global surface in order to provide a “spatially disaggregated population layer that is compatible with data sets

Table 1 Comparison of global gridded population datasets

Dataset	Population type	Method and ancillary data	Resolution	Year(s) available
Global Human Settlement Layer (GHS-POP)	Resident-based population	Binary dasymetric; built-up areas (based on GHS-BU)	250 m, 1 km (World Mollweide projection)	1975, 1990, 2000, 2015
Gridded Population of the World (GPW)	Nighttime population	Areal weighted; no ancillary data	30 arc seconds	2000, 2005, 2010, 2015, 2020
Global Rural Urban Mapping Project (GRUMP)	Nighttime population	Binary dasymetric; urban areas, nighttime lights	30 arc seconds	1990, 1995, 2000
LandScan Global Population database 2000–2017	Day time (ambient) population	Multi-variable dasymetric; land cover, roads, slope, urban areas, village locations	30 arc seconds	Annually,
WorldPop 2000–2020	Residential population	Random Forest-based dasymetric redistribution; land cover, elevation, slope, nighttime lights, climate, roads, waterways, settlements, protected areas, facilities	3 arc seconds	Annually,

Based on comparison table from POPGRID Data Collaborative (CIESIN, 2018a) and (Leyk et al. 2019)

from social, economic, and Earth sciences disciplines, and remote sensing” for use in “research, policy-making and communications” (Center for International Earth Science Information Network (CIESIN) 2018b, p. 3). The first release of GPW was in 1995; however, here, the current release, version 4.11, will be discussed. Over time, the methodology and inputs to GPW have been improved and refined.

The GPWv4 population distribution surfaces are available both in population counts (persons per pixel) and population density (persons per km²) for 2000, 2005, 2010, 2015, and 2020. The spatial resolution is 30 arc seconds (which corresponds to approximately 1 km at the equator); however, aggregated datasets are also available at 2.5 arc minutes, 15 arc minutes, 30 arc minutes, and 1 degree for faster processing. The data are published in WGS1984 geographic coordinate system and is available in ASCII, GeoTiff, and NetCDF formats. Each year of data includes one dataset which is consistent with national population censuses and registers and another which has been adjusted to match United Nations (UN) estimates of population taken from the UN World Population Prospects, 2015 Revision.

The GPWv4 population surfaces are generated using areal weighting. The input population data are gathered through an intensive process of data collection from national statistics offices and other organizations. In GPWv4, population data correspond to the highest spatial resolution available from censuses undertaken between 2005 and 2014 (CIESIN, 2018b). Specifically for Sweden, the population inputs are from the 2010 population register at administrative level 3 (corresponding to församling). This is the highest spatial disaggregation of population data freely available to the public. Input population data are matched to administrative boundaries collected from national agencies and other organizations. The priority is to use boundaries from the census used by each country; however, when these are not available, other sources are used. The sources of both population and boundary data are available on the GPW website. The yearly population estimates are then disaggregated to the 30 arc second grid using areal weighting. A water mask is applied before disaggregation in order to ensure that population is not assigned to waterbodies and ice-covered areas (CIESIN, 2018b; Doxsey-Whitfield et al. 2015).

A main benefit of the fact that GPW does not use any ancillary data in its modeling approach is that the outputs can be incorporated with other data without worrying that they might be endogenous to the modeled population (CIESIN, 2018b; Doxsey-Whitfield et al. 2015). However, one of the limitations with this approach is the assumption that population is evenly distributed within administrative boundaries (Deichmann et al. 2001). This translates into a limitation with change in precision over space due to the fact that smaller input units produce higher accuracy results than larger input units (Deichmann et al. 2001; Doxsey-Whitfield et al. 2015). Additionally, the confidence in the estimates varies across countries as a function of the currency, spatial resolution, and accuracy of the input data (CIESIN, 2018b; Doxsey-Whitfield et al. 2015) and can be affected by the temporal interpolation, especially in regions where population changes dramatically under a short period (Deichmann et al. 2001).

GRUMP

The GRUMP database is an attempt to better model differences in population distribution across the urban rural divide. It consists of several data products; however, here,

only the population grid will be discussed. The GRUMP population grid builds on GPW data (version 3) to generate an improved population grid that reallocates population more accurately between urban and rural areas (Balk et al. 2005). The GRUMP population grid is available for 1990, 1995, and 2000 and is based on the 2000 round of Population and Housing Censuses. Like GPW, GRUMP population is provided in two sets, one based on unadjusted census estimates and one based on census estimates adjusted to UN population estimates and includes both population counts and population densities. GRUMP population data are produced in WGS1984 geographic coordinate system and has a spatial resolution of 30 arc seconds. It is available in BIL, ASCII, and GRID (ESRI) formats.

The GRUMP population datasets use a binary dasymetric methodology to reallocate population to urban and rural areas. They are based on two vector inputs: administrative areas with population counts built on data from GPWv3 and urban extents along with urban population counts based on national statistics and nighttime lights. The inputs are overlaid in order to generate a vector dataset that delineates urban and rural areas in each administrative unit. This dataset becomes the input to GRUMPe, an algorithm that, working on a country basis, reallocates the total population of each administrative unit to rural and urban areas based on the area and population of the urban area and total administrative area, the area of the region where the administrative and urban areas overlap and UN national estimates for the percentage of the population in urban and rural areas. The output of the GRUMPe reallocation algorithm is a vector dataset which is converted to a raster grid to produce the GRUMP dataset (D. Balk et al. 2005; Balk et al. 2010).

