Applying the Degree of Urbanisation

A methodological manual to define cities, towns and rural areas for international comparisons

2020 edition

[Blank: inside cover]

Applying the Degree of Urbanisation

A methodological manual to define cities, towns and rural areas for international comparisons

2020 edition

Printed by Imprimerie Bietlot in Belgium

Manuscript completed in November 2020

European Union identifiers

Print:	ISBN 978-92-76-20305-6	doi: 10.2785/71043	KS-02-20-499-EN-C
PDF:	ISBN 978-92-76-20306-3	doi: 10.2785/706535	KS-02-20-499-EN-N

FAO identifier
ISBN 978-92-N-NNNNN-N

UN-Habitat identifiers
ISBN 978-92-1-132866-0 HS number: HS/019/20E

Luxembourg: Publications Office of the European Union, 2020

Theme: General and regional statistics

Collection: Statistical books

Abstract

Applying the Degree of Urbanisation — A methodological manual to define cities, towns and rural areas for international comparisons has been produced in close collaboration by six organisations — the European Commission, the Food and Agriculture Organization of the United Nations (FAO), the United Nations Human Settlements Programme (UN-Habitat), the International Labour Organization (ILO), the Organisation for Economic Co-operation and Development (OECD) and The World Bank.

This manual develops a harmonised methodology to facilitate international statistical comparisons and to classify the entire territory of a country along an urban-rural continuum. The degree of urbanisation classification defines cities, towns and semi-dense areas, and rural areas. This first level of the classification may be complemented by a range of more detailed concepts, such as: metropolitan areas, commuting zones, dense towns, semi-dense towns, suburban or peri-urban areas, villages, dispersed rural areas and mostly uninhabited areas.

This manual is intended to complement and not replace the definitions used by national statistical offices (NSOs) and ministries. It has been designed principally as a guide for data producers, suppliers and statisticians so that they have the necessary information to implement the methodology and ensure coherency within their data collections. It may also be of interest to users of subnational statistics so they may better understand, interpret and use official subnational statistics for taking informed decisions and policymaking.

Editorial team

Lewis Dijkstra (European Commission, Directorate General for Regional and Urban Policy), Teodora Brandmüller (European Commission, Eurostat), Thomas Kemper (European Commission, Joint Research Centre), Arbab Asfandiyar Khan (FAO), INFORMA s.à. r.l. and Paolo Veneri (OECD).

Production and layout

This publication was produced by INFORMA s.à r.l.

Contact details

Eurostat
Bâtiment Joseph Bech
5, rue Alphonse Weicker
2721 Luxembourg
E-mail: estat-user-support@ec.europa.eu

For more information please consult

Eurostat's website: https://ec.europa.eu/eurostat

Acknowledgements

The editors of this publication would like to thank the following colleagues — Bianka Fohgrub, Oliver Müller, Ruxandra Roman Enescu and Jane Schofield (European Commission, Eurostat), Sergio Freire, Luca Maffenini, Michele Melchiorri and Marcello Schiavina (European Commission, Joint Research Centre), Eva Panuska Jandova and Antoine Malherme (European Commission, Publications Office of the European Union), Pietro Gennari (FAO), Robert Ndugwa (UN-Habitat), Monica Castillo (ILO), Ellen Hamilton (The World Bank), Simon Allen and Andrew Redpath (INFORMA s.à r.l.) — who were involved in its preparation.

Contents

1. Introduction

A United Nations Resolution adopted in September 2015, *Transforming our world: the 2030 Agenda for Sustainable Development* (UN (2015)) includes several indicators for sustainable development goals (SDGs) that should be collected for cities or for urban and rural areas. So far, however, no global methodology or international standard has been proposed to delineate these areas. The broad array of different criteria applied in national definitions of urban and rural areas poses serious challenges to cross-country comparisons (ILO (2018)). The *Action Framework of the Implementation of the New Urban Agenda* (UN-Habitat (2017)) and the *Global Strategy to improve Agricultural and Rural Statistics* (IBRD-WB (2011)) both highlight the need for a harmonised methodology to facilitate international comparisons and to improve the quality of urban and rural statistics in support of national policies and investment decisions.

This is why six organisations — the European Commission, the Food and Agriculture Organization of the United Nations (FAO), the United Nations Human Settlements Programme (UN-Habitat), the International Labour Organization (ILO), the Organisation for Economic Co-operation and Development (OECD) and The World Bank — have been working closely together over the past four years to develop a harmonised, simple and cost-effective methodology. This new methodology allows statistics to be compiled by degree of urbanisation, identifying cities, towns and semi-dense areas, and rural areas at level 1 of the classification. By using three classes instead of only two (urban and rural), it captures the urban-rural continuum. To improve the international comparability of urban and rural indicators for sustainable development goals (SDGs), it is recommended to produce these by degree of urbanisation.

The first level of the degree of urbanisation classification may be extended in two ways. The first extension, called level 2 of the degree of urbanisation classification, is a more detailed territorial typology: it identifies, cities, towns, suburban or peri-urban areas, villages, dispersed rural areas and mostly uninhabited areas. The second extension, defines functional urban areas (otherwise referred to as metropolitan areas), covering cities and the commuting zones around them. In order to produce SDG indicators by level 2 of the degree of urbanisation classification or by functional urban area, it is necessary to use surveys with large samples. As a result, it will not always be feasible to produce SDG indicators for these two extensions.

To highlight the interest and the feasibility of producing SDG indicators by degree of urbanisation, this manual includes examples of indicators from 12 of the 17 goals for a range of countries across the globe. The indicators tend to have a clear urban gradient with cities at one end, rural areas at the other and with towns and semi-dense areas in between. In some cases, cities tend to fare better, for example in terms of access to education, in others, rural areas tend to do better, for example in terms of personal safety.

This methodological manual is meant to complement and not replace the already existing definitions used by NSOs and ministries. Indeed, these national definitions typically rely on a much wider set of criteria which may have been refined to take into account specific characteristics, context and policy objectives.

The manual has been designed principally as a practical guide for data producers, suppliers and statisticians so that they have the necessary information to implement the methodology and ensure coherency and consistency within their data collections and analyses. It may also be of interest to users of subnational statistics — such as policymakers, the private sector, research institutions, academia — so that they may better understand and interpret official subnational statistics.

The manual was produced at the request of the 51st session of the UN Statistical Commission (UNSC), which 'endorsed the methodology for delineation of cities and urban and rural areas for international and regional statistical comparison purposes, [] and urged the release of a technical report on the implementation of the methodology for delineation of cities and urban and rural areas as early as possible' ([1]).

Table 1.1: Milestones on the way to the endorsement by the UN Statistical Commission

October 2016	**UN-Habitat III conference, Quito** The European Commission's Commissioner for Regional and Urban development announced a joint voluntary commitment with the OECD and The World Bank to develop a global, people-based definition of cities and settlements.
March 2017	**UN Statistical Commission (UNSC), New York** Presentation of the work plan, first results and discussion on next steps in two dedicated side events.
April 2017	**UN-Habitat Expert Group meeting, Brussels** The Expert Group Meeting on Geospatial Definitions for Human Settlements Indicators of the SDGs concluded that a standard definition of a city is needed for global reporting and monitoring of the SDGs.
November 2017	**UN Statistical Division (UNSD) survey** The UNSD sent a questionnaire to 20 countries to gather feedback on the proposed methodology. At least three quarters of the respondents stated that the methodology was useful for international comparisons and to compile indicators for the UN's SDGs.
January 2018	**Food and Agriculture Organization of the United Nations (FAO) Expert Group meeting, Rome** The Expert Group meeting on *Improving Rural Statistics: Rural Definition and Indicators* reviewed and made recommendations on the methodology.
March 2018	**UN Statistical Commission (UNSC), New York** The interim results were presented at a side event of the UNSC, which highlighted the interest and support for this global development. Further consultations and communication to raise the awareness and understanding of this new methodology were planned.
December 2018	**FAO and the Global Strategy to improve Agricultural and Rural Statistics (GSARS) published its findings on pilot tests** FAO and the GSARS tested the definition (at level 1 and level 2) for seven countries in their national contexts. The report also assessed the countries' capacity and capability to report on a subset of core SDG indicators, applying the methodology and using existing data collection mechanisms.
October 2018 – October 2019	**UN-Habitat regional workshops** UN-Habitat organised seven regional workshops to present the methodology and discuss how it could be improved and applied nationally. A total of 85 countries participated in these workshops (see Figure 9.5 for a complete list).
January 2019	**UN Expert Group meeting, New York** An Expert Group meeting on the *Statistical Methodology for Delineating Cities and*

([1]) UN Statistical Commission (UNSC), *Report on the fifty-first session (3-6 March 2020)*, Economic and Social Council, Official Records, 2020, Supplement No. 4, E/2020/24-E/CN.3/2020/37, 51/112 paragraph (i-j).

	Rural Areas (UN (2019)) concluded that both the degree of urbanisation and functional urban area classifications were useful to monitor the SDGs and should be used in parallel with national definitions of urban and rural areas.
March, 2019	**UN Statistical Commission (UNSC), New York** The UNSC welcomed the work on developing the methodology for the delineation of urban and rural areas and the definition of cities based on the degree of urbanisation classification, and requested the submission of the final assessment, to be prepared in consultation with Member States, on the applicability of this methodology for international and regional comparison purposes to the Commission at its fifty-first session (see E/2019/24-E/CN.3/2019/34, Decision 50/118, paragraph (d)).
March 2020	**UN Statistical Commission, New York** **The UNSC 'endorsed the methodology for delineation of cities and urban and rural areas for international and regional statistical comparison purposes'.**

References

FAO and GSARS (2018), *Pilot tests of an international definition of urban–rural territories*, Technical Report no. 37, Rome.

IBRD-WB (2011), *Global Strategy to improve Agricultural and Rural Statistics*, International Bank for Reconstruction and Development/The World Bank, Economic and Sector Work, Report No. 56719-GLB, Washington D.C.

UN (2015), *Transforming our world: the 2030 Agenda for Sustainable Development*, United Nations, General Assembly, A/RES/70/1, New York.

UN (2019), *Expert Group Meeting on Statistical Methodology for Delineating Cities and Rural Areas*, United Nations Statistics Division, New York.

UN-Habitat (2017), 'New Urban Agenda', *United Nations Conference on Housing and Sustainable Urban Development (Habitat III)*, United Nations, General Assembly, A/RES/71/256, New York.

UN-Habitat (2017), *Expert Group Meeting on Geospatial Definitions for Human Settlements Indicators of the SDGs*, Brussels.

2. The legal and policy framework

Designing effective policies requires a good understanding of the socioeconomic conditions that exist in cities and in urban and rural areas, which in turn depends on building a solid base of knowledge about people, their activities, communities, well-being and their interaction with the environment. Reliable, timely and internationally comparable datasets for different urban and rural areas can only be produced on the basis of coherent and harmonised methodology that delineates cities, urban and rural areas in a consistent manner.

2030 Agenda for Sustainable Development

In 2015, the United Nations General Assembly adopted the *2030 Agenda for Sustainable Development* (UN (2015)). At the core of the agenda, there is a set of 17 sustainable development goals (SDGs), which provides a global policy framework for stimulating action until the year 2030 in areas of critical importance related to people, the planet, prosperity, peace and partnership. A global list of 232 indicators was developed to measure progress towards 169 targets across these 17 goals from the 2030 agenda. Cities, urban and rural areas play a crucial role for many policy areas underlying the SDGs such as eradicating poverty and hunger, housing, transport, infrastructure, land use or climate change. Beyond SDG 11 — make cities and human settlements inclusive, safe, resilient and sustainable — which focuses explicitly on cities and communities, an estimated two thirds of the 169 targets can be measured and analysed for cities and urban and rural areas which can help shape sustainable development policies from the ground up and provide support to help reach the targets set in the 2030 agenda.

New Urban Agenda

Urbanisation is a phenomenon that impacts economies, societies, cultures and the environment. It is projected that 55 % of the world’s population will be living in cities by 2050 (OECD and European Commission (2020)). Not only is there a growing level of interest in the rapid growth and shape of urban developments, but also in the linkages that exist between individual cities and between urban and rural areas. One particular area of policy interest is that of mega cities and large metropolitan areas that benefit from economies of agglomeration, industrial clustering and innovation, while at the same time facing significant challenges with respect to sustainable urban development (for example, congestion or environmental impacts).

A United Nations Conference on Housing and Sustainable Urban Development (Habitat III) in Quito, Ecuador, on 20 October 2016 adopted the *New Urban Agenda*; it was subsequently endorsed by the United Nations General Assembly on 23 December 2016 (UN-Habitat (2017)). The *New Urban Agenda* seeks to provide a vision for a more sustainable future by promoting a new model of urban development, based on the premise that cities can be the source of solutions to, rather than the cause of, many global challenges. It provides standards and principles for the planning, construction, development, management, and improvement of urban areas following five main pillars: national urban policies, urban legislation and regulations, urban planning and design, local economy and municipal finance, and local implementation.

Rural development policies

Rural areas are intrinsically important and fundamentally different from urban areas and thus (often) require a different set of interventions and policies that aim to improve the livelihood of their populations. Research and empirical evidence show that rural areas are characterised by: slow dynamics of farm productivity, widespread income inequality and volatility of agricultural income; considerable outward migration flows to urban areas that result in depopulation of rural areas; a lack of efficient physical, technological and information technology (IT) infrastructures; public and private services that are more costly to provide and more difficult to access than in urban areas (OECD (2020)).

Despite their importance, rural statistics on income and livelihoods are sparse and uncommon, mainly due to the fact that there is no consistent international definition of rural areas. Rural areas are usually defined based on national policy objectives; sometimes, as a residual, once urban areas are defined, or sometimes based on a combination of multiple criteria, for example, population size and density, the presence of agriculture, remoteness from urban areas and a lack of infrastructure and/or basic social services.

It is important to highlight that rural statistics are territorial in nature, in contrast to sectoral statistics that focus on a single activity. People in rural areas are typically engaged in several different economic activities beyond agriculture, fisheries and forestry, for example mining and quarrying, as well as in craft production. Some of the main challenges facing rural areas include: malnutrition, food insecurity, poverty, limited adequate health and education services, a lack of access to other basic infrastructure and the under-utilisation of labour.

In formulating a rural development policy, the FAO draws on issues identified in the *2030 Agenda for Sustainable Development*, while acknowledging that rural areas have particular characteristics that present unique challenges. These include, among others: the dispersion of rural populations; topographical features (terrain and landscapes) that may act as a barrier for the efficient provision of infrastructure; an (over) reliance on the agricultural sector; ensuring that natural resources and environmental quality are protected.

International statistics differentiating between urban and rural areas

The idea of differentiating between urban and rural areas for international statistics dates back several decades. In 1991, the European Union labour force survey introduced a variable to indicate the characteristics of the areas where respondents lived. However, its results had limited comparability internationally.

In 2012, the OECD together with the European Commission developed a new way to measure metropolitan areas (OECD (2012), later extended in Dijkstra et al. (2019)). It seeks to ensure that statistics on urban development are made more robust through the provision of an internationally recognised definition of cities and their commuting zones as functional economic units that may guide policymakers better in areas such as planning, infrastructure, transport, housing, education, culture and recreation.

The European Commission’s Directorate-General for Regional and Urban Policy (DG REGIO) published *A harmonised definition of cities and rural areas: the new degree of urbanisation* (Dijkstra and Poelman (2014)). It describes the degree of urbanisation classification and distinguishes three different classes: cities, towns and suburbs, and rural areas (or densely, intermediate and thinly populated areas) that are based on information for population grids to provide more robust data (greater comparability and availability).

Prior to 2017, territorial typologies and their related methodologies within the European Statistical System (ESS) did not have any legal basis. On 12 December 2017, an amending Regulation (EU) 2017/2391 of the European Parliament and of the Council was adopted as regards territorial typologies (Tercet), followed on 18 January 2018 by a consolidated and amended version of Regulation (EC) No 1059/2003 of the European Parliament and of the Council on the establishment of a common classification of territorial units for statistics (NUTS). The main objectives of Tercet include: establishing a legal recognition of territorial typologies for the purpose of European statistics by laying down core definitions and statistical criteria; integrating territorial typologies into the NUTS Regulation so that specific types of territory may be referred to in thematic statistical regulations or policy initiatives, without the need to (re-)define terminology such as cities and urban or rural areas; ensuring methodological transparency and stability, by clearly promoting how to update the typologies.

As part of the Global Strategy to improve Agricultural and Rural Statistics (GSARS), the FAO published *Guidelines on defining rural areas and compiling indicators for development policy* (FAO (2018)). These guidelines provide a definition of which territories should be considered as rural and a more detailed breakdown of different types of rural places to promote like-for-like comparison internationally. The guidelines seek to provide information on concepts and methods to improve the quality, availability and use of rural statistics.

The United Nations Statistics Division (UNSD) plays a pivotal role in the coordination of the world population and housing census programme and, in 2017, the United Nations published *Principles and Recommendations for Population and Housing Censuses* (UN (2017)). In a similar vein, the *Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing* was published by the United Nations Economic Commission for Europe (UNECE (2015)), providing a set of recommendations tailored specifically to the needs of European statisticians. Both documents provide guidance and assistance in the planning and execution of censuses and, among others, aim to facilitate improvements in the comparability of subnational data. Two different approaches are identified for the coding of housing or population units: the first is based on coding units to their lowest-level enumeration area, while the second is based on a coordinate or grid-based system. European countries were urged to adopt the use of grid data and identifiers for coordinate references so that the results of their next censuses could potentially provide a wide spectrum of spatial analyses.

References

Dijkstra, L. and H. Poelman (2014), 'A harmonised definition of cities and rural areas: the new degree of urbanisation', *Regional Working Paper 2014, WP 01/2014*, European Commission Directorate-General for Regional and Urban Policy.

Dijkstra, L., H. Poelman and P. Veneri (2019), 'The EU-OECD definition of a functional urban area', *OECD Regional Development Working Papers, No. 2019/11*, OECD Publishing, Paris.

Eurostat (2019), *Methodological manual on territorial typologies — 2018 edition*, Publications Office of the European Union, Luxembourg.

FAO (2018), *Guidelines on defining rural areas and compiling indicators for development policy*, Food and Agriculture Organization of the United Nations (FAO), Rome.

ILO (2018), ‘Rural-urban labour statistics’, *20th International Conference of Labour Statisticians*, International Labour Office, ICLS/20/2018/Room document 3/Rev. 3, Geneva.

OECD (2012), *Redefining "Urban": A New Way to Measure Metropolitan Areas*, Organisation for Economic Co-operation and Development, OECD Publishing, Paris.

OECD (2020), *Rural Well-Being: Geography of Opportunities*, forthcoming, OECD Publishing, Paris.

OECD and European Commission (2020), ‘Cities in the World: A New Perspective on Urbanisation’, *OECD Urban Studies*, OECD Publishing, Paris.

UN (2015), *Transforming our world: the 2030 Agenda for Sustainable Development*, United Nations, General Assembly, A/RES/70/1, New York.

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.

UNECE (2015), *Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing*, United Nations Economic Commission for Europe, United Nations, New York and Geneva.

UN-Habitat (2017), ‘New Urban Agenda’, *United Nations Conference on Housing and Sustainable Urban Development (Habitat III)*, United Nations, General Assembly, A/RES/71/256, New York.

3. Rationale

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 a **definition 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 lead to bias in 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 (SDGs) in both urban and rural areas.

The proposed methodology may be used to compile statistics according to the degree of urbanisation. The methodology 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). The methodology may be extended in two ways. A first extension provides a further breakdown, identifying medium-sized and small settlements, in other words, towns and villages in more detail (level 2 of the degree of urbanisation classification). Both level 1 and level 2 are exhaustive, insofar as they classify the entire territory of a region/country. A 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. Data for functional urban areas can be particularly useful to support policymaking in a number of domains, including urban planning, such as transport infrastructure and services, as well as economic development.

Some of the main advantages of the degree of urbanisation classification are that it:

- fulfils the principles of official statistics (see Chapter 4);
- reduces the bias generated by the different shapes and sizes of small spatial units (see Chapter 5);
- captures the urban-rural continuum (see Chapter 6);
- is relatively ’easy’ to implement (see Chapter 9).

4. How the principles of official statistics and classifications are fulfilled

This chapter reviews the methodology that is used to compile statistics by degree of urbanisation according to the 10 principles specified in *Best Practice Guidelines for Developing International Statistical Classifications* (UN (2013)).

- **Conceptual basis:** the degree of urbanisation classification relies on population density and size. Population size is also used in most national definitions of urban and rural areas. The functional urban area classification additionally uses commuting data, which is often used for national definitions of metropolitan areas. Each of these elements is clearly defined. Tests have shown that the methodology captures settlements of different sizes and economic relations between cities and their surrounding commuting zones.
- **Classification structures:** the degree of urbanisation classification is hierarchical with two levels, the functional urban area classification has a single level.
- **Classification types:** the methodology proposes two international reference classifications. As a result, the classifications may require some adaption to meet country specific conditions.
- **Mutual exclusivity:** the classes at each level (levels 1 and 2) of the degree of urbanisation classification and the functional urban area classification are mutually exclusive.
- **Exhaustiveness:** levels 1 and 2 of the degree of urbanisation classification are exhaustive, in other words, they classify the entire territory of a country. The functional urban area classification is also exhaustive, insofar as it covers metropolitan and non-metropolitan areas that together make up the entire territory of a country.
- **Statistical balance:** estimates based on the *Global Human Settlement Layer (GHSL)* population grid show that the classifications produce classes where the populations are not too disparate in size. As a result, they will allow for effective cross-tabulation of data.
- **Statistical feasibility:** the classifications were kept simple so as to make them feasible to apply across all countries of the world. The degree of urbanisation classification requires a population grid, which has already been estimated globally. A growing number of countries have produced or are planning to produce such a grid. The functional urban area classification also requires commuting data, which are not widely available across countries. However, auxiliary data sources such as from mobile telephones or employment registers can help to fill this gap.
- **Classification units/statistical units:** the classifications propose simple classes (such as cities, towns and semi-dense area, rural areas or metropolitan areas) which can be used with a wide variety of statistical units such as people, jobs, enterprises, buildings, farms, land use, and so on.
- **Time-series comparability:** estimates based on the GHSL population grid show that data using the degree of urbanisation classification capture changes over time, but are not too volatile.

References

UN (2013), '*Best Practice Guidelines for Developing International Statistical Classifications*', Expert Group on International Statistical Classifications, Department of Economic and Social Affairs, Statistics Division, United Nations, New York.

5. Constructing a population grid

A population grid is a powerful tool to analyse issues that require a consistently high spatial resolution, such as access to public transport, exposure to flooding or patterns of urbanisation. Census enumeration areas provide a high level of spatial resolution in urban areas, but usually a much coarser resolution in rural areas, which makes them less suitable for this type of analysis.

Because a population grid is so useful, a number of organisations are promoting their production and use, including the United Nations Global Geospatial Information Management (UN GGIM), the United Nations Population Fund (UNFPA) and the POPGRID Data Collaborative initiative ([2]).

Population grids have a number of important advantages:

- grid cells all have the same size allowing for easy comparison;
- grids are stable over time;
- grids integrate easily with other data (for example, meteorological or air quality data);
- grid cells can be assembled to form areas reflecting a specific purpose and study area (mountain regions, water catchment areas, metropolitan areas).

The first modern population grids were produced in Scandinavia based on geo-coded population registers in the 1970s. Today, over 30 countries have an official population grid, including Brazil and all the countries in the European Statistical System (ESS). In addition, a substantial number of countries have recently conducted a geo-coded census or are preparing one. Such a census can produce a high quality official population grid (see Subchapter 5.1).

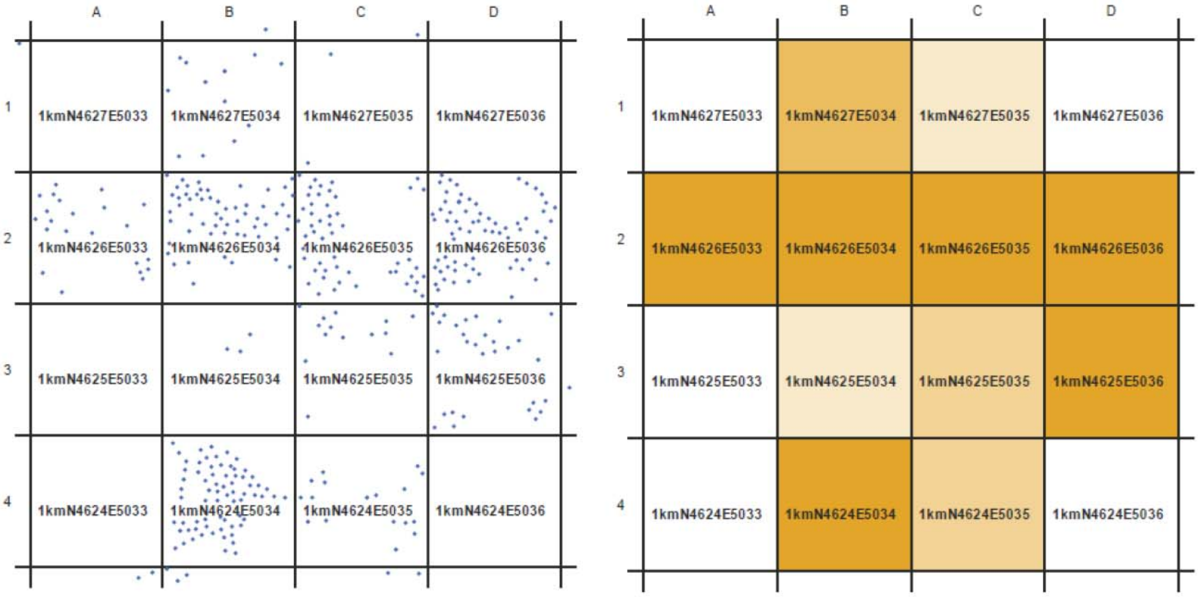
In the absence of a geo-coded census or population register, a disaggregation grid can be created by combining the population of census units (enumeration areas) with high-resolution land use data from national or global sources (see Subchapter 5.2). If census population data for an entire country are not available, models can estimate grid cell population data for areas not covered by the census (see Subchapter 5.3). Finally, a number of emerging sources of big data from mobile phones or social media can also be used to estimate a population grid, although these sources pose a number of issues of reliability and stability over time (see Subchapter 5.4).

5.1 A grid based on the aggregation of point data

Ideally, a population grid is based on a geo-referenced point dataset with a high spatial accuracy (see Figure 5.1). This guarantees a high quality grid and avoids any need for estimations or disaggregations. These points can be derived from a variety of sources. A growing number of countries have or will conduct a digital census where the exact geographical location of each household is recorded ([3]). Countries with a geo-coded cadastre, a building register or an address register can use these to generate a set of points with population data. Once the point data have been created, they can simply be aggregated to square grid cells.

([2]) POPGRID Data Collaborative initiative (https://www.popgrid.org/).
([3]) United Nations Statistics Division, *Guidelines on the use of electronic data collection technologies in population and housing censuses* (https://unstats.un.org/unsd/demographic/standmeth/handbooks/data-collection-census-201901.pdf).

Figure 5.1: Example of point-based data overlaid on a statistical geo-coded grid of 1 km² (left) and population counts in shades of orange according to population density per 1 km² cell (unpopulated grid cells in white) for aggregated point-based information (right)

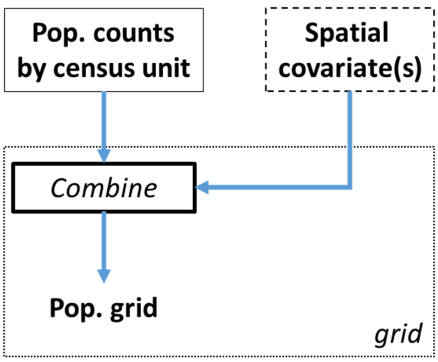


The exact location of each household is considered confidential. However, aggregating these data to grid cells of 1 km^2 is often sufficient to address confidentiality concerns. Some countries also apply a limited amount of record swapping to provide an even higher guarantee of confidentiality (Eurostat (2019) and GEOSTAT 1B ([4])).

5.2 A grid based on the disaggregation of population data

In the absence of point data, a population grid can be produced by disaggregating population data from census enumeration areas or administrative units (such as municipalities, districts or provinces) using auxiliary data with a higher spatial resolution, such as land cover or built-up area data, that are linked to the presence of people (see Figure 5.1).

Figure 5.1: Simplified workflow for population grid creation by disaggregation of existing counts



In a disaggregation grid, the total population of a census unit or administrative unit is distributed across the grid cells covering that unit based on other data that are linked to the presence of people. This disaggregation can be done in a variety of ways. The simplest method relies on a single

([4]) European Forum for Geography and Statistics (EFGS), GEOSTAT 1B (https://www.efgs.info/geostat/1B/).

covariate and allocates the population proportionally to that covariate. *GHS-POP R2019A* (Florczyk et al. (2019)) is a good example of such an approach ([5]).

A slightly more complex method uses multiple covariates. For example, the population may be allocated proportionally to all built-up areas with the exception of non-residential areas and roads and railways. The *European Settlement Map* (Corbane et al. (2019)) is an example that distinguishes between residential and non-residential buildings and excludes roads ([6]).

