UN-GGIM: Europe | Work Group on Data Integration | subgroup 2

The territorial dimension in SDG indicators: the contribution of geospatial data and analysis and its combination with statistical data

11.3.1 | Ratio of land consumption rate to population growth rate [tier II indicator]

Brief discussion

At the ***global level***, UN-Habitat and other partners such as the Global Human Settlement Layer (GHSL) team and ESRI will support various components for reporting on this indicator.

Conceptually, the indicator requires defining the two components of population growth and land consumption rate:

> Population growth rate is proposed to be measured by the increase of a population in a country during a period, usually one year, expressed as a percentage of the population at the start of that period. It reflects the number of births and deaths during a period and the number of people migrating to and from a country;
>
> Land consumption can include a) the expansion of built-up area which can be directly measured; b) the absolute extent of land that is subject to exploitation by agriculture, forestry or other economic activities; and c) the over-intensive exploitation of land that is used for agriculture and forestry. The percentage of current total urban land that was newly developed (consumed) is proposed to be used as a measure of the land consumption rate. The fully developed area is also sometimes referred to as built up area.

The indicator should be disaggregated by location (intra-urban), income level and urban typology and monitoring is targeted to be repeated at regular intervals of five years (starting in 2017), allowing for three reporting points until the year 2030. The periods for both urban expansion and population growth rates should be at comparable scale. At the global level, the suggested method to calculate the indicator should consider the following steps:

a) Estimate the land consumption rate (LCR):

$$LCR = \frac{LN\left(\frac{Urb_{t+n}}{Urb_t}\right)}{(y)}$$

b) Estimate the population growth rate (PGR):

$$PGR = \frac{LN\left(\frac{Pop_{t+n}}{Pop_t}\right)}{(y)}$$

c) Estimate the ratio of land consumption rate to population growth rate (LCRPGR).

$$LCRPGR = \left(\frac{LN\left(\frac{Urb_{t+n}}{Urb_t}\right)}{(y)}\right) \Big/ \left(\frac{LN\left(\frac{Pop_{t+n}}{Pop_t}\right)}{(y)}\right)$$

where: Urb_t = Total areal extent of the urban agglomeration in km² for past/initial year; Urb_{t+n} = Total areal extent of the urban agglomeration in km² for current year; Pop_t = Total population within the city in the past/initial year; Pop_{t+n} = Total population within the city in the current/final year; y = The number of years between the two measurement periods.

This indicator is categorized under Tier II, meaning the indicator is conceptually clear and an established methodology exists but data on many countries is not yet available. Data for this indicator, at the global level, is available for all cities and countries (UN DESA - United Nations Department of Economic and Social Affairs population data) and satellite images from open sources. Data regarding the size of the city is usually available from the urban planning units of the cities, but new options using remote sensing techniques have also been developed to estimate the land that is currently developed or considered as built up areas out of the total city land. The Global Human Settlement Layer (GHSL) [see Box 1] technology open framework is proposed for global open spatial baseline data production (built-up and population grids).

Box 1 - Global Human Settlement Layer

The Global Human Settlement Layer (GHSL) is a new free and open access tool, operating through data and methods free access policy, aiming to measure the human presence on the planet. It is supported by the Joint Research Centre (JRC) and the DG for Regional Development (DG REGIO) of the European Commission, together with the international partnership of the GEO Human Planet Initiative. This tool includes a review of previous efforts to map settlements at various scales and using different datasets in order to produce new global spatial information, evidence-based analytics and knowledge for describing the human presence on the planet. The paradigm underlying the GHSL is the design and implementation of new spatial data mining technologies for automatic processing, analysis and knowledge integration from heterogeneous data (i.e. global, multiple fine-scale satellite image data streams, census data and volunteering geographic information sources). The general methodology behind GHSL data introduces concepts of GHS built-up, GHS population, and the GHS settlement model [Figure 1].