Some issues with the GRUMP population dataset are related to the estimation of urban extents using nighttime lights, which have known issues with “blooming” and an inability to detect areas with low or no electrification (D. Balk et al. 2005; Elvidge et al. 2004). However, GRUMP uses a combination of nighttime lights and buffered settlement extents based on population size in dimly lit areas, which helps reduce this limitation. Additionally, issues relevant to GPW, including reliability of input population estimates, are also relevant for GRUMP. Finally, GRUMPe has observed spatial differences in accuracy depending on the quality of input data. Where the input administrative data and urban extents data are good, GRUMPe does not add much to the areal weighting of GPW. Where the administrative population data are moderate or poor but the urban extent data are good, GRUMPe performs very well. Finally, where both the administrative and urban extents data are poor, then GRUMPe does not perform well, since the outputs can only be as good as the inputs (D. Balk et al. 2005). Since the release of GPWv4, users are recommended to use that dataset instead, since it supersedes GPWv3 and thus also GRUMP.

GHS-POP

The Global Human Settlement Layer is a group of datasets which describes the human presence on the planet based on Earth observation satellite sensors, national statistical surveys, and crowd sources and is processed by “exploiting novel spatial data analytics tools allowing to handle their complexity, heterogeneity and large volume” (Pesaresi et al. 2016, p. 10). Again, only the population grid dataset will be discussed here, although reference to other layers will be made in that they contribute to its creation.

The population grid (GHS-POP) is available for 1975, 1990, 2000, and 2015. It is provided in a World Mollweide projection (an equal area projection) with a 250-m and 1-km spatial resolution. The grid cell value represent both the count and density of the population in the cell (density since each cell has a consistent size). The datasets are distributed in GeoTiff format.

GHS-POP also uses a binary dasymetric methodology to disaggregate population. Its population data input comes from the GPWv4 UN-WPP-adjusted dataset. These administrative area level population data are disaggregated to built-up areas based on the GHS built-up area (GHS-BU) layer, a grid dataset which shows the percentage of each cell that is covered by built-up areas. GHS-BU is generated using automated analysis of Landsat satellite imagery to identify the location and density of built-up areas (Freire et al. 2016; Pesaresi et al. 2016). The population data are reallocated in one of three ways, although most of the population (95%) is allocated in the first way (Freire et al. 2016). If the administrative area is large enough to generate 250 m grids and contains built-up areas, then the population for that administrative area is assigned in proportion to the density of the built-up areas. If the administrative area is large enough to generate 250 m grids but does not contain any built-up areas, then the population is allocated using an areal weighting approach. If a cell is located on the border of an administrative area, it is assigned to the administrative area its centroid falls in. And finally, if the administrative area is smaller than a 250-m grid cell, then a centroid is generated for the area and the population of all centroids found within a cell is added (European Commission Joint Research Center (JRC) 2017; Freire et al. 2016). Because of the use of centroids to identify the location of cells and administrative units, shifts in the location of population to neighboring cells may occur (Freire et al. 2016).

The benefit of GHS-POP is that it restricts population to built-up areas and makes its density directly proportional to the density of built-up areas (Freire et al. 2015). However, because the reallocation of population in GHS-POP is based on the density of built-up areas, population may be allocated to “non-residential” areas such as commercial, industrial, and recreational areas. For this reason, the output of GHS-POP is said to depict the “resident-based population,” which is not the population at its place of residence, but rather residential population allocated to built-up areas (Freire et al. 2016). In order to generate population counts at the place of residence, land use information would need to be considered (Freire et al. 2016).

LandScan

LandScan gridded population data was first generated in 1998 and, for over 15 years, the methodology and data sources have been improved. LandScan population grids are freely available for researchers and students after registration and are available online yearly from 2000 to 2018 (as explained below, LandScan data are not, however, a consistent time series). The population grids are produced in WGS1984 geographic coordinate system and have a resolution of 30 arc seconds. They are provided in GRID and binary raster formats (both ESRI formats). The cell values for LandScan data represent integer population counts of ambient population. Ambient population, in contrast to resident population, is the average population over a 24-h period for typical days, weeks, and seasons (Dobson et al. 2000). Therefore, it represents not only where people live but also where they work and travel (Dobson et al. 2000).

LandScan uses a highly modeled approach to disaggregate subnational census data to the grid level through a series of dynamically adaptable algorithms using a combination of spatial data, satellite imagery, and derived products and manual corrections. The algorithms used by LandScan are proprietary and therefore not available publicly. Each grid cell in LandScan is assigned a “likelihood” coefficient for the presence of population. The subnational census population is then assigned proportionally to each cell based on this likelihood coefficient. Likelihood coefficients are based on the relationship between population and ancillary data such as land cover, roads, slope, urban areas, village locations, and high-resolution imagery. Relationships vary locally and are drawn from socio-economic and cultural understanding of the area (Oak Ridge National Laboratory (ORNL) 2019; Rose and Bright 2014).

The LandScan algorithm has been refined and improved over time. Additional and higher accuracy data have also been incorporated every year. For this reason, time series comparisons using LandScan data are discouraged, since changes cannot be solely attributed to changes in population but may instead reflect changes in the input data or algorithms. An additional weakness of LandScan is that the population and ancillary data inputs are not documented specifically, making it more difficult to evaluate fitness for use in various situations. Direct comparisons with other population datasets are also cautioned because ambient population is of course not directly comparable with resident population (Rose and Bright 2014). For example, the ambient population in city centers is expected to be higher than the resident population because of the presence of a high level of employment uses (Dobson et al. 2000).

WorldPop

The WorldPop project was initiated in 2013 to unite a number of different initiatives focused on producing population distribution and composition maps. Apart from population counts, the WorldPop database includes a number of different demographic indicators at the grid level. WorldPop population count datasets are available at the country level globally. They are available yearly from 2000 to 2020 (unlike LandScan, this does represent a time series) at a spatial resolution of 3 arc seconds (approximately 100 m at the equator) and in geographic coordinate system WGS1984.