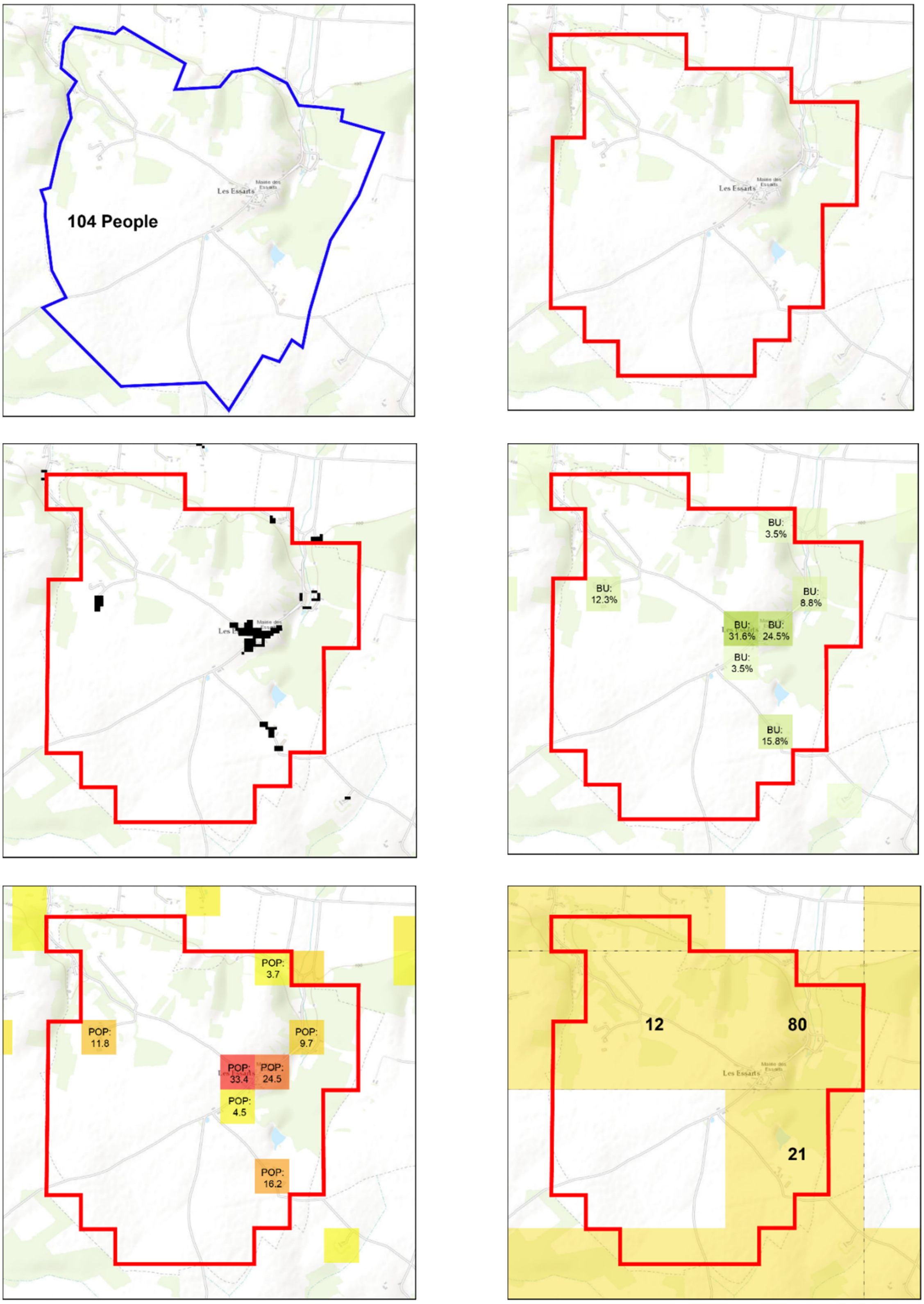
A more complex method uses multiple co-variates combined with a 'random forest' estimation technique to determine the weights to distribute the population. *WorldPop* (Tatem (2017)) is a good example of such an approach ([7]).

Regardless of the disaggregation method selected, two key issues will determine the quality of the resulting population grid. First, the size (area) of the units for which population data are available: the smaller the spatial unit, the higher the quality of the grid. Second, the quality of the covariate: a covariate that is closely linked to the presence of people and that avoids errors of omission and commission will produce a higher quality grid. For example, a geospatial layer of built-up areas or building footprints with high spatial resolution is considered to be highly suitable for such a purpose. Such sources are often based on remote sensing, which may not detect all built-up areas or buildings (omission) or may mistakenly identify some areas as built-up or as covered by a building (commission). Several organisations offer open access global layers based on remote sensing data, including the *Global Human Settlement Layer (GHSL)* produced by the European Commission's Joint Research Centre (JRC).

To allocate proportionally the population within a census unit based on a single covariate involves a number of steps that are presented in Figure 5.2. The first map shows a census unit and its population (p). The second map shows the boundary of this census unit rasterised using a 250 m grid. Through this process each 250 m cell is assigned to one and only one census unit. This process can also be done at a finer resolution (100 m or smaller) to ensure a closer match between the original census unit and the assigned cells, although this requires a more powerful computer. The third map shows the built-up areas (b), which are mapped at 30 m resolution in binary fashion, in other words, built-up or not. The fourth map shows, for each 250 m cell, the built-up area within that cell as a share of the total built-up area within the census unit (b % = b in cell / b in census unit). The fifth map shows the population that has been allocated proportionally based on the share of the built-up area (POPcell = p * b %). Because the sum of the shares of built-up areas in all the cells in a census unit is 100 %, the sum of the population in these cells will exactly match the population of the census unit. The sixth map shows the population for a set of 1 km grid cells (in yellow). Note that, unlike for the 250 m cells, the sum of the three 1 km^2 grid cells (113 people) is higher than the population of the census unit (104 people) because these three grid cells include the population of a few 250 m cells that belong to neighbouring census units.

([5]) Joint Research Centre, *Global Human Settlement Layer* (https://ghsl.jrc.ec.europa.eu).
([6]) Copernicus, *European Settlement Map* (https://land.copernicus.eu/pan-european/GHSL/european-settlement-map).
([7]) *WorldPop* (https://www.worldpop.org/).

Figure 5.2: Example of the process used to generate the GHS-POP layer (extract from a location in France)



Note: Esri, HERE, Garmin, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), swisstopo, © OpenStreetMap contributors and the GIS User Community. Processed by JRC.

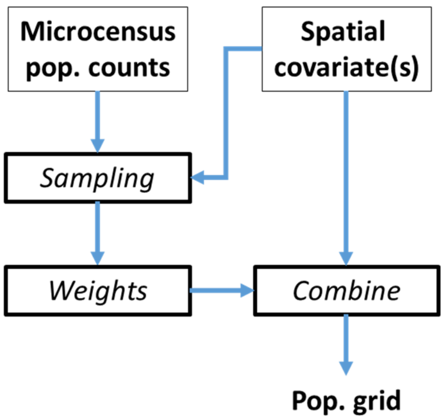
The GHS-POP (Florczyk et al. (2019) is produced in this way. It disaggregates residential population estimates for four target years using the best available census units, adjusted to UN WPP estimates (the population input is the *Gridded Population of the World* v4.10 (CIESIN (2018)). The disaggregation is done using the built-up areas as detected by the GHSL.

5.3 Extrapolating a population grid based on a partial micro-census

Comprehensive and accurate population data for small areas can be costly and logistically challenging to collect, but they represent a fundamental basis for government decision and policymaking. In resource-constrained settings, national population and housing census data can be outdated, inaccurate, or missing specific groups, while registry data can be lacking or incomplete. In addition, certain areas of a country may not be included in national data collections due to conflict, inaccessibility or cost limitations. In such cases, a different approach is needed to produce a complete population grid.

When a geo-referenced census is not available or it is considered unsuitable due to a lack of completeness, freshness, or reliability, a different approach can be employed to create a population grid. This technique is more challenging as it does not start from pre-existing population counts for the entire country; instead, the total is estimated using a population distribution model. Such an approach requires the availability of detailed and reliable data from a micro-census or survey which does not cover the entire country to develop a model. This technique estimates a count — at the level of grid cells — through combining sampling with ancillary data, typically remotely-sensed (for example, the density of buildings, urban areas). Given such a spatial covariate covering the whole country and surveys (micro-census) for a subset of the country, these data are combined to derive parameters or weights in a statistical model characterising the population's distribution. This model is then used to predict the population's distribution in non-surveyed areas (see Figure 5.4) under the assumption that the surveyed area is representative of the whole area.

Figure 5.4: Simplified workflow for population grid creation in the absence of census counts



Recent advances in the availability of detailed satellite imagery, geo-positioning tools for field surveys, statistical methods and computational power are providing opportunities to complement traditional collection methods for data on population by modelling and estimation into areas that were missed from enumeration (Wardrop et al. (2018)). Bayesian geostatistical modelling approaches to predict population numbers and age/sex structures from small area micro-census surveys, or incomplete census enumeration, have been developed and applied for multiple countries where instability, funding or other obstacles have limited recent national data collection exercises.

Using a set of spatially complete datasets as covariates, including satellite-derived building footprints, along with a spatial covariance structure makes it possible for models to predict population by age and sex in unobserved areas across a country, together with associated uncertainty metrics (Wardrop et al. (2018)). Cross-validation typically shows high model accuracies at subnational levels ([8]). This technique has the potential to fill gaps where enumeration could not be undertaken and to provide contemporary, regularly-updated and accurate population information to support decision-making and development in challenging contexts ([9]). Datasets built using these approaches for Nigeria, Zambia and the Democratic Republic of the Congo are available from *WorldPop* ([10]).

5.4 Alternative and emerging data sources for creating population grids

In recent years, a number of emerging data sources and technologies have been explored for direct mapping of the population or as alternative proxies for its disaggregation; at present, this work has mainly been carried out as a proof-of-concept. Examples include data for mobile phones (Deville et al. (2014)), crowdsourcing/volunteered geographic information (Bakillah et al. (2014)) and location-based social media (Aubrecht et al. (2011) and (2017)). For example, in countries with a high mobile phone penetration rate and many mobile phone towers, the night-time location of mobile phones could be used to generate a high-resolution population grid. Some promising approaches involve the integration of conventional with unconventional data sources, for example, combining official statistics with big data from remote sensing, volunteered geographic information, social media and mobile phones (Aubrecht et al. (2018)).

However promising, there are a number of issues concerning these types of data and technologies, for example, the sustainability of such approaches, data access and ownership, privacy and anonymity of social media users, or representation bias (Zhang and Zhu (2018)). The main challenge for developers is how to scale-up highly localised approaches to wide geographical areas (continents, the world) to provide datasets that are open and free (in a sustainable way). Given these as yet unsolved challenges, such data cannot currently be used as a reliable substitute for an official population and housing census that — in addition to complying with strict technical and statistical specifications — collects a wealth of additional information on population characteristics and living conditions.

References

Aubrecht, C., D. O. Aubrecht, J. Ungar, S. Freire and K. Steinnocher (2017), 'VGDI — advancing the concept: volunteered geo-dynamic information and its benefits for population dynamics modeling', *Transactions in GIS*, Volume 21, Issue 2, pp. 253-276.

Aubrecht, C., J. Ungar and S. Freire (2011), 'Exploring the potential of volunteered geographic information for modeling spatio-temporal characteristics of urban population: a case study for Lisbon metro using foursquare check-in data', *Proceedings of the 7th International Conference on Virtual Cities and Territories*, pp. 57-60.

([8]) For example, the United Nations Population Fund (https://www.unfpa.org/resources/new-methodology-hybrid-census-generate-spatially-disaggregated-population-estimates).

([9]) For example, the United Nations Population Fund or GRID3 (https://grid3.org/solution/high-resolution-population-estimates).

([10]) *WorldPop* Open Population Repository (https://wopr.worldpop.org/).

Aubrecht, C., J. Ungar, D. O. Aubrecht, S. Freire and K. Steinnocher (2018), 'Mapping land use dynamics using the collective power of the crowd', *Earth Observation Open Science and Innovation*, ISSI Scientific Report Series, Volume 15, pp. 247-253.

Bakillah, A., S. Liang, A. Mobasheri, J. J. Arsanjani and A. Zipf (2014), 'Fine-resolution population mapping using OpenStreetMap points-of interest', *International Journal of Geographical Information Science*, Volume 28, Issue 9, pp. 1 940-1 963.

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

Corbane, C., F. Sabo, V. Syrris, T. Kemper, P. Politis, M. Pesaresi, P. Soille and K. Osé (2019), 'Application of the Symbolic Machine Learning to Copernicus VHR Imagery: the European Settlement Map', *IEEE Geoscience and Remote Sensing Letters*, Volume 17, Issue 7, pp.1 153-1 157.

Deville, P., C. Linard, S. Martin, M. Gilbert, F. R. Stevens, A. E. Gaughan, V. D. Blondel, A. J. Tatem (2014), 'Dynamic population mapping using mobile phone data', *Proceedings of the National Academy of Sciences of the United States of America*, Volume 111, No. 45, pp. 15 888-15 893.

Eurostat (2019), *Methodological manual on territorial typologies — 2018 edition*, Publications Office of the European Union, Luxembourg.

Florczyk, A., C. Corbane, D. Ehrlich, S. Freire, T. Kemper, L. Maffenini, M. Melchiorri, M. Pesaresi, P. Politis, M. Schiavina, F. Sabo, and L. Zanchetta (2019), *GHSL Data Package 2019*, JRC 117104, EUR 29788 EN, Publications Office of the European Union, Luxembourg.

Tatem, A. (2017), 'WorldPop, open data for spatial demography', *Scientific Data 4*, Article No. 170004.

Wardrop, N. A., W. C. Jochem, T. J. Bird, H. R. Chamberlain, D. Clarke, D. Kerr, L. Bengtsson, S. Juran, V. Seaman, A. J. Tatem (2018), 'Spatially disaggregated population estimates in the absence of national population and housing census data', *Proceedings of the National Academy of Sciences*, Volume 115, No. 14, pp. 3 529-3 537.

Zhang, G. and A-X. Zhu (2018), 'The representativeness and spatial bias of volunteered geographic information: a review', *Annals of GIS*, Volume 24, Issue 3, pp. 151-162.

6. Methodology for producing statistics by level 1 of the degree of urbanisation

This chapter presents the key methodological recommendations for the production of internationally comparable statistics along an urban-rural continuum. It describes how to compile statistics for level 1 of the degree of urbanisation classification — providing analyses for cities, and for urban and rural areas — and is recommended as the basis for a territorial classification of indicators on sustainable development goals (SDGs).

6.1 Terminology

Two sets of terms have been developed to describe level 1 of the degree of urbanisation classification. The first set uses short and simple terms such as cities and rural areas. The second set uses more technical and neutral language. The second set can be helpful to avoid overlaps with the terms used in national definitions.

Table 6.1: Short and technical terms for classifying grid cells by degree of urbanisation

Short terms	Technical terms
Urban centres	High density clusters
Urban clusters	Moderate density clusters
Rural grid cells	Mostly low density cells

Small spatial units can be administrative units — such as municipalities — or statistical areas — such as census units (enumeration areas).

Table 6.2: Short and technical terms for classifying small spatial units by degree of urbanisation

Short terms	Technical terms
Cities	Densely populated areas
Towns and semi-dense areas	Intermediate density areas
Rural areas	Thinly populated areas

6.2 Short description

Level 1 of the degree of urbanisation classifies small spatial units as (i) cities or densely populated areas, (ii) towns and semi-dense areas or intermediate density areas and (iii) rural areas or thinly populated areas. This is done using 1 km² grid cells, classified according to their population density, population size and contiguity (neighbouring cells). Each small spatial unit belongs exclusively to one of the three classes.

Urban areas consist of cities plus towns and semi-dense areas. Because level 1 of the degree of urbanisation classification was developed to capture the urban-rural continuum, it is recommended to report indicators for all three classes instead of only for the urban-rural dichotomy. This is important because towns and semi-dense areas may differ significantly both from cities and from rural areas. Semi-dense areas in low- and middle-income countries are often described as peri-urban areas. In high-income countries, they are usually described as suburbs. In both cases, these areas have a moderate density and are at the transition between a rural area and a city or town.

Within national statistical systems, there is generally a high level of agreement concerning the two outermost classes: cities are typically classified as being urban, while villages and sparsely-populated areas are typically classified as being rural. By contrast, the classification of intermediate areas is less clear-cut: some countries prefer to classify them as urban, others as rural, with a third group of countries choosing to create an intermediate class between these two extremes. The degree of urbanisation classification tries to accommodate these intermediate areas and different points of view to emphasise that towns and semi-dense areas are partway between a city and a rural area. This is important because policymaking that is uniformly applied across the three classes may not be suitable and could benefit from being tailored to the specific requirements of cities, towns and semi-dense areas or rural areas.

6.3 Grid cell classification

The basis for the degree of urbanisation classification is a 1 km² population grid (for more details on how to construct a population grid, see Chapter 5). Each grid cell has the same shape and surface area, thereby avoiding distortions caused by using units varying in shape and size. This is a considerable advantage when compared with alternative approaches such as those based on the use of population data for local administrative units (for example municipalities).

The use of relatively small (1 km²) and uniform grid cells means that the basic concept underlying the methodology is to look inside larger local administrative units to detect the presence of individual cities, towns and semi-dense areas as well as rural areas. This makes it possible to create a more accurate classification.

Understanding contiguous cells

Before looking at the identification of the three different cluster types, it is necessary to understand the concept of contiguous cells. Figure 6.1 shows an array of nine grid cells, with the focus on the central cell which is surrounded by eight others, numbered 1 to 8.

Figure 6.1: Contiguous grid cells

1	2	3
4		5
6	7	8

Two types of contiguous grid cells can be identified:
(i) **four-point contiguity**, which is a **narrower definition excluding diagonals** — all cells that touch each other excluding those cells that only touch each other on a diagonal; only cells numbered 2, 4, 5 and 7 are contiguous to the central cell in Figure 6.1 according to this narrower definition.

(ii) **eight-point contiguity**, which is a **broad definition including diagonals** — all cells that touch each other in any way, including cells that are linked only on a diagonal; all cells numbered 1 to 8 are contiguous to the central cell in Figure 6.1 according to this broader definition.

Stage 1: classifying grid cells

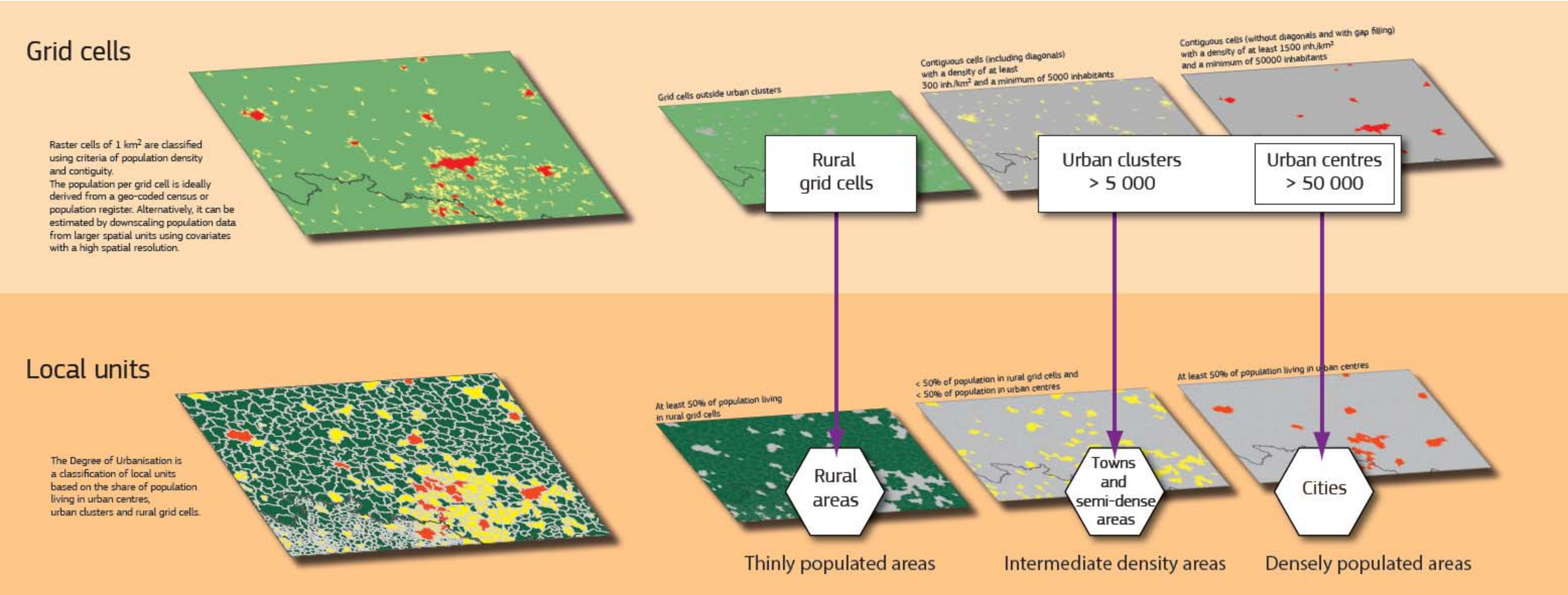
Each cluster type is identified by classifying 1 km² population grid cells according to characteristics that are based on their total population and population density.

Groups of 1 km² population grid cells are plotted in relation to their neighbouring cells to identify:

- **An urban centre (high-density cluster)** — a cluster of contiguous grid cells of 1 km² (using four-point contiguity, in other words, excluding diagonals) with a population density of at least 1 500 inhabitants per km² and collectively a minimum population of 50 000 inhabitants after gap-filling; if needed, cells that are 50 % built-up may be added (see Subchapter 7.3).
- **An urban cluster (moderate-density cluster)** — a cluster of contiguous grid cells of 1 km² (using eight-point contiguity, in other words, including diagonals) with a population density of at least 300 inhabitants per km² and a minimum population of 5 000 inhabitants.
- **Rural grid cells (or mostly low-density cells)** — grid cells that are not identified as urban centres or as urban clusters.

Note that a grid cell may belong to an urban centre and an urban cluster as their definitions are not mutually exclusive. To create a grid layer where every cell is allocated to one and only one class, urban centre cells can be excluded from the urban clusters. This mutually exclusive grid layer can be used to classify the small spatial units without any modification of the definitions. This mutually exclusive grid layer will also closely match the classification of the spatial units, especially when they are small.

Figure 6.2: Schematic overview of the degree of urbanisation classification



Note: for more information, see http://ec.europa.eu/regional_policy/sources/docgener/work/2014_01_new_urban.pdf.

Source: Directorate-General Regional and Urban Policy, based on data from Eurostat, JRC, national statistical authorities

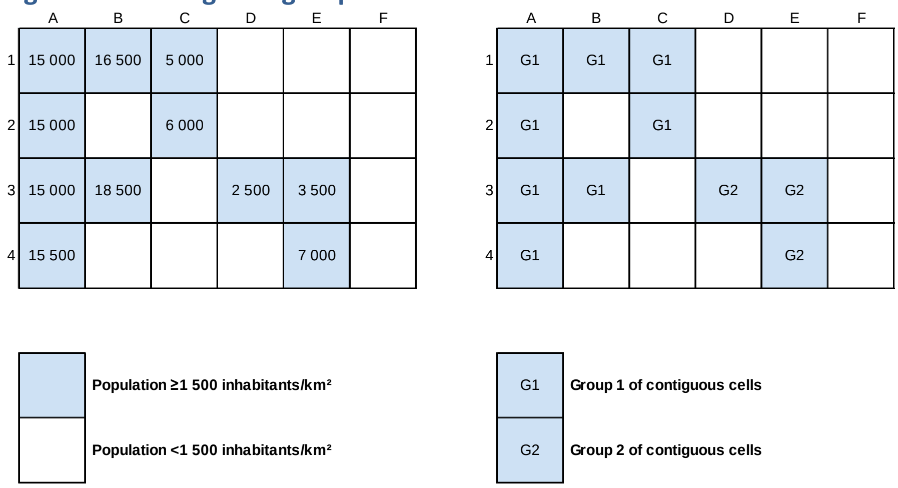
6.3.1 Urban centres (high-density clusters)

The identification of urban centres (high-density clusters) is done in three steps. The first step involves identifying groups of contiguous cells:

- all cells with a population density of at least 1 500 inhabitants per km² are plotted (light blue shading in Figure 6.3);
- groups of contiguous grid cells are identified (groups G1 and G2 in Figure 6.3). If needed, cells that are 50 % built-up may be added (see Subchapter 7.3).

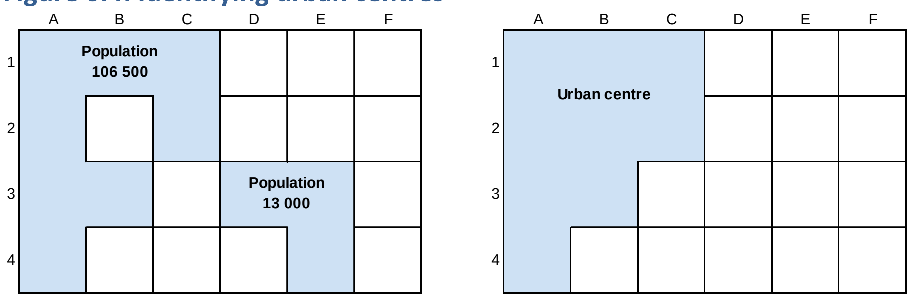
Contiguous cells are grouped together, however, when identifying urban centres diagonal contiguity is excluded. As such, in the example of Figure 6.3, cells C2 and D3 are not considered as contiguous; rather, they are each part of different groups (G1 and G2).

Figure 6.3: Contiguous groups for urban centres



In a second step, each group of contiguous grid cells is analysed in relation to its total number of inhabitants and only those groups of contiguous cells with collectively 50 000 inhabitants or more are selected (see Figure 6.4). Continuing with the same example, Group G1 is considered an urban centre as it has a population of 106 500 inhabitants, as shown in Figure 6.4, while G2 is not an urban centre as its population is only 13 000 inhabitants.

Figure 6.4: Identifying urban centres



The third step for identifying urban centres is taken to fill gaps and smooth borders. This is done by applying an iterative majority rule.

The iterative 'majority rule'

If five or more of the (eight) cells surrounding a particular cell belong to the same unique urban centre, then that cell is also considered to belong to the same urban centre; this process is repeated (iteratively) until no more cells may be added.

Note that the criterion for gap-filling following the majority rule includes cells that are linked only on a diagonal. For example, cell B2 on the left-hand side of Figure 6.4 has seven of its eight surrounding cells that belong to the same urban centre. This cell should therefore subsequently be added to the urban centre to smooth borders (as shown on the right-hand side of Figure 6.4).

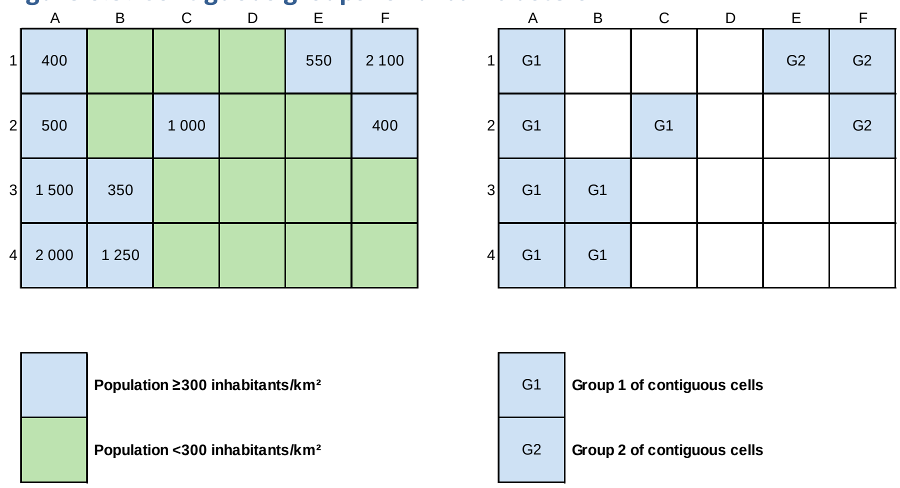
6.3.2 Urban clusters (or moderate-density clusters)

The technique used to identify urban clusters (moderate-density clusters) is similar to that used for urban centres (high-density clusters). Rather than using a threshold of at least 1 500 inhabitants per km², the identification of urban clusters is based on grid cells with a population density of at least 300 inhabitants per km² (see Figure 6.5).

The initial identification of urban clusters is done in two steps:

- all cells with a population density of at least 300 inhabitants per km² are plotted (light blue shading in Figure 6.5);
- groups of contiguous grid cells are identified (groups G1 and G2 in Figure 6.5); note that contiguous grid cells may include cells that are linked only on a diagonal (eight-point contiguity) — as shown, for example, by cell C2.

Figure 6.5: Contiguous groups for urban clusters



Thereafter, each group of contiguous grid cells is analysed in relation to its number of inhabitants and those groups of contiguous cells with collectively 5 000 inhabitants or more are selected; these are urban clusters. Continuing with the same example, Group G1 is considered an urban cluster as it has a population of 7 000 inhabitants, as shown in Figure 6.6, while G2 is not an urban cluster as its population is only 3 050 inhabitants.

