Cities 166 (2025) 106293

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities

Assessing the impacts of gridded population model choice on degree of urbanisation metrics

Wen-Bin Zhang[a,*], Dorothea Woods[a], Iyanuloluwa Deborah Olowe[a], Marcello Schiavina[b], Weixuan Fang[c], Graeme Hornby[a], Maksym Bondarenko[a], Joachim Maes[d], Lewis Dijkstra[e], Andrew J. Tatem[a], Alessandro Sorichetta[f]

[a] WorldPop, School of Geography and Environmental Science, University of Southampton, Southampton SO17 1BJ, UK
[b] NTT DATA, Rue de Spa 8, 1000 Bruxelles, Belgium
[c] School for policy studies, University of Bristol, Bristol BS8 1TZ, UK
[d] European Commission, Directorate-General for Regional and Urban Policy, 1049 Brussels, Belgium
[e] Directorate for Fair and Sustainable Economy, European Commission Joint Research Centre (JRC), Ispra, Italy
[f] Dipartimento di Scienze della Terra "Ardito Desio", Università degli Studi di Milano, Via Mangiagalli 34, 20133 Milan, Italy

ARTICLE INFO

Keywords:
Degree of urbanisation
Gridded population data
Urbanisation metrics
Sensitivity analysis
Settlement classification

ABSTRACT

Defining urban and rural areas is crucial for assessing the accessibility of services and opportunities that impact people worldwide. The Degree of Urbanisation framework, endorsed by the UN Statistical Commission, primarily uses population grids to classify areas through a harmonised, population-centric approach, enabling international comparisons. However, variations in the distribution of population counts at the grid-cell level across different population datasets can significantly influence the resulting patterns. We applied the Degree of Urbanisation to 16 countries in Africa and the Caribbean, using four population grids to evaluate these effects. It shows that differences primarily occur in the classification of urban cluster. On average, 27.5 % of the population falls into mixed classes, with 17.5 % in mixed rural and urban cluster areas and 7.8 % in mixed urban cluster and urban centre areas. Population grids that only model populations within satellite-detected settlements show limited disagreement, with mixed rural and urban cluster population classifications decreasing by 5.6 percentage points and mixed urban cluster and urban centre populations by 1.4. Population modelling approaches that distribute populations more broadly, including outside of detected built-up areas, substantially reduce settlement identifications, resulting in 42.3 % fewer urban centres and 66.2 % fewer dense urban clusters than the average across the study countries. Our analyses highlight the potential sensitivity of Degree of Urbanisation to gridded population modelling assumptions and provide guidance on its implementation.

1. Introduction

Defining and delineating urban areas, as well as characterizing the entire rural-urban continuum, is essential for effective urban planning, resource allocation, decision- and policy-making. For example, both planners and geographers need it to map population (Rozenfeld et al., 2008), organise and design transport systems (El-Bouayady et al., 2024), characterise LULC changes (AlDousari et al., 2022; Shohan et al., 2024) and make comparisons across cities and countries (Taubenböck et al., 2024). The UN World Urbanisation Prospects (United Nations, 2019a) offers a comprehensive overview of global urbanisation trends, highlighting the significant variations in "urban" definitions across countries, based on population size, density, economic function, administrative boundaries, and infrastructure. For example, Denmark uses a population size threshold of 200 inhabitants to define urban areas, while in Japan an urban area is defined by having a population of at least 50,000 inhabitants, with over 60 % of households engaged in commerce, services, or manufacturing. Moreover, 37 countries (including India, Bangladesh, Russia, Tanzania and Zimbabwe) use agricultural employment shares in their national urban definition, making it impossible to reliably monitor the changes of agricultural employment in such urban and rural areas, as fluctuating agricultural employment may cause areas being reclassified to a different class.

Definitions of cities are typically aligned with local administrative

* Corresponding author.
E-mail address: wb.zhang@soton.ac.uk (W.-B. Zhang).

https://doi.org/10.1016/j.cities.2025.106293
Received 27 December 2024; Received in revised form 16 April 2025; Accepted 14 July 2025
Available online 21 July 2025

units (LAUs), such as in the Rural-Urban Continuum Codes (Butler, 1990) and the Urban Influence Codes (Ghelfi & Parker, 1997), developed and used by the US Department of Agriculture to distinguish metropolitan from non-metropolitan counties. Additionally, municipalities can be designated as urban areas by administrative decree, such as in Brazil and China. There are also many academic urban definitions, mainly based on density measures (De Bellefon et al., 2021; Niu & Silva, 2021), extent of built-up area (Almulhim et al., 2024; Angel et al., 2016), classification of satellite imageries (Alajizah & Altuwaijri, 2024), or complex agglomeration indexes (Uchida & Nelson, 2010) – most of which are applied to a single country or to an exogenous and unharmonised list of cities, and therefore are not globally applicable. This diversity of definitions and LAUs with various shapes and sizes across countries hinders the possibility of performing consistent and comparable urban analyses at the global scale, especially considering that LAUs can change over time within the same country, leading to further inconsistencies in comparative analysis (Taubenböck et al., 2022).

The 2030 Agenda for Sustainable Development requires reporting of progress towards the Sustainable Development Goals (Klopp & Petretta, 2017; Moyer & Hedden, 2020) and comparison of socio-economic conditions for urban and rural regions separately (Menon et al., 2000; Stanny et al., 2021), which in turn has highlighted the need for a globally harmonised definition of rural-urban typologies and the development of new cross-national methods. To this end, the Degree of Urbanisation (Dijkstra et al., 2021; European Commission: Eurostat, 2021) (DEGURBA) was initiated and developed by the European Union, the Food and Agriculture Organization of the United Nations (FAO), the International Labour Office (ILO), the Organization for Economic Cooperation and Development (OECD), UN-Habitat and the World Bank in 2016, and endorsed by the United Nations Statistical Commission (UN-SC) in 2020 (United Nations Statistical Commission, 2020). The DEGURBA approach uses primarily population density and population size to categorise 1 km grid cells as urban centre grid cells, urban cluster grid cells, and rural grid cells. Adopting grid cells as its base unit (instead of LAUs) to define a cross-national methodology can minimize the impact of substantially different LAUs in terms of shape and size across countries (Hijmans, 2009; Runfola et al., 2020) (i.e., the Modifiable Areal Unit Problem (Openshaw, 1984); MAUP), facilitating international statistical comparisons. Beside population density and size, DEGURBA uses grid cell class contiguity to form clusters of adjacent grid cells sharing a minimum density and having a total cluster population size, which determines, according to different criteria, their DEGURBA class. An urban centre is defined as contiguous grid cells with a density above 1500 residents per km^2 of permanent land and at least 50,000 people within the centre; an urban cluster is defined as contiguous grid cells with a density above 300 residents per km^2 of permanent land and at least 5000 people in the cluster; and rural areas are all grid cells with a density below 300 residents per km^2 of permanent land and all the other grid cells outside urban clusters and centres. The DEGURBA methodology proposes a more refined classification of the urban-rural continuum with a nested second level classification delineating the outline of urban centres, dense and semi-dense urban clusters (nested in the urban clusters) and rural clusters (within rural areas) (European Commission: Eurostat, 2021).

According to this definition, the DEGURBA methodology relies heavily on the input 1 km gridded population dataset, which in turn can vary significantly based on the type of population data and modelling strategy used to produce it (Leyk et al., 2019). Despite a difference of 10 people per km^2 may be negligible for many applications, such variations, if close to the pre-defined population density thresholds of the DEGURBA, can significantly influence the resulting settlement classification and consequently the policy decisions informed by it. However, there have been limited investigations of these variations on the DEGURBA outputs, especially for low- and middle-income settings, where the availability of recent, detailed, and reliable small area demographic data, to support the production of accurate gridded population datasets, is often lower than in other settings worldwide (Mikkelsen et al., 2015; Pesaresi et al., 2024; Randall & Coast, 2016).

Here, we assessed the uncertainty of the DEGURBA approach by using four different gridded population datasets, including a set of 1) WorldPop Bespoke National (WPB) datasets, 2) Constrained WorldPop Global (WPC) datasets, 3) Unconstrained WorldPop Global (WPU) datasets, and 4) Global Human Settlement multitemporal population grid (GHS-POP) datasets. First, we conducted a comparison at the grid cell level to assess the uncertainty of population share by DEGURBA using different population datasets. Second, we examined the spatial entities identified by different population datasets in terms of quantity, area, and population. Third, we quantified and analysed the uncertainty inherent in the DEGURBA approach. The analysis was conducted on a selection of 16 countries (see Fig. 1): Burkina Faso (BFA), Cameroon (CMR), DR Congo (COD), Dominican Republic (DOM), Ghana (GHA), Guinea (GIN), Mali (MLI), Mozambique (MOZ), Namibia (NAM), Niger (NER), Nigeria (NGA), Senegal (SEN), Sierra Leone (SLE), South Sudan (SSD), Uganda (UGA), and Zambia (ZMB).

2. Data and methods

In this section, we first provide a detailed description of the study population datasets, and the auxiliary data used to produce DEGURBA outputs alongside each grid population dataset. Then, we outline the harmonisation process applied across the study population datasets. It ensures consistency and comparability in the DEGURBA outputs by aligning datasets with varying spatial resolutions, coordinate reference systems, and national boundaries.