WorldPop population grid datasets are created using Random Forest–based dasymetric redistribution. This method uses a Random Forest model to generate population density predictions based on ancillary data, which WorldPop refers to as covariates. These covariates include land cover, elevation data and derived slope estimates, nighttime lights, climactic spatial variation, roads, waterways, settlements, protected areas, and facilities such as schools, hospitals, and health clinics (Stevens et al. 2015) and can vary by country based on data availability and the relative importance for population estimation at each location. The population inputs are taken from a GIS-linked database of census and official population estimates constructed through the WorldPop initiative and based on GPWv4. Details of the Random Forest model are available in Stevens et al. (2015). The predicted population densities for each country are then used to disaggregate population estimates using a dasymetric mapping approach weighted by the population density prediction. The covariates and estimation algorithms are all freely available from the WorldPop website.

The WorldPop Random Forest methodology is said to improve on other methods because it incorporates many ancillary datasets with little tuning or supervision. However, some important challenges include the standardization of these multiple inputs that come in multiple scales and resolutions, the fact that many of the inputs are highly correlated and the presence of many non-linear interactions (Stevens et al. 2015). One limitation of the data is that Random Forest predictions are limited to the range of population densities of the inputs. This can have an effect in larger administrative areas where population is concentrated in one location (Stevens et al. 2015). However, it is found that the Random Forest methodology is no worse in these areas than other approaches that use less ancillary data (Stevens et al. 2015).

Candidate population grid inputs

Grid data for population counts from each of the above described datasets were downloaded between 2019-03-11 and 2019-03-12. Table 2 provides a summary of the datasets used as well as the temporal and spatial resolution used in the study.

We maintained the native projection for all datasets and the native resolution for all except WorldPop (we used only the 1-km resolution for GHS-POP). For WorldPop, we aggregated the 3 arc second resolution to 30 arc seconds. The data was clipped to the extent of the Swedish border, with the exception of the WorldPop dataset which was downloaded already clipped. Grid population cells with NoData were assigned a value of zero in order to ensure that all cells could be compared to the known population grid.

The datasets used in this analysis were chosen because they were identified as the most commonly used datasets in current research. We would like to make two notes regarding the inclusion of certain datasets in this analysis. First, GRUMP has been superseded by the latest version of GPW (version 4) and its use is no longer

Table 2 Description of datasets used in analysis

Dataset (and version)	Source	Year(s) used	Resolution used
GHS-POP R2015A	(European Commission & Columbia University, 2015)	1990, 2000, 2015	1 km
GPW version 4.11, population count	(CIESIN, 2018d)	2000, 2005, 2010, 2015	30 arc seconds
GPW version 4.11, UN WPP-adjusted population count	(CIESIN, 2018c)	2000, 2005, 2010, 2015	30 arc seconds
GRUMP version 1, population count	(CIESIN, International Food Policy Research Institute (IFPRI), The World Bank, & (CIAT), 2011)	1990, 1995, 2000	30 arc seconds
GRUMP version 1, UN WPP-adjusted population count	(CIESIN et al., 2011)	1990, 1995, 2000	30 arc seconds
LandScan	(ORNL, 2018)	2000, 2005, 2010, 2015	30 arc seconds
WorldPop, Global High-Resolution Denominators Project	(WorldPop 2018)	2000, 2005, 2010, 2015	30 arc seconds

recommended; however, we included it in this analysis since it is still sometimes used in research. Second, LandScan and GHS-POP represent population which is not equivalent to the “night-time” or resident-based population which may affect comparison results, as population comparisons are made against known residential statistics. However, both have been included in this analysis because of the temporal and population type comparisons which were also done. The results which compare the datasets over time and in different situations are still highly relevant. Furthermore, comparing these datasets to a resident-based population highlights the differences between resident-based and ambient populations and shows the importance of choosing the correct dataset for the application in question.

Known population data input

The known population was provided by the Swedish Statistical Bureau (Statistiska centralbyrån, SCB) and consisted of a 100 m by 100 m vector grid containing the count of the population which resides in each grid. These data are available only to researchers upon specific request to SCB. For this reason, they are not currently used in the generation of the candidate population grids which we examine in this research and are thus relevant for use as a comparison. The gridded data are based on information from the Swedish population register. Included in the Swedish population register are all registered residents of Sweden, which comprises all Swedish citizens and all non-Swedish citizens with a residence permit for a minimum of 12 months. The Swedish population register does not include temporary migrants, undocumented migrants or asylum seekers who have not yet received a residence permit (SCB [n.d.](#)). To generate the 100 m by 100 m grid data, each person in the population register is geocoded to the location of residence (fastighet) and this information is then generalized to the grid level based on the location of the centroid of each residence building. The grid data are highly accurate; during the 6 study years, an average of approximately 5300 people are missing from the dataset per year, as compared to official population totals from SCB. This amounts to approximately 0.06% of the population.

The known population data were obtained for 1990, 1995, 2000, 2005, 2010, and 2015. For 1990–2010, the data were in an RT90 25 gon väst projection, while for 2015, they were in SWEREF99 TM. First, the known population of each grid cell was assigned to the cell’s centroid and the centroids were then re-projected to both WGS1984 and World Mollweide projections to match the projection of the candidate population grids.

Methods

Overlay

To compare the known and candidate population distributions, each candidate population grid was overlaid with the known population grid centroids for the corresponding year and appropriate projection. The total known population within each candidate population cell was then calculated as the sum of the known population centroids which fell within each candidate population cell. This produced an aggregated population grid with cell by cell totals for candidate and known populations.

Comparison statistics

Comparison statistics were then calculated between the known (k) and candidate (g) populations for each dataset and year. The comparison statistics used were the percent mean absolute error (%MAE), shown in Eq. (1), the percent root mean square error (%RMSE), shown in Eq. (2), Pearson's r , the percentage of cells correctly identified as either populated or unpopulated, and the relative differences, shown in Eq. (3).

$$\%MAE = \frac{\frac{1}{n} \sum_{j=1}^n |k_j - g_j|}{\frac{\sum_{j=1}^n (k_j)}{n}} \quad (1)$$

$$\%RMSE = \frac{\sqrt{\frac{1}{n} \sum_{j=1}^n (k_j - g_j)^2}}{\frac{\sum_{j=1}^n (k_j)}{n}} \quad (2)$$

$$\text{relative difference} = \frac{(g - k)}{(g + k)} \quad (3)$$

The %MAE and %RMSE measure the absolute fit between the known and candidate populations, with the %RMSE “penalizing” outliers since they stand out more than when looking only at the %MAE. The percent values are a standardization of the MAE and RMSE values by the average known population within the cells. They were deemed necessary in order to allow for cross-dataset comparisons, since the cell size differs between datasets (due to different projections and resolutions) and thus the sizes of the study regions differed.