GSH built-up (BU) grids were produced based on Landsat imagery (1975, 1990, 2000 and 2015) and on automatic analysis of satellite imagery, by exploiting texture, morphology and pattern to derive a 'built-up presence index'. The distribution of built-up areas is expressed as their proportion of occupied area in each cell. GSH built-up were produced for a 250m^2 and a 38m^2 resolution. GHS population (POP) grid was produced based on national available layers on census data and administrative polygons for the years 1975-1990-2000-2015 with a 250m^2 resolution. The combined information results in a new layer that represents the presence and density of population. Built-up area is typically expressed with continuous values representing the proportion of building footprint area within the total size of the cell. Population grid cell value represents the numbers of inhabitants and the GSH settlement model (S-MOD) aims at classifying human settlements according to certain rules of population and built-up density and contiguity of grid cells, namely by taking into consideration the DEGURBA framework.

Figure 1 - Global Human Settlement – general methodology

Source: European Commission, [Global Human Settlement]

The analysis of the WG members regarding the metadata on this indicator has pointed out three main dimensions that need further development:

- At the conceptual level

A clear definition regarding the underlying concepts for the operationalization of this indicator is needed, namely:

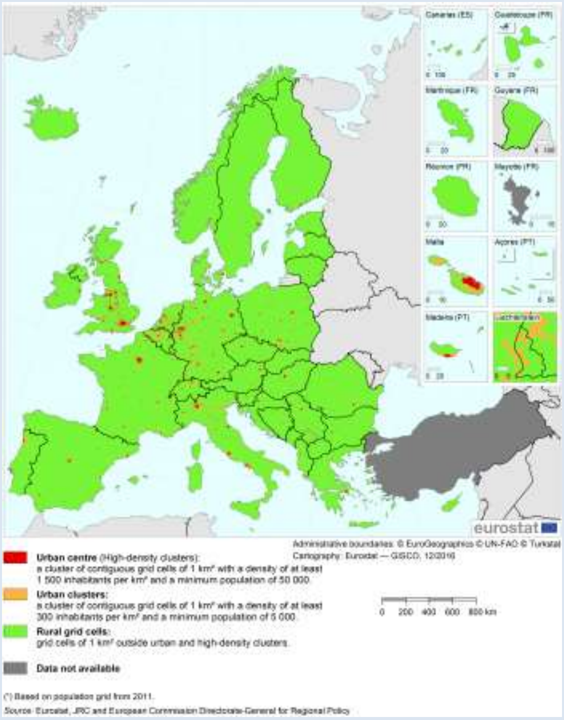
- Urban area / city and built-up area: these are two different concepts and the metadata presentation must be more precise on the use of these terms. In particular the 'urban' concept may be based on a normative approach (zoning) or a *de facto* approach and the *de facto* approach can be based either on morphological or on functional definitions; additionally city can be assessed based on an administrative perspective or a statistical definition. Built-up area should be used as a metric to capture the artificial land and the expansion of land consumption and not be used as an alternative of the former. Urban area/city concepts are therefore a mean to define the territorial aim to apply in this indicator, while built-up areas are the object of the indicator operationalization as a way to capture artificial land (and land consumption over time).

- In this context, urban area/city operationalization should be based on an international statistical definition to deal with issues of comparability. The EU definition (TERCET regulation) can be taken as a reference to a worldwide definition following the discussions in the context of Habitat III [see Box 2]. Having in mind the TERCET regulation, we can refer either to the concept of European 'city' or the 'functional urban area' (Urban Audit) or the urban area from DEGURBA, which correspond to 'cities' + 'towns and suburbs'. Additionally, countries may also have national classifications. The indicator could also account for the entire territory and not be limited to the urban perimeter.

Box 2 - Territorial typologies at the European level relevant for urban delimitation

