Applying the degree of urbanisation manual - Rationale and advantages

Statistics Explained

3. Rationale and advantages

This article forms part of an online methodological manual, *Applying the Degree of Urbanisation – A methodological manual to define cities, towns and rural areas for international comparisons: 2021 edition* .

Different countries use different criteria to define urban and rural areas which reflect their various perspectives as to what constitute urban and rural areas. It is clear that individual countries need to have their own national definitions that can be implemented in their statistical systems and used to disaggregate indicators by urban and rural areas for their own national policy purposes. Nonetheless, in order to have meaningful international comparisons of statistical indicators by urban and rural areas there is also an undisputed need for adefinition that is nationally relevant and internationally comparable at the same time

.

Such a definition was lacking for international official statistics and international statistical standards. Without a harmonised global methodology, comparisons of the level of urbanisation and indicators for urban and rural areas were difficult to interpret as the differences in definitions could affect the results.

The proposed solution was to develop a global definition of cities, urban and rural areas that could be used generally across the world based on the same delineation criteria for all regions/countries. This proposal should result in a harmonised and universal mapping of cities, towns and semi-dense areas and rural areas. Having internationally comparable statistical information is fundamental for solid evidence-based policymaking and measuring progress towards the sustainable development goals in both urban and rural areas.

This new methodology has been designed not to replace national definitions, but to complement them with a definition that is both nationally relevant and internationally comparable.

There are six clear advantages of the new methodology, namely that it:

- captures the urban-rural continuum through three different classes at level 1 of the degree of urbanisation classification and through seven different classes at level 2 (see Chapter 6 and Chapter 7);
- uses the same population size and density thresholds across the globe (see Chapter 6 and Chapter 7);
- starts from a population grid to reduce the bias of using spatial units with different shapes and sizes (see Chapter 5);
- measures population clusters directly instead of indirectly by using building clusters as an approximation of population clusters (see Chapter 6 and Chapter 7);
- defines areas independently from their access to services to ensure that this access can be monitored reliably, in other words, without interference from the definition;
- proposes a relatively cost-effective approach that can be applied to existing data collections (see Chapter 5 , Chapter 9 and Chapter 10).

3.1 Captures the urban-rural continuum in harmonised manner

The UN's World Urbanization Prospects[1] presents data for urban areas and for rural areas. Many countries, however, use an approach with multiple classes to better capture the urban-rural continuum. For example, the 2011 census in India defined three types of urban areas: statutory towns, census towns and outgrowths. The United States census used urbanized areas, urban places outside of urbanized areas, and rural places and territory. The census in Portugal used predominantly urban areas, medium urban areas, and predominantly rural areas, while South Africa used three geography types: urban areas, rural areas and traditional areas.

The degree of urbanisation classifies the entire territory of a country along an urban-rural continuum. It combines population size and population density thresholds to capture three mutually exclusive classes: cities, towns and semi-dense areas, and rural areas (level 1 of the degree of urbanisation classification). Comparing level 1 of the degree of urbanisation classification to the traditional urban-rural dichotomy, and depending on the country under consideration, national definitions may include towns and semi-dense areas in an urban class or a rural class (see Figure 3.1). For example, the population of towns and semi-dense areas is almost entirely classified as urban according to the national definitions employed in Portugal, Brazil, France and the United States, while in Uganda and India the population of towns and semi-dense areas is generally classified as rural.

By creating a separate class for areas where there is often no general agreement within national definitions, the degree of urbanisation classification proposes a compromise which acknowledges both approaches and enhances international comparability.

There are two principal extensions to the methodology. The first (level 2 of the degree of urbanisation classification) provides a further breakdown for towns and semi-dense areas and for rural areas, with each divided into three separate subclasses (see Chapter 7).

The second extension, defines functional urban areas (also referred to as metropolitan areas). These complement the degree of urbanisation classification by extending the concept of a city to include its surrounding commuting zone. This provides a more economic perspective of the urban-rural continuum. It can also be combined with level 1 of the degree of urbanisation classification to distinguish rural areas inside and outside a metropolitan area.