2.1. Data collection

2.1.1. Gridded population datasets

The main data inputs consist of the four 2020 gridded population datasets used to produce the DEGURBA outputs for testing the DEGURBA sensitivity to varying gridded population dataset inputs; they are described below.

GHS-POP (https://human-settlement.emergency.copernicus.eu/ghs_pop2023.php) depicts the residential population at global level (Schiavina et al., 2023) derived from a population dynamics model combining the administrative unit-based population counts, collected and used by the Center for International Earth Science Information Network (CIESIN) for producing the Gridded Population of the World (GPWv4.11), and two ancillary population time series (i.e., the UN World Population Prospects 2022 and the UN World Urbanisation Prospects 2018 population time series at country and urban agglomeration levels) (Freire et al., 2018; Pesaresi et al., 2024). Population counts are disaggregated from administrative units to 100x100m and 1x1km grid cells, informed by the distribution of residential built-up volume as mapped in the Global human Settlement Built-up volume grid (GHS-BUILT-V, Pesaresi and Politis (2023)). The 2020 1x1km resolution version, in a Mollweide projection (ESRI:54009), was used in the framework of this study.

Unconstrained WorldPop Global (https://hub.worldpop.org/geodata/listing?id=69) is produced by using the administrative unit-based population counts, estimates, and projections collected and used by the CIESIN for producing the Gridded Population of the World (GPWv4.11). Administrative unit-based population counts, estimates and projections are disaggregated to ~100x100m grid cells using the Random Forest-based approach described in Stevens et al. (2015) and Sorichetta et al. (2015). The term 'unconstrained' refers to the assumption that no settlement dataset is accurate enough to identify all residential settlements/buildings globally and therefore, although used as a covariate in the Random Forest model, the settlement map is not used as a mask to set predicted population numbers to zero in areas without mapped settlements. The 2020 3 × 3 arc-second resolution

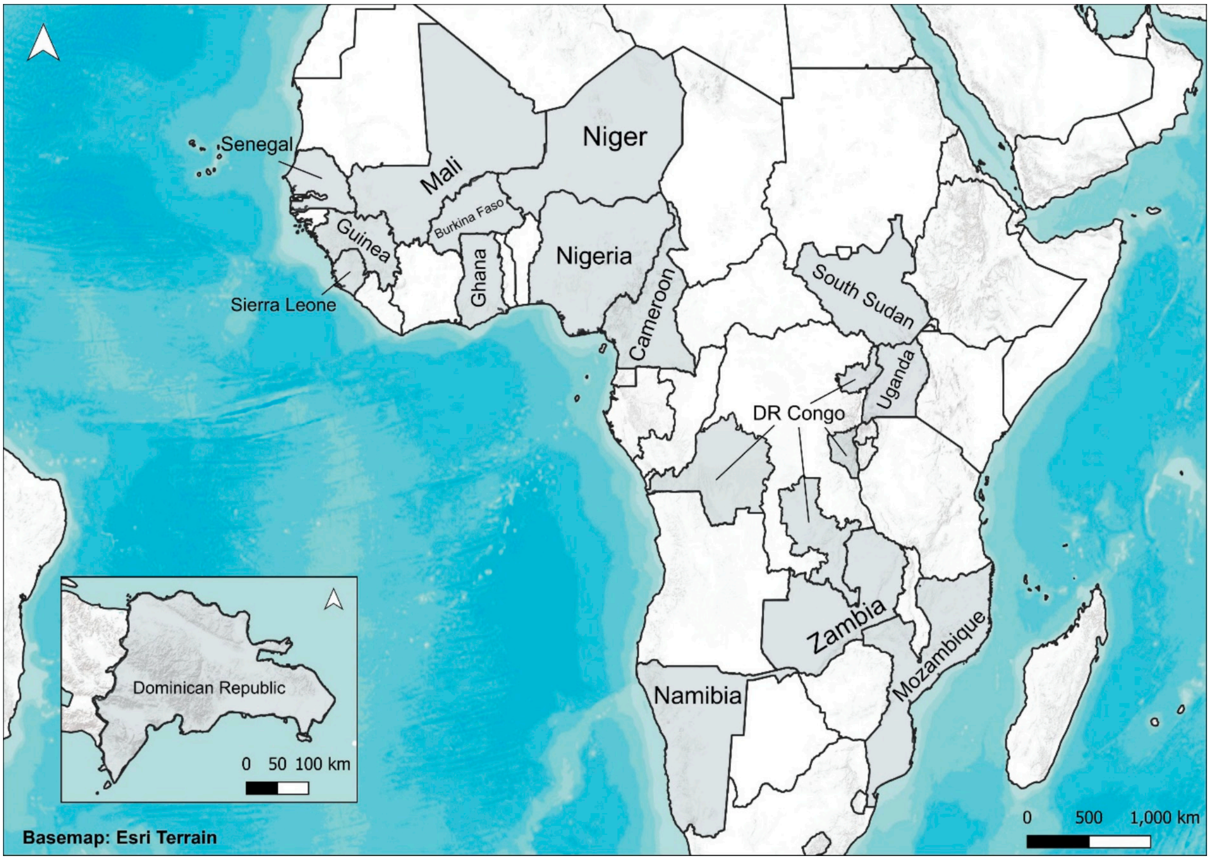

Fig. 1. The 16 study countries. These countries are selected due to the potential high variability of the available gridded population datasets and the lack of official gridded data. Noting that only four provinces of DR Congo were included in this study due to the availability of WorldPop bespoke datasets. Source: The base map is Esri Terrain; the national boundaries are published by the UN World Food Programme. The outputs of the Degree of Urbanisation using different population datasets can be found in Supplementary Information Section A.

version, in a GCS-WGS84 projection (EPSG:4326), was used in the framework of this study.

Constrained WorldPop Global (https://hub.worldpop.org/geodata/listing?id=79) uses the same population data used to produce the Unconstrained WorldPop Global datasets but uses additional satellite-derived building footprint data (Ecopia.AI & Maxar Technologies, 2020) as a mask for 51 African countries, and the outputs of a built settlement growth model for the remaining countries (Nieves et al., 2020) (https://hub.worldpop.org/project/categories?id=15). The Random Forest based modelling approach (Stevens et al., 2015) is then applied to disaggregate population to only those grid cells identified as settled. The 2020 3 × 3 arc-second resolution version, in a GCS-WGS84 projection (EPSG:4326), was used in the framework of this study.

WorldPop Bespoke National (https://wopr.worldpop.org/?/Population) represents a collection of bespoke national estimates produced in collaboration with governments. These represent a mixture of top-down and bottom-up models, each tailored to the needs and data availability of the corresponding country. For the top-down models, these are typically disaggregation of recent censuses or national projections using the Random Forest-based approach referenced above. For the bottom-up models, these are based on either field data collection of small area 'microcensus' data, recent household surveys, or incomplete census data. The models built for each country utilise geospatial covariates including building footprints to estimate populations for 3 × 3 arc-second grid cells with complete national coverage, in a GCS-WGS84 projection (EPSG:4326). Details on the type of bespoke method used for each country can be found in Table 1. Note that a hybrid approach (Wardrop et al., 2018) was used for Ghana and Guinea. Specifically, it used household-level survey data, which do not have full coverage of the country, and projected census totals to constrain population estimates.

2.1.2. Auxiliary data

In addition to the population data, gridded national boundaries are required to provide a common extent for each country, for this purpose the openly accessible rasterised boundaries (https://hub.worldpop.org/geodata/listing?id=62), in a GCS-WGS84 projection (EPSG:4326), with a spatial resolution of 3 × 3 arc-second, used during the production of the WorldPop population datasets were employed. The second auxiliary input is a land share grid describing the proportion of water and land in each grid cell. The land/water grid (Woods et al., 2024), in a GCS-WGS84 projection (EPSG:4326) with a spatial resolution of 3 × 3 arc-second, used in the production of the WorldPop population datasets was used for producing the DEGURBA outputs based on WorldPop's population data. In contrast, the DEGURBA outputs based on GHS-POP were produced using the GHS-LAND layer embedded in the GHS-DUG tool.

2.2. Implementation of degree of urbanisation

All gridded national boundary datasets were first projected to the Projected Coordinate Reference System (PCRS) Mollweide (ESRI:54009) and then resampled to 1 km spatial resolution. This is the required format by the GHS-DUG tool (Maffenini, Schiavina, Melchiorri, et al., 2023). The land grid is also projected to the same PCRS and converted from a binary presence-absence grid to a raster dataset showing the land share of each grid cell of the 1 km gridded national boundaries (Fig. 2). This is important so that during the running of the GHS-DUG tool with WorldPop population datasets later, populations are allocated correctly

Table 1
WorldPop Bespoke National models by country.