Figure 6.6: Identifying urban clusters

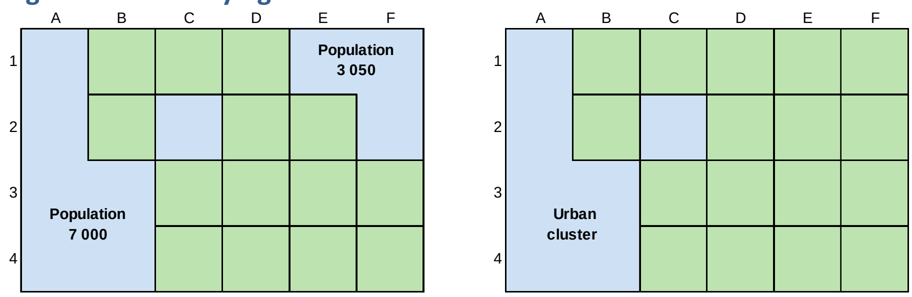
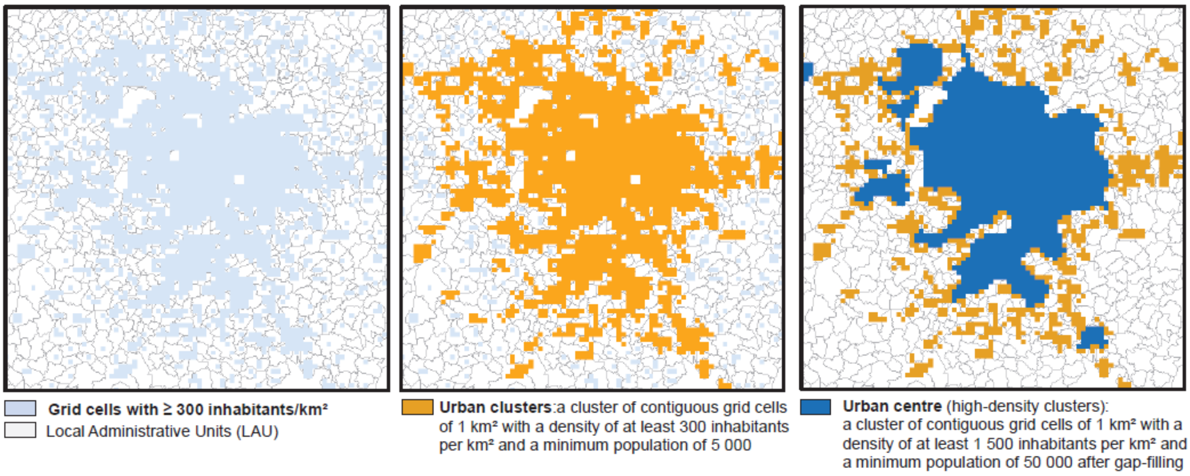


Figure 6.7 shows a schematic overview from grid cell classification through to the identification of urban centres. In the first image, grid cells with a population density of at least 300 inhabitants per km² are identified. The second image overlays these grid cells showing urban clusters (moderate-density clusters) that are composed of contiguous grid cells linked by eight-point contiguity and at least 5 000 inhabitants. The final image overlays the information on urban clusters by identifying an urban centre — a set of contiguous grid cells that have a population density of at least 1 500 inhabitants per km² and at least 50 000 inhabitants (after gap-filling according to the iterative 'majority rule').

Figure 6.7: Schematic overview of identifying urban clusters and urban centres



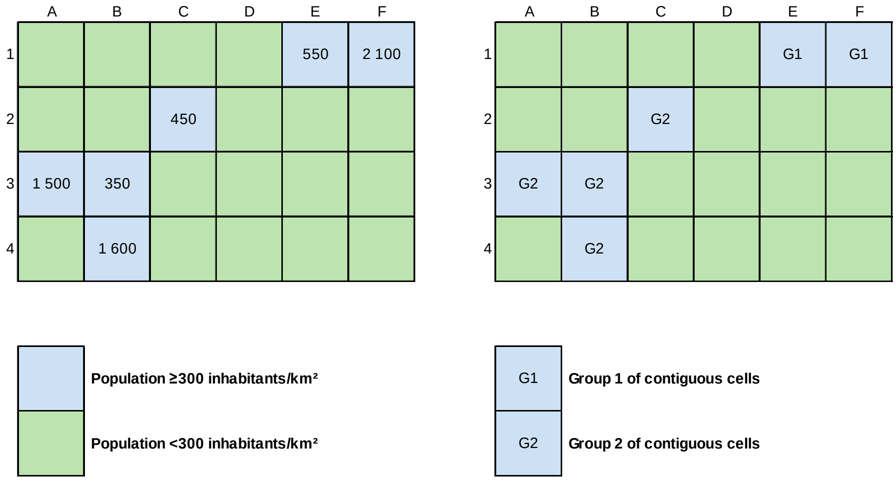
Source: Eurostat, JRC and European Commission, Directorate-General Regional and Urban Policy and Directorate-General Agriculture and Regional Development

6.3.3 Rural grid cells

Rural grid cells are those cells that are not identified as urban centres or as urban clusters. The majority of rural grid cells have a population density that is less than 300 inhabitants per km², although this is not necessarily the case. Some rural grid cells may have a higher number of inhabitants if they do not form part of a cluster that meets the criteria for an urban centre or an urban cluster.

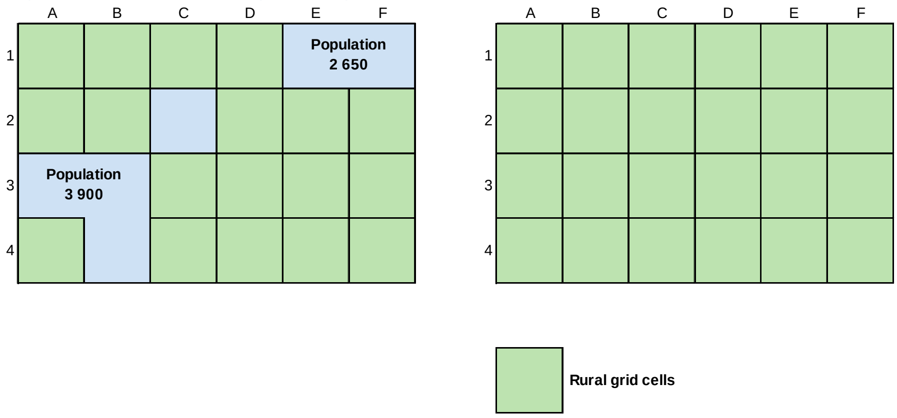
In Figure 6.8, cells A3, B4 and F1 each meet the population criterion for an urban centre (at least 1 500 inhabitants per km²), while cells B3, C2 and E1 each meet the population criterion for an urban cluster (at least 300 inhabitants per km²).

Figure 6.8: Detecting rural grid cells



Each group of contiguous grid cells (groups G1 and G2 in the right-hand side of Figure 6.8) may be analysed in relation to their total number of inhabitants and those groups of contiguous cells with collectively 5 000 inhabitants or more are selected. In Figure 6.9, it can be seen that neither group G1 with a total population of 3 900 inhabitants nor group G2 with a total population of 2 650 inhabitants reaches the population threshold for an urban cluster. As such, each cell in these two groups is classified as a rural grid cell, as shown on the right-hand side of Figure 6.9.

Figure 6.9: Identifying rural grid cells



Note also, as mentioned above, that it is possible for grid cells with a population density of less than 300 inhabitants per km² to be classified as part of an urban centre, due to gap-filling or as a result of adding cells that are 50 % built-up (see Subchapter 7.3).

6.4 Classifying small spatial units

Stage 2: classifying small spatial units by degree of urbanisation

Once all grid cells have been classified and urban centres, urban clusters and rural grid cells identified, the next step concerns overlaying these results onto small spatial units, as follows:

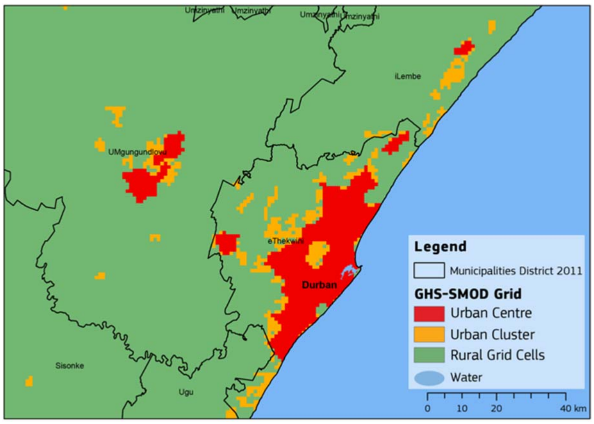
- **cities (or densely populated areas)** — small spatial units that have at least 50 % of their population in urban centres;
- **towns and semi-dense areas (or intermediate density areas)** — small spatial units that have less than 50 % of their population in urban centres and less than 50 % of their population in rural grid cells;

- **rural areas (or thinly populated areas)** — small spatial units that have at least 50 % of their population in rural grid cells.

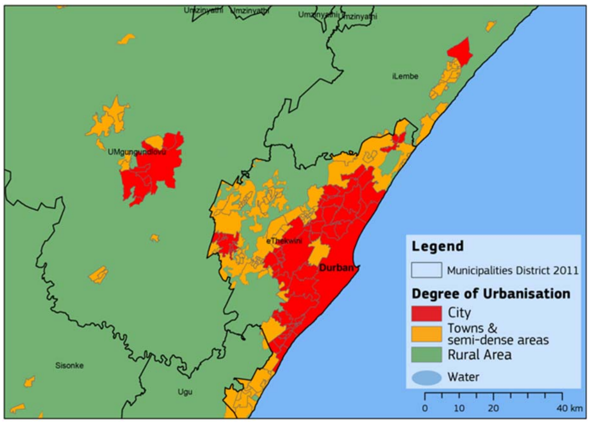
Cities (or densely populated areas) consist of one or more small spatial units with at least 50 % of their population in an urban centre. A small spatial unit can be either an administrative unit or a statistical area. Examples of administrative units include a municipality, a district, a neighbourhood or a metropolitan area. Some of these administrative units also have a political role as electoral districts or in terms of local government. Statistical areas can be census units/enumeration areas, census blocks, census tracts, wards, super output areas, named places or small areas.

Map 7.1 shows the grid cell classification for Durban in South Africa and Map 7.2 shows the classification of small spatial units.

Map 6.1: Grid cell classification around Durban, South Africa



Map 6.2: Classification of small spatial units around Durban, South Africa

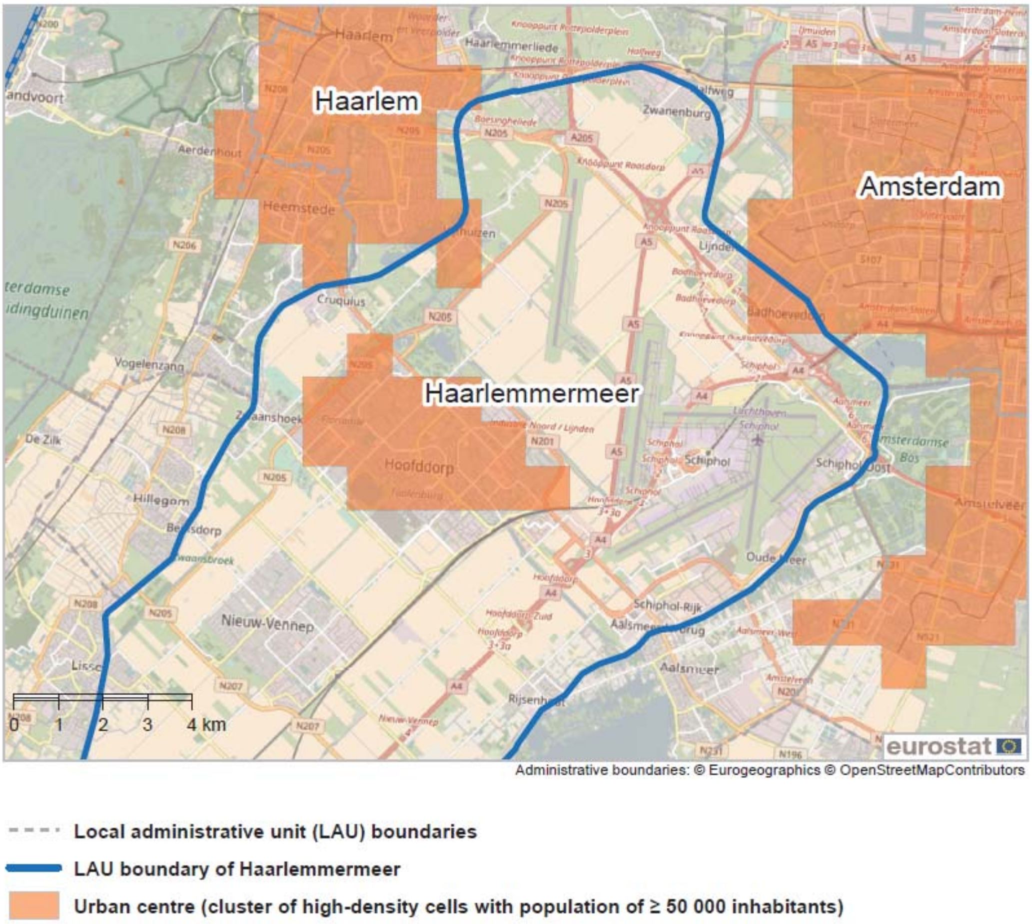


Source: Florczyk et al. (2019)

Note that each small spatial unit should be classified to one and only one of the three classes within level 1 of the degree of urbanisation classification. However, in order to classify small spatial units based on the population grid, these units have to be transformed into a raster as well, which can lead to some situations which require case-by-case solutions (see Subchapters 7.2.4 and 7.2.6 for more information on different types of adjustments that may be made).

Map 6.3 shows that, when classifying small spatial units as cities, it may be necessary to consider more than one urban centre. In this example, there were 65 593 people living in the urban centre of Haarlemmermeer in the Netherlands, which equated to just 46 % of the total population of the small spatial unit for Haarlemmermeer (below the threshold of 50 % that is required to identify a city). Nevertheless, as shown in the example, there were two adjacent small spatial units — Amsterdam and Haarlem — and their urban centres spill over into Haarlemmermeer. Aggregating the total population of the three urban centres that are located within the boundaries of Haarlemmermeer results in the share of those living in urban centres rising to some 54 % of the total population; as such, Haarlemmermeer is classified as a city.

Map 6.3: More than one urban centre needed to define a city — an example for Haarlemmermeer, the Netherlands



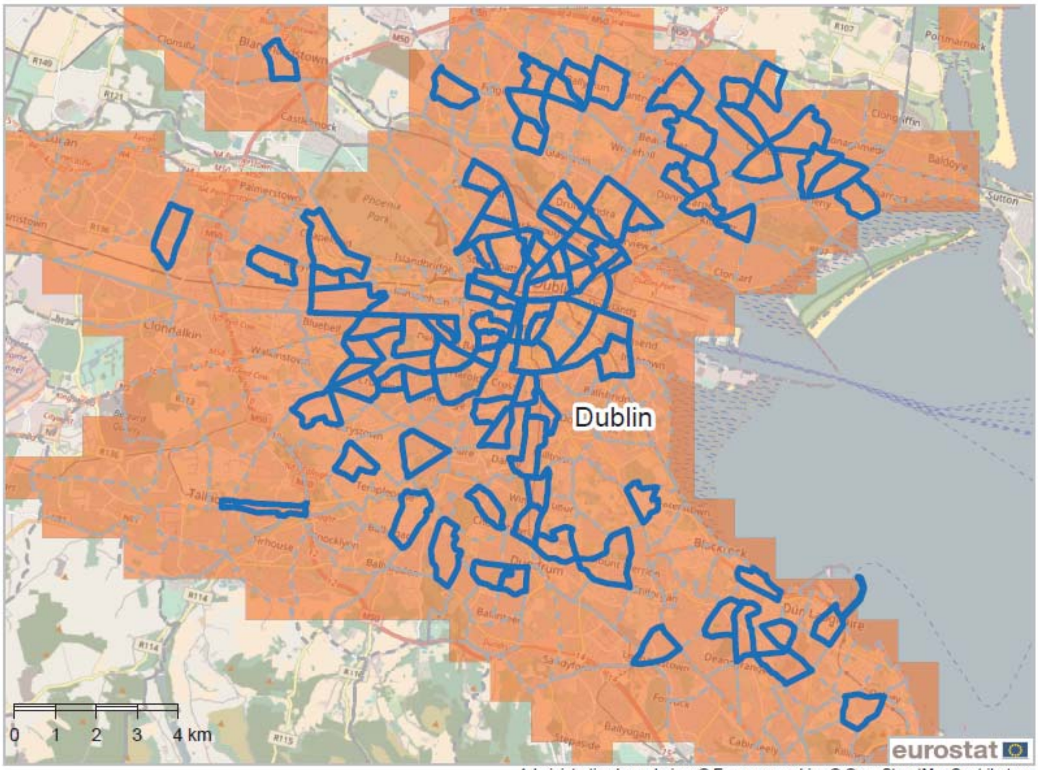
Note: GEOSTAT population grid from 2011 and small spatial units for 2016.
Source: Eurostat, JRC and European Commission, Directorate-General Regional and Urban Policy and Directorate-General Agriculture and Regional Development

Small spatial units with no population in the raster equivalent

Some small spatial units will be too small to have a 1 km² grid cell equivalent. When determining their class within level 1 of the degree of urbanisation, these small spatial units are not assigned any population as they are physically too small (smaller than one grid cell); as such, they are given no initial classification.

After the initial classification, these remaining small spatial units can be selected. For each small spatial unit a centroid falling within its boundaries should be determined. These centroids can be used to classify the remaining small spatial units. They should be spatially joined to the grid-based typology, whereby the small spatial unit gets the classification of the grid type in which its centroid falls. In the EU, such small spatial units were found to be exclusively in urban centres. An example is provided for Dublin in Ireland (Map 6.4).

Map 6.4: Small spatial units with no population in the raster equivalent — an example for Dublin, Ireland



Administrative boundaries: © Eurogeographics © OpenStreetMapContributors

- - - - Local administrative unit (LAU) boundaries

LAU without raster equivalent

Urban centre (cluster of high-density cells with population of ≥ 50 000 inhabitants)

Note: GEOSTAT population grid from 2011 and small spatial units for 2016.

Source: Eurostat, JRC and European Commission, Directorate-General Regional and Urban Policy and Directorate-General Agriculture and Regional Development

6.5 Changes over time that impact on the classification given to each small spatial unit

The classification given to each small spatial unit according to level 1 of the degree of urbanisation classification should be updated to reflect any changes to the underlying sources of information that are used to determine their class. As such, the classes may be updated to reflect: changes to small spatial unit boundaries or changes to population distributions for 1 km² grid cells. The frequency of such updates varies according to the source of information.

Changes to the classification given to each small spatial unit resulting from a revision of population distributions for 1 km² grid cells are less common and these may be expected every 5 or 10 years, when new census data become available.

Annual updates of the degree of urbanisation classes assigned to small spatial units should be made to reflect changes to small spatial unit boundaries. These modifications can be implemented in two ways: applying the methodology for the degree of urbanisation classification as described above for the new layer of small spatial units; or estimating the degree of urbanisation based on changes to small spatial unit boundaries. The first approach is more labour intensive, while the second is particularly suitable if boundary changes for small spatial units are relatively minor or consist principally of merging small spatial units, especially if these have the same class at level 1 of the degree of urbanisation classification.

Updating to reflect changes in small spatial unit boundaries

Small spatial unit boundaries may change over time in three different ways: small spatial units may merge, they may undergo a boundary shift, or they may be split. The most common change for small spatial units within the EU in recent years has been for two or more small spatial units to be merged; boundary shifts have been less common, while splitting small spatial units has been rare.

Case 1: small spatial unit mergers

Merging two small spatial units with different degrees of urbanisation may be resolved by giving precedence to the more densely populated spatial unit:

- when merging small spatial units composed of a city and a town or semi-dense area, reclassify the new small spatial unit as a city;
- when merging small spatial units composed of a town or semi-dense area and a rural area, reclassify the new small spatial unit as a town or semi-dense area.

Such a process may be further refined by taking into account the relative population sizes of the two small spatial units.

Case 1a: small spatial unit mergers involving the same degree of urbanisation

The degree of urbanisation classification is additive, meaning that if two small spatial units classified as rural areas are subsequently merged into a single small spatial unit then they will remain a rural area; this is also true for the other classes in the classification.

Case 1b: small spatial unit mergers involving rural areas and towns and semi-dense areas

These mergers can be addressed in two simple ways: using the population of the urban cluster or using the population of the small spatial units.

In the first case, if the population of the relevant urban cluster(s) is available then add the population inhabiting the urban cluster for each of the small spatial units and divide this by the total population of the new small spatial unit to determine the new degree of urbanisation class. If more than 50 % of the population of the new small spatial unit lives in an urban cluster, the new small spatial unit should be classified under towns and semi-dense areas. If the population share is less than 50 %, then the new small spatial unit should be classified under rural areas.

In the second case, if the population living in the urban cluster cannot be identified, then the degree of urbanisation class may be determined based on the population distribution between the small spatial units. If more than 50 % of the population of the new small spatial unit comes from rural areas, the new small spatial unit should be classified under rural areas. If more than 50 % of the population of the new small spatial unit comes from towns and semi-dense areas, the new small spatial unit should be classified under towns and semi-dense areas.

Case 2: small spatial unit boundary shifts

Whereas mergers can be dealt with using simple techniques, boundary shifts cannot always be as reliably addressed. Indeed, in some rare cases, boundary shifts between small spatial units that have the same degree of urbanisation class can lead to a change in the classification given to the small spatial units. Such complexity means that a simple rule of thumb is often the preferred and most efficient approach.

A simple rule may be established whereby if a small spatial unit loses less than 25 % of its previous population or gains less than 50 % of its population due to boundary shifts, then the degree of urbanisation class does not change. This rule of thumb is likely to cover 90 % of all boundary shifts and ensures continuity. If this is not the case, then further investigation is required, as described below.

Case 2a: changes in the degree of urbanisation classification due to boundary shifts are excluded

For each small spatial unit, the share of population in the three different types of population grids cells is known. For example, if as the result of a boundary shift the population of a small spatial unit that has 100 % of its population in rural grid cells shrinks, then it will remain classified under rural areas. Equally, if a boundary shift for a small spatial unit that has 100 % of its population in rural grid cells rises, then the new small spatial unit would need to more than double its population before it could (potentially) be classified under towns and semi-dense areas. As a result, if a boundary shift leads to a change in population that is too small to tip the population share of the revised small spatial unit below 50 % of the relevant grid cells, it remains in the same degree of urbanisation class.

Case 2b: changes in the degree of urbanisation classification due to boundary shifts are unlikely (but cannot be excluded)

If the boundary shift leads to a change in population that is theoretically sufficient to tip the population share of the revised small spatial unit below or above 50 %, but the shift is between small spatial units with the same classification by degree of urbanisation, then the same class should be maintained.

Case 2c: changes in the degree of urbanisation classification due to boundary shifts are likely

In some cases, changes in the degree of urbanisation class are likely. As an example, if a city were to gain part of a suburb (classified under towns and semi-dense areas) as a result of a boundary shift. The city gains a small number of additional inhabitants (which does not have an impact on its classification by degree of urbanisation). The suburb loses some of its population (that is reclassified to the city). As a result, the population in the revised small spatial unit covered by the suburb may have less than 50 % of its population living in an urban cluster in which case it should subsequently be reclassified under rural areas.

Case 3: splitting small spatial units

This type of change is relatively rare. Therefore, the main recommendation is one of continuity; in other words, maintain the same degree of urbanisation class. If a small spatial unit is split, the new small spatial units should have the same degree of urbanisation class as the old small spatial unit. If there are concerns that the new small spatial units may have a different degree of urbanisation class, the same approaches as described for boundary shifts may be used.

References

Florczyk, A., C. Corbane, D. Ehrlich, S. Freire, T. Kemper, L. Maffenini, M. Melchiorri, M. Pesaresi, P. Politis, M. Schiavina, F. Sabo, and L. Zanchetta (2019), *GHSL Data Package 2019*, JRC 117104, EUR 29788 EN, Publications Office of the European Union, Luxembourg.

7. Extensions to level 1 of the classification

The first two sections of this chapter describe possible extensions to level 1 of the degree of urbanisation classification: how to compile statistics for level 2 of the degree of urbanisation classification and how to compile statistics for functional urban areas (otherwise referred to as metropolitan areas). Both of these extensions have the potential to provide additional useful insight into the spatial structure of a territory/country. The final section details how specific geographic issues should be addressed from a methodological standpoint and provides information on further possible extensions.

7.1 Level 2 of the degree of urbanisation

The three classes assigned under level 1 of the degree of urbanisation provide an important first step to assess the urban-rural continuum. Cities are clearly defined settlements which can be organised by population size. The other two classes, however, are quite heterogeneous and do not identify specific types of settlement. The level 1 class of towns and semi-dense areas includes towns, but it does not separate them from semi-dense areas. Equally, rural areas contain villages, but the degree of urbanisation level 1 does not separate them from other thinly populated areas. Therefore, a second level or sub-classification has been introduced to capture the full settlement hierarchy of large, medium and small settlements or, in simpler terms, cities, towns and villages.

7.1.1 Terminology

Two sets of terms have been developed to describe level 2 of the degree of urbanisation classification. The first set uses simple and short terms such as city, town and village. The second set uses more neutral and technical language. The second set can be helpful to avoid overlaps with the terms used in national definitions.

Table 7.1: Short and technical terms for classifying grid cells for levels 1 and 2 of the degree of urbanisation classification

	Grid classification	
Level	Short terms	Technical terms
1	**Urban centre**	**High density cluster**
2	Urban centre	Dense, large cluster
1	**Urban cluster**	**Moderate density cluster**
2	Dense urban cluster	Dense, medium cluster
2	Semi-dense urban cluster	Semi-dense, medium cluster
2	Suburban or peri-urban grid cells	Semi-dense grid cells
1	**Rural grid cells**	**Mostly low density cells**
2	Rural cluster	Semi-dense, small cluster
2	Low density rural grid cells	Low density grid cells
2	Very low density rural grid cells	Very low density grid cells

Small spatial units can be administrative units — such as municipalities — or statistical areas — such as census units (enumeration areas).

Table 7.2: Short and technical terms for classifying small spatial units to levels 1 and 2 of the degree of urbanisation classification

	Local Unit Classification	
Level	Short terms	Technical terms
1	**City**	**Densely populated area**
2	City	Large settlement
1	**Town & semi-dense area**	**Intermediate density area**
2	Dense town	Dense, medium settlement
2	Semi-dense town	Semi-dense, medium settlement
2	Suburban or peri-urban area	Semi-dense area
1	**Rural area**	**Thinly populated area**
2	Village	Small settlement
2	Dispersed rural area	Low density area
2	Mostly uninhabited area	Very low density area

Semi-dense areas in low- and middle-income countries are often described as peri-urban areas. In high-income countries, they are usually described as suburbs. In both cases, these areas have a moderate density and are at the transition between a rural area and a city or town.

7.1.2 Short description

Level 2 of the degree of urbanisation classification is a hierarchical sub-classification of level 1. It was created to identify medium and small settlements, in other words, towns and villages. Practically, it splits two classes into six sub-classes.

- Towns and semi-dense areas are split into three subclasses:
  - (i) dense towns;
  - (ii) semi-dense towns;
  - (iii) suburban or peri-urban cells; and
- Rural areas are split into three subclasses:
  - (i) villages;
  - (ii) dispersed rural areas;
  - (iii) mostly uninhabited areas.

Level 2 of the degree of urbanisation classification is implemented with the same two-stage approach as level 1 of the classification. Firstly, grid cells are classified based on population density, population size and contiguity. Subsequently, small spatial units are classified according to the type of grid cells in which their population resides.

7.1.3 Grid cell classification

Stage 1: classifying grid cells

An urban centre is identified in the same manner as for the degree of urbanisation level 1.

- **An urban centre** consists of contiguous (using four-point contiguity) grid cells with a density of at least 1 500 inhabitants per km^2. An urban centre has a collective population of at least 50 000. Gaps in this cluster are filled and edges are smoothed. If needed, cells that are 50 % built-up may be added (see Subchapter 7.3).

The urban cluster cells that are not part of an urban centre can be subdivided into three types.

- **A dense urban cluster** consists of contiguous (using four-point contiguity) grid cells with a density of at least 1 500 inhabitants per km^2, with a collective population of at least 5 000 and less than 50 000 in the cluster.
- **A semi-dense urban cluster** consists of contiguous (using eight-point contiguity) grid cells with a density of at least 300 inhabitants per km^2 and has a collective population of at least 5 000 (in other words, an urban cluster) and this cluster is neither contiguous with nor within 2 km of a dense urban cluster or an urban centre ([11]).
- **Suburban or peri-urban cells** are the remaining urban cluster cells, in other words those not part of a dense or semi-dense urban cluster. These grid cells are part of an urban cluster that is contiguous (using eight-point contiguity) or within 2 km of a dense urban cluster or an urban centre.

Rural grid cells can be categorised into three types.

- **A rural cluster** consists of contiguous (using eight-point contiguity) grid cells with a density of at least 300 inhabitants per km^2 and a collective population between 500 and 5 000 in the cluster.
- **Low density rural grid cells** are rural grid cells with a density of at least 50 inhabitants per km^2 and are not part of a rural cluster.
- **Very low density rural grid cells** are rural grid cells with a density of less than 50 inhabitants per km^2.

7.1.4 Classifying small spatial units

Stage 2: classifying small spatial units

For level 2 of the degree of urbanisation classification, small spatial units are classified as cities in the same manner as in level 1.

- **A city consists** of one or more small spatial units that have at least 50 % of their population in an urban centre.

Within level 2 of the classification, small spatial units classified as towns and semi-dense areas can be divided into three subclasses.

- **Dense towns** have a larger share of their population in dense urban clusters than in semi-dense urban clusters (in other words, they are dense) and have a larger share of their population in dense plus semi-dense urban clusters than in suburban or peri-urban cells (in other words, they are towns).
- **Semi-dense towns** have a larger share of their population in semi-dense urban clusters than in dense urban clusters (in other words, they are semi-dense) and have a larger share of their population in dense plus semi-dense urban clusters than in suburban or peri-urban cells (in other words, they are towns).
- **Suburban or peri-urban areas** have a larger share of their population in suburban or peri-urban cells than in dense plus semi-dense urban clusters.

([11]) Measured as outside a buffer of three grid cells of 1 km^2 around dense urban clusters and urban centres.