[1] United Nations Department of Economic and Social Affairs: Population Dynamics .

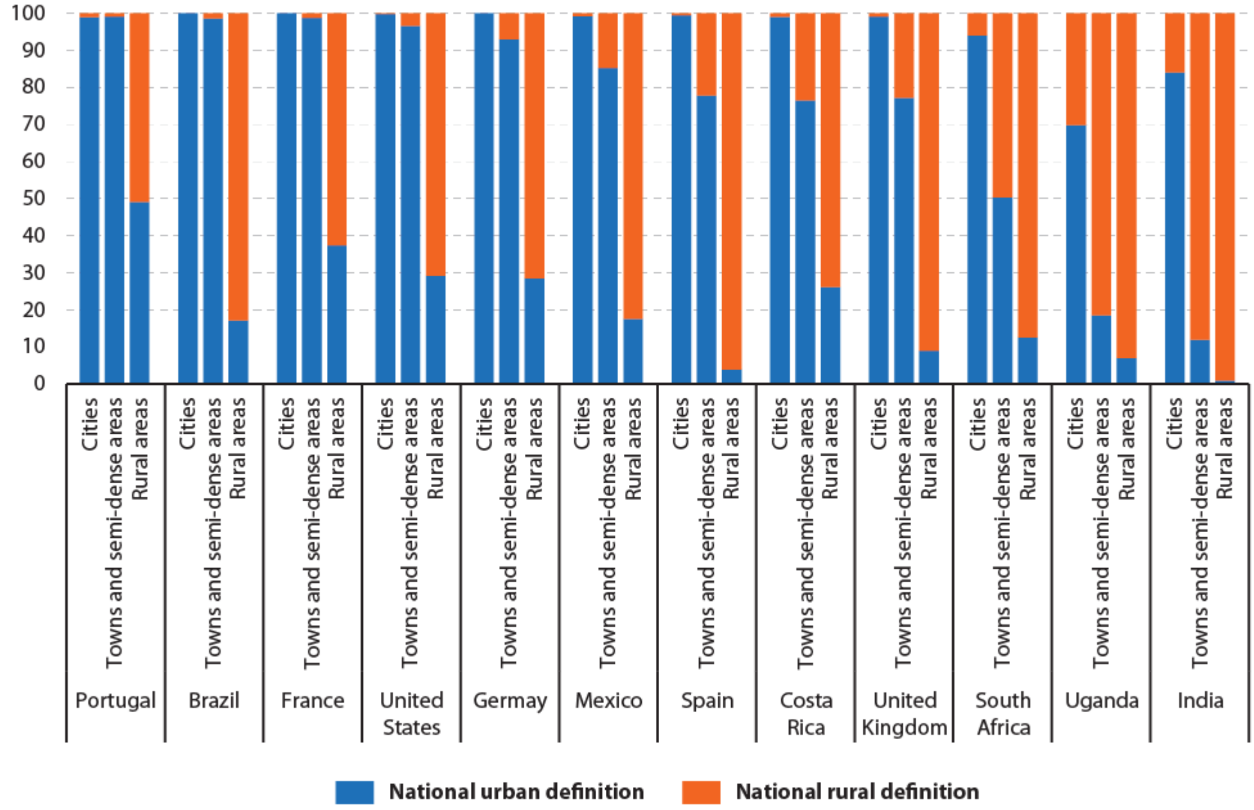

Figure 3.1: Share of total population according to the degree of urbanisation classification and national urban-rural definitions, selected countries, mixed reference years (%) Note: this figure shows for each degree of urbanisation the proportion of the population that was classified as urban and as rural according to national definitions. Countries are ranked by the share of their population living in towns and semi-dense areas that are classified as urban according to national definitions. Reference years vary from 2010 to 2018 depending on the selected country.