	Country	Details of WorldPop bespoke
Africa	Burkina Faso	Bottom-up model based on both census enumeration in 2019 and statistical modelling, where enumeration was not feasible.
	Cameroon	Bottom-up model based on listings from household surveys.
	DR Congo	Bottom-up model for 12 of the 26 provinces based on three rounds of micro-census surveys in 2017, 2018 and 2021.
	Ghana	Hybrid model
	Guinea	Hybrid model
	Mali	Bottom-up model based on a combination of total population counts at enumeration area level and modelled population counts created using a Bayesian statistical model for areas that could not be covered by the cartography team because of security issues. The main input data for the model are the cartography data collected in the safe part of the country in 2019–2020.
	Mozambique	Top-down model based on the projected 2022 population totals for Districts derived from the 2017 Population and Housing Census.
	Namibia	Not available in WorldPop Bespoke National data
	Niger	Top-down model based on Digital Commune boundary and official projected population totals for 2021 based on the 2012 Population and Housing Census.
	Nigeria	Bottom-up model generated using an approach that linked population, observed at 1142 micro-census survey locations collected in 2016–2017.
	Senegal	Top-down model based on the official projected population totals for 2020 derived from the 2013 Population and Housing Census.
	Sierra Leone	Top down based on the 2015 Population and Housing Census.
	South Sudan	Top-down model based on projections from the 2008 Population and Housing Census adjusted for displacement and accounting for refugee camps.
	Uganda	Top-down model based on official projections for 2022 based on the 2014 Population and Housing Census.
	Zambia	Bottom-up model based on household listing from three surveys in 2017, 2018 and 2019.
Caribbean	Dominican Republic	Not available in WorldPop Bespoke National data

within grid cells with available land area (i.e., grid cells partially covered by water or partially belonging to a neighbouring country can only receive population counts proportionally to the available land share).

For each country, each population grid is subjected to a slightly different harmonisation process to ensure that all four grids are consistent in terms of spatial resolution, extent, alignment and population total, enabling comparability of the derived DEGURBA outputs. Fig. 2 gives an overview of the workflow carried out for each input population grid.

As a first step all three population grids from WorldPop were projected to the PCRS Mollweide (ESRI:54009) using the volume preserving POPWARP v3 tool (Maffenini, Schiavina, Carneiro Freire, et al., 2023), with the corresponding gridded national boundary dataset set as the target grid to ensure the spatial alignment of the three output grids. GHS-POP is supplied already in the Mollweide projection (ESRI:54009). For large countries this step proves processing intensive and has to be carried out on a High Performance Cluster (HPC). The HPC was used for Mozambique, Nigeria, Sierra Leone, South Sudan, Uganda, Mali, Namibia, Senegal and DR Congo. In these cases, the built-in option “snap and warp to target grid” of the POPWARP tool could not be used and the population grid was projected to World Mollweide without the use of a target grid.

The second step of the workflow, to ensure spatial alignment and identical pixel extents of all population grids in the Mollweide projection (ESRI:54009), consists into extracting and snapping all population grids to the exact extent of the corresponding gridded national boundary.

The third step only applies to the Unconstrained WorldPop population grid, which by construction should have population distributed across all grid cells with available land share. However, due to the different tools used to project and resample the Unconstrained WorldPop population grid and the WorldPop land/water grid (i.e., POPWARP for the population grid vs. manual project and aggregate for the land share grid), this condition is no longer maintained across the full extent of the study area. Therefore, in such cases, unpopulated grid cells with available land share are “nibbled” to replace the NoData with valid values of their nearest neighbours.

The fourth step is carried out for all three WorldPop population grids and sets populated grid cells to NoData, where there is no land available according to the corresponding land share grid. Also in this case, such mismatch is due to the different tools used to project and resample the

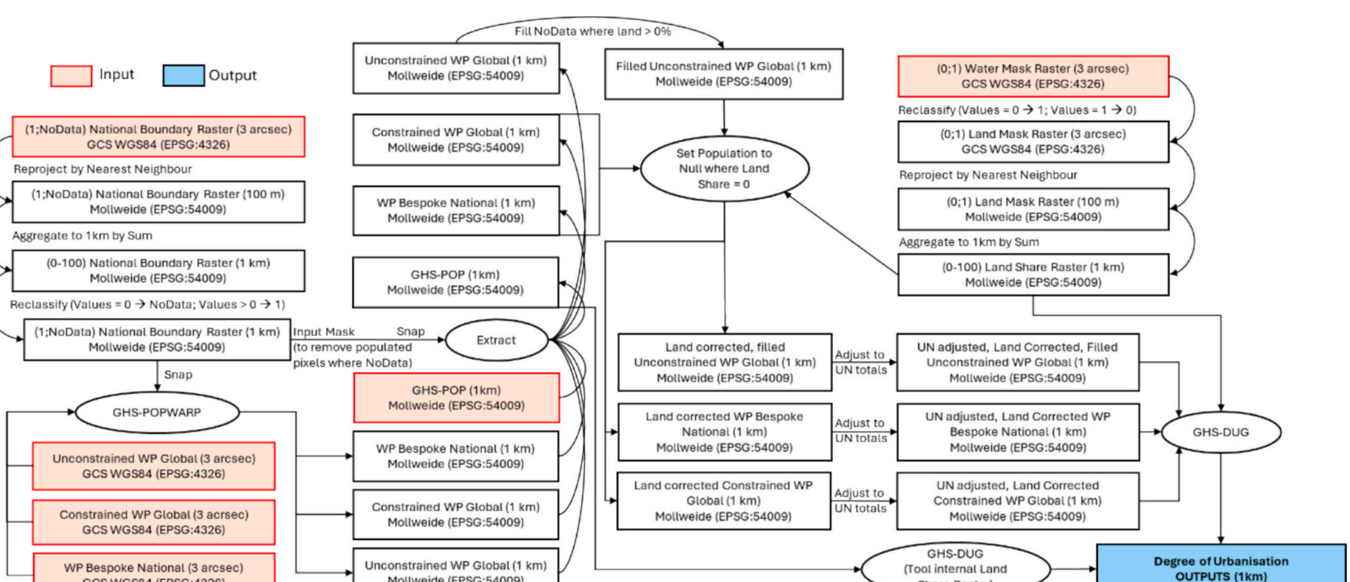

Fig. 2. Harmonisation process of population grids as preparation for production of Degree of Urbanisation.

Unconstrained WorldPop population grids and the WorldPop land/water grid.

The fifth step consists in adjusting, by proportional redistribution, the value of each populated pixel to match the 2020 UN national total (United Nations, 2019b). This is done to ensure that differences in the DEGURBA outputs are solely due to different distributions of the same population total, and not to different distributions of different population total.

Finally, the GHS-DUG v6.1 tool is run with each harmonised population grid as input to obtain comparable DEGURBA grids and spatial entity vector datasets. As already mentioned in section 2.1.2, for GHS-POP the land layer internal to the GHS-DUG tool is used, while for the three WorldPop population grids the land share grid, derived from the WorldPop land/water grid used during the production of the WorldPop population grids, is applied.

3. Results

3.1. Degree of urbanisation profile

When using the four different types of gridded population datasets to produce the DEGURBA outputs, each single grid cell, within the four DEGURBA output datasets, falls into one of the seven classes shown in Fig. 3a, including 1) Urban Centre by all datasets; 2) Urban Cluster by all datasets; 3) Rural by all datasets; 4) Mix of rural and urban centre, classified as both rural and urban centre by different datasets; 5) Mix of rural and urban cluster, classified as both rural and urban cluster by different datasets; 6) Mix of urban cluster and urban centre, classified as both urban cluster and urban centre by different datasets; and 7) Mix of all three classes, where the grid cell is classified as rural, urban cluster, and urban centre simultaneously by different datasets.

The respective population shares by DEGURBA using different population datasets are shown in Fig. 3b. The difference in the total share of urban centre and urban cluster population between WPB and WPC is modest, with an average of 3.6 (IQR: 1.4 to 5.6) percentage points difference, indicating a high degree of similarity between the two models. GHS-POP generally returning the highest urban centre and urban cluster total populations. On average, GHS-POP returns 9.5 (IQR: 6.2 to 11.7) percentage points above the combined average of WPB and WPC. The most extreme county is South Sudan where GHS-POP returns 25 percentage points over the combined average of WPB and WPC. In contrast, WPU produces the lowest population shares for urban centre and urban cluster in almost all the countries. On average, WPU leads to a population share that is 11 (IQR: 7 to 13.6) percentage points below the combined average of WPB and WPC. The most extreme country is still South Sudan where WPU leads to a population share that is 38 percentage points below the combined average of the WPB and WPC.