Pearson's r was used as a measure of the linear association between the known and candidate populations.

In order to calculate the percentage of cells correctly identified as populated or unpopulated, the known and candidate datasets were converted to Boolean datasets, where cells with populations greater than zero were deemed populated and cells with populations equal to zero were deemed unpopulated. The percent correctly populated is the number of cells that were identified as populated in both the candidate and known datasets, divided by the number of known populated cells, while the percent correctly unpopulated cells is the same calculation with unpopulated cells replacing populated cells. The total percent correct is the sum of the number of cells identified as populated in both datasets and the number identified as unpopulated in both divided by the total number of cells.

The relative differences were used in order to evaluate the differences between the known and candidate populations in each cell relative to the size of the population in each cell, rather than looking at the absolute differences. The relative difference ranges from -1 to 1 . Negative relative difference values indicate that the candidate population underestimates the known population; that is, the candidate population is lower than the known population. Positive values indicate that the candidate population overestimates the known population. Values of -1 occur where the candidate population is estimated to be zero, but the cell is in fact populated based on the known population. Values of 1 are the opposite; in these cells, the known population is zero, but the candidate population has assigned the cell as populated. A relative difference of zero indicates that the known and candidate populations are equal. The relative differences were only used for mapped comparisons of the datasets.

Population density comparison

The population density for each cell in the known population was calculated as the population count divided by the area of the cell in km^2 . This is, of course, an approximation of population density, since cells can be partially covered by non-land areas. The cells were then divided into three groups: high, low, and zero population density. The high and low population density groups were divided using three different cutoffs as a robustness check. The cutoffs were determined as follows:

- Average population density for Sweden for the period from 1990 to 2015: 22 people/ km^2
- Average global population density for the period from 1990 to 2015: 50 people/ km^2
- Average of the average country population density for the period from 1990 to 2015: 382 people/ km^2

The spatial distribution of high- and low-density cells for each of these cutoffs is consistent with Swedish population centers (tätort) and divisions between urban and rural areas. Figure 2 shows the population distribution for the known population for the medium cutoff of 50 people/ km^2 during 2015.

The comparison statistics were then calculated for the high, low, and zero density groups separately.

Population change comparison

Known population change was calculated for all but the first year of data for each dataset. Population change was calculated as the difference between the population in each cell relative to the first year available for that dataset. It was determined that 70 to 95% of cells experienced no change over time. The cells were then divided into growth, decline, and stable cells corresponding to cells with change above zero, below zero, and equal to zero. Comparison statistics were then calculated separately for each of the three change groups. Additionally, known population cells with zero population were removed and the comparison statistics were calculated for the three groups using only these known populated cells.

Fig. 2 Reference map of Sweden with population density. The medium cutoff of 50 people/km² for population density classification is shown

Results

Here, we present selected results from our analysis, where we identify the most common trends and, in many cases, aggregate findings to the dataset level. However, we present the full comparison statistics (by dataset and year) in the Supplementary Materials.

Cell by cell comparison by candidate dataset

Table 3 shows the average comparison statistics for each candidate dataset. The full results (by dataset and year) can be seen in Table A1. Our analysis revealed that, of the candidate datasets studied, GHS-POP, LandScan, and WorldPop performed better than GRUMP and GPWv4 on a cell by cell basis. The lowest errors (%MAE and %RMSE) were found in GHS-POP, followed by WorldPop, LandScan, GPWv4, and finally GRUMP.

Figure 3 illustrates the association between the known and candidate populations by dataset. Pearson's r ranged from 0.81 to 0.85 for GHS-POP, from 0.65 to 0.68 for GPWv4, from 0.45 to 0.46 for GRUMP, from 0.61 to 0.75 for LandScan, and from 0.80 to 0.83 for WorldPop. For all datasets, Pearson's r increased slightly over time, showing that the linear association between the known and candidate populations was stronger for more recent years. The low association seen with GRUMP clearly illustrates the fact that the GRUMP methodology uses less spatially detailed population inputs and a limited set of ancillary data resulting in less refined population estimates per cell. Essentially, high population densities are captured less accurately since population is disaggregated over often larger areas, resulting in a lower maximum population per cell.

The %RMSE and %MAE also revealed that LandScan 2000 was an outlier as compared to the other LandScan datasets studied, with significantly higher values. Because the LandScan algorithm is modified and improved over time, this indicates a possible change in the algorithm sometime between 2000 and 2005. Figure 4 shows the four LandScan datasets studied. A visual examination of the LandScan 2000 dataset as compared to the 2005 and later datasets suggests that fewer zero values were assigned in 2000 since a greater number of known unpopulated cells were overestimated and that roads played a more prominent role, possibly leading to increased error in the 2000 dataset. Relative difference maps for the other four datasets can be seen in Figure A1.