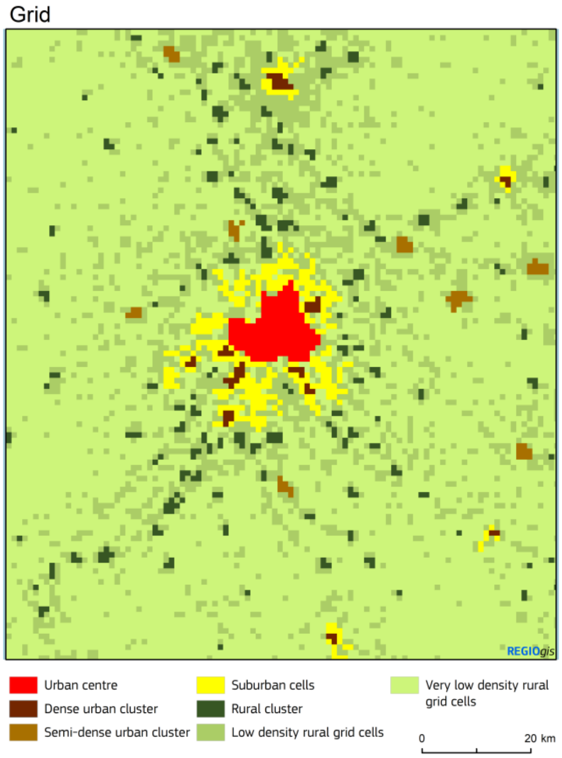
Dense and semi-dense towns can be combined into towns. This reduces the number of classes that are identified for level 2 of the classification and may be useful especially if the population share in semi-dense towns is low.

In a similar vein to towns and semi-dense areas, within level 2 of the classification small spatial units classified as rural areas can be divided into three subclasses.

- **Villages** have the largest share of their rural grid cell population living in a rural cluster.
- **Dispersed rural areas** have the largest share of their rural grid cell population living in low density rural grid cells.
- **Mostly uninhabited areas** have the largest share of their rural grid cell population living in very low density rural grid cells.

Map 7.1 and Map 7.2 show the application of the methodology to Toulouse and its surroundings.

Map 7.1: Grid cell classification around Toulouse, France for level 2 of the degree of urbanisation classification



Map 7.2: Small spatial unit classification around Toulouse, France for level 2 of the degree of urbanisation classification

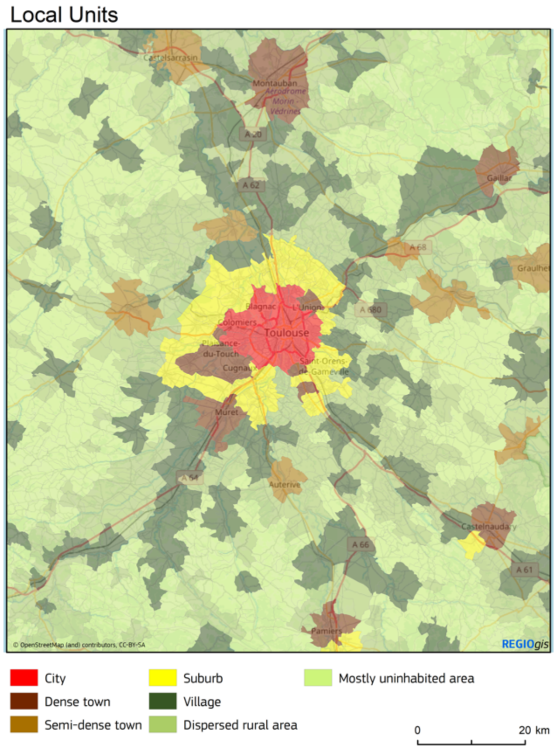


Figure 7.1 provides a simplified and schematic overview of level 2 of the degree of urbanisation classification.

Figure 7.1: Schema for the grid cell classification for level 2 of the degree of urbanisation classification

		Population size thresholds of the cluster of cells (settlement size)			No population size criterion (not a settlement)
		>= 50,000	5,000 - 49,999	500 - 4,999	
Population density of cells, inhabitants per km^2	>= 1500	Urban centre	Dense urban cluster		
	>= 300		Semi-dense urban cluster*	Rural cluster	Suburban or peri-urban grid cells
	>= 50				Low density rural grid cells
	<50				Very low density rural grid cells

* Semi-dense urban clusters can have a population of more than 49,999

7.2 Defining functional urban areas

The degree of urbanisation classification may be complemented by a classification of functional urban areas (FUAs) ([12]). A functional urban area (or metropolitan area) is composed of a city plus its surrounding, less densely populated spatial units that make up the city's labour market, its commuting zone. This commuting zone generates a daily flow of people into a city and back (home to their dwelling). Such areas are often referred to as 'functional' because they capture the full economic function of a city. A functional urban area classification is particularly useful to inform policymaking in a number of domains, including transport, economic development and planning. Several national statistical authorities, including those of Brazil, Italy, Japan and the United States, complement their urban and rural area classifications with a classification of metropolitan areas.

The functional urban area classification and the degree of urbanisation classification are linked because they use exactly the same concept of a city. The functional urban area classification is exhaustive, in other words it covers all of the small spatial units in a territory, as those areas that are not classified as functional urban areas (metropolitan areas) are classified as areas outside a functional urban area (non-metropolitan areas).

It should be noted that not all of the areas within a functional urban area need to be classified as urban areas and that, as such, a functional urban area may contain rural areas if these belong to the commuting zone of a city. In a similar vein, it is possible for an urban area to form part of an area outside a functional urban area if it does not belong to the commuting zone of a city.

7.2.1 Terminology

This section summarises the terms that are necessary to distinguish the different concepts that are used to define functional urban areas.

([12]) This subchapter is adapted from Dijkstra et al. (2019).

Table 7.3: Terminology related to functional urban areas

Preferred term	Synonym	Geographic level
Urban centre	High-density cluster (HDC)	Grid
City	Densely populated area	Small spatial unit
Commuting zone		Small spatial unit
Functional urban area (FUA)	Metropolitan area	Small spatial unit
Area outside a functional urban area (non-FUA)	Non-metropolitan area	Small spatial unit

7.2.2 Short description

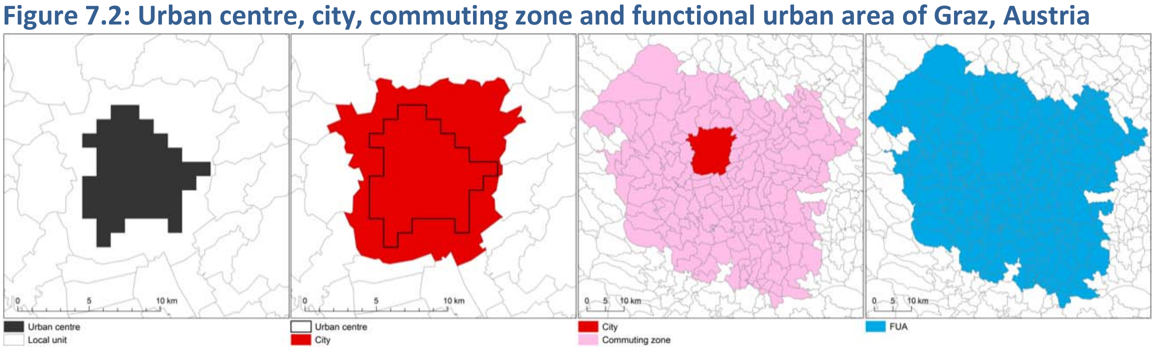
A **functional urban area (metropolitan area)** can be defined in four steps:

- Identify an urban centre — a set of contiguous grid cells with a density of at least 1 500 inhabitants per km² and with a collective population of at least 50 000.
- Identify a city — one or more small spatial units that have at least 50 % of their population in an urban centre.
- Identify a commuting zone — a set of contiguous small spatial units that have at least 15 % of their employed residents working in a city.
- A functional urban area (metropolitan area) is the combination of a city with its commuting zone.

Consequently, within the functional urban area classification, all the areas of a territory outside of cities and their commuting zones may be considered as areas outside a functional urban area (non-metropolitan areas).

Figure 7.2 shows visually the different concepts that are used in the classification of functional urban areas, notably the urban centre, the city, and the commuting zone.

Figure 7.2: Urban centre, city, commuting zone and functional urban area of Graz, Austria



The following data sources are required to compile statistics for functional urban areas:

- a residential population grid with the number of inhabitants per km² of land area (in other words, excluding water bodies);
- digital boundaries for small spatial units;
- commuting flows between the small spatial units and the number of employed residents per small spatial unit.

How to estimate commuting flows?

Several countries do not collect commuting data as part of their census. Other sources such as linked population and employment registers or mobile phone data could be used to estimate such flows.

Estonia offers an illustrative example where — as reported in two studies commissioned by the Ministry of the Interior and conducted by the Mobility Lab of the University of Tartu (Ahas and Silm (2013); Ahas et al. (2010)) — mobile positioning data made it possible to delineate functional urban areas (metropolitan areas). The movements between individuals' anchor points (in other words, residence, work, and so on) are aggregated at the level of small spatial units (in other words, municipalities) in order to produce a matrix of flows. Such a matrix has the benefit of providing an estimation of mobility patterns for the entire population, rather than for employees only, at a highly disaggregated spatial scale.

The Netherlands has also produced a flow matrix between all small spatial units within its territory using mobile phone data (Van der Valk et al. (2019)).

7.2.2.1 Definition of an urban centre

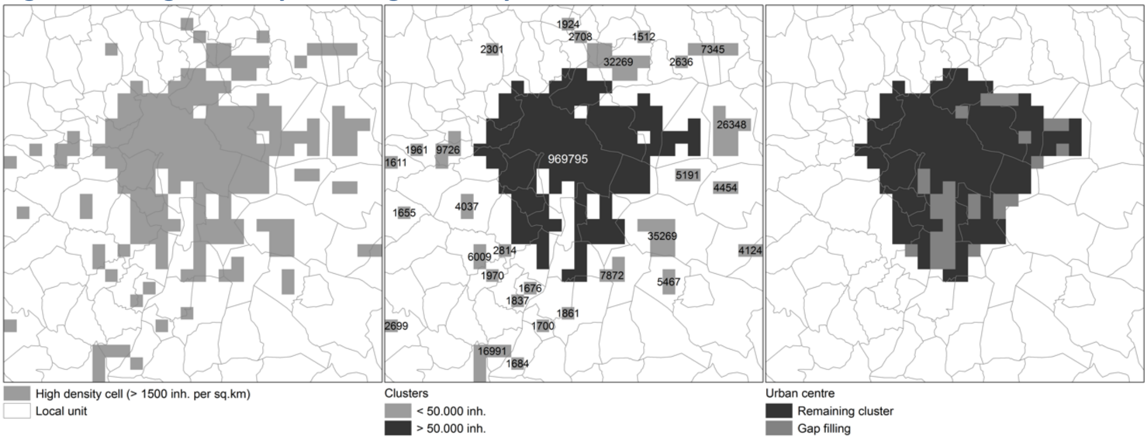
The first step focuses on the concentration of population in space, which is the simplest and most uncontroverted feature of a city — the starting point for this definition. The idea of a city as a place with a relatively high concentration of population in space is common to many disciplines that describe a city including economic, social, cultural and geographical ones.

Many national definitions of a city rely on the population size and density of a small spatial unit. This causes two types of problems. A big city in a relatively large spatial unit may have a very low or rural population density. For example, Ulaanbaatar, the capital city of Mongolia, has a population of 1.4 million but a density of only 270 inhabitants per km². The population of a city is difficult to determine when it is spread out over multiple small spatial units. For example, how many people live in Paris?

An urban centre, as defined in this methodological manual, relies on a population grid which can identify spatial concentrations of population independently from political or administrative boundaries, using spatial units of the same shape and size. An urban centre or high-density cluster is a spatial concept based on grid cells of 1 km². It is defined in three steps, as indicated below and represented in Figure 7.3.

- **Step 1:** all grid cells with a density of at least 1 500 inhabitants per km² of land are selected. If needed, cells that are 50 % built-up may be added (see Subchapter 7.3).
- **Step 2:** contiguous high-density cells are then clustered. Only those clusters with at least 50 000 inhabitants are kept. To avoid over-aggregation, four-point contiguity is used (in other words, cells with only the corners touching are not considered).
- **Step 3:** gaps in each cluster are filled separately and its edges smoothed.

Figure 7.3: High-density cells, high-density clusters, urban centre of Toulouse, France

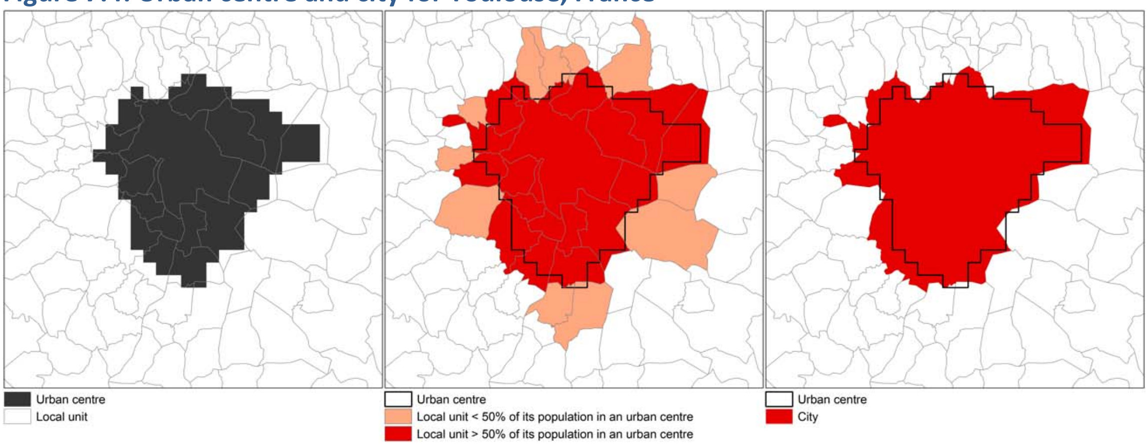


7.2.2.2 Definition of a city

A city consists of one or more small spatial units with at least 50 % of their population in an urban centre. A small spatial unit can be either an administrative unit or a statistical area. Examples of administrative units include a municipality, a district, a neighbourhood or a metropolitan area. Some of these administrative units also have a political role as electoral districts or in terms of local government. Statistical areas can be census units/enumeration areas, census blocks, census tracts, wards, super output areas, named places or small areas. Examples of small spatial units used in OECD countries include communes in France, municipalities in Italy, *sigungu* in South Korea and census subdivisions in Canada.

The best small spatial unit for this definition is the smallest unit for which commuting data are available ([13]). Figure 7.4 shows the process through which a city is identified by intersecting the grid-based urban centre with small spatial units.

Figure 7.4: Urban centre and city for Toulouse, France



7.2.2.3 Definition of a commuting zone

Once all cities have been defined, commuting zones can be identified using the following steps:

- if 15 % of employed persons living in one city work in another city, these cities are treated as a single city — this step is referred to as a ‘polycentricity check’;
- all small spatial units with at least 15 % of their employed residents working in a particular city are identified as part of the commuting zone for that city (see Figure 7.5, second panel);

([13]) In principle, commuting data at grid level would be another usable option, if available.

- enclaves, in other words, small spatial units entirely surrounded by other small spatial units that belong to a commuting zone or a city are included and exclaves or non-contiguous small spatial units are excluded (see Figure 7.5, third panel).

It can happen that, due to a low intensity of commuting flows, there is no commuting zone for a specific city. In this case, there is a perfect correspondence between the functional urban area and the city. The delineation of functional urban areas is summarised in Figure 7.5.

Figure 7.5: City, commuting zone and functional urban area for Genova, Italy

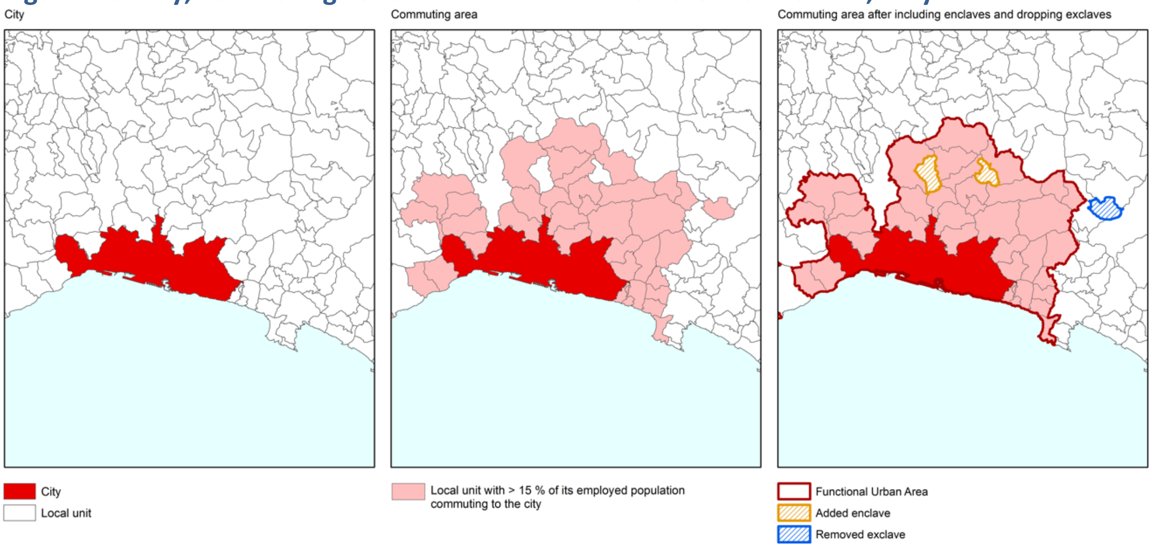
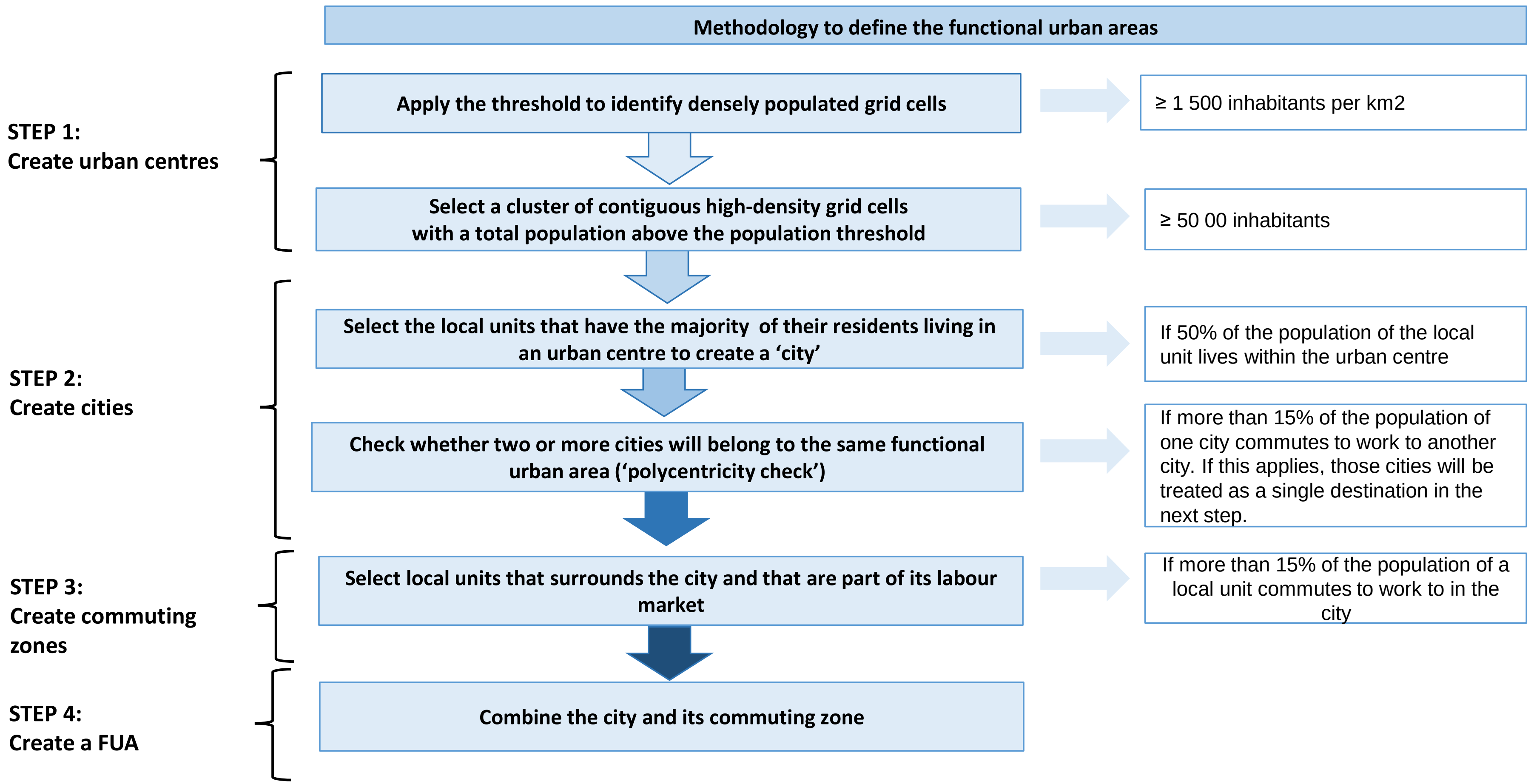


Figure 7.6: Defining a functional urban area



7.2.3 Defining an urban centre

The approach to identify an urban centre as part of the functional urban area classification is identical to that described for level 1 of the degree of urbanisation classification (see Subchapter 6.3.1). To identify an urban centre (high-density cluster):

- Select all 1 km² grid cells with a density of at least 1 500 inhabitants per km² of land area (in other words, for each cell the density should be calculated by excluding bodies of water); if needed, cells that are 50 % built-up may be added (see Subchapter 7.3).
- Cluster all contiguous cells above this density threshold using only four points of contiguity and keep those clusters with at least 50 000 inhabitants (high-density clusters); remove any clusters that have less than 50 000 inhabitants.
- Fill any gaps and smooth borders using the 'majority rule' iteratively until no more cells may be added.

The identification of an urban centre is based on a population grid. Several statistical authorities already produce their own population grids. For example, the 2011 GEOSTAT grid covers all EU Member States ([14]). Australia, Brazil, Colombia and Egypt either have their own grid or are developing one. Other national statistical authorities plan to produce an official population grid by geo-coding their next census. Because these grids are based on points, they are called 'bottom-up' grids. In other words, the grid is created from the bottom-up using data with a higher spatial resolution. Various institutions provide modelled global population grids that are publicly available (see Chapter 5).

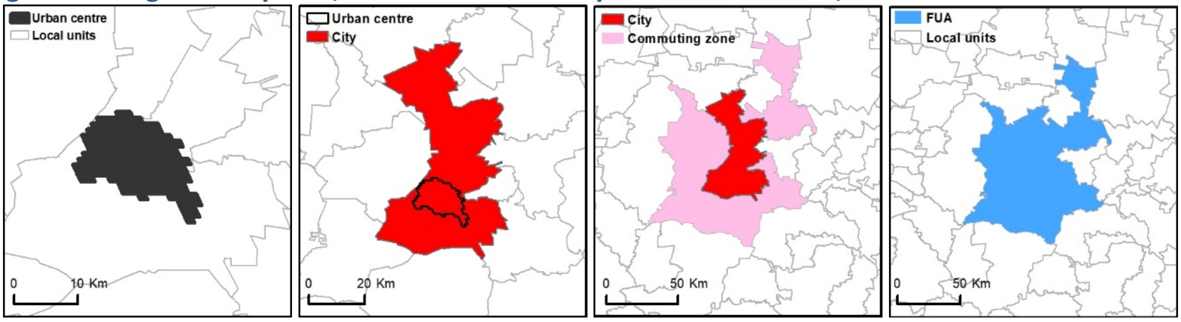
In countries with relatively low-density urban development, a very accurate population grid and a strong separation of land uses, this approach may lead to an excessive fragmentation of urban centres. In such places, grid cells with shopping centres, transport infrastructure or business parks will not reach the residential density threshold to be included in the urban centre and this has the potential to create breaks between adjacent areas. The quality of the population grid also plays a role. In a disaggregation grid, some population would still be attributed to commercial or industrial areas, whereas in a bottom-up grid this would not be the case. Therefore, fragmentation is less likely to occur when using a disaggregation grid. To resolve this issue, grid cells that are 50 % built-up may be added to the urban centre. This resolves the problem in this specific type of city and has little to no impact on higher-density cities, as virtually all the cells that are 50 % built-up have a high enough population density or are added as part of the gap-filling process.

7.2.4 Defining a city

In most cases, defining a city is simple. There is a single urban centre located in a single small spatial unit. This means that all of the urban centre population is located in that small spatial unit and the share of its population in that urban centre is very high (see Figure 7.6).

([14]) GEOSTAT 2011 (http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/population-distribution-demography/geostat).

Figure 7.6: High-density cells, urban centre and city for San Luis Potosí, Mexico

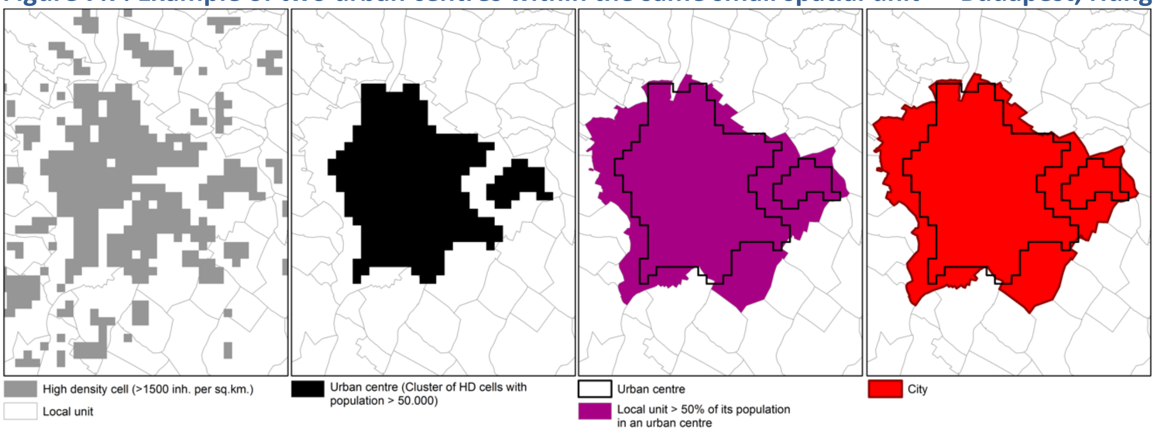


However, in some cases the relationship is more complex. Two cases are discussed below: (i) if a city contains more than one urban centre; and (ii) if an urban centre covers two distinct cities.

7.2.4.1 A city contains more than one urban centre

It may be that a wide river, a steep slope or an industrial area has led to a split in the urban centre. In this case, the small spatial unit simply represents both urban centres. For example, Budapest has two separate urban centres (Buda on the west bank of the Danube and Pest on the east bank). They both fall within the same small spatial unit (see Figure 7.7).

Figure 7.7: Example of two urban centres within the same small spatial unit — Budapest, Hungary



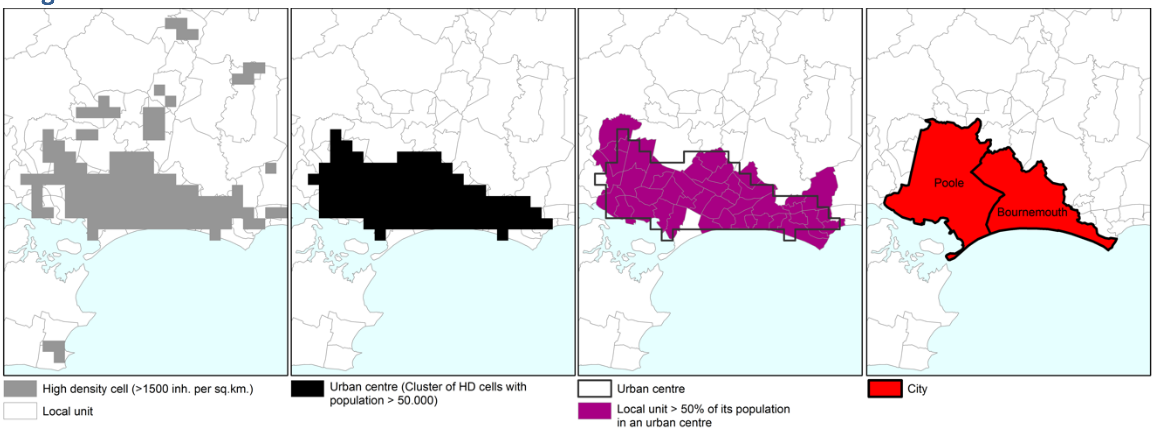
7.2.4.2 An urban centre covers two distinct cities

Some urban centres cover two (or more) distinct cities, in the sense of two distinct urban settlements with their own centre and their own name. This can happen because these cities have grown towards each other but remain functionally distinct. If the population grid is estimated, this situation might occur because the estimated population is often more evenly distributed than the actual population.

In some cases, an urban centre can become too big to be plausible as the centre of a daily urban system, meaning that it is too large to be considered as a space encompassed by the daily movements of people between residence and place of work.

When a single urban centre covers two or more distinct cities, a national statistical authority can choose to create multiple cities. For example, Poole and Bournemouth in the United Kingdom share a single urban centre (see Figure 7.8) but are two separate cities. However, each of these cities should have a population of at least 50 000 inhabitants. If there is at least a one-way commuting flow of more than 15 % between these two cities, they should have a joint commuting zone and therefore be part of the same functional urban area. If, instead, the flow of commuting between the two cities is less than 15 %, then each city should have its own commuting zone and its own functional urban area. In addition, the urban centre can also be split into two parts along the border between the two cities.

Figure 7.8: Example of two cities with a single urban centre — Poole and Bournemouth, the United Kingdom



7.2.4.3 What is a greater city?

In some situations, an urban centre may stretch far beyond the boundaries of the small central spatial unit that gives it its name. This is often the case for (large) capital cities that have outgrown their central spatial unit, such as Athens, Copenhagen, Paris or Valletta. To avoid confusion, the prefix 'greater' is often added to their name. This is already common practice in several countries, for example Greater London, Greater Dublin, Grand Paris and so on.