Fig. 3c illustrates the agreement between the DEGURBA outputs

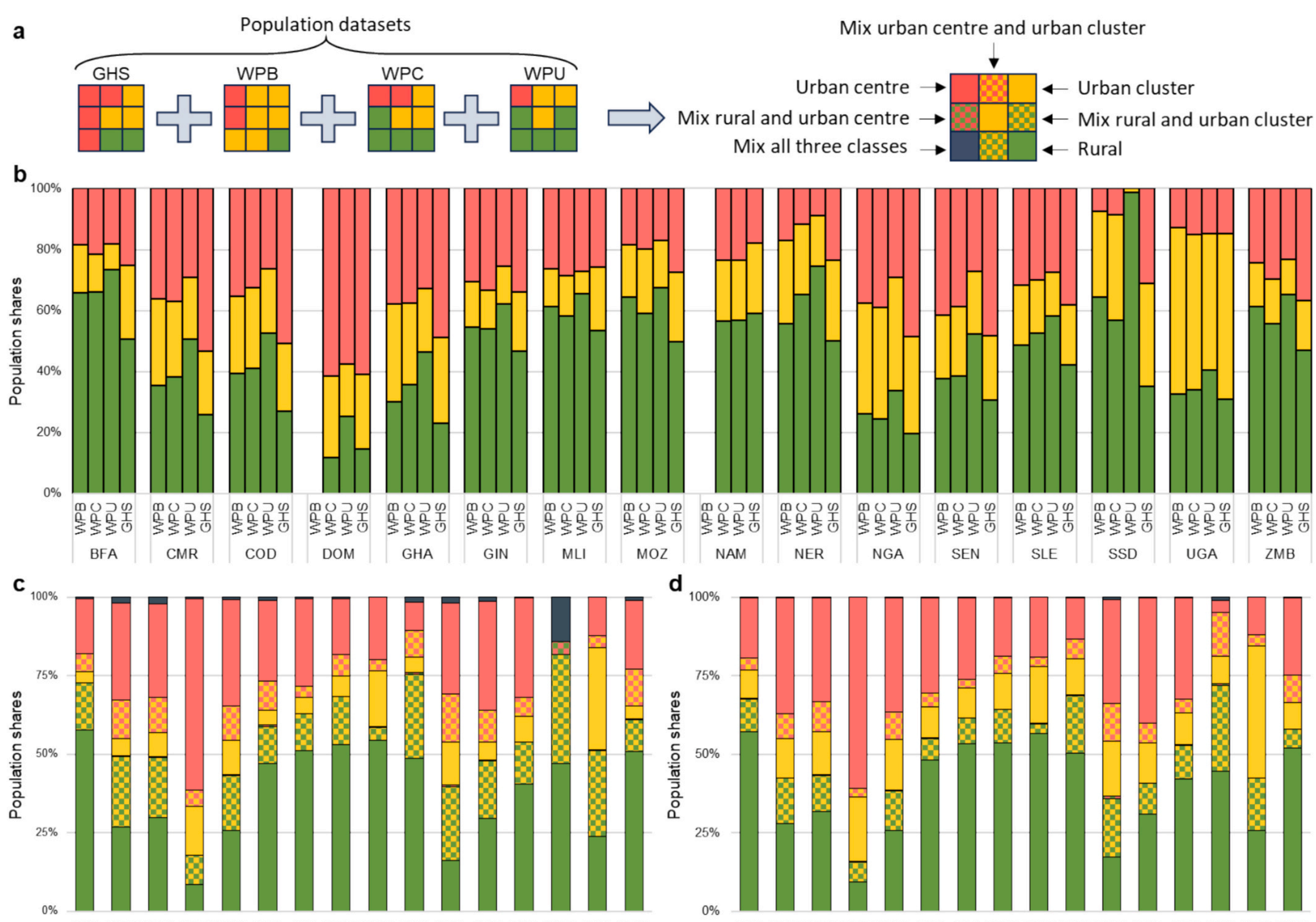

Fig. 3. Country specific Degree of Urbanisation results produced from different population datasets. a. Illustration of seven classes a single grid cell can fall into when using different population datasets to produce Degree of Urbanisation. **b.** Population shares by Degree of Urbanisation within each country using different population datasets. **c.** Population shares by Degree of Urbanisation based on the seven grid classes using the four population datasets. **d.** Population shares by Degree of Urbanisation based on the seven grid classes using three population datasets without Unconstrained WorldPop Global. WPB is WorldPop Bespoke National, WPC is Constrained WorldPop Global, WPU is Unconstrained WorldPop Global, and GHS is GHS-POP.

obtained using the four gridded population datasets in terms of population shares. In these 16 countries, the share of population in mixed classes is 27.5 percentage points on average with IQR 20.9 to 33.8 percentage points. Most disagreements happen in mix rural and urban cluster with 17.5 (IQR: 11.7 to 23.2) percentage points on average, notably making up for Nigeria's 45.1 million (23.4 percentage points) and Uganda's 11.9 million inhabitants (27.5 percentage points). The second large disagreements appear in mix urban cluster and urban centre with 7.8 (IQR: 4.2 to 11.2) percentage points on average. This shows relatively higher sensitivity of DEGURBA in defining urban cluster when comparing the outputs produced using the four different gridded population datasets.

Fig. 3d shows a reduction in the proportion of mixed categories when WPU is excluded. On average in these 16 countries, the share of population in mixed classes is 18.9 percentage points which is 8.7 percentage points lower than using all four population datasets. Mixed classes like mix rural and urban cluster and mix urban cluster and urban centre decrease 5.6 (IQR: 3.7 to 7.9) and 1.4 (IQR: 1.2 to 3.3) percentage points on average, respectively. For example, in Nigeria, mix rural and urban cluster reduces from 23.4 percentage points to 18.5 percentage points of the total population. Mix of all three classes is even further reduced to a negligible proportion, with the highest 1.1 (14.3 with WPU) percentage points in South Sudan across the 16 countries. The reduced presence of mixed classes indicates that, without WPU, the different DEGURBA output datasets align more closely to each other, leading to a clearer separation of grid cells into distinct urban centre, urban cluster, and rural classifications. This suggests, in turn, less variability of the DEGURBA outputs produced using constrained gridded population datasets.

3.2. Uncertainty quantification and analysis

In addition to the population share by DEGURBA, the landscape of DEGURBA using different population datasets may also be different even with the same area shares by DEGURBA. For example, there are two urban centre (red) grid cells in WPB and WPC as shown in Fig. 3a, but these urban centre grid cells are not at the same place, resulting in two different DEGURBA landscapes. In Fig. 4, we represent the uncertainty of landscapes and population shares using the positions of national dots on Cartesian axes. The analysis of uncertainty in DEGURBA classifications reveals distinct patterns among countries when including or excluding WPU. In general, South Sudan exhibits high sensitivity to the gridded population dataset of choice regarding both the resulting population shares and urbanisation landscape. In contrast, countries such as Namibia and Mali have stable DEGURBA classifications, showing minimal sensitivity to the gridded population dataset of choice. Most countries, like Nigeria and Guinea, show high variability in the resulting population shares, with moderate to high agreement in the resulting urbanisation landscape.

Excluding WPU significantly reduces uncertainty across most countries, resulting in more stable classifications. Nigeria's uncertainty in terms of population shares decreases by approximately 64.7 %, and Senegal's by 76.7 %. South Sudan experiences a dramatic reduction of 92.5 %, indicating its high sensitivity to WPU. Ghana shows a 44.8 % reduction, confirming less variability in classifications when WPU is not used. These changes suggest that, for most countries, removing WPU enhances the stability of the DEGURBA classification and reduces its sensitivity to the gridded population dataset of choice. However, Namibia's increase of 60.8 % suggests some exceptions where WPU may

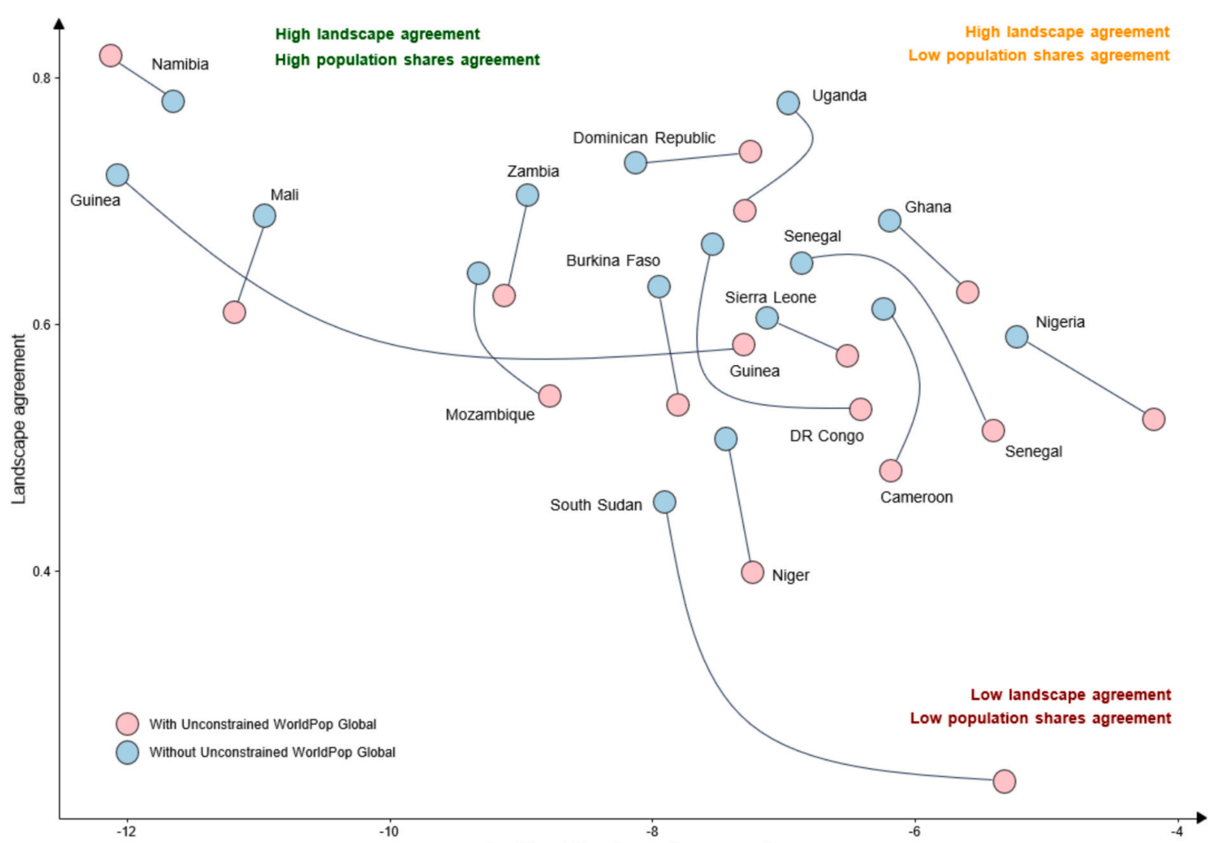

Fig. 4. The overall nation-scale sensitivity of Degree of Urbanisation to different population datasets as input. Population shares disagreement refers to the averaged difference between every two population distributions by Degree of Urbanisation using different population datasets. Landscape agreement refers to the averaged agreement (measured by Cohen's Kappa coefficient) between population datasets to produce Degree of Urbanisation L1 categories at each grid cell. More details and the national uncertainty measurements can be found in Supplementary Information Section B.

provide stability. With respect to the urbanisation landscape sensitivity, results are generally more consistent when implementing DEGURBA with different population datasets apart from WPU. Indeed, when not using WPU, Senegal's landscape agreement increases by 26.5 %, while South Sudan shows a nearly 99 % increase. These increments confirm that WPU may not be intrinsically suited for DEGURBA which relies on population concentrations for the identification of different type of settlements.