3.2 Uses the same population size and density thresholds across the globe

National definitions often use very different population size and density thresholds (see Figure 3.2), which can potentially reduce the international comparability of the resulting data. The degree of urbanisation classification uses the same thresholds across the globe. These harmonised population size and density thresholds drew inspiration from national definitions:

- out of the 103 countries that use a minimum population size threshold to define urban areas, 84 use a threshold of 5 000 inhabitants or less – this minimum threshold of 5 000 inhabitants was employed to define urban clusters;
- Japan uses a minimum population size threshold of 50 000 inhabitants – this criterion was employed to define urban centres;
- China and the Seychelles use a minimum population density threshold of 1 500 inhabitants per km2– this criterion was employed to define urban centres.

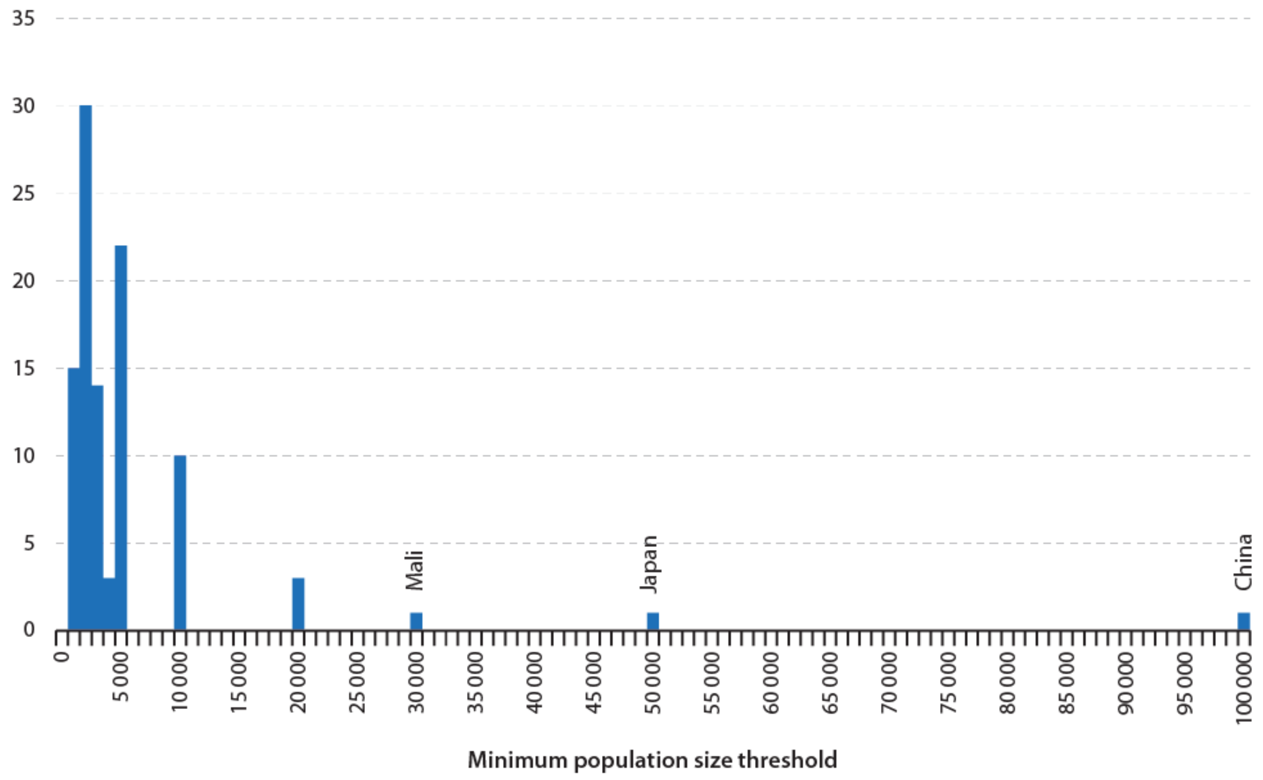

Figure 3.2: Distribution of minimum population size thresholds used to define urban areas (count of countries) Source: UN World Urbanization Prospects

An extensive sensitivity analysis was performed both on official grids in the EU and two global grids (GHS-POP[2]and WorldPop[3]).