The functional urban area classification ensures that the most comparable boundaries are selected. It does this by first defining an urban centre independently from administrative boundaries and only in a second step identifying the administrative boundaries that correspond best to this urban centre. In this way, it is possible to ensure that comparisons are not made between, for example central Paris (within the confines of the périphérique) and the full urban sprawl of Berlin or London. Countries with relatively small spatial units, such as France and Switzerland, are more prone to this problem of 'underbounding' (see also Subchapter 7.2.6).

In short, a greater city is a city. The addition of the term 'greater' functions only as a warning to the data users that this definition of the city contains more small spatial units than the central spatial unit which gives this city its name.

7.2.5 Defining a commuting zone

7.2.5.1 Checking for connected cities: the polycentricity check

The delineation and definition of a commuting zone starts with the polycentricity check, in other words, a check to see if two or more cities are linked by strong commuting flows. If city A has 15 % of its employed residents commuting to city B, then these two cities will share a single commuting zone. Note: it is sufficient that the flow of commuters reaches 15 % in a single direction. For example, if city B has a commuting flow of less 15 % to city A, it will still share the same commuting zone.

The polycentricity check is applied only once; it is not an iterative rule. For example: City C has a commuting flow of 20 % to city D. City E has a commuting flow of 10 % to city C and 10 % to city D.

Then cities C and D will have a shared commuting zone, but city E will have its own commuting zone because the commuting flow to each individual city is too small. If city H and city I both have a commuting flow over 15 % to city J, then all three cities will share a single commuting zone.

7.2.5.2 Creating the commuting zone

The next step is to identify all small spatial units with at least 15 % of their employed residents working in a single city (or both cities in the case of cities linked by commuting flows). If a small spatial unit has a commuting flow of more than 15 % to two different cities, it will become part of the commuting zone of the city to which the flow is biggest. If a small spatial unit has a commuting flow of 20 % to city K and 17 % to city L, it will be classified as part of the commuting zone of city K.

Enclaves, in other words, small spatial units surrounded by a single functional urban area, are included and exclaves (or non-contiguous small spatial unit) are excluded. An enclave is defined as a small spatial unit that shares 100 % of its land border with the functional urban area (city or commuting zone); water borders are not considered. An exclave is defined as a small spatial unit that does not share any border with the functional urban area (city or commuting zone); in other words, it is a non-contiguous spatial unit.

The city destination for commuting flows should be the best approximation of the urban centre, in other words, all the units with at least 50 % of their population in the urban centre. If the city boundary is adjusted by adding or dropping a few small spatial units or shifted to a higher administrative level (see next section), this adjusted city should not be used for the commuting analysis; the only exception is where a single urban centre covers multiple cities.

7.2.6 Which small spatial units to use?

The population grid helps to address what is referred to as the modifiable areal unit problem ([15]). However, when these grid concepts are used to classify small spatial units, the problem that different shapes and sizes of spatial units will lead to different results reappears.

Many countries have more than one local administrative level and more than one potential type of statistical area that might be chosen as the small spatial units to delineate cities and functional urban areas. Smaller spatial units will normally lead to a closer match between an urban centre and a city. However, national statistical authorities may not be able to provide annual data for many indicators at such a detailed level. Furthermore, smaller spatial units, such as wards or districts, may not have as strong a political role as larger spatial units (such as municipalities).

This section describes some of the issues encountered when using different sizes of spatial units and proposes, where feasible, options for how to address them.

([15]) The modifiable areal unit problem (or MAUP) highlights that using different boundaries can produce different results. For example, altering the boundaries of electoral districts can change the outcome in first-past-the-post systems. When using larger spatial units, the degree of urbanisation classification tends to categorise fewer people as living in rural areas and cities and more people as living in towns and semi-dense areas. The MAUP was originally identified by Gehlke and Biehl (1934) and further developed by Openshaw (1984).

7.2.6.1 *Large spatial units may lead to the over-, under- or non-representation of an urban centre by a city*

The population of an urban centre and that of a city can differ by a considerable amount if a country has relatively large spatial units. Below are three types of issues that may potentially arise when using relatively large spatial units to define a city.

Overrepresentation

A city can have almost double the population of an urban centre. For example, an urban centre of 50 001 inhabitants in a spatial unit of 100 000 would mean that this spatial unit will be defined as a city. This is a tricky problem to solve as the only alternative to the overrepresentation is non-representation, in other words, by not defining this spatial unit as a city.

Underrepresentation

A city can also have a much smaller population than the urban centre it represents. Take for example, an urban centre of 200 000 inhabitants that is split across four spatial units. One spatial unit (A) has a population of 50 000 and all of its inhabitants live in the urban centre. The other three spatial units (B, C, D) each have a population of 150 000 inhabitants of which respectively 60 000, 50 000 and 40 000 live in that urban centre. As a result, the city will consist of just the one spatial unit (A) with a population of 50 000 inhabitants and not the other three spatial units (B, C or D).

This underrepresentation can be reduced by adding the spatial unit with the highest share of its population in that urban centre to the city (spatial unit B with 60 000 of its 150 000 inhabitants in the urban centre). This would bring the population of the city up to 200 000 inhabitants, of which 110 000 would be living in the urban centre.

Non-representation

The most extreme form of under-representation is non-representation. For example, a spatial unit with a population of 200 000 inhabitants with a single urban centre of 75 000 inhabitants will not be classified as city. As a result, this urban centre will not be represented by a city, in other words, non-representation, something which is more likely to happen for small urban centres.

In a country where all the spatial units are relatively large, it is likely that all of the small urban centres will not be represented by cities. This would create a quite skewed representation of urban centres as all small urban centres would be missing. One option to address this problem is that for half of the small urban centres without a city, their spatial unit is classified as a city even though their share of population in an urban centre is less than 50 %.

7.2.6.2 *Small spatial units may lead to a loss of the link to local government or to less statistical data*

In a country with relatively large spatial units, most cities will consist of a single spatial unit. As a result, each city will have a single local government. This makes it easier to communicate indicators to local politicians/representative groups and helps to ensure good inputs for policymaking.

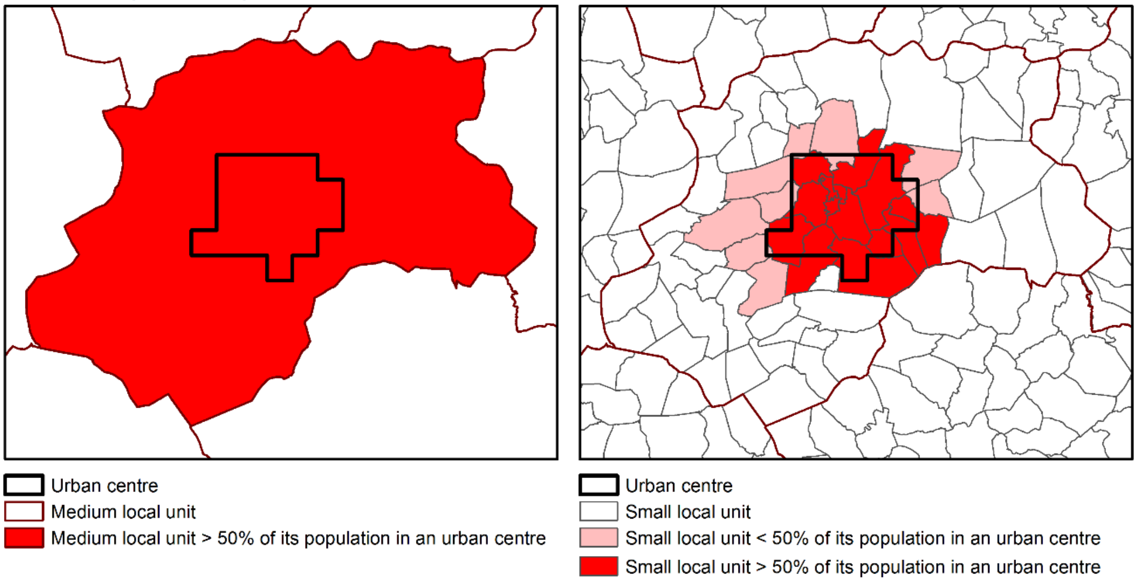
In countries with relatively small spatial units, most cities will consist of multiple spatial units. These small spatial units will ensure that there is a close match between the population in the urban centre and the population in the city. The trade-off is that the city will not match a single local government, which makes it more complicated to communicate data to local politicians/representative groups.

This effect can be shown in Portugal, which has both municipalities (*municipio* or *concelho*) and parishes (*freguesia*). If the urban centre of Braga in Figure 7.9 is used to define the municipal level (left panel), there is a simple one-to-one relationship; the local government of Braga is organised at the municipal level. If the urban centre is used to define a city at the parish level (right panel), the relationship becomes a more complicated one-to-many relationship; the simple link with the local government of Braga is also lost.

When statistical areas are used as building blocks to define a city and/or a functional urban area, the latter can be adapted *ex post* to the closest local administrative units. For example, cities and their commuting zones in the United States have been delineated using census tracts as building block units, but subsequently adapted to the closest county boundaries, by including the counties where the share of population living in cities and functional urban areas was higher than 50 %.

The imperfect match between the cities and functional urban areas and their respective urban centres can be informative for policymakers. Administrative boundaries of cities often remain unchanged for decades, while cities can expand or shrink. Many OECD countries, following the urban expansion that occurred in the last few decades, have created new levels of government for large cities encompassing multiple spatial units. For example, France has created *métropoles* to help govern its 21 biggest cities.

Figure 7.9: Example of the influence of the choice of type of spatial unit — municipal and parish levels, Braga, Portugal



7.2.6.3 Adjusting the city to ensure a better representation of the urban centre or a better link to local government

If a country wishes to adjust the delineation of its cities to get a better link between a city and its urban centre or a city and its local government, it can add or drop a spatial unit as long as the two following rules are respected:

- **Rule 1:** a spatial unit with less than 50 % of its population in an urban centre can be added to a city if at least 50 % of the population of this expanded city lives in an urban centre.

- **Rule 2:** a spatial unit with at least 50 % of its population in an urban centre can be excluded from a city as long as at least 75 % of the population of that urban centre lives in a city after excluding the spatial unit.

These two rules were designed to provide statistical limits to these optional changes that can be made.

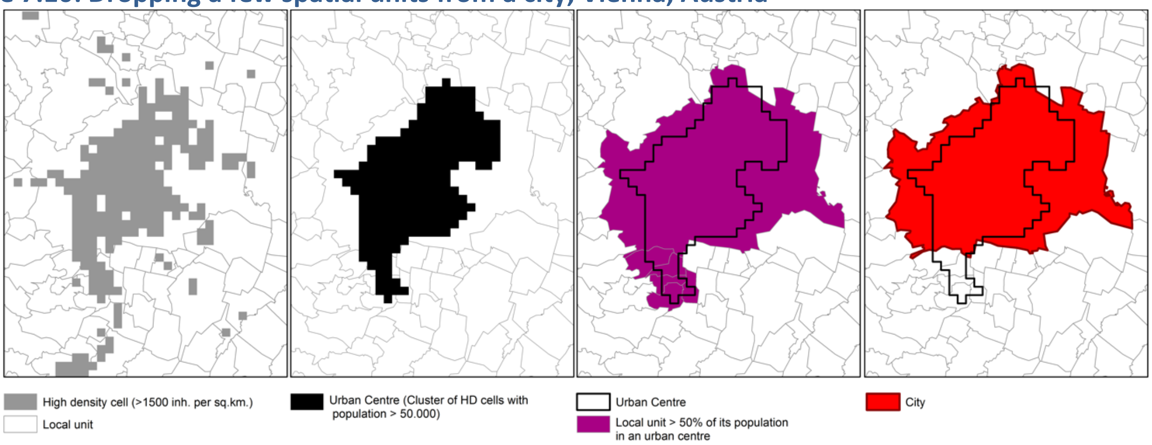
City adds a few spatial units

Returning to the example of Braga in Portugal: if the urban centre is used to define the city at the parish level it would only contain some of the parishes in the municipality of Braga. Defining Braga at the municipal level amounts to adding these surrounding parishes to the city. As still more than 50 % of the population of the municipality of Braga lives in the urban centre, this complies with rule 1; it also ensures a direct link to Braga's local government.

City drops a few spatial units

An example of the application of rule 2 is presented for Vienna in Austria. A number of small spatial units just south of the city of Vienna have 50 % or more of their population in the urban centre of Vienna. As more than 75 % of the population of the urban centre lives in the city of Vienna, these smaller spatial units can be excluded without significantly compromising the comparability of the results (see Figure 7.10).

Figure 7.10: Dropping a few spatial units from a city, Vienna, Austria



Cities without an urban centre

The definition that has been developed provides an estimate of the population of an urban centre. Two elements may reduce the accuracy of this estimate: (i) geographic features and (ii) the source of the population grid data.

The definition does not take into account the specific geography of a city. Some geographic features, such as steep slopes, cliffs or bodies of water may lead to an underestimation of the population of an urban centre. This affects in particular cities with a small centre.

The definition works best when a bottom-up grid (based on point data) or a high-resolution, hybrid grid (based on a mixture of points and smaller statistical areas) is available, which ensures that the density of the population (per km²) is very accurate. In countries where such a grid is not yet available, the population of a small spatial unit has to be disaggregated based on a given criterion, such as land use data in the case of the GHS-POP grid produced by the European Commission's Joint

Research Centre (JRC). This is called a top-down approach, which is generally less accurate. It tends to underestimate the population cells with a moderate to high-density and overestimate population in those grid cells with a low population density. Due to this imprecision, there remains a margin of error, especially for smaller centres.

Therefore, a national statistical authority may opt to classify a small spatial unit as a city when it lacks an urban centre of more than 50 000 inhabitants, but fulfils the following two conditions:

- the presence of an urban centre of 50 000 inhabitants, which the definition does not capture due to geographic features or population grid estimation techniques;
- the small spatial unit has a population of more than 50 000 inhabitants.

For example, a small spatial unit which has two clusters of high-density cells separated by a river or a bay which together have a collective population of 50 000 inhabitants can be argued to have an undetected urban centre. A small spatial unit with a high-density cluster of 49 000 inhabitants based on a top-down population grid can be argued to have an undetected urban centre (see Subchapter 7.3 for more details).

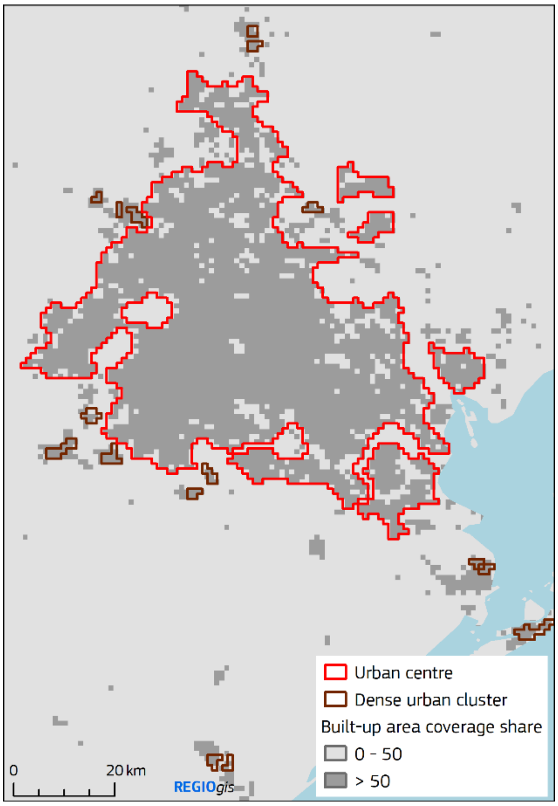
7.3 Specific geographic issues for the degree of urbanisation and functional urban area classifications

7.3.1 Railways, highways, malls, office parks and factories

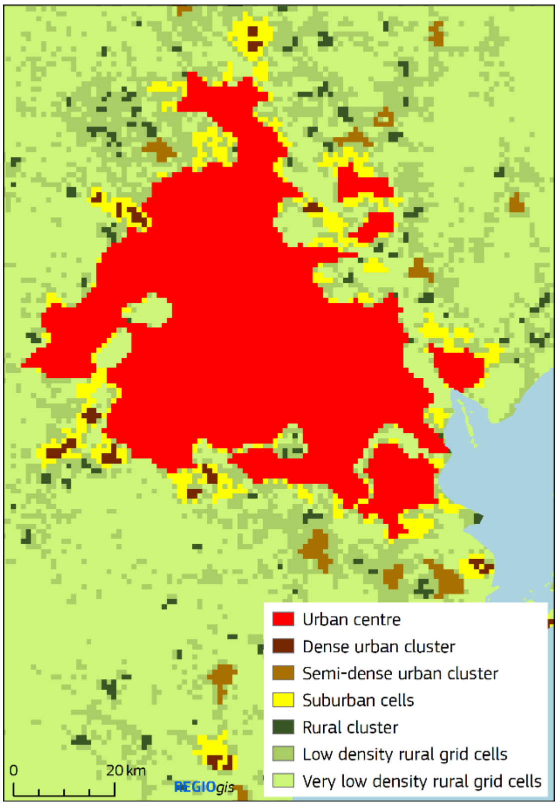
In a few countries with a strong separation of land use functions and relatively low-density urban developments, the methodology may generate multiple urban centres for a single city. For example, Houston in the United States has nine urban centres if the methodology is applied without considering cells that are at least 50 % built-up (see Map 7). This is often because highways, railways, shopping centres, office parks and factories typically have little or no residential population and can occupy enough of a single grid cell that it does not reach the population density threshold of at least 1 500 inhabitants per km². Although many people may use these areas during the daytime, the methodology is designed to be applied to the residential population, broadly speaking the night-time population. As a consequence, areas which are intensively used by city residents during the day but which have few, if any, residents might not be considered to be part of a city.

Creating urban centres using both criteria — cells with a density of at least 1 500 inhabitants per km² and cells that are at least 50 % built-up — resolves this issue. For example, in Houston the nine separate urban centres are all connected by cells that are at least 50 % built-up (see Map 7.4). When the urban centre is defined using both of these criteria, the nine separate urban centres become one (see

Map 7.5: Built-up cells, urban centres and dense urban clusters considering built-up cells, Houston, United States

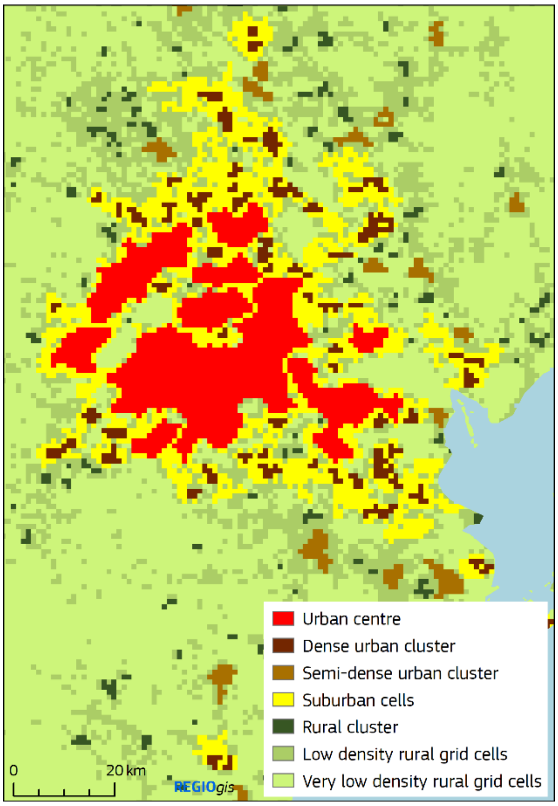


Map 7.6: Grid cell classification considering built-up cells, Houston, United States

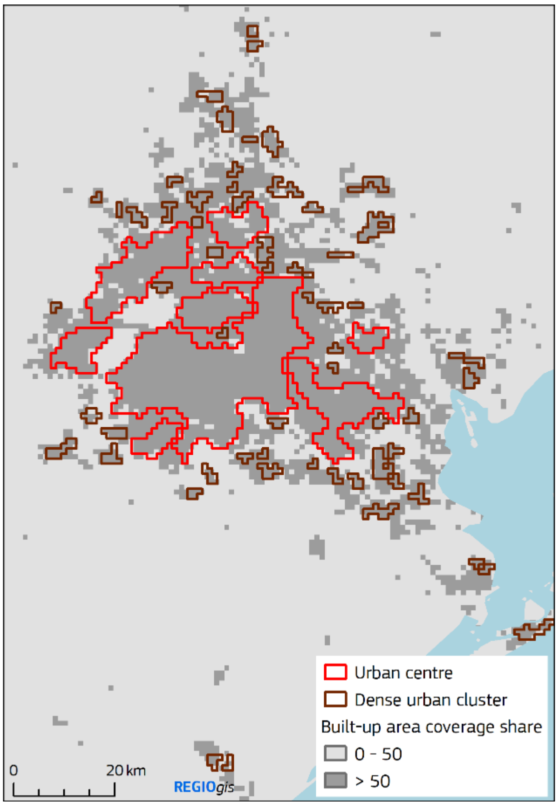


). In addition, a few separate dense urban clusters are also combined such that they reach the 50 000 population threshold and become an urban centre (see Map 7.6).

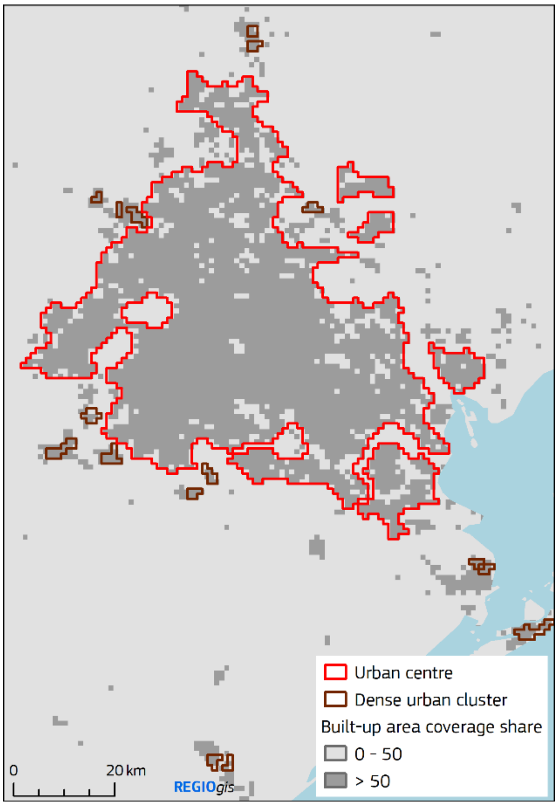
Map 7.3: Gird cell classification without considering built-up cells, Houston, United States



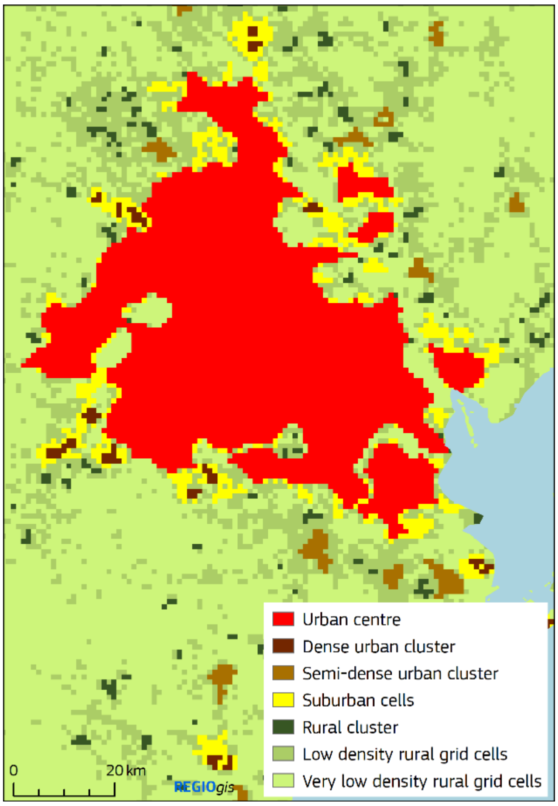
Map 7.4: Built-up cells, urban centres and dense urban clusters without considering built-up cells, Houston, United States



Map 7.5: Built-up cells, urban centres and dense urban clusters considering built-up cells, Houston, United States



Map 7.6: Grid cell classification considering built-up cells, Houston, United States



7.3.2 Water bodies, steep slopes and parks in a city

The presence of water bodies, steep slopes and parks may have an impact on the capacity of the methodology to identify a city. These elements can lead to gaps or separations which result in a single urban centre being fragmented into multiple centres or — when these fail to reach the minimum population threshold of 50 000 inhabitants — multiple dense urban clusters.

To overcome these problems, the methodology can be adapted to address gaps or separations that are due to the presence of waterways, parks and/or areas with steep slopes. This optional process should be applied to clusters of high-density grid cells before evaluating the minimum population of urban centres. Hence, the initial input of the workflow are clusters of contiguous grid cells characterised by a population density threshold of at least 1 500 inhabitants per km², without any criterion for the total population of the cluster.

For the purpose of this process description, they are called sHDCs (small high density cells, as no minimum population threshold was applied). Each of these sHDCs is stored as a polygon and receives its unique number, which is required in further steps of the workflow. Additional spatial data are needed to represent the areas that will be taken into account in a special exercise to fill gaps in or separations between sHDCs:

- Waterways should ideally be portrayed as polygon features. If these are not available, waterway line features should be buffered to model the actual width of the waterway. Furthermore, waterway polygons can (optionally) be buffered by a limited width (for instance, a maximum of 50 metres) to portray adjacent zones which are assumed not to be suitable for the construction of buildings.

- Zones with steep slopes should be retrieved from a layer with appropriate spatial detail. Usually this will be a selection of raster cells, with resolution equal to or higher than 1 km². The selection of steep areas should be converted to polygons.
- Parks will also be represented by polygons; these should be retrieved from dedicated thematic layers.

The polygons representing waterways, steep slopes and parks are merged into a common polygon layer. Next, only the areas in the close neighbourhood of sHDCs should be taken into account for this special potential gap or separation filling.

To assess this spatial relationship, each of the sHDCs is expanded by applying a buffer. The size of this buffer should be between 500 and 2 000 metres depending on the local circumstances (in other words, depending on the size of the water bodies, areas with steep slopes and parks). Then the common polygon layer for waterways, steep slopes and parks is intersected with the expanded sHDCs. Hence, the aim is to keep only those parts of waterways, steep slopes and parks that are located close to an sHDC. The selected waterways, steep slopes and parks are converted to 1 km² grid cells by selecting those cells that are at least 50 % covered by the common polygon layer for waterways, steep slopes and parks.

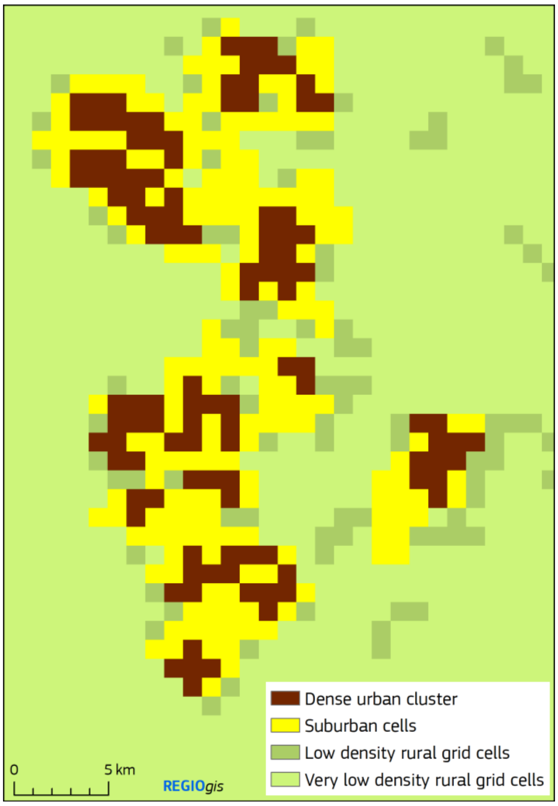
In the next step, the grid cells of selected waterways, steep slopes and parks are merged with the sHDC grid cells. If this results in changes to the boundaries of the sHDCs, the result can be twofold:

- two or more sHDCs are linked by the grid cells added for waterways, steep slopes and parks;
- the coverage of a single sHDC has been expanded by adding adjacent grid cells for waterways, steep slopes and parks.

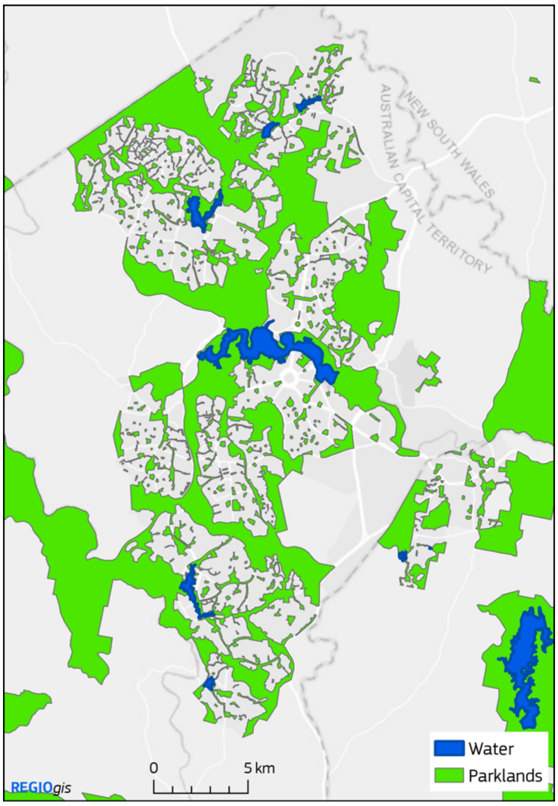
The goal of this adapted methodology is to capture only the first case when overlaying the adjusted sHDCs with the original ones. If an adjusted sHDC contains more than one original sHDC then the adjustment should be kept; a new sHDC has been created, covering two or more original sHDCs. If the adjusted sHDC only contains a single original sHDC then the adjustment should be discarded, reverting to the original classification of grid cells (as there is no need to expand the sHDC by adding nearby waterways, steep slopes or parks).