Among the 16 considered countries, the DEGURBA approach shows the highest sensitivity to the gridded population dataset of choice for South Sudan. Fig. 5 shows the town of Kapoeta, located in the Eastern Equatoria State of South Sudan, where WPU predominantly classifies grid cells as rural. This pattern extends across almost the whole South Sudan, with WPU classifying about 99 % of grid cells as rural and <1 % as urban, reflecting its limitations when used for implementing the DEGURBA. In contrast, GHS-POP and WPC classified Kapoeta as a urban centre, while the WPB, aligning with South Sudan official definitions, categorized it as a urban cluster since it did not meet the 50,000-resident threshold for an urban centre. The predominance of mixed classes in South Sudan (see Fig. 3) reflects these classification discrepancies and the challenges posed by using outdated population estimates and the presence of widespread displacement (Jordan, 2018). These factors contribute to high variability of the gridded population datasets and highlight the unique challenges of accurately classifying settlements in such a dynamic and evolving context.

3.3. Variation of spatial entities

The DEGURBA method, in particular the GHS-DUG tool, also generates spatial entities (i.e., vector polygons) based on the clusters of grid cells of the same DEGURBA class. Fig. 6 compares the spatial entities created by different population datasets with regards to quantity, area, and population. GHS-POP generally shows the highest estimates for urban areas, especially Urban Centres and Dense Urban Clusters, with positive deviations from the average in quantity (64.0 %) and population (32.8 %). In contrast, WPU frequently underestimates the quantity of these urban clusters, with substantial negative deviations, such as −42.3 % for Urban Centres and − 66.2 % for Dense Urban Clusters. For Semi-dense Urban Clusters, WPC identifies a larger number of clusters, with positive deviations of 26.8 % in quantity and 19.9 % in population, while WPU assigns them the largest area (37.5 %) despite showing a lowest cluster count (−30.3 %). Regarding Rural Clusters, WPB reports the highest values, with positive deviations in quantity (39.6 %), area (36.7 %), and population (38.6 %), whereas WPU again underestimates with deviations of −73.8 % in quantity and − 71.5 % in population. Overall, WPU distribute population over broader areas results in lower representation of dense urban clusters. In contrast, GHS-POP and WorldPop constrained models provide more consistent identification of concentrated urban populations, while WPB better aligns with rural distributions.

In Fig. 7, it is possible to observe that GHS-POP produces the most

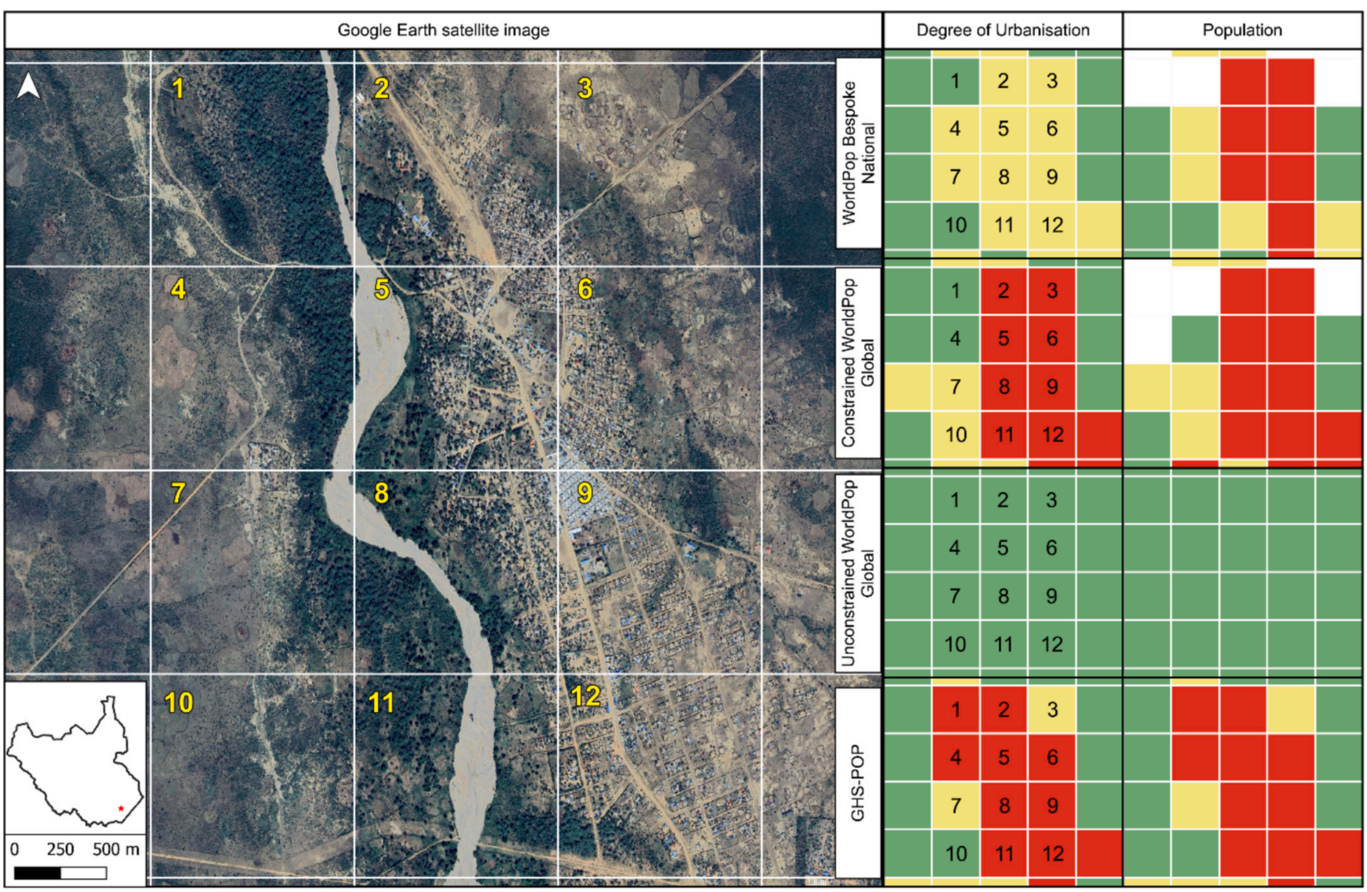

Fig. 5. Example maps for Kapoeta in southeastern South Sudan, demonstrating how using different population datasets can produce different Degree of Urbanisation classifications. Red grids refer to urban centre; yellow grids refer to urban cluster; and green grids refer to rural in the Degree of Urbanisation. The colour in the population column represents whether the population count at each grid meets the threshold for urban centre (inhabitants per km^2 > 1500, red), urban cluster (300 < inhabitants per km^2 ≤ 1500, yellow), and rural (inhabitants per km^2 ≤ 300, green). The location of the area was represented by the red star in the inset national map. The four population datasets estimate gridded population for the year of 2020, whilst the Google Earth satellite image was taken in 2023. Note: the image has no impact on the results and is just for reference. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