Using GHS-POP, the combination of a density threshold of 1 500 inhabitants per km2and a minimum population size threshold of 50 000 inhabitants identified at least one city in every country of the world that had at least 250 000 inhabitants[4], with Vanuatu as the only exception. Using GHS-POP, all small island developing states (SIDS) either have a city or a town. Those SIDS estimated to have a town with at least 5 000 inhabitants were Antigua and Barbuda, Dominica, Grenada, Kiribati, the Marshall Islands, Micronesia, Nauru, Palau, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Samoa, Seychelles, Tonga, Tuvalu and Vanuatu.

3.3 Starts from a population grid to reduce the bias generated by the different shapes and sizes of spatial units

More than half of the countries in the world have a national definition with a minimum population size threshold to classify urban areas. However, applying these thresholds to spatial units that differ in shape and size will influence the results and reduce international comparability. Furthermore, the application of national thresholds can lead to some small rural areas being classified as urban only because they are part of a large(r) administrative unit. For example, Plockton in Scotland has just 387 inhabitants, but it is part of the Highland Council which has more than 230 000 inhabitants. Using a population size threshold would classify Plockton as a rural area, but Highland Council

[2]A spatial raster dataset with the distribution and density of population, expressed in terms of the number of inhabitants per cell (https://ghsl.jrc.ec.europa.eu/data.php).

[3]Spatially detailed information on the number of inhabitants mapped to administrative boundaries (https://www.worldpop.org/focus_areas).

[4]Testing the degree of urbanisation at the global level (https://ghsl.jrc.ec.europa.eu/CFS.php).

as urban, while both are perceived as rural.

To avoid classifying rural areas as urban, some national definitions add a population density requirement. However, a large city can have a very low population density if it is part of an administrative unit with a very large area. For example, Ulaanbaatar in Mongolia has 1.4 million inhabitants, but it has a relatively low population density of only 270 inhabitants per km2. This bias created by the shape and size of spatial units is called the modifiable areal unit problem. It can be addressed by using spatial units of the same shape and size, such as the population grid. Figure 3.3 shows how there is only one settlement identified when using the population density of the administrative units, whereas the population grid reveals that there are actually two settlements (circled in red) when analysed across identical spatial units (a grid).

The method proposed here starts with a classification of a 1 km2population grid. This creates a classification which is independent from the administrative units of a country and is typically far more detailed. For example, the European Union has around 120 000 local administrative units, but more than 4 million grid cells of 1 km2. Some national definitions are applied census enumeration areas, which are typically much smaller than local administrative units. As they are designed to capture roughly the same number of households, they tend to be (very) small in urban areas and (very) large in rural areas. As a result, the population density of enumeration areas will be higher in urban areas and lower in rural areas as compared with units of the same shape and size. For example, the Australian mesh block vary in size by a factor of one billion from 0.0001 km2to more than 100 000 km2. Such large differences in size are bound to have a significant impact on population density figures and thus also on a definition that relies on population density. The benefit of using the grid is that all the cells have the same shape and size and their borders are stable over time. This produces a classification which is more comparable across space and more stable over time.

The second step of this method classifies administrative or statistical spatial units, which reintroduces the problem of working with units of varying shapes and sizes. Therefore, it is recommended to use small administrative or statistical spatial units; this should ensure a good match with the grid classification. Applying this method to very large units, such as regions, may significantly alter population shares when compared with the grid classification.