Only those new sHDC which reach a minimum population threshold of 50 000 inhabitants are kept. Thereafter, the normal smoothing and gap-filling process is applied to turn them into an urban centre.

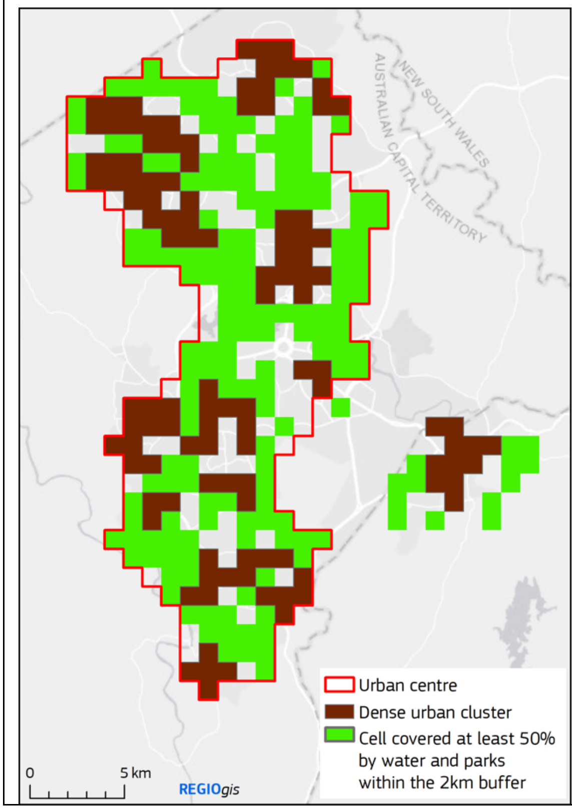
Map 7.7: Grid cell classification, Canberra, Australia



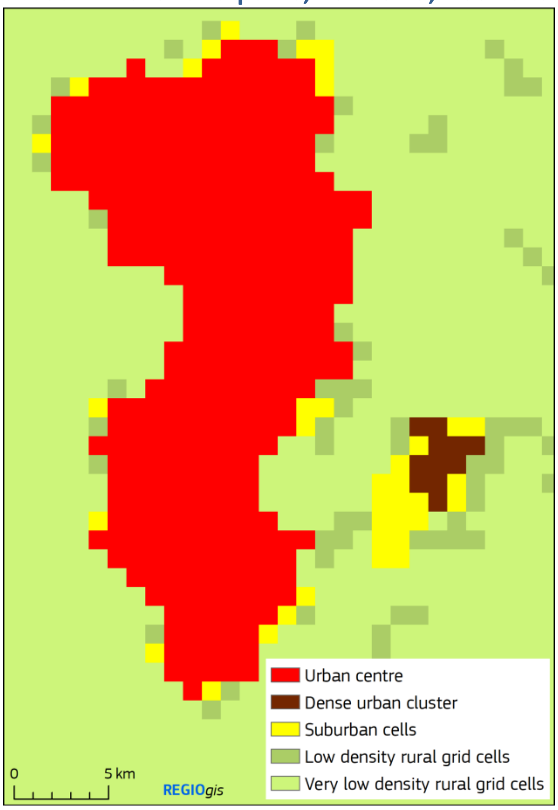
Map 7.8:Water and parks, Canberra, Australia



Map 7.9: Dense urban clusters and cells covered by water and/or parks, Canberra, Australia



Map 7.10: Grid cell classification taking into account water and parks, Canberra, Australia



Adjusting the results for cities

As the degree of urbanisation classification and the functional urban area classification share a common definition of cities, any changes that are made to the delineation of cities should be adopted for both of these classifications (using the same rules). More information on adjustments that might be made when delineating cities is provided in Subchapters 7.2.4 and 7.2.6.

7.4 Other possible extensions to the methodology: remoteness and land cover

Global Strategy's, *Guidelines on defining rural areas and compiling indicators for development policy*, published in 2018 ([16]) identified three dimensions of 'rurality': (i) sparse settlement; (ii) remoteness from urban areas; and (iii) land cover or use. While consideration of all three dimensions is potentially useful for policy design and analysis, it is the dimension of sparse settlement (population size and density) that is captured by the degree of urbanisation classification.

Sparse settlement reflects the idea that at one end of a continuum (as measured by population size or density) there are rural areas that are more sparsely populated and settled, while at the other there are urban areas (that are the most populous and densely settled parts of a country). Remoteness affects the opportunities people have to gain access to markets (for goods, services and labour) and to public services. It is most often represented by the difficulty of physical travel to places where markets and services are more (widely) available. Land cover is the physical cover on the land including vegetation (either planted or naturally occurring) and any buildings, other structures or features constructed by humans. Land cover reflects and determines land use, which is related to the human activities that take place there.

Remoteness

In general, remoteness (or distance from urban areas) is considered an important dimension of rurality. In combination with low population density, remoteness characterises rural areas that face particular challenges concerning their development. Remote areas are generally those where population densities are low, markets of all kinds are thin, and the unit costs of delivering most social services and many types of infrastructure are high. Additionally, in these areas that are distant from urban centres, farm-gate (or factory-gate) prices of outputs are often low and prices of inputs are often high, while it is usually difficult to recruit skilled personnel to work in public services or private enterprises. In contrast to remote areas, urban areas are characterised by agglomeration economies, in other words benefits that come when firms and people locate near to one another in cities and industrial clusters, effectively lowering the costs of transporting goods and sharing knowledge. More specifically, remoteness signifies the extent of opportunity people have to gain access to markets.

The dimension of remoteness can be included in the methodology for analytical purposes, though the identification of an empirical measure of remoteness to be used depends on the context in each country. Instinctively, a remote area is far from a city in terms of distance or the time it takes to travel physically from one place to another. The mode and speed of transportation, however, would be expected to vary depending on terrain and on the presence or absence of infrastructure. Travel by road or train might be the most common means of transport in one place, but travel by water or foot may be more common in another. While the variables chosen might be different across countries, or even within countries, the underlying supposition is that physical access to a city is key, however it is achieved.

([16]) Global Strategy (http://gsars.org/wp-content/uploads/2018/12/GS-GUIDELINES-RURAL-AREAS-EN-FINAL-2018.pdf).

Physical distance is not a perfect proxy, however, as distance may not restrict access to one service (for example, access to online education) but it may be a significant barrier to another (for example, access to a surgery at a regional hospital). Some disadvantages of remoteness may be overcome by telecommunication or internet services, as for example with the provision of health care services through satellite video. However, remoteness in terms of travel time is likely to be the most expedient approach when selecting a variable.

Concept of remoteness: an example for small regions

The OECD has used the concept of remoteness in a classification of small regions based on their access to functional urban areas (Fadic et al. (2019)). Based on this, a small subnational region (or territorial level 3 region, TL3) is classified as either a 'metropolitan region' — if at least half of its population lives in a functional urban area of at least 250 000 inhabitants — or as a 'non-metropolitan region'. The concept of remoteness is thereafter used to further characterise non-metropolitan regions, more specifically:

- if at least half of the population in a non-metropolitan region cannot reach a functional urban area within a one hour drive, then that region is sub-classified as 'remote';
- if at least half of the population in a non-metropolitan region can reach a functional urban area within a one hour drive, then that region is sub-classified depending on the size of the functional urban area, as a non-metropolitan region:
    - 'with access to a metropolitan region' (for functional urban areas of at least 250 000 inhabitants); or
    - 'with access to a small functional urban area' (for functional urban areas with less than 250 000 inhabitants).

In short, though the concept of remoteness seems straightforward, it is not always clear how to represent it with data. For example, is remoteness always a function of physical distance? Or might this barrier be reduced/removed, for example by access to telecommunications that allow commercial transactions to take place virtually or social services like health care to be delivered remotely? Furthermore, data on road networks and their use are hard to come by on a global scale, although there have been recent attempts at improvement. The use of mass/public transportation might also be complicated to measure. In any case, remoteness might be considered less a permanent aspect of rurality and more a condition to be addressed by taking steps to improve access to markets and services in rural areas themselves. If that is the case, then a definition of remoteness should not include any elements that themselves are policy targets.

Land cover

Land cover consists of vegetation (occurring naturally or cultivated), buildings, roads and other man-made features and describes cover by forest, grassland, impervious surfaces, cropland and other land and water types (such as wetlands and open water). This is in contrast to land use that defines what people do on the landscape (for example, work in factories, live in houses, use parks and gardens for recreation, graze cattle on agricultural land) with the intention of getting benefit from its use. A given type of land cover, say tree cover, may support multiple land uses: for example, recreation, logging and/or conservation. For rural development policies and analytical purposes, countries may use land cover as an additional dimension to further enrich their understanding of rural areas and augment rural development policies (Global Strategy (2018)).

References

Ahas, R. and S. Silm (2013), 'Regionaalse Pendelrände Kordusuuring' (Re-study of regional commuting), *Regionaalministri Valitsemisala*.

Ahas, R., S. Silm, O. Järv, E. Saluveer, and M. Tiru (2010), 'Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones', *Journal of Urban Technology*, Volume 17, Issue 1, pp. 3-27.

Dijkstra, L., H. Poelman and P. Veneri (2019), 'The EU-OECD definition of a functional urban area', *OECD Regional Development Working Papers,* No. 2019/11, OECD Publishing, Paris.

Fadic, M., J. E. Garcilazo, A. I. Moreno-Monroy and P. Veneri (2019), 'Classifying small (TL3) regions based on metropolitan population, low density and remoteness', *OECD Regional Development Working Papers*, No. 2019/06, OECD Publishing, Paris.

Gehlke, C. E. and K. Biehl (1934), 'Certain Effects of Grouping upon the Size of the Correlation Coefficient in Census Tract Material', *Journal of the American Statistical Association Supplement,* Volume 29, Issue 185A, pp. 169-170.

Global Strategy (2018), *Guidelines on defining rural areas and compiling indicators for development policy*.

Openshaw, S. (1984), *The Modifiable Areal Unit Problem*, CATMOG 38, Geo Books, Norwich.

Van der Valk, J., M. Souren, M. Tennekes, S. Shah, M. Offermans, E. De Jonge, J. Van der Laan, Y. Gootzen, S. Scholtus, A. Mitriaieva, B. Sakarovitch, S. Hadam, M. Zwick, M. Rengers, A. Kowarik, M. Weinauer, J. Gussenbauer, M. Debusschere and A. Termote (2019), *City data from LFS and Big Data*, Publications Office of the European Union, Luxembourg.

8. Selected indicators for sustainable development goals by degree of urbanisation and functional urban area

The methodology described in this manual has been developed to facilitate the international comparison of cities and urban and rural areas. The UN's sustainable development goals (SDGs) include numerous indicators that should be compiled for individual cities or for urban and rural areas. This chapter shows that many of these indicators can already be calculated by degree of urbanisation using a wide variety of sources. These examples not only show the feasibility of this approach, but also underscore its interest. In particular, they show the benefit of compiling data separately for cities, towns and semi-dense areas, and rural areas. In most countries, these indicators follow a clear urban gradient with an increasing or decreasing performance as one moves from one end of the continuum, through towns and semi-dense areas, to the other end of the continuum.

The degree of urbanisation classification can be used with a wide variety of data sources. It can be integrated into household surveys: for example, the European Union labour force survey (EU-LFS) codes its respondents according to level 1 of the degree of urbanisation classification using the municipality in which the respondent lives. Face-to-face interviews are increasingly geo-coded, which makes the application of the degree of urbanisation even easier. For example, recent Demographic and Health Surveys (USAID/WHO) and the face-to-face World Poll (Gallup) are all geo-coded.

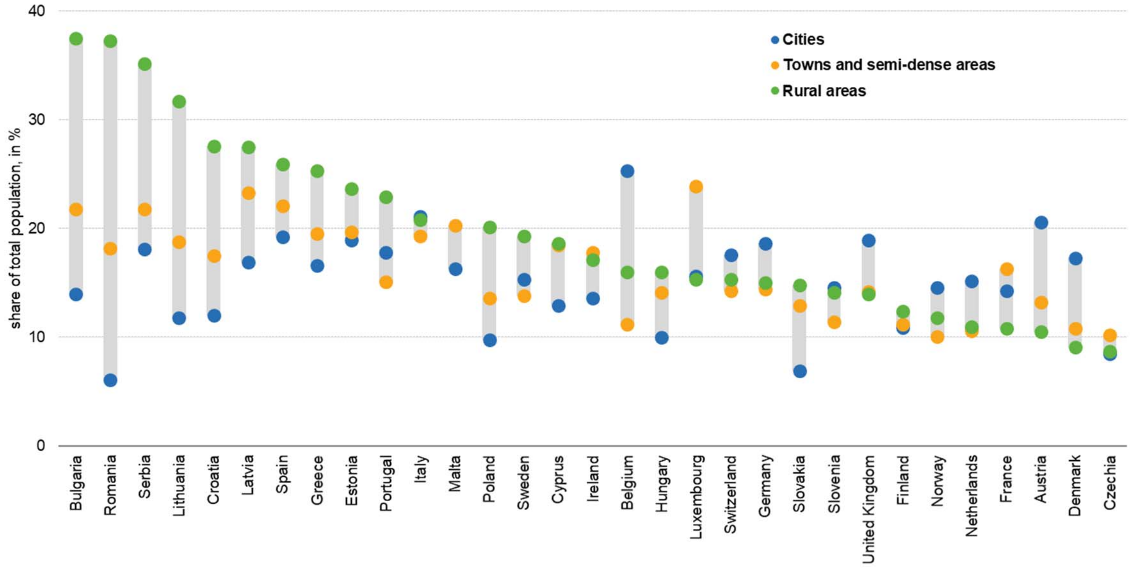
To ensure robust results, these surveys should have a large enough sample in each of the degree of urbanisation classes. As a result, it is easier to produce data by level 1 of the degree of urbanisation classification using surveys than by level 2 or by individual functional urban area. Therefore, producing SDG indicators by degree of urbanisation level 1 is considered the most suitable approach for international comparisons.

The degree of urbanisation classification can also be used with geospatial data, such as remote sensing and point locations. For example, air pollution, changes in the built-up area and the distance to the nearest health facility can all be calculated by degree of urbanisation. The examples below are organised by SDG and include one or more examples for most, but not all, goals. One of the many benefits of geospatial data is that they typically cover the entire territory. As a result, indicators can reliably be provided not only for level 1 of the degree of urbanisation classification, but also for level 2 and even for individual cities and functional urban areas.

SDG 1 — End poverty in all its forms everywhere

Error! Reference source not found. shows the share of the population at risk of poverty for a number of European countries. A household is classified as being at risk of poverty if its income is below 60 % of the national equivalised median income after taxes and transfers. This is an example for SDG indicator 1.2.1: it reveals significant disparities in the situation along the urban-rural continuum. In around 40 % of European countries, the poverty rate was (considerably) higher in rural areas than in cities. This was most notably the case in countries with relatively low ratios of GDP per inhabitant, for example Bulgaria and Romania. In several western and northern European countries with higher levels of GDP per inhabitant, the risk of poverty was higher in cities than it was in towns and semi-dense areas, or rural areas. This was the case in Austria, Belgium, Denmark, the United Kingdom, the Netherlands, Germany, Norway and Switzerland.

Figure 8.1: Share of the population at risk of poverty, by degree of urbanisation, selected European countries, 2017
(%)



Source: Eurostat (online data code: ilc_li43)

SDG 2 — End hunger, achieve food security and improved nutrition and promote sustainable agriculture

Statistics on moderate or severe food insecurity are based on the food insecurity experience scale (FIES), as developed by the Food and Agriculture Organization of the United Nations (FAO). An FIES survey module forms part of the World Poll (Gallup), from which national estimates of the prevalence of moderate and severe food insecurity may be produced. For each country, this indicator was computed on combined sub-samples for each year in which geo-referenced data were available. Therefore, the statistics presented are not intended to be representative of the population by degree of urbanisation.

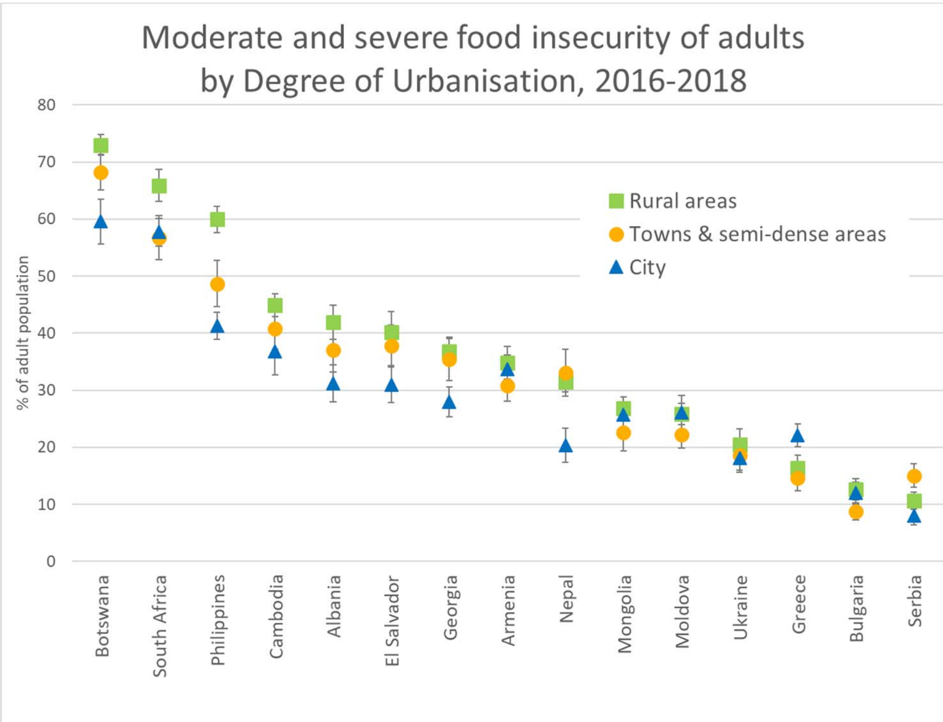
Food insecurity is principally, but not exclusively, a rural problem: rural areas are often found to be significantly more food insecure than cities. Across the seven most food insecure countries shown in Figure 8.2, the prevalence of food insecurity at a moderate or severe level for the adult population living in rural areas was, on average, 11 percentage points higher than for the corresponding share recorded for people living in cities. For example, 73 % of the adult population living in rural areas of Botswana experienced this type of food insecurity during the period 2016-2018, compared with 60 % of adults who were living in cities.

Rural areas were not systematically more food insecure than urban areas. For example, in Armenia, Mongolia, Bulgaria and Moldova there was little or no difference in the prevalence of food insecurity between adults living in cities and those living in rural areas. By contrast, food insecurity was significantly higher across the adult population living in the cities of Greece (22 %) than it was for the rural population (16 %).

Among countries with a high overall prevalence of food insecurity, the share of adults living in towns and semi-dense areas facing food insecurity was generally situated between the extremes observed for people living in cities and those living in rural areas. Food insecurity for adults living in towns and semi-dense areas was lower than the share recorded for people living in rural areas for seven of the countries shown in Figure 8.2, while there were nine where the prevalence of food insecurity among adults living in towns and semi-dense areas was higher than the share recorded for people living in cities.

Across the three classes of the degree of urbanisation, the prevalence of food insecurity was lowest for adults living in towns and semi-dense areas of six of the countries shown. By contrast, adults living in towns and semi-dense areas of Serbia were considerably more likely to face food insecurity (than those living in cities or in rural areas); this pattern was repeated (although it was far less pronounced) in Nepal.

Figure 8.2: Share of the adult population aged 15 years or over facing moderate or severe food insecurity, by degree of urbanisation, 2016-2018
(%)



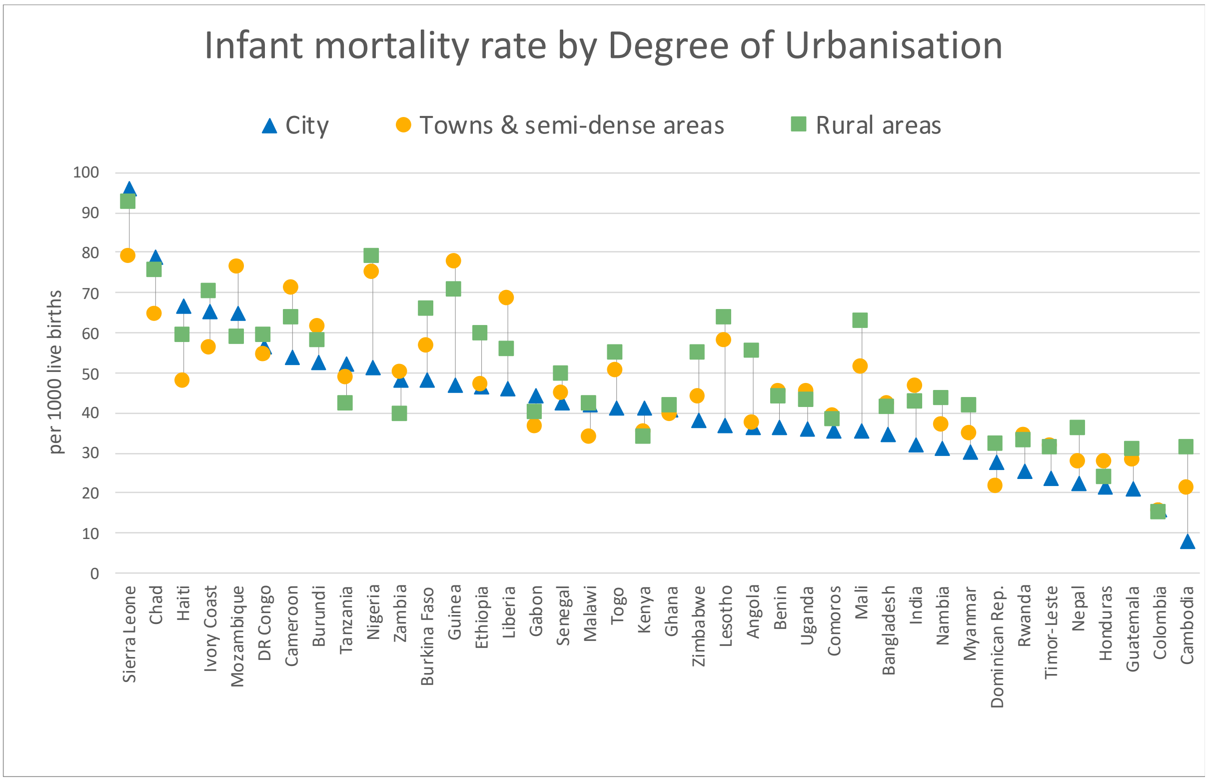
Note: each data point is shown with error bars that indicate the 95 % confidence interval; in those cases where error bars by degree of urbanisation overlap, the differences between point estimates are not statistically significant.
Source: FAO

SDG 3 — Ensure healthy lives and promote well-being for all at all ages

In most countries covered by the Demographic and Health Survey (USAID), infant mortality is notably higher in rural areas than in cities (see **Error! Reference source not found.**). In six countries (Mali, Nigeria, Lesotho, Guinea, Cambodia and Angola) the infant mortality rate was at least 20 deaths per 1 000 live births higher in rural areas than it was in cities. In a few countries, cities had a higher infant mortality rate, but the difference tended to be smaller. In five countries (Mozambique, Haiti, Kenya, Zambia and Tanzania), the infant mortality was between 5 and 10 deaths per 1 000 live births higher in cities than in rural areas.

Note: this is not an SDG indicator, but it is closely linked to the under-5 mortality rate and the neo-natal mortality rate (respectively SDG 3.2.1 and SDG 3.2.2).

Figure 8.3: Infant mortality rate, by degree of urbanisation, selected countries, 2012-2016
(per 1 000 live births)



Note: the infant mortality rate is defined as the probability of a child dying before their first birthday and is expressed per 1 000 live births; the sample is limited to births that took place between one and five years prior to the interview.
Source: Demographic and Health Survey as calculated by Henderson et al. (2020)

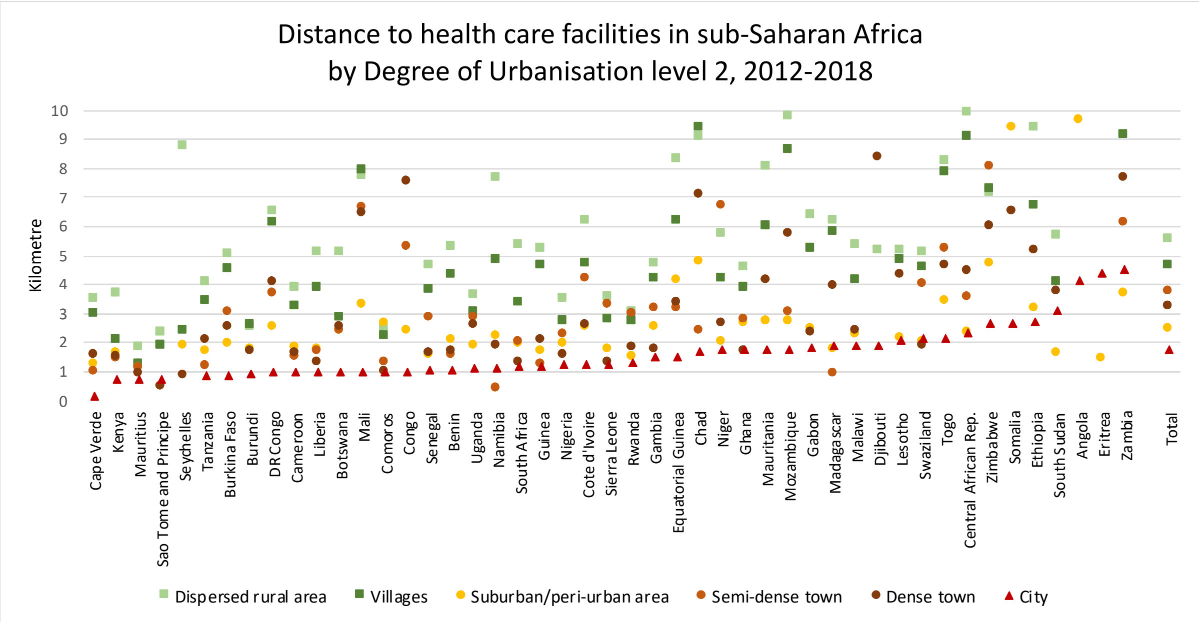
Infant mortality may be influenced by the distance to the nearest health facility, which tends to be larger in rural areas than in cities; **Error! Reference source not found.** shows this distance for a selection of sub-Saharan countries.

As these data are very comprehensive, data can be calculated for level 2 of the degree of urbanisation classification. This reveals a very clear urban-rural gradient with distances increasing from cities to suburbs, to towns, to villages and so on. In cities, the nearest health facility was, on average, only 1.7 km away, less than a 30-minute walk. People living in suburbs were generally closer to a health facility (on average 2.5 km) than people living in dense and semi-dense towns (3.2 km and 3.8 km respectively). Within rural areas, those living in villages tended to live closest to the nearest health facility (4.7 km) followed by people living in dispersed rural areas (5.6 km), while people living in mostly uninhabited areas had the furthest distance to travel (12 km), equivalent to a three-hour walk.

Note: this is not an SDG indicator but it is closely linked to health worker density and distribution (SDG 3.c.1) and the proportion of health facilities that have a core set of relevant essential medicines available and affordable on a sustainable basis (SDG 3.b.3).

Figure 8.4: Average distance to the nearest health care facility, by degree of urbanisation, sub-Saharan countries, 2012-2016

(km)



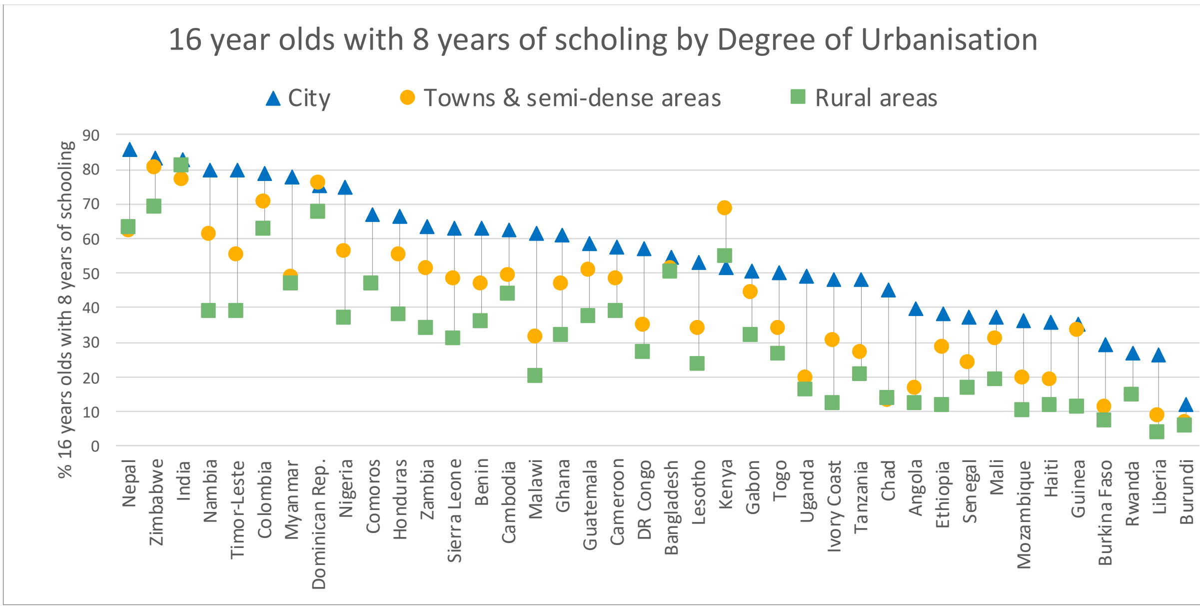
Source: JRC calculation using GHS-POP and Maina et al. (2019)

SDG 4 — Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all

In virtually all of the countries shown in **Error! Reference source not found.**, 16-year-olds living in cities are far more likely to have completed eight years of schooling compared with those living in rural areas. Across the selected countries that are shown, 55 % of 16-year-olds living in cities had completed eight years of schooling compared with only 31 % in rural areas. The share of 16-year-olds living in towns and semi-dense areas that had completed eight years of schooling was in between, at 41 %. The only exceptions (among those countries shown) to the pattern described above were: India and Bangladesh where the differences by degree of urbanisation were very small; Kenya where 16-year-olds living in towns and semi-dense areas were most likely to have completed eight years of schooling, followed by those living in rural areas with a slightly lower share recorded for those living in cities.