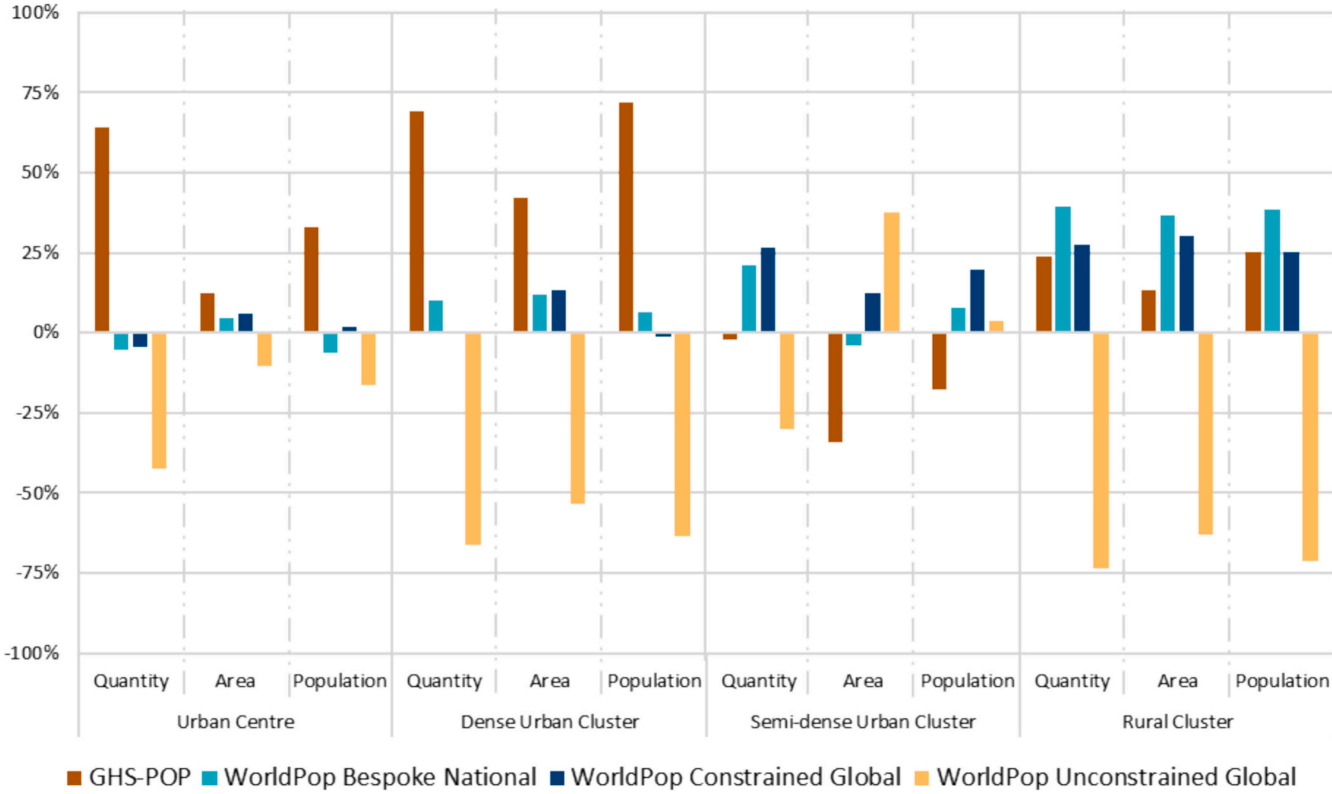

Fig. 6. Impact of gridded population dataset choice on spatial entities in terms of quantity, area, and population across all countries. For each population dataset, the y-axis represents the relative metric deviation in percentage from the corresponding mean referring to the urban centres, dense urban clusters, semi-dense urban clusters, and rural clusters of all considered countries (each metric has been first calculated for each single country and then averaged across all considered countries). The raw data can be found in Supplementary Information Section C.

Urban Centres around Ouagadougou, Burkina Faso. WPU identifies only the largest Urban Centre cluster, whereas WPC and WPB each identifies an additional Urban Centre cluster. This indicates that WPU is only able to identify major metropolitan areas (i.e., city 1) due to its generally low population estimates per grid cell. Consequently, fewer Urban Centres are identified, though the overall population and area remain similar to those identified by the other datasets. Additionally, WPU tends to produce more Semi-dense Urban Clusters. As a result, it consistently identifies significantly fewer Dense Urban Clusters and Rural Clusters in terms of quantity, area, and population compared to the other three gridded population datasets.

4. Discussion

The choice of focusing on GHS-POP and WorldPop gridded population datasets is based on their unique advantages and capabilities compared to the other available datasets, such as LandScan and GPW. WorldPop's datasets cover a wide range of modelling techniques, including census-independent methods, and those co-developed with National Statistical Offices (NSOs), ensuring comprehensive and up-to-date population estimates. GHS-POP use information about the distribution, volume, and classification of built-up areas to redistribute administrative unit-based population counts within built-up areas. Both WorldPop and GHS-POP provide consistent and well-documented methodologies for their datasets production, allowing us to explore any observed differences in their DEGURBA outputs. LandScan is excluded because it primarily captures ambient populations rather than actual residential populations (Dobson et al., 2000), and its documentation on inputs and methods is somewhat limited (Rose & Bright, 2014), making it challenging to discern the underlying drivers of any observed differences in our context. Similarly, GPW's simple areal weighting approach does not incorporate the distribution of built-up areas (CIESIN, 2017), which can result in less accurate representations of residential populations in both urban and rural contexts. Our used WorldPop Unconstrained Global dataset also did not restrict population to be redistributed to built-up area only, but the population distribution across cell is informed by several proxy layers making it more representative than GPW. These factors led us to focus on GHS-POP and WorldPop datasets, which better capture the nuances of residential population distribution for our analysis.

DEGURBA shows less sensitivity with respect to the use of WPB, WPC, and GHS-POP datasets, with relatively high agreement among the corresponding results. These results generally align closely, resulting in minimal variation across different types of settlements. Besides, the similarity between WPB and WPC in producing DEGURBA results suggests that DEGURBA is relatively insensitive to the methodological variations used in producing constrained gridded population datasets. This consistency, in turn, highlights their suitability to be used for implementing the DEGURBA, as they provide a robust basis for settlement identification and classification without significant discrepancies. Rural Clusters, in particular, show near-identical classifications across the three constrained datasets, highlighting DEGURBA's stability in classifying well-defined settlement types. Even for transitional areas, such as dense urban clusters and semi-dense urban clusters, DEGURBA generally produce similar classification results, regardless of the constrained dataset of choice. This overall consistency highlights DEGURBA's reliability and demonstrates that it can produce stable settlement patterns that are not highly dependent on the specific constrained gridded population dataset of choice, making it a valuable tool for urban and rural analysis.

In contrast, the results obtained using WPU consistently deviates from the ones obtained using the constrained datasets. Compared to the

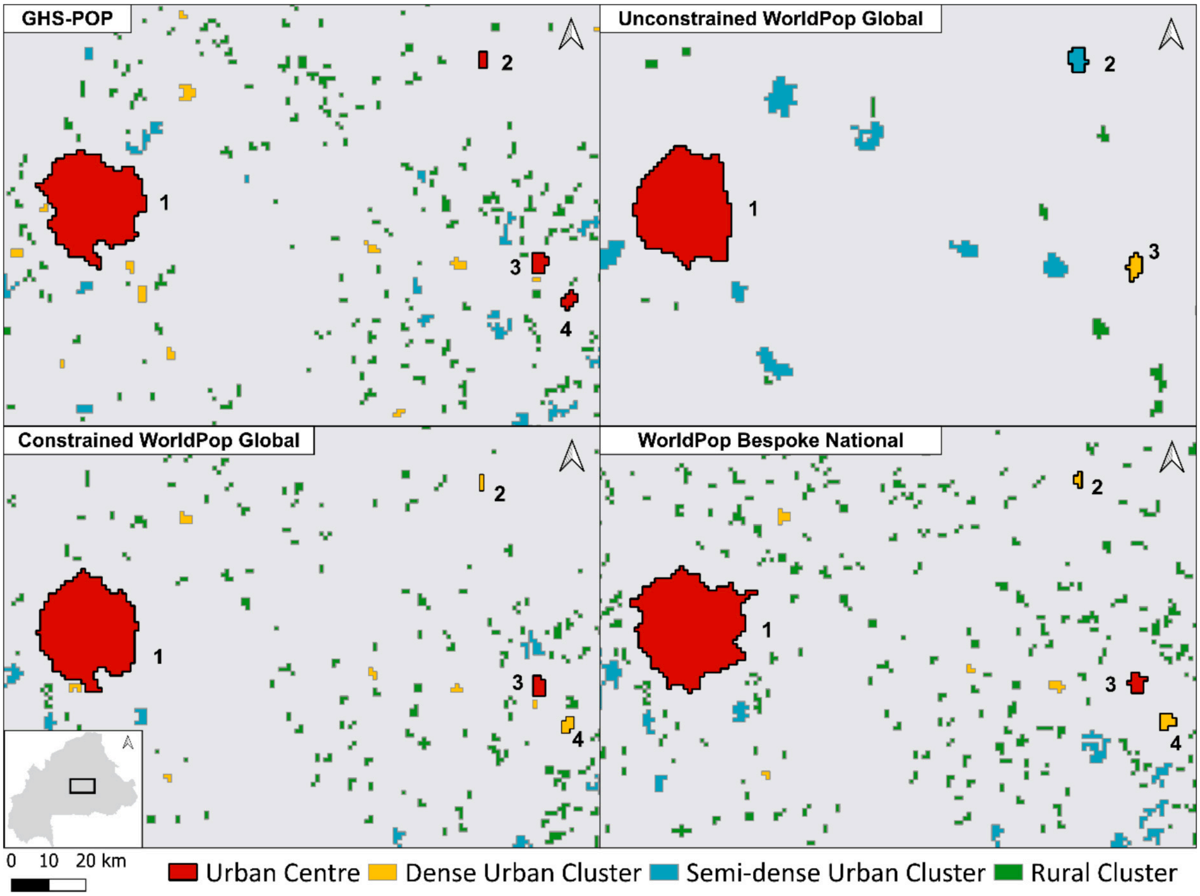

Fig. 7. The spatial entities around Ouagadougou, Burkina Faso identified by the four population datasets. The spatial entities are labelled based on the Urban Centres, i.e., cities, identified by GHS-POP. City 1 is Ouagadougou, the capital of Burkina Faso.

study constrained datasets, WPU allocates populations more broadly within the whole study area without constraining it to built-up areas (Lloyd et al., 2019; Stevens et al., 2015). This broad distribution may lead to the misclassification of urban centres as rural, particularly when the urban clusters just above the thresholds. Such misclassifications highlight the limitations of using datasets reallocating population without the consideration of built-up areas for implementing DEGURBA. The smoothing effect characterizing the population distribution depicted by the WPU datasets makes them less suited for identifying smaller settlements based on population density, which is crucial for DEGURBA's classification.