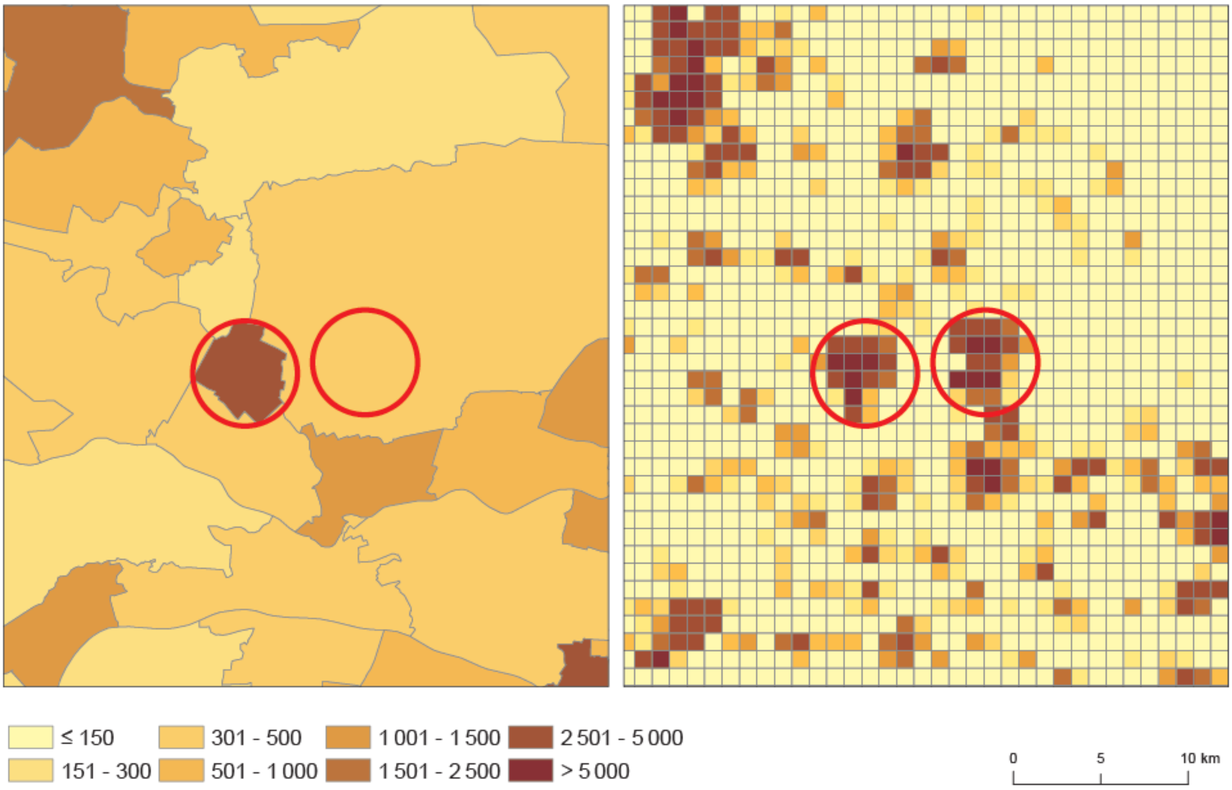

Figure 3.3: Population density of administrative units and grid cells in Veenendaal, the Netherlands, 2011 (inhabitants per km^2) Source: Eurostat (GEOSTAT 2011)

3.4 Measures population clusters directly

The United Nations *Principles and Recommendations for Population and Housing Censuses* (UN (2017)) defines a locality or settlement as a distinct population cluster (Section 1.8, p. 187). In the past, however, it was not possible to measure where people were clustered, although buildings were often mapped at a much higher spatial resolution than the population. For example, a cadastral map with the outline of each building has a very high spatial resolution and can be used to identify which buildings are within 200 m of each other. Population data, however, was only available at a much coarser spatial resolution. Therefore, some national and academic definitions used clusters of buildings to identify settlements.

Today, however, far more precise information is available on the distribution and location of populations. With the advent of geo-coded censuses, geo-referenced population registers and high resolution population grids, the spatial resolution of population data has increased dramatically and allows the direct identification of population clusters. As a result, it is no longer necessary to approximate a population cluster by using a cluster of buildings.

Measuring population concentrations directly makes them more comparable across different levels of (economic) development. Cities in high-income countries tend to have far more built-up area per inhabitant than cities in low-income countries (for example, because cities in high-income countries tend to have bigger houses, as well as more spacious offices and shops). Using only the built-up area to define cities would mean that a high-income country would have more cities and each city would be bigger (in terms of area) than for a low-income country, even if they had exactly the same urban structure in terms of population clusters.

Measuring population concentrations directly also makes them more comparable over time. In many countries, the amount of built-up land grows faster than the size of the population. This means that over time, less and less people would be needed for a certain size built-up area to be reached. As a result, definitions based on built-up areas are likely to inflate the share of the urban population over time, whereas people-based definitions are not affected by this problem.