Note: this is not an SDG indicator, but it is closely linked to the proportion of children and young people (a) in grades 2/3; (b) at the end of primary education; and (c) at the end of lower secondary education achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex (SDG 4.1.1).

Figure 8.5: Share of 16-year-olds having completed eight years of schooling, by degree of urbanisation, selected countries, 2012-2016
(%)

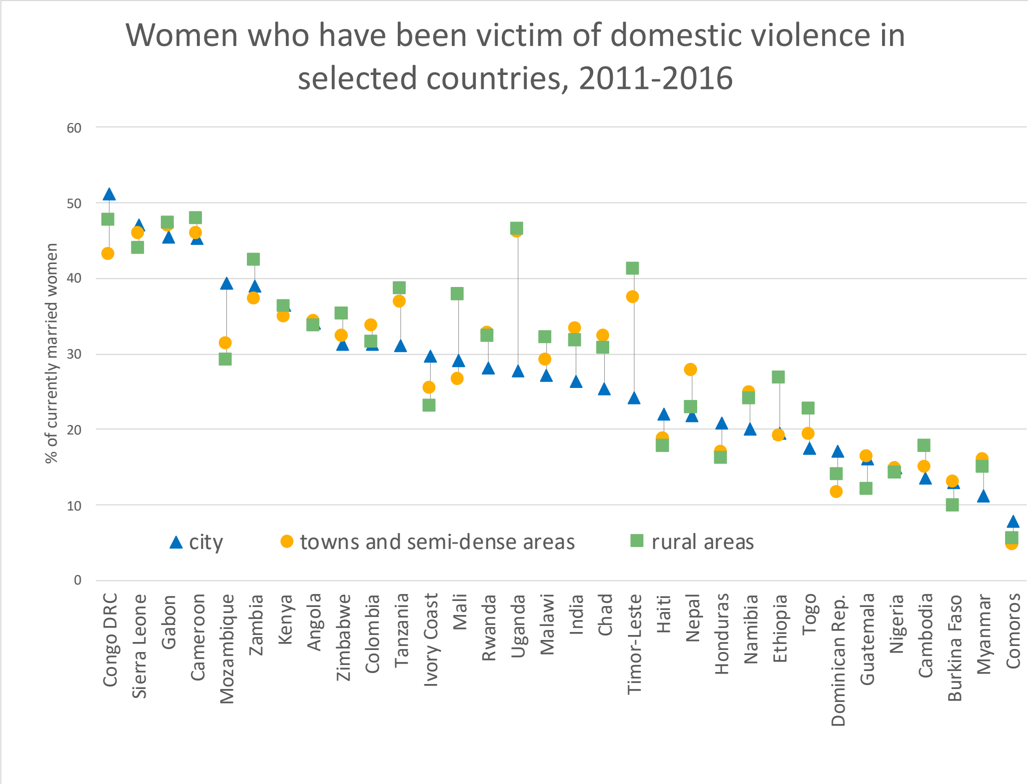


Source: Demographic and Health Survey as calculated by Henderson et al. (2020)

SDG 5 — Achieve gender equality and empower all women and girls

Among the countries shown in **Error! Reference source not found.**, on average 29 % of married women living in rural areas had experienced domestic violence, compared with 28 % for married women living in cities and 27 % for towns and semi-dense areas. In some countries, the share of married women having experienced domestic violence was considerably higher for those living in rural areas compared with those living in cities, for example in Uganda the difference was 19 percentage points and in Timor-Leste it was 17 points. In Mozambique, however, this pattern was reversed as married women living in cities were more likely to have experienced domestic violence than those living in rural areas (with a gap of 10 percentage points). This indicator captures SDG 5.2.1 with the only difference being that it does not ask if the domestic violence experienced by married women occurred during the 12 months prior to the Demographic and Health Survey.

Figure 8.6: Share of married women who have been the victim of domestic violence, by degree of urbanisation, selected countries, 2012-2016
(%)

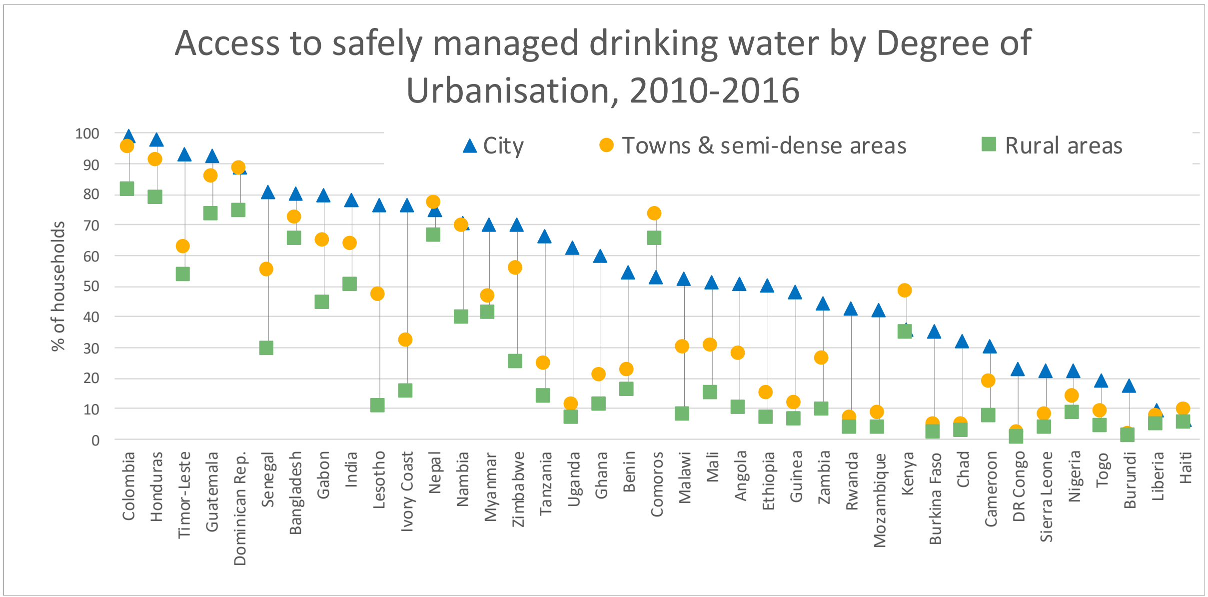


Source: Demographic and Health Survey as calculated by Henderson et al. (2020)

SDG 6 — Ensure availability and sustainable management of water and sanitation for all

Error! Reference source not found. shows that in most countries included in the Demographic and Health Survey a higher share of households in cities had access to safely managed drinking water than the share recorded for households in towns and semi-dense areas, which in turn had a higher share than for households in rural areas. On average, across all of the countries shown, 56 % of households in cities had access to safely managed drinking water compared with 26 % of households in rural areas, while households in towns and semi-dense areas had an intermediate share (37 %). This indicator corresponds to SDG 6.1.1.

Figure 8.7: Share of households having access to safely managed drinking water, by degree of urbanisation, selected countries, 2010-2016
(%)

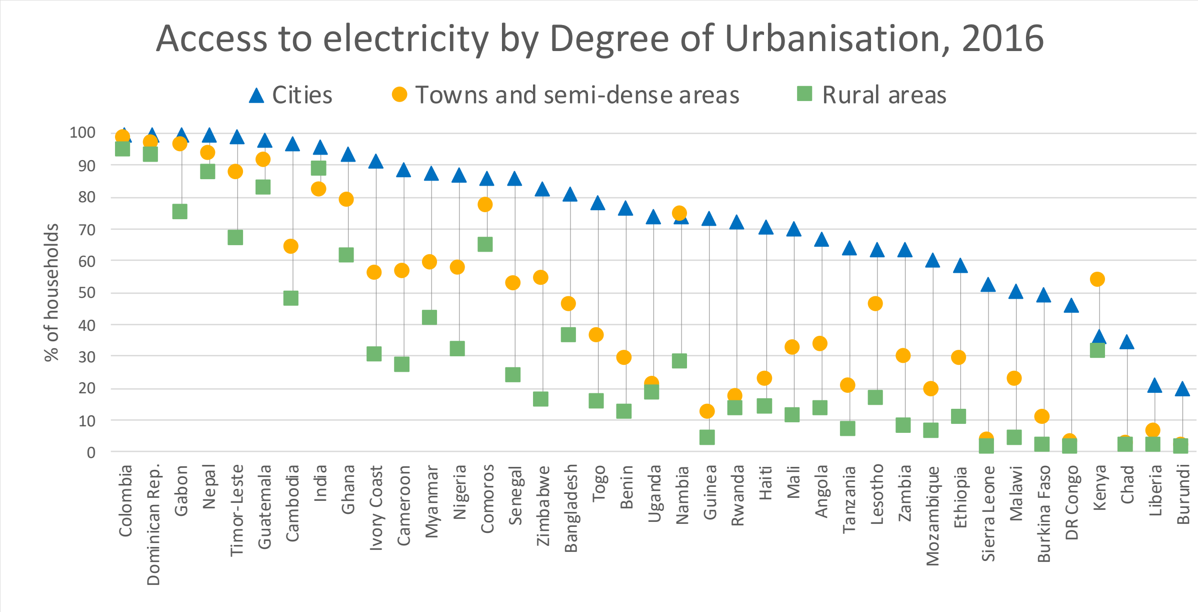


Note: safely managed drinking water is defined by the DHS-WHO Joint Monitoring Programme as all improved water sources that take zero minutes to collect or are on the premises; improved water sources encompass all piped water and packaged water, as well as protected wells or springs, boreholes, and rainwater.
Source: Demographic and Health Survey as calculated by Henderson et al. (2020)

SDG 7 — Ensure access to affordable, reliable, sustainable and modern energy for all

The share of households in cities with access to electricity was generally much higher than that recorded for households in rural areas. On average, across all of the countries shown in **Error! Reference source not found.**, 73 % of households in cities had access to electricity compared with 31 % in rural areas. Households in towns and semi-dense areas had an intermediate share (45 % had access to electricity). In 11 out of the 39 countries shown in Figure 8.8, the share of households in rural areas with access to electricity was within the range of 0-10 %. This indicator corresponds to SDG 7.1.1.

Figure 8.8: Share of households having access to electricity, by degree of urbanisation, selected countries, 2016
(%)



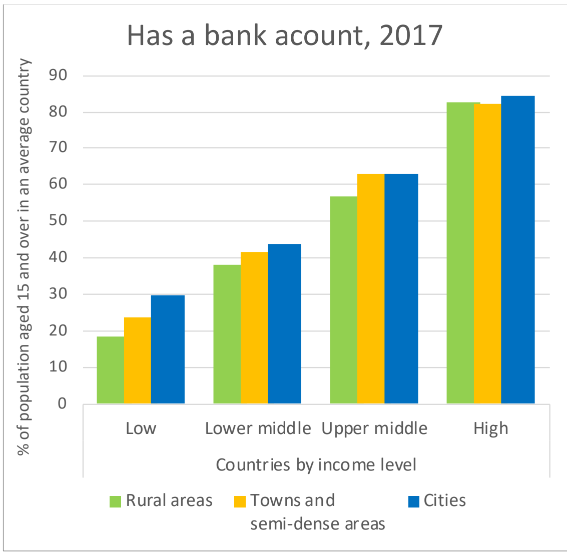
Source: Demographic and Health Survey as calculated by Henderson et al. (2020)

SDG 8 — Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all

Financial services can help people to escape poverty: for example, they can make it possible for people to invest in education, to finance health care or to start a business. Having a bank account is a first important step to accessing such services or taking such initiatives. A bank account also makes it easier to manage payments safely.

However, most people in low-income countries do not have a bank account. The share of the adult population (persons aged 15 years or over) living in low-income countries with a bank account was highest in cities (30 % of adult city-dwellers had a bank account; see **Error! Reference source not found.**). A much lower share (18 %) of the adult population in rural areas of low-income countries had a bank account. By contrast, the share of the population with a bank account in high-income countries was above 80 % for all three classes by degree of urbanisation. In the two groups of middle income countries, adults living in rural areas were also less likely to have a bank account than people living in towns and semi-dense areas or in cities.

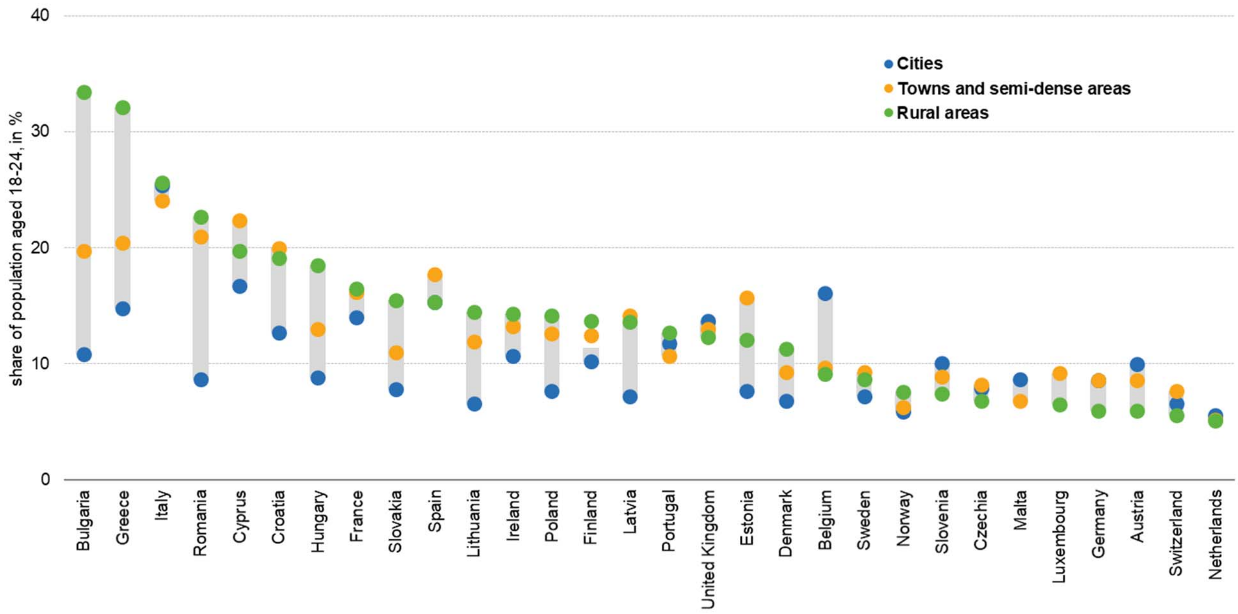
Figure 8.9: Share of the population aged 15 years or over with a bank account, by degree of urbanisation and income group, 2017
(%)



Source: Global Findex (2017)

Error! Reference source not found. shows that in most European countries the share of young people (aged 15-24 years) neither in employment nor in education or training (the NEET rate) was often considerably higher for young people living in rural areas than it was for those living in cities; this was most notably the case in Bulgaria, Greece, Romania and Hungary. However, in six of the countries shown, the NEET rate was higher for young people living in cities than it was for young people living in towns and semi-dense areas or in rural areas; this was most notably the case in Belgium and Austria, and was also observed in Slovenia, Malta, the United Kingdom and the Netherlands. This indicator corresponds to SDG 8.6.1.

Figure 8.10: Share of young people (aged 15-24 years) neither in employment nor in education or training, by degree of urbanisation, selected European countries, 2018
(%)

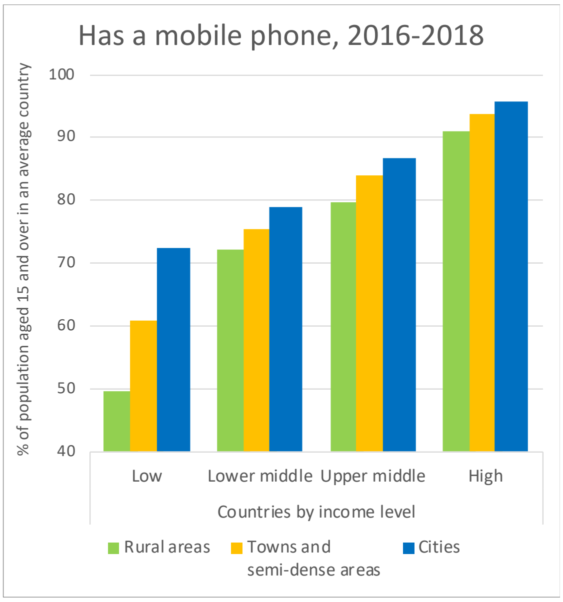


Note: the values for cities in Germany and Luxembourg are hidden by the values for towns and semi-dense areas.
Source: Eurostat (online data code: edat_lfse_29)

SDG 9 — Proportion of population covered by a mobile network, by technology

Mobile phone ownership has increased over the last few decades. Nevertheless, only half the rural population living in low-income countries owned a mobile phone, compared with almost three quarters of city-dwellers living in low-income countries (see **Error! Reference source not found.**). The gap in mobile phone access between rural areas and cities narrowed as average income levels increased. Nevertheless, in high-income countries there remained a 5 percentage point gap in mobile phone ownership in favour of city-dwellers.

Figure 8.11: Share of the population aged 15 years or over with a mobile phone, by degree of urbanisation and income level, 2016-2018
(%)



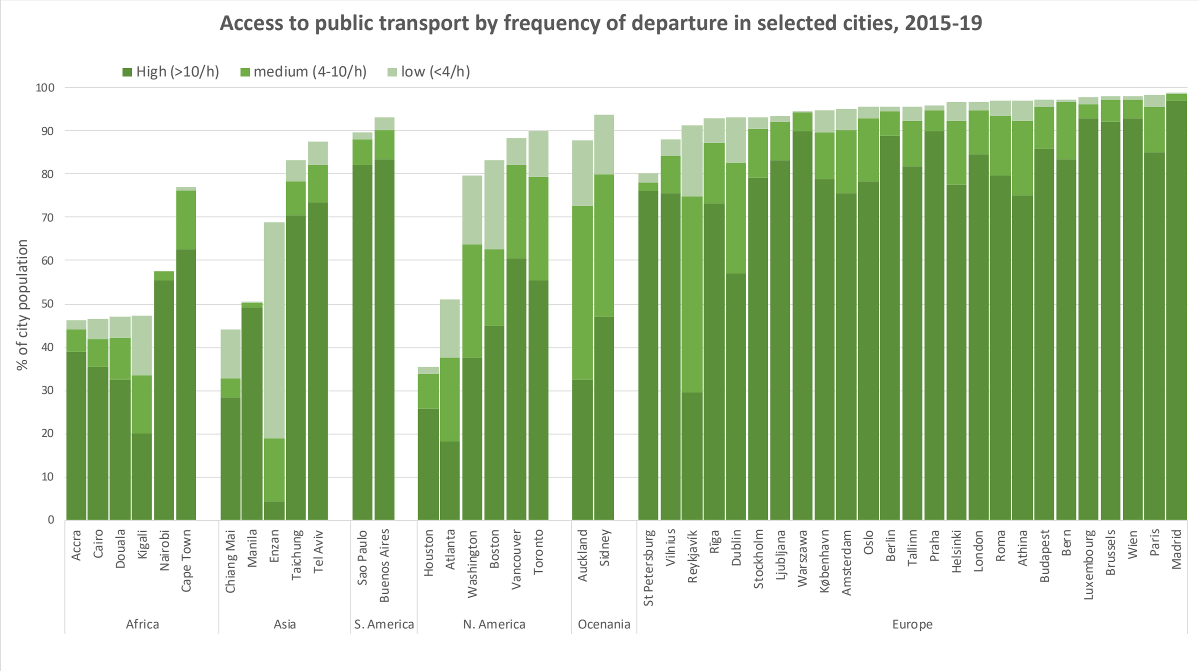
Source: Gallup World Poll

SDG 11 — Make cities and human settlements inclusive, safe, resilient and sustainable

Access to public transport in cities is considered critical to encourage low-carbon mobility and ensure that people can get where they need or want to go. This is especially the case for those people who cannot drive, do not want to drive or cannot afford to drive. The core SDG indicator 11.2.1 measures the share of city-dwellers living within 500 metres walking distance of a transport stop. A secondary indicator takes into account the frequency of departures and expands the distance under consideration so that transport stops within a 1 km radius by foot are taken into account if they provide access to a faster mode of transport (such as bus rapid transit, metro or rail). **Error! Reference source not found.** shows this secondary indicator.

The selected South American cities and most of the selected European cities had a relatively high level of access to public transport with a high frequency of departures. In the selected cities of North America and Oceania, access to public transport was somewhat lower (in particular in Houston and Atlanta), while the frequency of departures was generally lower than in European or South American cities. In the selected cities of Africa and Asia, the situation was more mixed. Some cities, including Cape Town, Taichung or Tel Aviv, offered a relatively high level of access combined with a relatively high frequency of departures. In most other cities selected for Africa and Asia, less than half the population had access to public transport.

Figure 8.12: Share of city-dwellers with access to public transport by frequency of departure, selected cities, 2015-2019
(%)

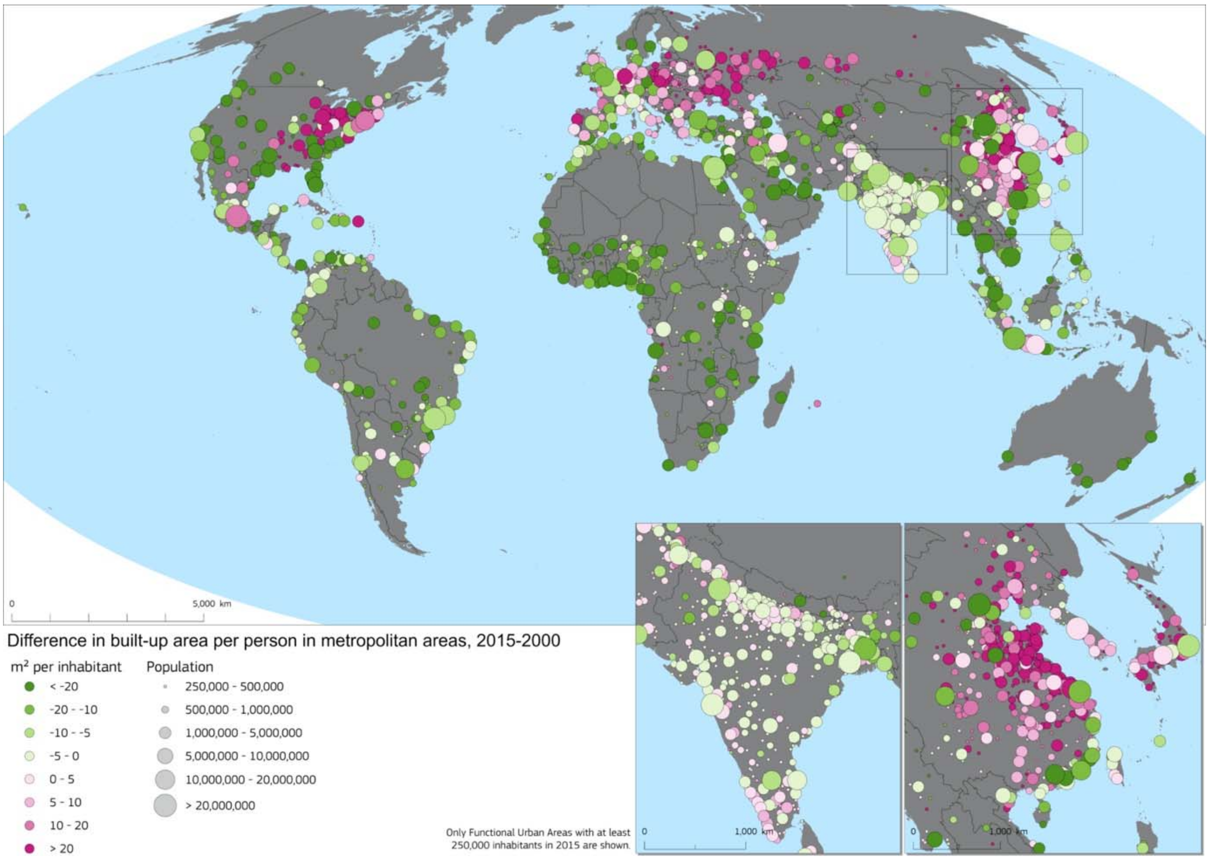


Source: European Commission and International Transport Forum calculated using General Transit Feed Specification (GTFS) data from various sources and population data from GHS-POP

To measure sustainable urbanisation, SDG indicator 11.3.1 is based on the ratio between land use change and population change. The methodology proposed for this indicator is rather complex (a unitless ratio of two logarithmic changes derived from boundaries that change over time). The indicator presented in Map 8.1 is simpler. It compares the amount of built-up land per person for two points in time using the most recent metropolitan boundary. This means that the indicator has a more understandable unit (built-up land in m^2 per person) and the changes can be compared with the amount of built-up land per person for the first reference period. The amount of built-up land is a secondary indicator for SDG 11.3.1.

Map 6.1 shows that most metropolitan areas in the world reduced their ratio of built-up land per inhabitant between 2000 and 2015 (those metropolitan areas shaded in green). Some metropolitan areas increased their amount of built-up land per inhabitant because their built-up land grew at a faster rate than their total number of inhabitants or because their total number of inhabitants declined, as was the case for many metropolitan areas of China, central Asia and eastern Europe. The data for metropolitan areas reducing their amount of built-up land per inhabitant should be interpreted cautiously and with regard to the initial level of built-up land. Those with very low amounts of built-up land per inhabitant may be characterised by low levels of infrastructure and high numbers of inhabitants living in crowded conditions.

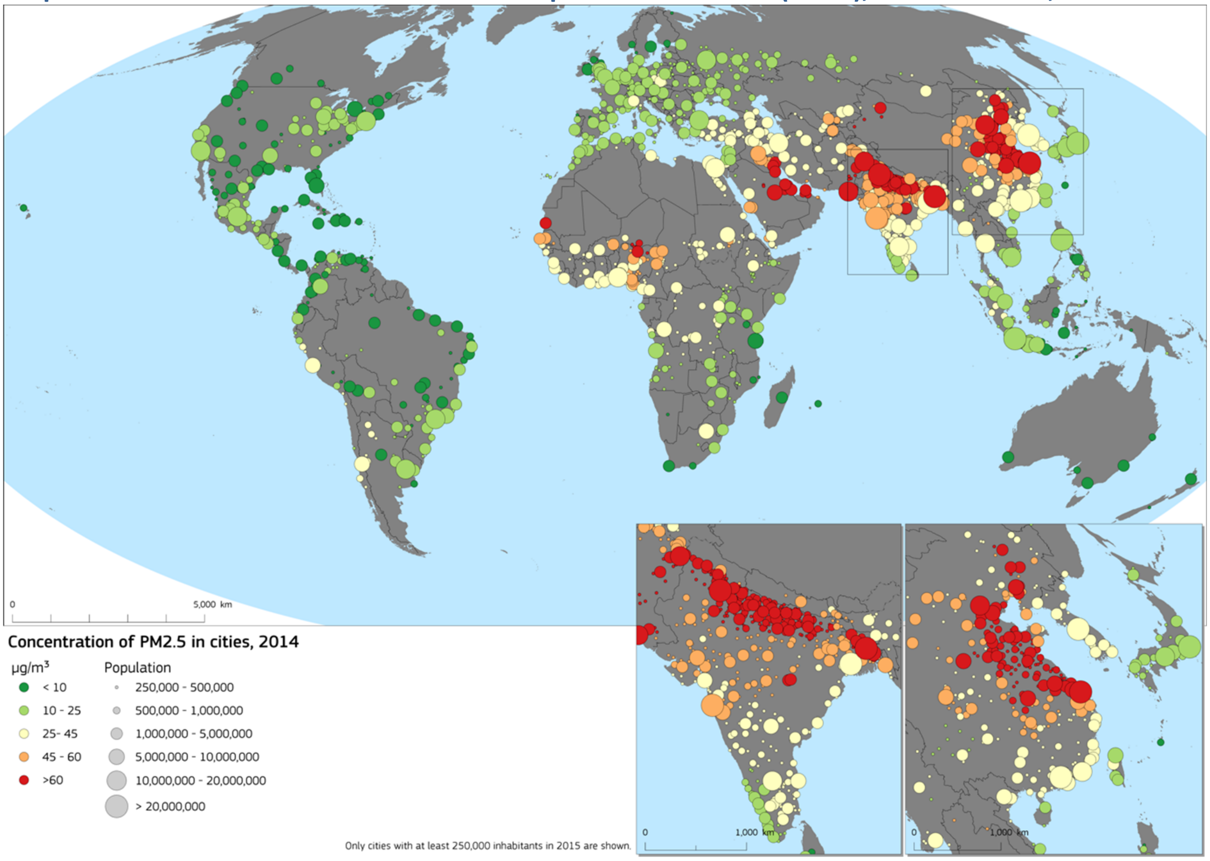
Map 8.1:5 Change in the ratio of built-up land per inhabitant, selected metropolitan areas, 2000-2015



Source: GHS-BUILT using boundaries from Moreno-Monroy et al. (2020)

The spatial concentration of people and economic activities in cities can lead to high levels of air pollution, which may potentially harm people's health and reduce their life expectancy, as well as having other consequences. Many cities in China and India had high concentrations of fine particulate matter ($PM_{2.5}$ — particles with a diameter of 2.5 micrometres (µm) or less) of at least 60 micrograms per cubic metre (µg/m^3), which was six times higher than the World Health Organisation's limit for protecting human health (10 µg/m^3).