On the other hand, if a settlement is not detected by the constrained approach, there will be missing predicting population at all in these places and overestimates numbers in those settlements that were detected as a result. This may lead DEGURBA to has more urban centres or urban clusters, and the small rural settlements will be missed. The constrained population datasets may use different settlements into which population is disaggregated. GHSL settlement is derived from Sentinel-2 composite and Landsat satellite imageries (Pesaresi et al., 2024), while WorldPop settlement is derived from building footprints which were extracted by Ecopia.AI using satellite imagery provided by Maxar Technologies (Ecopia.AI & Maxar Technologies, 2020). The settlement layers were mainly produced by interpreting satellite imageries instead of field surveys, meaning the population counts could be less accurate in areas or countries where there are insufficient imageries to detect settlements. Particularly the building footprint data were produced from satellite imagery from a mosaic of dates, and that for some areas the imagery may be many years old and therefore missing recent development. Additionally, built-up detection fails in places like Brazilian Amazonia as most of settlements are below tree canopy, and in desert areas some settlements are built below ground, or the building materials do have a reflectance not much distinguishable from that of the surrounding environment, e.g., sand and rocks. Of note that advances in settlement and building footprint mapping, driven by high-resolution satellite imagery and machine learning, are rapidly minimizing these issues.

The ideal population grid for DEGURBA applications is an official 1 km grid based on census data, such as the EU GEOSTAT 1 km grids (GEOSTAT 1 A, 2012), the Japanese official 1 km grid (Statistics Bureau of Japan, 2020), and the 1 km grid showing the 2023 estimated resident population for Australia produced by the Australian Bureau of Statistics (Australian Bureau of Statistics, 2022–23), as these provide the most accurate 1 km gridded population distribution data based on official government statistics. In the absence of such official grids, both top-down and bottom-up modelled datasets are widely used as proxies of the actual population distributions with different levels of accuracy. For top-down grids, accuracy is likely to be highest when the Average Spatial Resolution of the input administrative units population counts is lower than 1 km (Balk et al., 2006), as seen in countries like Portugal (0.17 km^2), Monaco (0.25 km^2), Italy (0.41 km^2), USA (0.81 km^2), and the UK (1 km^2). For bottom-up grids, reliable results can be achieved through the aggregation of geocoded census points, as in the recent census round in Mexico (Orozco, 2023). Where neither official census grids nor fine-resolution top-down or geocoded bottom-up data are available, or where the accuracy and completeness of census data are unclear, modelled top-down or bottom-up grids where predictions are constrained to a complete, consistent, and accurate map of buildings or settlement extent can provide a reliable and suitable approximation of the population distribution. The choice between these alternatives depends on the quality and spatial resolution of available dataset for top-down grids or the availability of country-specific bespoke modelled outputs for bottom-up grids.

5. Conclusion

The DEGURBA is a global, gridded, population-based definition for consistently classifying the urban-rural continuum enabling international statistical comparisons. The population count associated with each grid cell in different gridded population datasets may vary substantially depending on the methods used to produce them (Leyk et al., 2019), leading to different DEGURBA results. Our analysis evaluated the impact of the choice of gridded population dataset on the resulting DEGURBA in a set of African and Caribbean countries. We found that the DEGURBA is sensitive to a small degree to the gridded population datasets of choice in the study 16 countries, as long as they are produced using methods where population estimates are constrained to a defined settled area. In contrast, the sensitivity increases significantly when considering both constrained and unconstrained gridded population datasets, with results often leading to a large share of mixed classifications of rural and urban cluster populations. Therefore, when multiple gridded population datasets are available, the results here suggest that a comparison of them to assess the consistency/variability of the resulting DEGURBA outputs (see Supplementary Information Section D) may be worth. The results may guide whether further investigations into the accuracy/quality of the available gridded population datasets is needed in order to identify the most appropriate and accurate one to use for implementing the DEGURBA when official gridded population data is missing.

Accurately identifying urban and rural areas, as well as defining and classifying the rural-urban continuum, is vital for assessing the accessibility of services and opportunities available in the lives of people everywhere. DEGURBA presents an approach for measuring and mapping this consistently across the world based on global gridded population datasets and here we have provided analyses and demonstrations of the stability of the approach to different methods of population grid data production. These can help inform users of DEGURBA in their considerations and understanding of uncertainties in outputs, and guide its application across multiple fields.

CRediT authorship contribution statement

Wen-Bin Zhang: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Dorothea Woods:** Writing – review & editing, Investigation, Formal analysis, Data curation. **Iyanuloluwa Deborah Olowe:** Writing – review & editing, Resources, Data curation. **Marcello Schiavina:** Writing – review & editing, Methodology, Investigation. **Weixuan Fang:** Writing – review & editing, Investigation, Formal analysis. **Graeme Hornby:** Writing – review & editing, Investigation, Formal analysis. **Maksym Bondarenko:** Writing – review & editing, Investigation, Conceptualization. **Joachim Maes:** Writing – review & editing, Investigation, Formal analysis. **Lewis Dijkstra:** Writing – review & editing, Supervision, Investigation, Formal analysis. **Andrew J. Tatem:** Writing – review & editing, Supervision, Project administration, Investigation, Formal analysis. **Alessandro Sorichetta:** Writing – review & editing, Supervision, Project administration, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was funded by the European Commission's Directorate-General for Regional and Urban Policy under grant agreement ‘Population and SDG indicators by Degree of Urbanisation’ (2022CE16BAT035). The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official position of the European Commission.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cities.2025.106293.

Data availability

All data utilised in this study are publicly available and can be accessed from the respective sources as outlined in the manuscript.

References

Alajizah, S. M., & Altuwaijri, H. A. (2024). Assessing the impact of urban expansion on the urban environment in Riyadh City (2000–2022) using geospatial techniques. *Sustainability, 16*(11), Article 4799.

AlDousari, A. E., Kafy, A. A., Saha, M., Fattah, M. A., Almulhim, A. I., Faisal, A.-A., … Rahman, M. M. (2022). Modelling the impacts of land use/land cover changing pattern on urban thermal characteristics in Kuwait. *Sustainable Cities and Society, 86*, Article 104107. https://doi.org/10.1016/j.scs.2022.104107

Almulhim, A. I., Al Kafy, A., Ferdous, M. N., Fattah, M. A., & Morshed, S. R. (2024). Harnessing urban analytics and machine learning for sustainable urban development: A multidimensional framework for modeling environmental impacts of urbanization in Saudi Arabia. *Journal of Environmental Management, 357*, Article 120705.

Angel, et al. (2016). *Atlas of Urban Expansion—2016 Edition, Volume 1: Areas and Densities*. New York: New York University: Nairobi: UN-Habitat, and Cambridge, MA: Lincoln Institute of Land Policy. https://www.lincolninst.edu/publications/other/atlas-urban-expansion-2016-edition.

Australian Bureau of Statistics (2022–23), Regional population methodology, ABS Website, accessed August 27, 2024.

Balk, D. L., Deichmann, U., Yetman, G., Pozzi, F., Hay, S. I., & Nelson, A. (2006). *Determining global population distribution: Methods, applications and data* (pp. 119–156). Elsevier. https://doi.org/10.1016/s0065-308x(05)62004-0

Butler, M. A. (1990). *Rural-urban continuum codes for metro and nonmetro counties (no. 9028)*. US Department of Agriculture, Economic Research Service, Agriculture and Rural Economy Division.

CIESIN. (2017). *Gridded population of the world, version 4 (GPWv4): Population density, revision 11 (version 4.11) [data set]. Palisades*. NY: Socioeconomic data and applications center (SEDAC). https://doi.org/10.7927/H49C6VHW

De Bellefon, M. P., Combes, P. P., Duranton, G., Gobillon, L., & Gorin, C. (2021). Delineating urban areas using building density. *Journal of Urban Economics, 125*, Article 103226. https://doi.org/10.1016/j.jue.2019.103226

Dijkstra, L., Florczyk, A. J., Freire, S., Kemper, T., Melchiorri, M., Pesaresi, M., & Schiavina, M. (2021). Applying the degree of urbanisation to the globe: A new harmonised definition reveals a different picture of global urbanisation. *Journal of Urban Economics, 125*, Article 103312. https://doi.org/10.1016/j.jue.2020.103312

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 and Remote Sensing, 66*(7), 849–857.

Ecopia.AI & Maxar Technologies. (2020). Digitize Africa Data — Building footprints. https://www.maxar.com/products/imagery-basemaps.

El-Bouayady, R., Radoine, H., El Faouzi, N. E., Tayi, S., & Ozkan, H. C. (2024). Assessing and modeling the impact of urbanization on infrastructure development in Africa: A data-driven approach. *Cities, 155*, Article 105486. https://doi.org/10.1016/j.cities.2024.105486

European Commission: Eurostat. (2021). *Applying the degree of urbanisation: A methodological manual to define cities, towns and rural areas for international comparisons* (2021 ed.). Publications Office of the European Union. https://doi.org/10.2785/706535

Freire, S., Schiavina, M., Florczyk, A. J., MacManus, K., Pesaresi, M., Corbane, C., & Sliuzas, R. (2018). Enhanced data and methods for improving open and free global population grids: Putting ‘leaving no one behind’ into practice. *International Journal of Digital Earth, 13*(1), 61–77. https://doi.org/10.1080/17538947.2018.1548656

GEOSTAT 1A. (2012). GEOSTAT 1A – Representing census data in a European population grid. GEOSTAT 1A project. EFGS and Eurostat.. https://www.efgs.info/wp-content/uploads/geostat/1a/GEOSTAT1A-final-report.pdf.