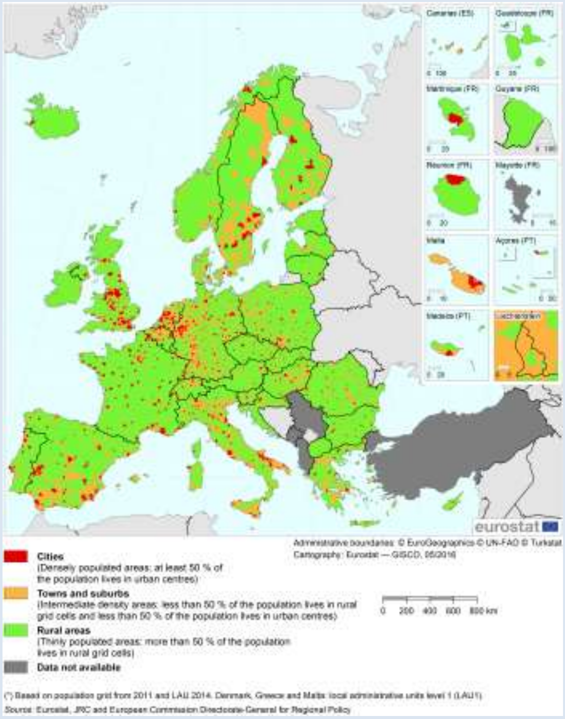
At the European level, the Regulation (EU) 2017/2391 of the European Parliament and of the Council of 12 December 2017, amending Regulation (EC) No 1059/2003, defines the territorial typologies (TERCET) to be used and published by the Commission (Eurostat), including typologies composed of territorial units at the levels of NUTS, LAU and grid cells. The **grid-based typology** (1 km^2) defines 'urban centres', 'urban clusters' and 'rural grid cells' [Figure 2]. At **LAU level** the following territorial typologies have been adopted: a) the *degree of urbanisation* (DEGURBA), which identifies 'Cities' or 'Densely populated areas', 'Towns and suburbs' or 'Intermediate density areas' and 'Rural areas' or 'Thinly populated areas' [Figure 3]; b) *functional urban areas* (FUA), which defines 'Cities' plus their 'Commuting zones'; and c) the *coastal areas*, which distinguishes between 'Coastal areas' and 'Non-coastal areas'. At **NUTS 3** level the following typologies have been approved: a) the *urban-rural typology*, which identifies 'Predominantly urban regions', 'Intermediate regions' and 'Predominantly rural regions'; b) the *metropolitan typology*, which defines 'Metropolitan regions' and 'Non-metropolitan regions'; and c) the *coastal typology*, which defines distinguishes between 'Coastal regions' and 'Non-coastal regions'.

Figure 2 - Grid-based typology (1km^2)



Source: Eurostat, JRC, EFGS, REGIO-GIS
(Eurostat - Statistics Explained – Territorial Typologies)

Figure 3 - Degree of urbanization for LAU level 2



Source: Eurostat, JRC and European Commission Directorate-General for Regional Policy (Eurostat - Statistics Explained – Territorial Typologies)

- At the data source level

Computation of built-up areas can rely on existing geospatial datasets or on procedures based on open source satellite data processing, which can be made available for countries to use. Satellite imagery data can be used to identify the areas to be considered as built-up areas. Additionally, at the European level, EU Copernicus Imperviousness HRL (20m) and CORINE Land Cover Map (CLC) could be possible data sources for built-up areas identification, but in the case of CORINE CLC spatial resolution is relatively coarse (25 ha). At country level, specific national products with higher resolution could also be used as data sources (e.g., in the case of Portugal, the Land Use and Land Cover Map (COS) which has a spatial resolution of 1 ha).

Therefore, regarding data sources: i) at the global level, the GHSL - Global Human Settlement Layer should be taken as a ready to use product and/or ESA Land Cover CCI but, additionally, special attention should be given to stimulate the European remote sensing derived products initiatives worldwide (Copernicus HRL: Imperviousness and CORINE Land Cover) and to encourage national initiatives on high quality land cover maps and urban cadastre; and ii) at the European level, for the sake of comparability, Copernicus HRL Imperviousness and CORINE Land Cover should be taken as main references, but national initiatives on high quality land cover maps and urban cadastre systems should be encouraged.

- At the geospatial processing level

Particular specifications are needed for geospatial processing, namely the spatial resolution for input and output geospatial data processing needs to be better identified.

- At the algorithm level

The metadata should be clear regarding the time intervals for the measurement of population and consumption areas although these may be depending on data availability.