Table 3 Average comparison statistics by dataset for cell by cell comparison

Dataset	%MAE	%RMSE	Pearson's r	Percent correct		
				Total	Populated	Unpopulated
GHS-POP	0.74	6.02	0.83	78.2	24.9	95.7
GPWv4	1.27	9.54	0.67	18.3	99.8	3.8
GRUMP	1.45	11.09	0.45	18.6	99.3	3.6
LandScan	0.9	9.18	0.7	81.1	74.4	82.2
WorldPop	0.84	7.41	0.82	17.6	99.9	3

Fig. 3 Association between known and grid estimated populations by dataset. Scatter plot of the known population and grid estimated population for each dataset with all years grouped together. The dashed line indicates a 1:1 linear association

Examining each candidate dataset's ability to correctly identify populated and unpopulated areas revealed that overall, WorldPop, GPWv4, and GRUMP datasets incorrectly identified on average only 18% of cells as either populated or unpopulated, while GHS-POP identified 78% and LandScan 81%. Here, again, LandScan 2000 stood out, with a prediction rate of 63% as compared to the other LandScan datasets which ranged from 85 to 88%. The low values for WorldPop, GPWv4, and GRUMP can be attributed to the fact that they assigned population values to most cells except for waterbodies, so while they predicted approximately 100% of populated cells, their ability to predict unpopulated cells was quite low. GHS-POP on average predicted only 25% of populated cells correctly but 96% of unpopulated cells correctly, while these values for LandScan were 74% and 82% respectively.

Taking these findings into account, we removed cells which were known to be unpopulated from the error analysis and found that the %RMSE and %MAE declined in all cases, but most significantly for WorldPop. Applying a threshold for population so that, for example, only cells with at least one estimated person per cell are counted as populated is therefore recommended. Additionally, aggregating grid populations to higher resolutions should also reduce this error.

Cell by cell comparison grouped by population density

To nuance these comparisons, we also compared the candidate datasets under different known population density conditions. First, we divided the cells into high, low, and zero density cells based on the known population, as explained in the “[Methods](#)” section. Table 4 shows the average comparison statistics for each candidate dataset using the medium cutoff of 50 people/km² to distinguish low and high population densities. The full results (by dataset and year as well as under all three cutoff values) can be seen in Table A2. For all datasets, we found that both the %RMSE and %MAE were highest for low-density areas and lowest for high-density areas. Pearson's r was highest in the high density cells and lowest in the low-density cells. This indicates that datasets are better at consistently estimating the population of high density cells than of low-density cells. Looking at averages for all years for each dataset, WorldPop had the lowest values of %RMSE and %MAE in low-density areas and GHS-POP in high-density areas.

Focusing on the percentage of cells correctly guessed as either populated or unpopulated, we found that GHS-POP and LandScan were much better at identifying populated cells in high density areas than in low-density areas; however, LandScan identified a higher percentage. GPWv4, GRUMP, and WorldPop identified almost

Fig. 4 Relative difference between LandScan estimated population and the known population. **a** LandScan 2000, **b** LandScan 2005, **c** LandScan 2010, **d** LandScan 2015

100% for both density levels. Both of these findings are consistent with previous findings for the entire dataset.

Cell by cell comparison grouped by population change

We also divided the cells into declining, growing, and stable population cells based on the known population, as outlined in the “[Methods](#)” section previously. We did this in order to test whether the candidate datasets were affected by changes in population. For example, whether fast growing regions could be accurately captured or whether population decline affected accuracy. We found that both the %RMSE and %MAE

Table 4 Average comparison statistics by dataset, year, and population density

Dataset	Density	%MAE	%RMSE	Pearson’s <i>r</i>	Percent correct
GHS-POP	High	0.54	1.1	0.83	78.2
	Low	1.77	8.18	0.08	17.1
GPWv4	High	0.76	1.68	0.7	100
	Low	1.99	7	0.08	99.8
GRUMP	High	0.84	1.99	0.47	99.5
	Low	2.38	6.82	0.11	99.2
LandScan	High	0.67	1.63	0.68	93
	Low	1.74	7.69	0.09	69.9
WorldPop	High	0.56	1.30	0.82	99.9
	Low	1.54	6.52	0.09	99.9

were lowest for growing cells, followed by declining cells and then stable cells in all datasets. Consistent with this finding, Pearson's r was lowest for stable cells, followed by declining cells and then growing cells. However, the difference (especially between growing and declining cells) was quite low. It therefore appears that population change (in Sweden at least) does not greatly affect the performance of the candidate datasets. The results of the population change comparison are shown in Table A3.

Discussion and conclusions

With the rising popularity of gridded population data and the wide variety of data providers and methodologies, it has become very important for researchers and data users to evaluate datasets in order to select those best suited for their needs, a fact that has begun to garner more and more attention in the literature (Kugler et al. 2019; Leyk et al. 2019). However, as previously discussed, there is still a lack of validation of gridded population data against known population data.

Leyk et al. (2019) provide an in-depth examination of many of the same gridded population datasets as are presented in this paper. They examine their fitness for use based on the methodologies used to generate each of the datasets and their input data. Their conclusions include that users should consider the required spatial and temporal resolution, whether the study examines urban or rural population, whether it examines residential or ambient population, and whether ancillary data from the gridded datasets are endogenous to the research question.

Our research supports their analysis by quantifying the similarities and differences between five gridded population datasets and known population data in Sweden over a 25-year period. This is a novel contribution in that such comparisons are rare in the literature. First, such temporal comparisons are missing and second, comparisons of several datasets are also missing. Additionally, pixel level comparisons are unusual, since such fine-grained reference data are often unavailable. By exploring pixel level differences, we were able to study both the ability of each candidate dataset to identify populated and unpopulated areas as well as their error, as compared to a known population distribution at their place of residence.

Our findings are consistent with the methodologies employed to create each of the studied datasets. Because they are more highly modeled, using a wider range of ancillary data as input, GHS-POP, LandScan, and WorldPop were able to more accurately estimate and model the known population. This is because they use more information when allocating the location of population and as such can more specifically assign population. However, the fact that they include more input ancillary data should be taken into account by users. These datasets should not be used to model phenomena which have gone into them as ancillary data, such as urban extents, since these variables would be endogenous in the input. GPWv4 was seen to be more reliable and accurate than GRUMP, as expected by the fact that the former has superseded the latter. While less accurate than GHS-POP, LandScan, and WorldPop in many situations, GPWv4 has the advantage that it does not use any ancillary data apart from population.