Ghelfi, L. M., & Parker, T. S. (1997). A county-level measure of urban influence. *Rural economy division, Economic Research Service, U.S. Department of Agriculture.* Staff paper no. 9702. Doi: 10.22004/ag.econ.278818.

Hijmans, R. J. (2009). GADM database of global administrative areas. https://gadm.org/index.html.

Jordan, L. (2018). Where the Only Certainty Is Uncertainty: Exploring the Demographics of Displacement in South Sudan. In *PAA 2018 Annual meeting. PAA*. April https://paa.confex.com/paa/2018/meetingapp.cgi/Paper/23160.

Klopp, J. M., & Petretta, D. L. (2017). The urban sustainable development goal: Indicators, complexity and the politics of measuring cities. *Cities, 63*, 92–97. https://doi.org/10.1016/j.cities.2016.12.019

Leyk, S., Gaughan, A. E., Adamo, S. B., de Sherbinin, A., Balk, D., Freire, S., … Frye, C. (2019). The spatial allocation of population: A review of large-scale gridded population data products and their fitness for use. *Earth System Science Data, 11*(3), 1385–1409. https://doi.org/10.5194/essd-11-1385-2019

Lloyd, C. T., Chamberlain, H., Kerr, D., Yetman, G., Pistolesi, L., Stevens, F. R., … Tatem, A. J. (2019). Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. *Big Earth Data, 3*(2), 108–139. https://doi.org/10.1080/20964471.2019.1625151

Maffenini, L., Schiavina, M., Melchiorri, M., Pesaresi, M., & Kemper, T. (2023). GHS-DUG user guide, publications Office of the European Union. *Luxembourg*. https://doi.org/10.2760/092009, JRC132760

Maffenini, L., Schiavina, M., Carneiro Freire, S., Melchiorri, M., & Kemper, T. (2023). GHS-POPWARP user guide, publications Office of the European Union. *Luxembourg*. https://doi.org/10.2760/24288, JRC132758

Menon, P., Ruel, M. T., & Morris, S. S. (2000). Socio-economic differentials in child stunting are consistently larger in urban than in rural areas. *Food and Nutrition Bulletin, 21*(3), 282–289. https://doi.org/10.1177/156482650002100306

Mikkelsen, L., Phillips, D. E., AbouZahr, C., Setel, P. W., De Savigny, D., Lozano, R., & Lopez, A. D. (2015). A global assessment of civil registration and vital statistics systems: Monitoring data quality and progress. *The Lancet, 386*(10001), 1395–1406. https://doi.org/10.1016/S0140-6736(15)60171-4

Moyer, J. D., & Hedden, S. (2020). Are we on the right path to achieve the sustainable development goals? *World Development, 127*, Article 104749. https://doi.org/10.1016/j.worlddev.2019.104749

Nieves, J. J., Bondarenko, M., Sorichetta, A., Steele, J. E., Kerr, D., Carioli, A., … Tatem, A. J. (2020). Predicting near-future built-settlement expansion using relative changes in small area populations. *Remote Sensing, 12*(10), Article 1545. https://doi.org/10.3390/rs12101545

Niu, H., & Silva, E. A. (2021). Delineating urban functional use from points of interest data with neural network embedding: A case study in greater London. *Computers, Environment and Urban Systems, 88*, Article 101651. https://doi.org/10.1016/j.compenvurbsys.2021.101651

Openshaw, S. (1984). The modifiable areal unit problem. In *Concepts and techniques in modern geography*.

Orozco, E. V. (2023). Use of new technologies and evidence-based decisions: Key factors in the strategy for the 2020 population and housing census in Mexico in the context of the COVID-19 pandemic. *Wiadomości Statystyczne. The Polish Statistician, 68*(06), 47–66.

Pesaresi, M., & Politis, P. (2023). *GHS-BUILT-V R2023A - GHS built-up volume grids derived from joint assessment of Sentinel2, Landsat, and global DEM data, multitemporal (1975-2030)*. European Commission, Joint Research Centre (JRC). PID http://data.europa.eu/89h/ab2f107a-03cd-47a3-85e5-139d8ec63283 https://doi.org/10.2905/AB2F107A-03CD-47A3-85E5-139D8EC63283.

Pesaresi, M., Schiavina, M., Politis, P., Freire, S., Krasnodębska, K., Uhl, J. H., … Florio, P. (2024). Advances on the global human settlement layer by joint assessment of earth observation and population survey data. *International Journal of Digital Earth, 17*(1), 2390454. https://doi.org/10.1080/17538947.2024.2390454

Randall, S., & Coast, E. (2016). The quality of demographic data on older Africans. *Demographic Research, 34*, 143–174. https://www.jstor.org/stable/26332031.

Rose, A. N., & Bright, E. A. (2014, January). The LandScan global population distribution project: Current state of the art and prospective innovation. In *Conference: Population Association of America Annual Meeting*. Boston, MA, USA https://www.osti.gov/biblio/1149746.

Rozenfeld, H. D., Rybski, D., Andrade, J. S., Batty, M., Stanley, H. E., & Makse, H. A. (2008). Laws of population growth. *Proceedings of the National Academy of Sciences, 105*(48), 18702–18707. https://doi.org/10.1073/pnas.0807435105

Runfola, D., Anderson, A., Baier, H., Crittenden, M., Dowker, E., Fuhrig, S., … Melville, G. (2020). geoBoundaries: A global database of political administrative boundaries. *PLoS One, 15*(4), Article e0231866. https://doi.org/10.1371/journal.pone.0231866

Schiavina, M., Freire, S., Carioli, A., & MacManus, K. (2023). GHS-POP R2023A–GHS population grid multitemporal (1975-2030). *European Commission, Joint Research Centre (JRC)*.. https://doi.org/10.2905/2FF68A52-5B5B-4A22-8F40-C41DA8332CFE

Shohan, A. A. A., Bindajam, A. A., Hang, H. T., Alshayeb, M. J., Alsulamy, S., & Mallick, J. (2024). Quantitative assessment of interplay between urbanization dynamics and land surface temperature variations using generalized additive model coupled PDP for sustainable urban planning and management. *Environmental Science and Pollution Research*, 1–22.

Sorichetta, A., Hornby, G. M., Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). High-resolution gridded population datasets for Latin America and the Caribbean in 2010, 2015, and 2020. *Scientific Data, 2*(1), Article 150045. https://doi.org/10.1038/sdata.2015.45

Stanny, M., Komorowski, Ł., & Rosner, A. (2021). The socio-economic heterogeneity of rural areas: Towards a rural typology of Poland. *Energies, 14*(16), Article 5030. https://doi.org/10.3390/en14165030

Statistics Bureau of Japan. (2020). Statistical Maps on Grid Square Basis - Grid Square Statistics of 2020 Population Census. https://www.stat.go.jp/english/data/mesh/map2020.html Accessed August 20, 2024.

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), Article e0107042. https://doi.org/10.1371/journal.pone.0107042

Taubenböck, H., Droin, A., Standfuß, I., Dosch, F., Sander, N., Milbert, A., Eichfuss, S., & Wurm, M. (2022). To be, or not to be 'urban'? A multi-modal method for the differentiated measurement of the degree of urbanization. *Computers, Environment and Urban Systems, 95*, Article 101830.

Taubenböck, H., Mast, J., Geiß, C., Wurm, M., Esch, T., & Seto, K. C. (2024). Global differences in urbanization dynamics from 1985 to 2015 and outlook considering IPCC climate scenarios. *Cities, 151*, Article 105117. https://doi.org/10.1016/j.cities.2024.105117

Uchida, H., & Nelson, A. (2010). Agglomeration index towards a new measure of urban concentration. In *World Development Report 2009*. https://doi.org/10.1093/acprof:oso/9780199590148.003.0003

United Nations. (2019a). *World urbanization prospects: The 2018 revision*. https://doi.org/10.18356/b9e995fe-en

United Nations. (2019b). 2019 Revision of World Population Prospects. https://population.un.org/wpp.

United Nations Statistical Commission. (2020). Report on the fifty-first session (3-6 march 2020). In *Economic and social council, official records*. supplement no. 4, E/2020/2024-E/CN.2023/2020/2037, 2051/2112 paragraph (i-j).

Wardrop, N., Jochem, W., Bird, T., Chamberlain, H., Clarke, D., Kerr, D., Bengtsson, L., Juran, S., Seaman, V., & Tatem, A. (2018). Spatially disaggregated population estimates in the absence of national population and housing census data. *Proceedings of the National Academy of Sciences, 115*(14), 3529–3537. https://doi.org/10.1073/pnas.1715305115

Woods, D., McKeen, T., Cunningham, A., Priyatikanto, R., Sorichetta, A., Tatem, A. J., & Bondarenko, M. (2024). *WorldPop high resolution, harmonised annual global geospatial covariates*. Version 1.0. Southampton, UK: University of Southampton. https://doi.org/10.5258/SOTON/WP00772