Additionally, with the proposed global indicator computation it may be difficult to capture the dynamics of cities with negative or zero population growth; or cities that due to severe disaster have lost part of their territories. To face this challenge, JRC has developed a tool to calculate the indicator 11.3.1 based on a proxy of Land Use Efficiency (LUE) with the Global Human Settlement Layer [see Box 1]. JRC tool proposes to adapt the formulation of the Land Use Efficiency indicator in order to measure the change rate of the built-up area per capita (Corbane *et al.,* 2016[1]):

$$Idx_t = \frac{Y_t - Y_{t+n}}{Y_t}$$

Where: $Y_t = BU_t / POP_t$; BU_t = built-up surface at t and POP_t = population at t.

The indicator can be estimated at different time intervals upon the availability of observations. In order to ensure the comparability of the results at different times, it is recommended to normalise the values to obtain the variation a 10-year average change which divides the indicator by n (the number of years that separate the observations) and then multiply by 10. The formula of the normalised indicator is:

$$Idx_t = \frac{Y_t - Y_{t+n}}{Y_t} * \frac{10}{n}$$

A script that can be installed in the toolbox of Quantum GIS (QGIS) has been made available.

[1] Corbane, C. *et al.* (2016). Assessment of Land Use Efficiency using GHSL derived indicators. *Atlas of the Human Planet 2016*. Publications Office of the European Union.

At the ***EU level***, the EU-SDG indicator set defined by Eurostat has not included an indicator that has a direct correspondence with the one defined at the global level. It included, however, until 2017, two similar indicators: *15.21 Artificial land cover per capita* and *15.24 Change in artificial land cover per year*. These two EU SDG indicators are based on the LUCAS survey and, hence, no spatial calculations and data analysis at raster or vector model are performed. Data is disseminated at NUTS 2 level, every three years. For the purpose of these two indicators, the work of the task team (ESTAT, GROW, EEA and JRC) on remote sensing for statistics to study the most appropriated data sources for EU SDG indicators on land use and land sealing would be relevant. The 2018 review of the EU-SDG indicator set resulted in the exclusion of indicator *Change in artificial land cover per year* from the indicator set.

In fact, other data sources could be used within the European context and segmentation of information by degree of urbanization and other typologies have been put forward by other departments of the Commission. The Copernicus Imperviousness HRL could be a potential data source. It is open and free and it captures the spatial distribution of artificially sealed areas, including the level of sealing of the soil *per* area unit. This includes road infrastructures and all other sealed surface. This data source is an operational product which already provides a time series spanning from 2006, being updated every three year cycles. It has a minimum mapping unit (MMU) of 20m (10 m from 2018 reference year) and uses Sentinel 2 from 2015 reference year. Another possibility is the CORINE Land Cover data source, which is also open and free, but spatial resolution is relatively coarse (25ha). Nevertheless, the next generation of CLC, CLC+, when available and starting with reference year 2018, will provide features with a predefined MMU of 0.5 ha. On the other hand, an alternative data source could be a very detailed cadastre that not only contains the boundaries of all land parcels, but also their land use and the size and shape of buildings, as well as spatial information on infrastructures used for transport. Geospatial processing and analysis would rely on the classification of land parcels based on a mapping of land use categories used in the cadastre that fall under the scope of the concept of artificial land. Cadastral parcels and transport network are available via the INSPIRE geoportal, but not necessarily as open data in all countries and building data is not yet available in all Member States as Annex III Theme.

At the ***national level***, from the cases analysed (Finland, Ireland, Italy and Portugal), it is possible to identify that this indicator has not been calculated, disseminated or reported by countries. National cases have identified the NSI as the agency responsible for the indicator and that the indicator would require specific articulation between NSI and NMCA. The national analysis of this indicator identified the relevance of data combination and of geospatial analysis techniques to calculate this indicator. Geospatial data and standards workflows could be made available at the European level to be used at national level and European procedures and methodologies could be considered as a global reference (e.g., JRC toolbox on LUE).

Cases have identified mainly national data sources available for both components (land consumption and population growth) and national territorial typologies on cities and urban areas. Data on population is produced regularly at national level, even if with limitations in terms of geospatial population datasets at a very detailed geographical level for more than one reference point in time, but data for land consumption is not produced so regularly. It was also pointed out that other data sources could be used to compute this indicator at national level. For example the case of

Italy, besides the national thematic layer of urban cadastre, points out global data from Sentinel-2 satellite images and European thematic layers (Copernicus HRL Imperviousness, CORINE Land Cover).