We showed that the datasets which employ dasymetric mapping methodologies (apart from GRUMP) have a stronger linear association with the known population and

consistently have lower estimation errors. WorldPop, GPWv4, and GRUMP excelled at identifying populated areas; however, this was at the expense of identifying unpopulated areas correctly. GHS-POP predicted unpopulated areas well; however, it tended to over-predict them, meaning that its accuracy in predicting populated areas decreased. Current LandScan datasets (after 2005) predicted both populated and unpopulated areas consistently well; however, those around the year 2000 were less accurate. Taking this into account, we showed that WorldPop had the lowest error, regardless of population density or population change over time, for populated areas. If, however, one included unpopulated areas, GHS-POP had a lower error in general and LandScan had a lower error in stable areas.

Emergency planning and resource allocation are two examples of applications where information about populated and unpopulated areas could be relevant. In such applications, the WorldPop dataset could be applied, along with a threshold, for determining populated areas and ultimately calculating affected population. Without the use of a threshold LandScan and GHS-POP would also be more appropriate choices to model population.

It is, of course, important to keep in mind when evaluating these results, the different methodologies and input ancillary data used for generating each of the candidate population datasets. Not all datasets, for example, represent nighttime population, which the known population does. Therefore, it might be expected that datasets such as LandScan and GHS-POP may have larger discrepancies than, for example, GPWv4. This is most likely in local analyses. For example, LandScan may perform worse than other datasets in a downtown area with a high amount of employment land uses but very little residential uses (when comparing to known nighttime population data). However, this does not appear to have affected the global statistics we have calculated, since GHS-POP, LandScan, and WorldPop have consistently shown higher accuracy in approximating the known population. Likewise, the algorithm used to produce LandScan gridded data is updated and refined yearly. As such, time-series evaluations should not use LandScan data regardless of the accuracy of the individual datasets in estimating true populations.

Care should be applied when working in low-density areas, since all datasets were shown to be less reliable in such situations. WorldPop performed best in low-density areas and would therefore be an appropriate choice in applications requiring population estimation in such areas. However, an understanding of the accuracy and reliability of the input ancillary data would be highly relevant as poor input data may have a disproportionate effect in low-density areas where population estimation was already less accurate even with good input data.

Population change did not appear to affect dataset accuracy extensively. However, it is important to note that the Swedish population is generally quite stable, so this finding may not be applicable in other regions. We hypothesize that population change may play a larger role in dataset accuracy in regions experiencing a high level of population growth or decline. This is especially relevant, we believe, for datasets with many ancillary inputs as these datasets can have a time lag in representing population change. For example, development and deterioration of infrastructure, tied to population growth and decline respectively may not be directly reflected in ancillary data (especially satellite data) and may therefore not be directly reflected in population allocation with more highly modeled dasymetric datasets. This is something we hope to come back to in future research.

A question that arises from this work and that we have touched on somewhat is whether the findings are also applicable to other countries, especially those with characteristics that differ from those of Sweden. This article has addressed model performance in a context with excellent input ancillary and population data. As such, we have been able to test the quality of input data from the five datasets and found their performance to vary. We posit that, in other contexts, where high-quality ancillary data on, for example, infrastructure does not exist or where the input population figures are not as accurate, these five datasets will not perform to the same level as they do in Sweden (although this is something we cannot prove here). The contribution of our findings, however, does not lie in the level of accuracy of each of the datasets, but in the comparison between them and in exploring how differing inputs and methodologies perform in comparison to one reference dataset. We would expect the rank order of our findings to hold in areas with similar quality input data; however, in areas with “poorer” input data, this order may change. For example, the accuracy of the more highly modeled datasets may decrease because of the limited availability of input data and this may lead to a situation where choosing a more highly modeled dataset may, in fact, be inappropriate. Our aim in future research is to take these findings further and examine whether the same rank order of performance is present in other contexts both at national and subnational scales.

It is also worth to point out that even in the Swedish context, these datasets can be improved. As we have mentioned throughout, the quality and type of input data have an effect on the accuracy of population estimation. The ability to, for example, detect built-up settlement areas accurately has an effect on the population allocation results of many of the datasets discussed here. Likewise, the accuracy of transportation network datasets or the ability to detect transportation networks can have an impact. Improved data inputs will lead to better outputs and such improvements are ongoing. The POPGRID Data Collaborative, mentioned in the introduction for this article, contains many great examples of developments within the field of settlement detection and population estimation. However, it is important to also remember that improvement of input ancillary data will only improve population estimation using areal interpolation techniques to the extent that the input population data is accurate.

Our study has important implications for the use of gridded population data in research, policy, and development. The Sustainable Development Goals identify quality disaggregated data as an integral component of progress measurement in order to meet their goal that no one is left behind. The five global gridded population datasets studied are such quality datasets, each with strengths and weaknesses, suited for certain situations or areas, but not others. This research provides a starting off point for an informed choice when it comes to disaggregated population data and should aid researchers in choosing the most appropriate dataset to meet their particular needs.

Acknowledgements Open access funding provided by Lund University.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not