Map 8.2: Annual mean concentration of fine particulate matter ($PM_{2.5}$), selected cities, 2014



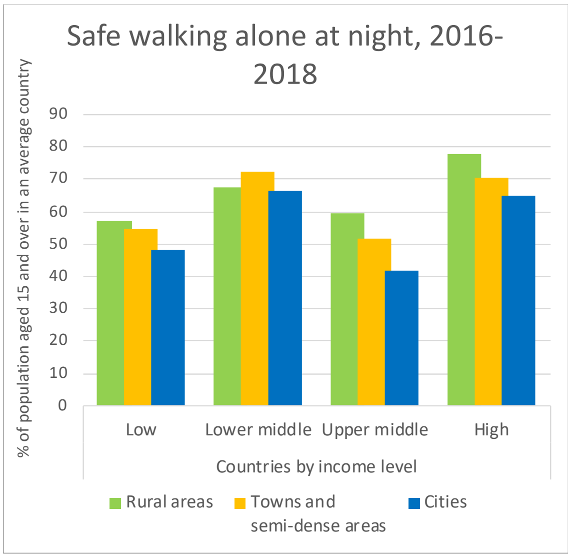
Source: JRC Urban Centre Database from Florczyk et al. (2019)

SDG 16 — Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels

People living in rural areas are more likely to feel safe when they are walking alone at night than city-dwellers. This information is covered by SDG indicator 16.1.4.

People living in rural areas felt safer walking alone at night than people living in cities for all four groups of countries based on average income levels (as shown in Figure 8.13). This urban gradient was clearly visible for low-, upper-middle and high-income countries. The gap in the proportion of people feeling safe between those living in rural areas and those living in cities was greater for high-income and upper-middle income countries than it was for low-income countries. In lower-middle income countries, people living in towns and semi-dense areas felt safer walking alone than people living in rural areas or in cities.

Figure 8.13: Share of the population aged 15 years or over who considered it was safe to walk alone at night, by degree of urbanisation and income group, 2016-2018
(%)

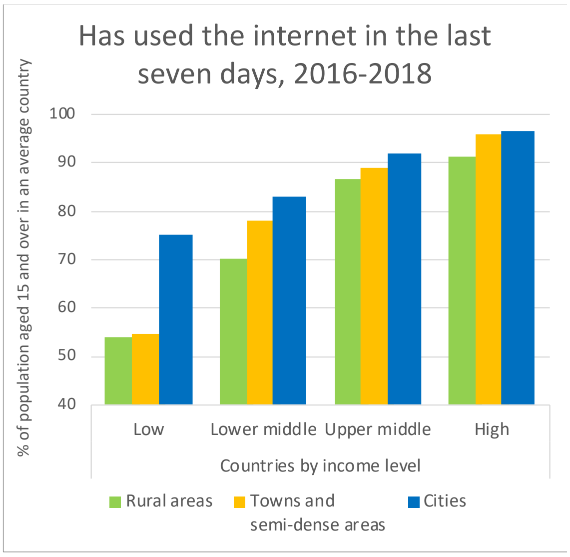


Source: Gallup World Poll

SDG 17 — Strengthen the means of implementation and revitalise the global partnership for sustainable development

SDG indicator 17.8.1 concerns use of the internet. Cities typically have a higher share of internet use than rural areas (see **Error! Reference source not found.**). The gap between cities and rural areas was biggest in low-income countries where, on average, 54 % of people aged 15 years or over in rural areas used the internet in the seven days prior to the Gallup World Poll survey, compared with 75 % in cities. As the average income level in a country goes up, the gap in internet use between rural areas and cities tends to narrow. Nevertheless, a 5 percentage point gap remained for high-income countries.

Figure 8.14: Share of the population aged 15 years or over having made use of the internet in the previous seven days, by degree of urbanisation and income group, 2016-2018
(%)



Source: Gallup World Poll

References

Dijkstra, L. and H. Poelman (2012), 'Cities in Europe: the new OECD-EC definition', *Regional Focus*, No. 01/2012, European Commission.

Eurostat (2019), '*Methodological manual on territorial typologies — 2018 edition*', Publications Office of the European Union, Luxembourg.

Florczyk, A., C. Corbane, M. Schiavina, M. Pesaresi, L. Maffenini, M. Melchiorri, P. Politis, F. Sabo, S. Freire, D. Ehrlich, T. Kemper, P. Tommasi, D. Airaghi and L. Zanchetta (2019), *GHS Urban Centre Database 2015, multitemporal and multidimensional attributes*, European Commission.

Global Findex (2017), *The Global Findex Database 2017*, The World Bank.

Henderson, J. V., V. Liu, C. Peng and A. Storeygard (2020), *Demographic and health outcomes by degree of urbanisation — perspectives from a new classification of urban areas*, Publications Office of the European Union, Luxembourg.

Maina, J., P. O. Ouma, P. M. Macharia, V. A. Alegana, B. Mitto, I. S. Fall, A. M. Noor, R. W. Snow and E. A. Okiro (2019), 'A spatial database of health facilities managed by the public health sector in sub Saharan Africa', *Scientific Data*, Volume 6(1).

Moreno-Monroy, A. I., M. Schiavina and P. Veneri (2020), 'Metropolitan areas in the world. Delineation and population trends', *Journal of Urban Economics*.

9. Tools and training

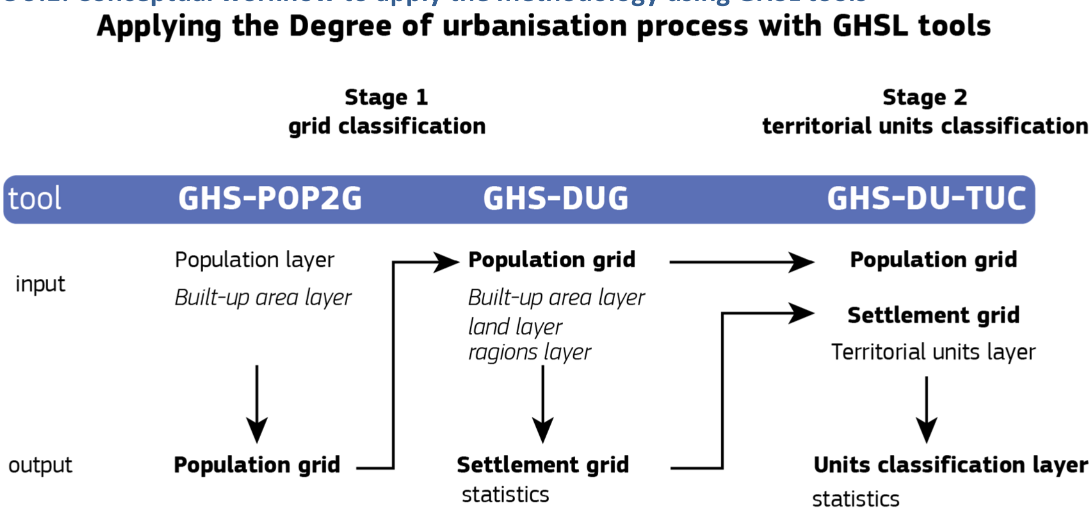
The degree of urbanisation classification is a geospatial concept that can be implemented in geographic information systems (GIS) — computer systems designed to analyse spatial data. However, this requires adequate expertise to operate the GIS in an appropriate way and the availability of the necessary population and, optionally, built-up density grids. There is a strong demand for ready-to-use tools that facilitate the application of the degree of urbanisation classification as well as for capacity building that assures a conscious implementation of the methodology on which it is based.

This chapter describes tools that are currently available and training materials that have been produced by the European Commission's Joint Research Centre (JRC) *Global Human Settlement Layer (GHSL)* project to support the development of a harmonised global definition of cities and settlements.

9.1 Tools

The tools described in this subchapter address three production steps that are described in the previous chapters. The first step is the construction of a regular-spaced population grid from given geospatial population data in the form of points or polygons (see Chapter 5). The second step is the application of the methodology to a given population grid and additional optional layers (see Chapters 6 and 7). Finally, in the last step, the derived grid cell classification is used to classify small spatial units into cities, towns and semi dense areas, or rural areas (see also Chapters 6 and 7). Figure 9.1 displays the workflow to operationalise the tools that have been produced within the framework of the GHSL.

Figure 9.1: Conceptual workflow to apply the methodology using GHSL tools



The tools described below are open and available free-of-charge from the GHSL tools website ([17]). They require the installation of MATLAB Runtime ([18]), which is a standalone set of shared libraries that enable the execution of compiled MATLAB applications. The tools are also available as an ArcGIS toolbox, compatible with ArcMap 10.6. The tools were developed to run on standard computers. They all run on Windows 10 operating systems with any processor and require at least 16GB RAM. It

([17]) *Global Human Settlement Layer* (https://ghsl.jrc.ec.europa.eu/tools.php).
([18]) Available from *MathWorks* (https://mathworks.com/products/compiler/matlab-runtime.html).

is important to note that more memory is required for processing larger data sets. More details can be found in the corresponding user manuals (see below for more information on specific user guides).

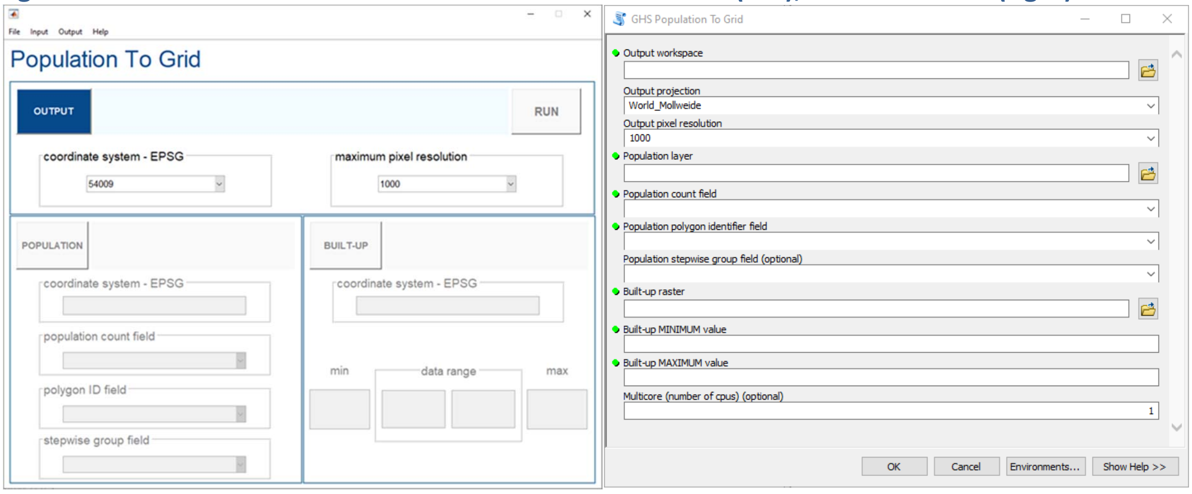
9.1.1 Construction of a population grid (population to grid tool — GHS-POP2G)

A population grid is the key input to produce the grid cell classification that is necessary in order to compile data by degree of urbanisation. A population grid is obtained by re-allocating population counts from points and/or polygons to gridded surfaces of regular and standardised grid cells or pixels. The population grid is produced through geospatial and geo-statistical processing of geo-coded population data (as available).

Population grids can be produced in alternative ways depending on the type of data available. One process is that of aggregation. The aggregation approach is generally used when micro-census source data have higher spatial detail (resolution) than the selected cell size of the population grid. A point-based micro-census is usually conducted at the building or census block level, and this high level of spatial detail should be the only one for which this aggregation technique should be deployed. Population grids are more generally produced through disaggregation of population counts attached to small spatial units — statistical areas or administrative units. The GHS population grid layers (GHS-POP) are produced through disaggregation (the population input is the *Gridded Population of the World* v4.10 (CIESIN (2018)). The disaggregation is driven by the density of built-up areas as a proxy for the location of the resident population.

To support the uptake of this methodology, the GHSL project has developed a population to grid tool — GHS-POP2G (version 2). This is a flexible tool to produce geospatial population grids in GeoTIFF format from census data. It operationalises the workflow developed for the production of the GHS-POP. GHS-POP2G offers the possibility to create population grids at 50 m, 100 m, 250 m and 1 km spatial resolutions, handling census data stored as point or polygon vector data (the latter case requires an additional covariate as input for dasymetric disaggregation); it is available as standalone software or as an ArcGIS toolbox (Figure 9.2). The principal purpose of the tool is to produce a population grid that may be used as an input for the degree of urbanisation grid tool (GHS-DUG) which has also been produced within the GHSL framework. However, potential uses of the tool and population grids extend far beyond this principal application. The GHS-POP2G user manual (Maffenini et al. (2020a)) explains all of the functionalities and requirements to run the tool.

Figure 9.2: GHS-POP2G interface window — standalone tool (left); ArcGIS toolbox (right)



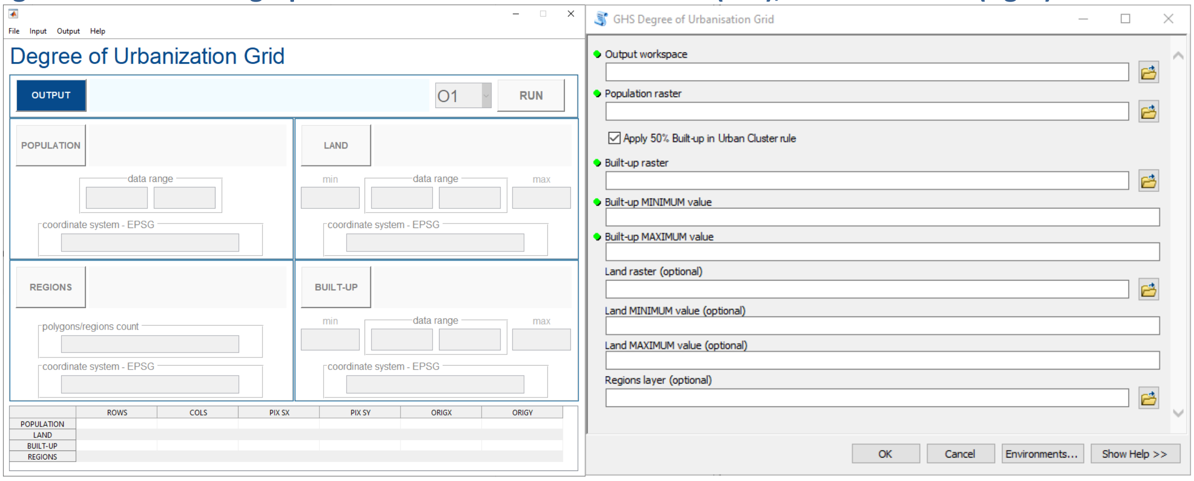
9.1.2 Classifying grid cells grid (degree of urbanisation grid tool — GHS-DUG)

The degree of urbanisation grid tool — GHS-DUG (version 4) is an information system to produce geospatial grids for degree of urbanisation classes and related statistics.

GHS-DUG 4 is designed as a scalable tool allowing the application of methodology to available population grids or to data made available in the *GHSL Data Package 2019* (Florczyk et al. (2019)).

The GHS-DUG implements the workflow developed for the production of the GHS-SMOD. It produces a grid cell classification for the entire area of interest in GeoTiff format at 1 km spatial resolution according to both level 1 and level 2 of the degree of urbanisation classification. GHS-DUG requires a population grid (at 1 km resolution) and optionally a built-up surface and land fraction layers. When a shapefile delimiting territorial units is provided, the tool calculates statistics by degree of urbanisation class. The principal purpose of the tool is the production of a classification of grid cells by degree of urbanisation. The GHS-DUG grid output is used to operationalise stage 2 of the methodology (the classification of small spatial units) and is used as an input for the degree of urbanisation territorial unit classifier tool (GHS-DU-TUC) also produced within the GHSL framework (see Subchapter 9.1.3). The GHS-DUG user manual (Maffenini et al. (2020b)) explains all of the functionalities and requirements to run the tool. Figure 9.3 shows the graphical interface of the GHS-DUG tool both for the standalone tool and for the ArcGIS toolbox.

Figure 9.3: GHS-DUG graphical interface — standalone tool (left); ArcGIS toolbox (right)

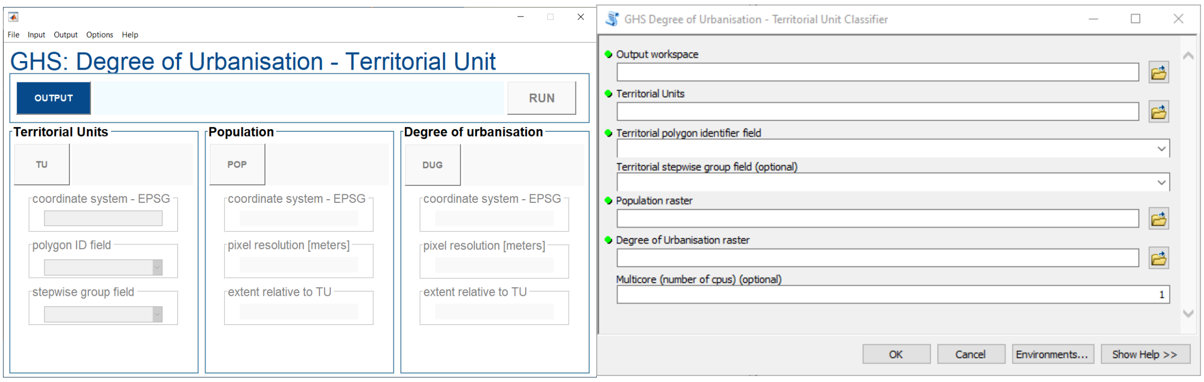


9.1.3 Classifying small spatial units (degree of urbanisation territorial unit classifier tool — GHS-DU-TUC)

The methodology classifies the entire territory of a country along the urban-rural continuum into regularly spaced grid cells. However, often it is required to classify small spatial units, for example a commune or municipality. The GHS-DU-TUC tool implements this transition from the grid cell classification to the classification of small spatial units based on the type of grid cells in which the majority of their population resides.

The degree of urbanisation territorial unit classifier tool — GHS-DU-TUC (version 1.0) is designed as an operational tool that classifies small spatial units based on the grid cell classification already derived using the GHS-DUG tool. It requires the following inputs: a classification of grid cells, a population grid and a geospatial layer containing the small spatial units. The input population grid must be the one used for the production of the grid cell classification through the GHS-DUG tool. GHS-DU-TUC produces a geospatial layer in vector format (a shapefile) that contains the classification of small spatial units according to levels 1 and 2 of the degree of urbanisation classification, plus a statistical table with the classification of the small spatial (territorial) units and their population counts. The GHS-DU-TUC user manual (Maffenini et al. (2020c)) explains all of the functionalities and requirements to run the tool. Figure 9.4 shows the graphical interface of the GHS-DU-TUC tool both for the standalone tool and for the ArcGIS toolbox.

Figure 9.4 GHS-DU-TUC graphical interface — standalone tool (left); ArcGIS toolbox (right)



9.2 Training

The tools described in the previous subchapter are distributed with detailed manuals that encourage autonomous use (see References at the end of this subchapter for further details). Nevertheless, additional training materials or training courses are available to expedite the correct selection and application of the different options. In preparation for the 51st session of the UN Statistical Commission, partner organisations supported a range of countries in different ways to increase their capacity to understand and implement the methodology.

UN-Habitat together with the European Commission organised seven regional workshops between 2018 and 2019 to present the methodology underlying the degree of urbanisation classification and discuss how this could be improved and applied to national data. A total of 85 countries participated in these workshops (see Figure 9.5).

Figure 9.5: Overview of regional workshops presenting the methodology

Abuja, Nigeria, 15-19 October 2018 with representatives from: Nigeria, Ghana, The Gambia, Sierra Leone, Kenya, Ethiopia, South Sudan, Liberia, Uganda	**Abidjan, Ivory Coast, 13-16 November 2018** with representatives from: Burundi, Burkina Faso, Central African Republic, Chad, Congo, Comoros, Democratic Republic of Congo, Madagascar, Djibouti, Mali, Niger, Senegal, Guinea, Togo, Ivory Coast
Lusaka, Zambia, 22-25 January 2019 with representatives from: Botswana, Malawi, Tanzania, Mauritius, Angola, Zimbabwe, Mozambique, South Africa, Eswatini, Lesotho, Namibia, Zambia	**Cairo, Egypt, 18-21 March 2019** with representatives from: Egypt, Morocco, Sudan, Tunisia, Bahrain, Iraq, Jordan, Kuwait, Lebanon, Oman, Palestine, Saudi Arabia, Syria, Yemen
Lima, Peru, 25-28 June 2019 with representatives from: Argentina, Bolivia, Brazil, Chile, Costa Rica, Colombia, Cuba, Dominican Republic, Ecuador, Mexico, Peru, Uruguay	**Delhi, India, 23-26 September 2019** with representatives from: Azerbaijan, Armenia, Bangladesh, Bhutan, India, Kyrgyzstan, Maldives, Nepal, Sri Lanka, Uzbekistan
Kuala Lumpur, Malaysia, 22-25 October 2019 with representatives from: Afghanistan, Australia, China, Iran, Kazakhstan, Lao PDR, Malaysia, Mongolia, Myanmar, New Zealand, Thailand, Timor-Leste, Vietnam	

As a follow-up to the workshops, the European Commission's Joint Research Centre (JRC) conducted dedicated training in the United Arab Emirates at the request of the Federal Competitiveness and Statistical Authority, and at the UN-Habitat Headquarters in Kenya for UN-Habitat staff. Further events and a comprehensive training package are under preparation.

The objective of training courses is to provide an overview of the data, methods and tools developed by the GHSL project, to provide examples of how data and tools can be used to apply the methodology, and which applications it can support. The course includes presentations and practical exercises. The presentations are targeted at a general audience with a background in regional and urban development and to those working for national statistical authorities; the practical exercises require some basic knowledge in GIS and spreadsheets and the installation of dedicated software prior to the exercise (see Subchapter 9.1 for more details of the specific requirements).

Training courses address four broad themes through presentations:

- The first module addresses the need for a global definition of urban and rural areas for international statistical comparisons.
- The second module explains the GHSL datasets: the built-up area spatial grids (GHS-BUILT), the population spatial grids (GHS-POP), the settlement model spatial grid (GHS-MOD) and the urban centre database (GHS-UCDB).
- The third module explains the JRC's solution to operationalise the degree of urbanisation classification into a settlement classification grid and into the classification of small spatial units by degree of urbanisation for urban and rural areas.
- The fourth module shows examples of GHSL data applications, to support policymaking with new findings on human settlements.

Those taking part in the practical training exercises can expect to learn the following skills:

- Construction of a population grid with the population to grid tool (GHS-POP2G).
- Classification of grid cells with the degree of urbanisation grid tool (GHS-DUG).
- Classification of small spatial units by degree of urbanisation with the degree of urbanisation territorial unit classifier tool (GHS-DU-TUC).
- Disaggregation of statistics and indicators according to the degree of urbanisation classification.
- Estimation of sustainable development goal (SDG) 11.3.1 for urban areas (LUE tool).

The production of stand-alone online courses and webinars is under preparation.

9.3 Online resources for the degree of urbanisation classification

In order to support the discussions with interested countries and stakeholders, the JRC has also created a dedicated web presence for the degree of urbanisation classification ([19]). The homepage contains everything that is needed to understand and implement the degree of urbanisation classification (see Figure 9.6).

([19]) *Global Human Settlement Layer* (https://ghsl.jrc.ec.europa.eu/degurba.php).

Figure 9.6: The degree of urbanisation website

The different sections include:

- An **introduction** to the methodology: why there is a need for a global, people-based definition of cities and urban and rural areas.
- A summary of the **methodology**.
- **Country fact sheets** summarise the application of methodology based on data from the GHSL and publicly available country borders.
- **Interactive maps** and the application of the methodology in the **Urban Centres Database.**
- A **data** section that provides open data for the global grids of the GHSL dataset including built-up area grids, population grids and settlement classification grids (GHS-SMOD layers).
- The **tools** section with links to the available set of tools for implementing the methodology.
- A list of essential **documents**.
- A section summarising materials and initiatives for **capacity building**.

References

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

Florczyk, A., C. Corbane, D. Ehrlich, S. Freire, T. Kemper, L. Maffenini, M. Melchiorri, M. Pesaresi, P. Politis, M. Schiavina, F. Sabo, and L. Zanchetta (2019), *GHSL Data Package 2019*, JRC 117104, EUR 29788 EN, Publications Office of the European Union, Luxembourg.

Maffenini, L., M. Schiavina, S. Freire, M. Melchiorri, M. Pesaresi and T. Kemper (2020a), *GHS-POP2G User Guide*, JRC 121485, Publications Office of the European Union, Luxembourg.

Maffenini, L., M. Schiavina, M. Melchiorri, M. Pesaresi, and T. Kemper (2020b), GHS-*DUG User Guide*, JRC 121484, Publications Office of the European Union, Luxembourg.

Maffenini, L., M. Schiavina, M. Melchiorri, M. Pesaresi, and T. Kemper (2020c), *GHS-DU-TUC User Guide*, JRC 121486, Publications Office of the European Union, Luxembourg.

10. Conclusions

The endorsement of the UN Statistical Commission in March 2020 of the methodology for the delineation of cities and urban and rural areas was a key milestone. However, work in this area is not over. As part of the endorsement process, the UN Statistical Commission made two additional requests. First, that a technical report on the implementation of the methodology for the delineation of cities and urban and rural areas was made available as quickly as possible; this manual responds to that request. Second, that the UN Statistics Division and the sponsoring organisations review the implementation of the methodology for the delineation of cities and urban and rural areas and report back to the UN Statistical Commission at one of its future sessions. As a result, the focus of the work will now shift to three different lines of action.

First, to encourage and support countries applying the methodology for compiling statistics by degree of urbanisation (level 1). The current census round presents an opportunity to apply this methodology using data with a high spatial resolution. In particular, countries that have conducted or will conduct a digital census and collect the GPS location of all households can produce a high-quality population grid. Such a population grid will create a highly robust and accurate classification of a country's settlements. This methodological manual presents a number of tools to make it easier to compile statistics by degree of urbanisation. Nevertheless, hands-on training and responding to specific questions will be needed to ensure that as many countries as possible apply the methodology in a consistent and coherent manner. Several of the organisations behind this work are ready to provide such training and technical support. This experience will then be summarised to report back on the implementation phase to the UN Statistical Commission.

Second, to improve and update global data. To support this work, the Joint Research Centre (JRC) of the European Commission has produced a global, estimated population grid for the years 1975, 1990, 2000 and 2015. Using new imagery, the SENTINEL 1 and 2 satellites and improved methods relying on artificial intelligence and cloud computing, the JRC will publish improved population grids and produce regular updates for free. This will ensure that national administrations, NGOs, the academic community and other interested parties have access to coherent, complete and up-to-date information. In addition, the JRC will explore how to project these population grids up to 2050 and even 2100 by incorporating the latest UN World Population Projections.

Third, to integrate this new methodology in the documentation of the relevant sustainable development goal (SDG) indicators. To facilitate the comparison of data for cities, towns and rural areas, the methodology should be included in the metadata of relevant SDG indicators. This will encourage more countries to produce the SDG indicators in such a way that they can be reliably compared across national borders. To this end, the organisations involved in this work will reach out to the custodian agencies of the various SDG indicators that might be produced by degree of urbanisation (level 1).

Getting in touch with the EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest to you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
– at the following standard number: +32 22999696 or
– by email via: https://europa.eu/european-union/contact_en

Finding information about the EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications at: https://op.europe.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: https://eur-lex.europa.eu

Open data from the EU

The EU Open Data Portal (https://data.europa.eu/euodp/en) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

Getting in touch with the Centre for Entrepreneurship, SMEs, Regions and Cities (CFE) — OECD

In person

The OECD headquarters are located in Paris.
More information available here: http://www.oecd.org/contact/

By email

You can contact: RegionStat@oecd.org or CFE.Contact@oecd.org

Finding information about the CFE

Online

You can find more information on these topics on the following websites: http://www.oecd.org/cfe/ or http://www.oecd.org/regional/regional-statistics/

Publications and datasets

You can consult our publications and datasets at: https://www.oecd-ilibrary.org/
If you wish to subscribe to our newsletter please visit: http://oe.cd/CFEnews

Social media

Follow us on:
LinkedIn: https://www.linkedin.com/company/oecd-local
Twitter: https://twitter.com/OECD_local

Applying the Degree of Urbanisation

A methodological manual to define cities, towns and rural areas for international comparisons

Applying the Degree of Urbanisation — A methodological manual to define cities, towns and rural areas for international comparisons has been produced in close collaboration by six organisations — the European Commission, the Food and Agriculture Organization of the United Nations (FAO), the United Nations Human Settlements Programme (UN-Habitat), the International Labour Organization (ILO), the Organisation for Economic Co-operation and Development (OECD) and The World Bank.

This manual develops a harmonised methodology to facilitate international statistical comparisons and to classify the entire territory of a country along an urban-rural continuum. The degree of urbanisation classification defines cities, towns and semi-dense areas, and rural areas. This first level of the classification may be complemented by a range of more detailed concepts, such as: metropolitan areas, commuting zones, dense towns, semi-dense towns, suburban or peri-urban areas, villages, dispersed rural areas and mostly uninhabited areas.

This manual is intended to complement and not replace the definitions used by national statistical offices (NSOs) and ministries. It has been designed principally as a guide for data producers, suppliers and statisticians so that they have the necessary information to implement the methodology and ensure coherency within their data collections. It may also be of interest to users of subnational statistics so they may better understand, interpret and use official subnational statistics for taking informed decisions and policymaking.

For more information

https://ec.europa.eu/eurostat/
https://ec.europa.eu/info/departments/regional-and-urban-policy_en
https://www.fao.org/home/en/
https://unhabitat.org/
https://www.oecd.org/
https://www.worldbank.org/

ISBN 978-92-76-20305-6

KS-02-20-499-EN-C

doi: 10.2785/71043