In the case of Ireland and Portugal, national geospatial data sources have been identified to measure land consumption, defined as ‘the expansion of built-up area’. In the case of Ireland, data relating to changes in built-up areas can be obtained from the PRIME2 database, maintained by Ordnance Survey Ireland (Irish NMCA) and corresponding to the central database of spatial information. The central premise behind PRIME2 is to have topologically consistent polygons that cover the surface of Ireland. These polygons are grouped into five broad categories: way, water, vegetation, artificial and exposed (non-vegetative ground such as sand and mud). For the built-up areas, artificial and way objects can be considered to be in scope. Artificial objects represent man-made ground cover such as concrete, tarmacadam, gravel, sloping masonry, rail bed, among others including gardens. Way objects in the PRIME2 database also represent all drivable and walkable roads and paths from motorways down to sidewalks. As for the case of Portugal, the indicator can be extracted from the Land Use and Land Cover Map (COS), which corresponds to a national product under the responsibility of the Directorate-General for Territory (Portuguese NMCA). Data series are available for four reference years - COS 1995, COS 2007, COS 2010 and COS 2015 - and corresponds to polygonal maps that represent homogenous land use/cover units. COS is based on a vector data model with a reference mapping unit corresponds to 1 hectare and a hierarchical system of 5-level classes. COS 2015 has a simplified nomenclature of 48 classes, which is compatible with previous editions at the first level. The built-up area concept is set to correspond to megaclass 1 of COS nomenclature - "artificial land", excluding the class 133 corresponding to “areas under construction” [see Box 3].

Box 3 - Calculation of the indicator based on the Land Use and Land Cover Map (COS): example for Portugal

A joint exercise between Statistics Portugal and the Directorate-General for Territory (the Portuguese NMCA) was carried out to calculate this indicator at municipality level for Portugal Mainland for the period 2010-2015. Two approaches were developed based on the following data and formulas:

- Data from the COS 2010 and COS 2015 with the selection of the class on 'artificial land' (megaclass 1), excluding the 'areas under construction';
- Data on population for the reference years of 2010 and 2015 based on annual resident population estimates calculated at municipality level and disseminated annually by Statistics Portugal.

The indicator was calculated based on the formula of land consumption rate to population growth rate (LCRPGR) as proposed by the global metadata and the alternative LUE formula proposed by JRC.

i) A more direct approach based on data on artificial land and population estimates at municipality level

Based on the data from COS 2010 and COS 2015, the area of artificial land occupied in each municipality was extracted based on a common territorial delimitation of municipalities as defined by the Official Administrative Map of Portugal. The available data on annual resident population estimates at municipality level was then directly used to calculate the indicator based on the two formulas – the LCRPGR and the LUE (JRC). The results for Portugal's mainland for the period 2010-2015 were -1.22 and -0.10 based on the LCRPGR and LUE formula, respectively. The use of the JRC formula allowed dealing with those situations with zero growth and, thus, provided consistent results for the different municipalities.

ii) A second approach based on the intelligent dasymetric mapping method

The intelligent dasymetric mapping corresponds to a downscaling methodology that allows the transference of attributes between areas of the same special domain, with the use of auxiliary information. This method has been recently used for the distribution of population based on information on land cover / land use (e.g., Silva *et al.*, 2013; Freire *et al.* 2016). This method was applied using the toolbox available for ArcGIS 10.3 (provided by the Environmental Protection Agency). The auxiliary information used to downscale the population from the municipality level to the polygon level was based on COS data with specific allocation rules being defined associated to each category of land use / land cover data. The results for Portugal's mainland for the period 2010-2015 based on this approach were also -1.22 and -0.10 for the LCRPGR and LUE formulas, respectively, and small differences were registered for the aggregation at municipality level – for the LCRPGR formula differences between the two approaches varied between -2.6 and 0.006 and for the LUE were very residual. The distribution of the results was, therefore, similar for both approaches [see Figure 4 and Figure 5].

Figure 4 – LCRPGR 2010-2015

Figure 5 – LUE 2010-2015

Source: Directorate-General for Territory, Land Use Land Cover Map (COS 2010 and 2015), Statistics Portugal, Annual Resident Population Estimates.