included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Balk, D., Pozzi, F., Yetman, G., Deichmann, U., & Nelson, A. (2005). The distribution of people and the dimension of place: methodologies to improve the global estimation of urban extents. In *Proceedings of the Urban Remote Sensing Conference*. Tempe, AZ: International Society for Photogrammetry and Remote Sensing.
- Balk, D. L., Yetman, G., & de Sherbinin, A. (2010). *Construction of gridded population and poverty data sets from different data sources*. Paper presented at the European forum for Geostatistics Conference, Tallinn, Estonia.
- Bhaduri, B., Bright, E., Coleman, P., & Urban, M. L. (2007). LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal*, 69(1), 103–117. <https://doi.org/10.1007/s10708-007-9105-9>.
- Calka, B., & Bielecka, E. (2019). Reliability analysis of LandScan gridded population data. The case study of Poland. *ISPRS International Journal of Geo-Information*, 8(5), 222.
- Center for International Earth Science Information Network (CIESIN). (2018a). *Compare data*. Retrieved from: <https://www.popgrid.org/compare-data>
- Center for International Earth Science Information Network (CIESIN). (2018b). *Documentation for the gridded population of the world, version 4 (GPWv4), revision 11 data sets*. Retrieved from Palisades NY: NASA Socioeconomic data and Application Center (SEDAC):
- Center for International Earth Science Information Network (CIESIN). (2018c). Gridded population of the world, version 4 (GPWv4): population count adjusted to match 2015 revision of UN WPP country totals, revision 11. Retrieved from: <https://doi.org/10.7927/H4PN93PB>.
- Center for International Earth Science Information Network (CIESIN). (2018d). Gridded population of the world, version 4 (GPWv4): population count, revision 11. Retrieved from: <https://doi.org/10.7927/H4JW8BX5>.
- Center for International Earth Science Information Network (CIESIN), International Food Policy Research Institute (IFPRI), The World Bank, & Centro Internacional de Agricultura Tropical (CIAT). (2011). Global rural-urban mapping project, version 1 (GRUMPv1): population count grid. Retrieved from: <https://doi.org/10.7927/H4VT1Q1H>.
- Cheng, Y., Chen, X., Ding, X., & Zeng, L. (2019). Optimizing location of car-sharing stations based on potential travel demand and present operation characteristics: the case of Chengdu. *J Adv Transp*, 2019, 1–13. <https://doi.org/10.1155/2019/7546303>.
- Deichmann, U. (1996). A review of spatial population database design and modeling (Technical Report 96–3). Retrieved from UC Santa Barbara: <https://escholarship.org/uc/item/6g190671>
- Deichmann, U., Balk, D., & Yetman, G. (2001). Transforming population data for interdisciplinary usages: from census to grid. Retrieved from SEDAC, CIESIN: <http://sedac.ciesin.org/downloads/docs/gpw-v3/gpwwdocumentation.pdf>
- Dobson, J. E., Bright, E. A., Coleman, P. R., Durfee, R. C., & Worley, B. A. (2000). LandScan: a global population database for estimating populations at risk. *Photogrammetric Engineering & Remote Sensing*, 66(7), 849–857. doi:0099-1112/00/6607-849\$3.00/0.
- Dou, Y., Huang, Q., He, C., Meng, S., & Zhang, Q. (2018). Rapid population growth throughout Asia's earthquake-prone areas: a multiscale analysis. *Int J Environ Res Public Health*, 15(9), 1893. <https://doi.org/10.3390/ijerph15091893>.
- Doxsey-Whitfield, E., MacManus, K., Adamo, S. B., Pistolesi, L., Squires, J., Borkovska, O., & Baptista, S. R. (2015). Taking advantage of the improved availability of census data: a first look at the gridded population of the world, version 4. *Papers in Applied Geography*, 1(3), 226–234. <https://doi.org/10.1080/23754931.2015.1014272>.
- Elvidge, C. D., Safran, J., Nelson, I. L., Tuttle, B. T., Hobson, V. R., Baugh, K. E., . . . Erwin, E. H. (2004). Area and positional accuracy of DMSP nighttime lights data. In *Remote Sensing and GIS Accuracy Assessment* (pp. 281–292): CRC Press.
- European Commission Joint Research Center (JRC). (2017). Documentation for the GHS population grid, derived from GPW4, multitemporal (1975, 1990, 2000, 2015) (GHS-POP). Retrieved from Ispra, Italy:

- European Commission, Joint Research Center (JRC): https://www.popgrid.org/popgrid_files/ghs_pop_v1_documentation.pdf
- European Commission, Joint Research Centre (JRC) & Columbia University, Center for International Earth Science Information Network (CIESIN). (2015). *GHS population grid, derived from GPW4, multitemporal (1975, 1990, 2000, 2015)*.
- Fisher, P. F., & Langford, M. (1995). Modelling the errors in areal interpolation between zonal systems by Monte Carlo simulation. *Environ Plan A*, 27, 211–224. <https://doi.org/10.1068/a270211>.
- Fisher, P. F., & Langford, M. (1996). Modeling sensitivity to accuracy in classified imagery: a study of areal interpolation by. *Prof Geogr*, 48(3), 299–309. <https://doi.org/10.1111/j.0033-0124.1996.00299.x>.
- Freire, S., Kemper, T., Pesaresi, M., Florczyk, A., & Syrris, V. (2015, 26–31 July 2015). *Combining GHSL and GPW to improve global population mapping*. Paper presented at the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS).
- Freire, S., MacManus, K., Pesaresi, M., Doxsey-Whitfield, E., & Mills, J. (2016). *Development of new open and free multi-temporal global population grids at 250m resolution*. Paper presented at the 19th AGILE Conference on Geographic Information Science, Helsinki, Finland.
- Haw, D. J., Cummings, D. A. T., Lessler, J., Salje, H., Read, J. M., & Riley, S. (2019). Differential mobility and local variation in infection attack rate. *PLoS Comput Biol*, 15(1), e1006600. <https://doi.org/10.1371/journal.pcbi.1006600>.
- Hjort, J., Karjalainen, O., Aalto, J., Westermann, S., Romanovsky, V. E., Nelson, F. E., et al. (2018). Degrading permafrost puts Arctic infrastructure at risk by mid-century. *Nat Commun*, 9(1), 5147. <https://doi.org/10.1038/s41467-018-07557-4>.
- Kugler, T. A., Grace, K., Wrathall, D. J., de Sherbinin, A., Van Riper, D., Aubrecht, C., et al. (2019). People and pixels 20 years later: the current data landscape and research trends blending population and environmental data. *Popul Environ*, 41(2), 209–234. <https://doi.org/10.1007/s11111-019-00326-5>.
- Leyk, S., Gaughan, A. E., Adamo, S. B., de Sherbinin, A., Balk, D., Freire, S., et al. (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. <https://doi.org/10.5194/essd-11-1385-2019>.
- Li, Y., & Liu, X. (2018). How did urban polycentricity and dispersion affect economic productivity? A case study of 306 Chinese cities. *Landsc Urban Plan*, 173, 51–59. <https://doi.org/10.1016/j.landurbplan.2018.01.007>.
- Linard, C., Gilbert, M., & Tatem, A. J. (2011). Assessing the use of global land cover data for guiding large area population distribution modelling. *GeoJournal*, 76(5), 525–538. <https://doi.org/10.1007/s10708-010-9364-8>.
- Lloyd, C. T., Sorichetta, A., & Tatem, A. J. (2017). High resolution global gridded data for use in population studies. *Scientific Data*, 4, 170001. <https://doi.org/10.1038/sdata.2017.1>.
- Melchiorri, M., Pesaresi, M., Florczyk, J. A., Corbane, C., & Kemper, T. (2019). Principles and applications of the Global Human Settlement Layer as baseline for the land use efficiency indicator—SDG 11.3.1. *ISPRS. International Journal of Geo-Information*, 8(2), 96. <https://doi.org/10.3390/ijgi8020096>.
- Mennis, J. (2003). Generating surface models of population using dasymetric mapping. *Prof Geogr*, 55(1), 31–42. <https://doi.org/10.1111/0033-0124.10042>.
- Mennis, J. (2009). Dasymetric mapping for estimating population in small areas. *Geogr Compass*, 3(2), 727–745. <https://doi.org/10.1111/j.1749-8198.2009.00220.x>.
- Mennis, J., & Hultgren, T. (2006). Intelligent dasymetric mapping and its application to areal interpolation. *Cartogr Geogr Inf Sci*, 33(3), 179–194. <https://doi.org/10.1559/152304006779077309>.
- Mondal, P., & Tatem, A. J. (2012). Uncertainties in measuring populations potentially impacted by sea level rise and coastal flooding. *PLoS One*, 7(10), e48191. <https://doi.org/10.1371/journal.pone.0048191>.
- Oak Ridge National Laboratory (ORNL). (2018). *High resolution global population data set*.
- Oak Ridge National Laboratory (ORNL). (2019). Documentation. *LandScan TM Geographic Information Science & Technology*. Retrieved from <https://landscan.ornl.gov/index.php/documentation>
- Openshaw, S. (1984). *The modifiable areal unit problem*. Norwich: Geo Abstracts.
- Pesaresi, M., Melchiorri, M., Siragusa, A., & Kemper, T. (2016). Atlas of the human planet 2016: mapping human presence on earth with the global human settlement layer. *JRC Science for Policy Report, EUR 28116 EN*. doi:<https://doi.org/10.2788/889483>.
- POPGRID Data Collaborative. (2020). Our mission. Retrieved from <https://www.popgrid.org/about-us>
- Rose, A. N., & Bright, E. (2014). *The LandScan global population distribution project: current state of the art and prospective innovation*. Paper presented at the Population Association of America 2014 Annual Meeting, Boston, MA.

- Siraj, A. S., Rodriguez-Barraquer, I., Barker, C. M., Tejedor-Garavito, N., Harding, D., Lorton, C., et al. (2018). Spatiotemporal incidence of Zika and associated environmental drivers for the 2015–2016 epidemic in Colombia. *Scientific Data*, 5, 180073. <https://doi.org/10.1038/sdata.2018.73>.
- Sordo-Ward, A., Granados, I., Iglesias, A., & Garrote, L. (2019). Blue water in Europe: estimates of current and future availability and analysis of uncertainty. *Water*, 11(3), 420. <https://doi.org/10.3390/w11030420>.
- Statistiska centralbyrån. (2019a). *Befolkningsstatistik: Nyckeltal för Sverige*.
- Statistiska centralbyrån. (2019b). *Invånare per kvadratkilometer efter kön och år [BE0101U1]*.
- Statistiska centralbyrån. (2019c). *Statistiska tätorter 2018: Antal, befolkning och arealer*.
- Statistiska centralbyrån. (n.d.). *Statistik om invandring – frågor och svar*. Retrieved from <https://www.scb.se/hitta-statistik/statistik-efter-amne/befolkning/befolkningens-sammansattning/befolkningsstatistik/produktrelaterat/Fordjupad-information/statistik-om-invandring%2D%2D-fragor-och-svar/>
- Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PLoS One*, 10(2), e0107042. <https://doi.org/10.1371/journal.pone.0107042>.
- Tatem, A. J., Campiz, N., Gething, P. W., Snow, R. W., & Linard, C. (2011). The effects of spatial population dataset choice on estimates of population at risk of disease. *Popul Health Metrics*, 9(1). <https://doi.org/10.1186/1478-7954-9-4>.
- Weiss, D. J., Nelson, A., Gibson, H. S., Temperley, W., Peedell, S., Lieber, A., et al. (2018). A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature*, 553, 333–336. <https://doi.org/10.1038/nature25181>.
- World Bank. (2018). World development indicators. Population growth (annual %).
- WorldPop. (2018). *Global High Resolution Population Denominators Project - Funded by The Bill and Melinda Gates Foundation (OPP1134076)*.
- Wu, S., Qiu, X., & Wang, L. (2005). Population estimation methods in GIS and remote sensing: a review. *GIScience & Remote Sensing*, 42(1), 80–96. <https://doi.org/10.2747/1548-1603.42.1.80>.
- Zandbergen, P. A., & Ignizio, D. A. (2010). Comparison of dasymetric mapping techniques for small-area population estimates. *Cartogr Geogr Inf Sci*, 37(3), 199–214. <https://doi.org/10.1559/152304010792194985>.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Maria Francisca Archila Bustos¹ · Ola Hall¹ · Thomas Nedomysl^{2,3} · Ulf Ernstson⁴

¹ Department of Human Geography, Lund University, Sölvegatan 10, 223 62 Lund, Sweden

² Department of Analysis and Coordination, Region Halland, Halmstad, Sweden

³ Department of Human Geography, Stockholm University, Stockholm, Sweden

⁴ School of Business, Economics and IT, University West, Trollhättan, Sweden