3.5 Defines areas to monitor access to services, not areas defined by access to services

The sustainable development goals include multiple indicators that monitor access to services or infrastructure. Examples include indicators measuring access to electricity, safely managed drinking water, a mobile phone network and all-weather roads. To properly monitor access to these services in urban and rural areas, they should not be part of the definition of such areas. For example, if the definition of an urban area includes a criterion that everyone should have access to electricity, this would mean, by definition, that the entire urban population would have to have access. This would make it impossible to monitor access to electricity in urban areas, as some large and dense settlements lacking electricity would not be classified as urban areas.

To avoid this problem, the degree of urbanisation does not use access to services or infrastructure as criteria. This means that it can be used to identify cities, towns and semi-dense areas, and rural areas that lack or have successfully acquired such a service. This can facilitate international policy exchanges on how to provide, for example, electricity to different types of areas.

Furthermore, the degree of urbanisation does not use the share of agricultural employment for both conceptual and empirical reasons. Rather, the methodology is people-based and this means that settlements of the same size are consistently classified in the same way. If a maximum threshold for agricultural employment was employed as part of the methodology to identify different areas, then settlements with the same population size could be classified either as urban or rural, undermining the central principle of the methodology.

Empirically, the share of employment in agriculture varies from more than 50 % to less than 1 % between different countries of the world. Using a fixed threshold for the share of agricultural employment would result in some countries being classified as entirely rural or entirely urban. This, in turn, would undermine the goal of facilitating international comparisons and measuring the sustainable development goals in a harmonised manner.

Because, agricultural employment is not part of the methodology, it may be distributed across all three classes. For example, in the EU-27, some 6 % of the people working in agriculture live in cities, 24 % live in towns and semi-dense areas and the remaining 69 % in rural areas. The presence of agricultural employment outside rural areas should not be seen as a problem, but rather as a benefit of this method. For example, farmers living in cities, towns and semi-dense areas will have better access to markets, allowing them to focus on more perishable and higher value added produce. They may also have more opportunities to combine farming with working in a different economic sector.

The United Nations*Principles and Recommendations for Population and Housing Censuses* (UN (2017)) mentions the lack of a single definition of urban and rural areas. It suggests that some countries may wish to use additional criteria including 'the percentage of the population engaged in agriculture, the general availability of electricity or piped water in living quarters and the ease of access to medical care, schools, recreation facilities and transportation'. The method presented here aims to fill the lack of a harmonised method to delineate cities, urban and rural areas. This method deliberately avoids the suggested additional indicators to ensure that a) settlements of the same size are classified in the same way and b) access to services can be monitored over time and space.

3.6 Proposes a cost-effective approach

This method is highly cost-effective for two reasons. First, a population grid can be created for a relatively low cost using existing data. Second, compiling statistics by degree of urbanisation can be done through aggregating existing data.

A population grid can be created using a geo-coded census or a geo-coded population register for little extra cost. These sources provide the exact location of the residents of a country. All that is further required is to add up the population per 1 km2grid cell and, if needed, treat the results to protect confidentiality. If the exact location of the population is not available, a population disaggregation grid can be created by combining the population of census enumeration areas with high resolution land use or land cover data; these data can be produced using remote sensing. Several organisations offer a free global layer, including the Global Human Settlement Layer .

Compiling data according to the degree of urbanisation can be relatively simple. If, for example, in a household survey, the location of where respondents live or the small spatial unit in which they live is available, then responses can be aggregated accordingly to compile statistics according to the degree of urbanisation. As the degree of urbanisation classification often has a quite balanced population distribution across its three classes, surveys will generally have a sufficiently large sample in each of the classes to produce reliable results. Other types of data, such as administrative data, can also be aggregated and compiled according to the degree of urbanisation as long as they are collected for small spatial units.

External links

- UN (2017), *Principles and Recommendations for Population and Housing Censuses – Revision 3* , ST/ESA/STAT/SER.M/67/Rev.3, Department of Economic and Social Affairs, Statistics Division, United Nations, New York.